

**UNIVERSIDAD
DE ANTIOQUIA**
1803



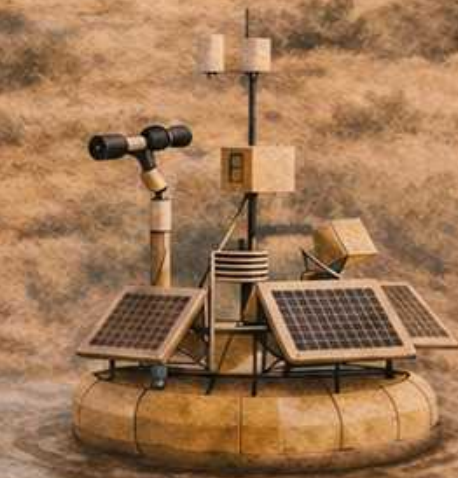
GeoLimna



Seminario GeoLimna–GEPAR

— *Volumen 3* —

Investigaciones 2024–2025



Coordinación

Juan Pablo Serna | David Fernández
Coordinadores de GeoLimna y GEPAR

Colección Seminario GeoLimna–GEPAR
Volumen 3. Investigaciones 2024–2025

© 2026 Universidad de Antioquia

© 2026 Facultad de Ingeniería, Universidad de Antioquia

© 2026 Grupos de investigación GeoLimna y GEPAR

ISSN Digital: *(por asignar)*

Queda hecho el depósito que ordena la Ley

(Ley 44 de 1993 / Decreto 460 de 1995)

Dirección editorial: Juan Pablo Serna y David Fernández

Dirección de colección: *Juan Pablo Serna y David Fernández*

Diseño de cubierta y diagramación: *Juan Pablo Serna y David Fernández*

Universidad de Antioquia – Facultad de Ingeniería

Calle 67 # 53–108

Medellín, Colombia

Código Postal: 050010

Correo electrónico: *jpablo.serna@udea.edu.co; david.fernandez@udea.edu.co*

Sitio web: *https://seminariogg.github.io/Seminario-Geolimna-GEPAR/*

No está permitida la reproducción total o parcial de este libro ni su tratamiento informático, ni la transmisión de ninguna forma o por cualquier medio ya sea electrónico, mecánico, por fotocopia, por registro u otros métodos, sin el permiso previo y por escrito de los titulares de derechos.

Contenido

Prefacio	IV
1 Detección asistida de puntos cefalométricos	1
2 Monitoreo remoto de variables ambientales	14
3 Automatización de la alimentación y clasificación de esquejes	28
4 Sistema móvil de visión artificial para control de calidad de ramos	44
5 Aplicación móvil para gestión y trazabilidad de videos	58
6 Conteo automatizado de plántulas en cama	74
7 Reconocimiento automático de placas vehiculares	88
8 Estimación automática de edad ósea	102
9 Detección y conteo automatizado de microalgas verdes	116
10 Estimación forense de edad mediante radiografías panorámicas	130
11 Reconocimiento móvil de placas vehiculares	146
12 Análisis multitemporal de coberturas terrestres	160
13 Detección automatizada de irregularidades en pavimento	178

Comité Académico

Facultad de Ingeniería - Universidad de Antioquia

Coordinadores

Juan Pablo Serna López

Profesor asociado, Departamento de Ingeniería Ambiental y Sanitaria.

`jpablo.serna@udea.edu.co`

David Fernández Mc Cann

Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones.

`david.fernandez@udea.edu.co`

Integrantes

Néstor Jaime Aguirre Ramírez

Profesor titular, Departamento de Ingeniería Ambiental y Sanitaria.

`nestor.aguirre@udea.edu.co`

Fabio de Jesús Vélez Macías

Profesor titular, Departamento de Ingeniería Ambiental y Sanitaria.

`fabio.velez@udea.edu.co`

Elkin Libardo Ríos Ortiz

Profesor titular, Departamento de Ingeniería Industrial. `elkin.rios@udea.edu.co`

Juan Pablo Urrea Duque

Profesor asociado, Departamento de Ingeniería Electrónica y Telecomunicaciones.

`juan.urrea@udea.edu.co`

Yarin Tatiana Puerta Quintana

Profesora asistente, Departamento de Ingeniería Ambiental y Sanitaria.

`tatiana.puerta@udea.edu.co`

Prefacio

Esta nueva entrega del *Seminario GeoLimna–GEPAR* prolonga una línea de trabajo consolidada como práctica académica, técnica e institucional. Si los fascículos anteriores permitieron organizar desarrollos iniciales, prototipos y experiencias de investigación aplicada, el presente tomo recoge una etapa de mayor madurez: proyectos con metodologías de validación, registros de software, manuales técnicos, interacción con usuarios reales y una intención cada vez más explícita de transferencia, uso y continuidad.

El fascículo *Investigaciones 2024–2025* muestra una convergencia amplia entre GeoLimna y GEPAR. En sus capítulos aparecen problemas ambientales, biomédicos, productivos, forenses, agrícolas y de infraestructura vial abordados desde una lógica común: observar un fenómeno, definir variables relevantes, construir un método, implementar una herramienta y contrastar los resultados con evidencia. La pregunta ambiental, biológica, médica o productiva no se resuelve al margen de la ingeniería, y la ingeniería tampoco se agota en el artefacto.

En este volumen se reúnen desarrollos relacionados con detección asistida de puntos cefalométricos, monitoreo remoto de variables ambientales, automatización en floricultura, control de calidad de ramos de flores, trazabilidad de videos para seguimiento de plántulas, conteo automatizado de plantas en cama, reconocimiento de placas vehiculares, estimación de edad ósea, análisis de microalgas, estimación forense de edad y análisis multitemporal de coberturas terrestres en zonas de minería aluvial. La diversidad temática no dispersa el sentido del fascículo; muestra una base instalada que se expresa en distintos dominios de aplicación.

Este libro conserva la filosofía editorial de la colección: transformar exposiciones, desarrollos, manuales, registros, tesis, artículos y resultados técnicos en una memoria estructurada. Los capítulos se presentan como reportes académicos de investigación y desarrollo; su valor reside en dejar trazabilidad de las decisiones, los criterios técnicos, los resultados alcanzados, sus límites y sus posibilidades de continuidad.

Conviene subrayar que esta compilación se sitúa en un punto intermedio especialmente

fértil del trabajo universitario. No todos los capítulos equivalen a un artículo científico tradicional ni pretenden reemplazarlo. En muchos casos, el seminario cristaliza conocimiento técnico que, de otro modo, quedaría disperso entre presentaciones, manuales, registros de software, tesis, informes, prototipos y documentos internos.

En el trasfondo persiste una convicción que atraviesa toda la colección: los problemas reales necesitan sistemas de observación cada vez más finos, y esos sistemas exigen ingeniería. A su vez, la ingeniería necesita horizonte, contexto y problema situado. Por eso en estas páginas aparecen sensores, algoritmos, instrumentación, procesamiento de imágenes, aprendizaje profundo, análisis de datos, automatización, interfaces, plataformas computacionales y teledetección, siempre ligados al propósito de comprender mejor un fenómeno, intervenir con mayor precisión un proceso o reducir la distancia entre la realidad observada y la decisión que debe tomarse.

La etapa 2024–2025 permite advertir un desplazamiento importante. Varios desarrollos ya no se entienden únicamente como pruebas de concepto, sino como piezas de una capacidad acumulativa. Los registros ante la Dirección Nacional de Derecho de Autor, los manuales, las metodologías de captura, las interfaces de operación, las bases de datos, los protocolos de validación y las articulaciones con docencia, investigación, extensión y sector productivo sugieren una producción más madura, reutilizable y transferible.

También es importante reconocer el papel formativo de estos desarrollos. En ellos participan estudiantes, profesores, grupos de investigación y aliados externos que aprenden a formular problemas, construir soluciones, documentar procesos, evaluar resultados y defender técnicamente una propuesta. Cada capítulo no solo describe un producto o una metodología, sino una experiencia de construcción de capacidades en ingeniería, ciencias ambientales, visión artificial, procesamiento de imágenes, inteligencia artificial y análisis aplicado.

Hacia adentro, el libro busca servir como memoria organizada de una etapa reciente del seminario, para que nuevas cohortes encuentren un punto de partida más alto. Hacia afuera, busca mostrar que la cooperación entre grupos, sostenida en trabajo serio y problemas bien elegidos, puede traducirse en soluciones con valor académico, experimental, institucional y aplicado. Si este tomo preserva esa lógica de construcción compartida y la hace legible para otros, habrá cumplido plenamente su función.

Grupos GeoLimna y GEPAR
Universidad de Antioquia
Volumen 3. Investigaciones 2024–2025

Capítulo 1

Detección asistida de puntos cefalométricos y cálculo automatizado del análisis de Bimler en radiografías laterales del cráneo

Autores

Sebastián González Herrera¹; David Stephen Fernández Mc Cann²; Elizabeth Llano Sánchez³

Resumen

El análisis cefalométrico constituye una herramienta central en el diagnóstico craneofacial, la planificación terapéutica y el seguimiento del crecimiento maxilofacial. Sin embargo, su aplicación cotidiana continúa dependiendo en buena medida de la localización manual de puntos anatómicos sobre radiografías laterales del cráneo, una tarea que exige experiencia clínica, demanda tiempo y presenta variabilidad entre observadores. En este capítulo se presenta un desarrollo orientado a automatizar, de manera asistida, el análisis cefalométrico simplificado de Bimler mediante la integración de visión por computador, aprendizaje profundo e interacción experta. La propuesta combina detección automática

¹Ingeniero electrónico. Universidad de Antioquia. Grupo de investigación GEPAR.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Profesora, Facultad de Odontología, Universidad de Antioquia.

inicial de puntos cefalométricos, interfaz de ajuste manual, cálculo geométrico de factores y medidas diagnósticas, y generación estructurada de reportes. El sistema fue implementado como aplicación de escritorio en Java e incorpora un modelo de inferencia basado en YOLOv8, integrado mediante una API, para proponer la localización aproximada de los puntos de referencia. Posteriormente, el especialista valida y corrige dicha propuesta antes de ejecutar el análisis completo. Este enfoque híbrido conserva la velocidad de la inteligencia artificial y la rigurosidad del juicio clínico. Los resultados reportados muestran la viabilidad de la metodología: el sistema logró identificar 19 de los 20 puntos cefalométricos de referencia y alcanzó una tasa de detección exitosa del 80% para un umbral de 2 mm. En conjunto, estos hallazgos indican que la automatización asistida puede reducir tiempos operativos, ordenar el flujo de trabajo y favorecer diagnósticos más reproducibles, sin desplazar el papel central del especialista [1, 2, 3].

Palabras clave: cefalometría; Bimler; visión por computador; aprendizaje profundo; radiografía lateral del cráneo; apoyo diagnóstico.

1.1. Introducción

La cefalometría traduce la morfología craneofacial a relaciones geométricas medibles. Sobre una radiografía lateral del cráneo, el clínico localiza puntos de referencia anatómicos y, a partir de ellos, construye factores, ángulos y distancias que sirven para interpretar el perfil facial, la relación sagital entre estructuras y la dirección predominante del crecimiento. Entre los distintos enfoques disponibles, el análisis de Bimler mantiene interés clínico porque no se limita a un conjunto de medidas aisladas, sino que propone una lectura estructural de la arquitectura facial y de su comportamiento funcional [1, 4].

Esa riqueza diagnóstica trae consigo una exigencia metodológica importante: la localización precisa de los puntos cefalométricos. Allí aparece una de las mayores dificultades prácticas del procedimiento. La identificación manual es laboriosa, consume tiempo y puede verse afectada por la superposición de estructuras, el contraste de la imagen y la variabilidad propia del observador. El trabajo que sustenta este desarrollo parte de esa necesidad: construir una herramienta que automatice la detección inicial de los puntos cefalométricos del método de Bimler y que, al mismo tiempo, mantenga abierta la posibilidad de verificación y ajuste por parte del especialista [1].

En este contexto, las técnicas recientes de aprendizaje profundo ofrecen una alternativa pertinente. La propuesta desarrollada emplea un esquema de detección automática inicial mediante YOLOv8, pero no clausura la intervención humana. La inteligencia artificial sugiere una ubicación aproximada y el especialista la valida, corrige y transforma en

análisis clínico útil dentro de una interfaz diseñada para ese propósito. El resultado no es un sustituto del juicio odontológico, sino un instrumento de apoyo más rápido, ordenado y trazable [2, 3, 5].

La Fig. 1.1 presenta la interfaz principal del sistema, desde la cual se accede al entorno de ajuste manual y al panel de análisis cefalométrico.

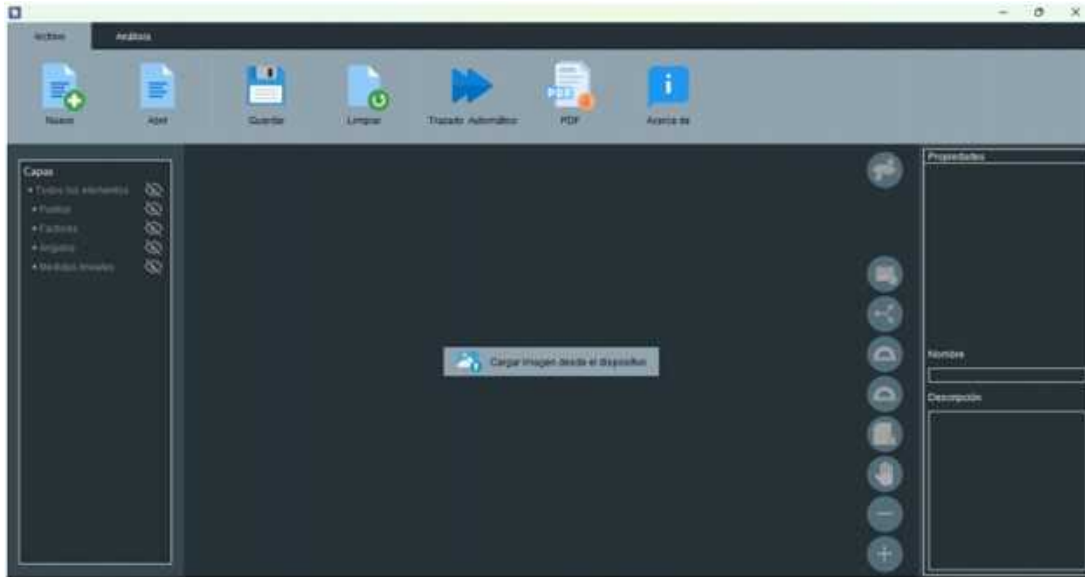


Figura 1.1: Interfaz principal del sistema para navegación entre el panel de ajuste manual y el panel de análisis cefalométrico.

1.2. Planteamiento del problema

Desde una perspectiva clínica, el problema puede describirse así: el análisis cefalométrico de Bimler es valioso, pero su ejecución manual exige tiempo, cuidado geométrico y experiencia especializada. Desde una perspectiva de ingeniería, el problema se descompone en tres etapas: localizar automáticamente los puntos de referencia sobre la radiografía, permitir su corrección experta y transformar esas coordenadas en factores, ángulos, medidas lineales e índices diagnósticos.

Sea una radiografía lateral digital del cráneo representada por la imagen

$$I : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R},$$

y sea el conjunto de puntos cefalométricos de interés

$$\mathcal{P} = \{P_1, P_2, \dots, P_N\},$$

donde, para este caso, $N = 20$. El objetivo del sistema es construir una función de estimación

$$\hat{\mathcal{P}} = F(I),$$

tal que $\hat{\mathcal{P}}$ constituya una localización inicial suficientemente próxima a la referencia clínica para ser corregida y validada por un usuario experto. A partir de esas coordenadas corregidas, el sistema debe calcular automáticamente el conjunto de mediciones

$$\mathcal{M} = \{m_1, m_2, \dots, m_q\},$$

donde \mathcal{M} agrupa factores, ángulos, medidas lineales e índices del análisis de Bimler [1, 3].

La novedad no reside únicamente en aplicar un detector de objetos sobre imágenes médicas. Lo relevante es integrar el detector dentro de un flujo diagnóstico completo. El desarrollo registrado ante la Dirección Nacional de Derecho de Autor se describe como una herramienta de visión por computador para análisis cefalométricos sobre radiografías craneofaciales laterales usando la metodología de Bimler, con detección aproximada por modelo de aprendizaje profundo, panel de ajuste manual e interfaz de visualización de resultados [5].

1.3. Fundamento conceptual

1.3.1. Cefalometría de Bimler

El método de Bimler organiza el diagnóstico a partir de puntos cefalométricos, sistemas de referencia, factores, ángulos y medidas lineales. Su interés clínico proviene de que ofrece simultáneamente una lectura morfológica y una lectura del crecimiento facial. El trabajo de grado que da soporte a este capítulo subraya, además, que aunque existen otros programas para cefalometría, muchas medidas de interés no suelen estar integradas, lo que obliga a procedimientos manuales adicionales. De allí que automatizar específicamente el análisis simplificado de Bimler represente un aporte concreto y no una simple reimplementación informática [1, 4].

1.3.2. Aprendizaje profundo para localización de puntos

La arquitectura elegida fue YOLOv8, implementada mediante una API de inferencia. En el manual técnico se indica que dicha API entrega predicciones con coordenadas centrales, dimensiones, clase y confianza, información suficiente para reconstruir la posición aproximada de los puntos cefalométricos sobre la imagen. Esta decisión tecnológica es coherente con un problema donde se requiere rapidez de respuesta y una ubicación inicial que luego pueda ser afinada por el usuario [3, 6, 7].

1.3.3. Interacción humano-máquina

El sistema no asume que la predicción automática sea definitiva. El manual de usuario establece que el trazado automático produce una localización aproximada de los puntos de referencia y que luego debe ser ajustada con precisión por el profesional. Esta decisión resulta especialmente adecuada en aplicaciones biomédicas: la inteligencia artificial acelera el trabajo, pero la confiabilidad clínica se preserva mediante validación experta [2].

La Fig. 1.2 muestra el panel de ajuste manual, donde el especialista revisa, desplaza y valida la localización de los puntos cefalométricos antes del cálculo diagnóstico.

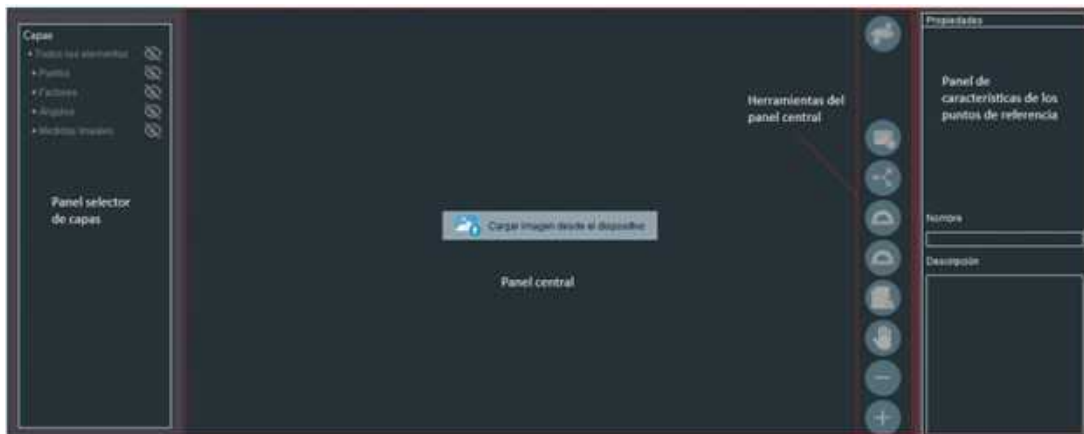


Figura 1.2: Panel de ajuste manual con herramientas geométricas, control de capas y zona de trabajo para validación experta de puntos cefalométricos.

1.4. Metodología

La metodología general se organizó en cinco fases: recopilación del material radiográfico, etiquetado de puntos de referencia, creación del conjunto de datos, entrenamiento y

selección del modelo de detección, y desarrollo del sistema de interacción y análisis [1].

1.4.1. Conformación del conjunto de imágenes

El conjunto de datos estuvo formado por 148 radiografías cefalométricas, distribuidas en 120 imágenes para entrenamiento, 19 para prueba y 9 para validación. Las imágenes originales tenían tamaño de 1935×2400 píxeles, con tamaño de píxel de $0,1 \times 0,1$ mm. Antes del entrenamiento se realizó ajuste de contraste, aumento de datos por reflejo y redimensionamiento a 640×640 píxeles. Estos pasos permitieron uniformar el flujo de inferencia y mejorar la capacidad de generalización del modelo [1].

1.4.2. Etiquetado y referencia experta

Las radiografías fueron marcadas manualmente por un especialista, identificando los 20 puntos cefalométricos relevantes del método. Esa marcación experta constituye la referencia contra la cual se evalúa el desempeño del sistema. Esto es importante porque las métricas de localización no comparan el modelo contra una noción abstracta, sino contra una localización anatómica clínicamente validada [1].

1.4.3. Modelo de inferencia

La salida del detector puede representarse como un conjunto de predicciones

$$\mathcal{D} = \{(x_i, y_i, w_i, h_i, c_i, s_i)\}_{i=1}^M,$$

donde x_i y y_i son las coordenadas del centro de la predicción i , w_i y h_i son el ancho y el alto del recuadro delimitador, c_i es la clase asociada y s_i el nivel de confianza. Esta estructura coincide con la documentación resumida en el manual técnico para la API de inferencia empleada [3].

1.4.4. Cálculo geométrico del análisis

Una vez establecidos los puntos de referencia, el sistema calcula distancias y ángulos que alimentan las tablas diagnósticas. Si dos puntos se representan como $P_1(x_1, y_1)$ y $P_2(x_2, y_2)$, la distancia euclidiana entre ellos se calcula mediante

$$d(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (1.1)$$

Si dos rectas se expresan por medio de vectores \mathbf{u} y \mathbf{v} , el ángulo entre ellas puede estimarse como

$$\theta = \cos^{-1} \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \right). \quad (1.2)$$

El manual técnico documenta funciones para el cálculo de distancias, pendientes, intersecciones y ángulos entre líneas, que sirven como base matemática del análisis automatizado [3].

La Fig. 1.3 presenta un ejemplo de localización de puntos cefalométricos sobre una radiografía lateral del cráneo.



Figura 1.3: Ejemplo de localización de puntos cefalométricos sobre la radiografía lateral del cráneo, base para el cálculo automático del análisis.

1.5. Arquitectura funcional del sistema

El sistema desarrollado puede entenderse como la integración de cuatro módulos funcionales. El primero administra proyectos, radiografías y persistencia de la información. El segundo procesa la imagen y solicita la inferencia del modelo para obtener los puntos aproximados. El tercero ofrece un entorno de ajuste manual con visualización multicapa,

herramientas geométricas y control de escala. El cuarto organiza y exporta los resultados diagnósticos [2, 3].

La aplicación fue construida como una herramienta de escritorio en Java 8. Este entorno permitió desarrollar una interfaz gráfica robusta para interacción local, mientras que la inferencia automática se integró por medio de una API externa. Esta separación entre interfaz y motor de predicción favorece la modularidad y la posible actualización futura de los modelos [3].

Desde el punto de vista estructural, el sistema puede representarse mediante la tupla

$$A = (U, I, V, R),$$

donde U representa la interfaz de usuario, I el módulo de inferencia automática, V el módulo de validación y edición experta, y R el módulo de resultados y reportes. Esta representación resume el carácter híbrido del desarrollo: un sistema que combina automatización inicial, control humano y persistencia documental del análisis.

La Fig. 1.4 muestra la interfaz de resultados, en la cual se organizan los factores, medidas angulares, medidas lineales e índices del análisis cefalométrico de Bimler.



Figura 1.4: Interfaz de resultados con factores, medidas angulares, medidas lineales e índices del análisis cefalométrico de Bimler.

1.6. Evaluación del desempeño

Para evaluar el sistema se emplearon tres métricas: error radial medio, desviación estándar y tasa de detección exitosa. Estas métricas, conocidas respectivamente como MRE, SD y SDR, cuantifican el error promedio de localización, la dispersión de ese error y el porcentaje de puntos correctamente localizados dentro de un umbral dado [1].

Sea (x_i, y_i) la ubicación de referencia del punto i y (\hat{x}_i, \hat{y}_i) su predicción. El error radial individual se define como

$$R_i = \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}. \quad (1.3)$$

A partir de allí, el error radial medio se obtiene como

$$\text{MRE} = \frac{1}{N} \sum_{i=1}^N R_i, \quad (1.4)$$

y la desviación estándar como

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^N (R_i - \text{MRE})^2}{N - 1}}. \quad (1.5)$$

La tasa de detección exitosa se expresa como

$$\text{SDR} = \frac{N_{\text{exitosos}}}{N_{\text{total}}} \times 100 \%. \quad (1.6)$$

Estas expresiones formalizan la precisión geométrica del sistema y permiten traducir el desempeño del modelo a términos clínicamente interpretables.

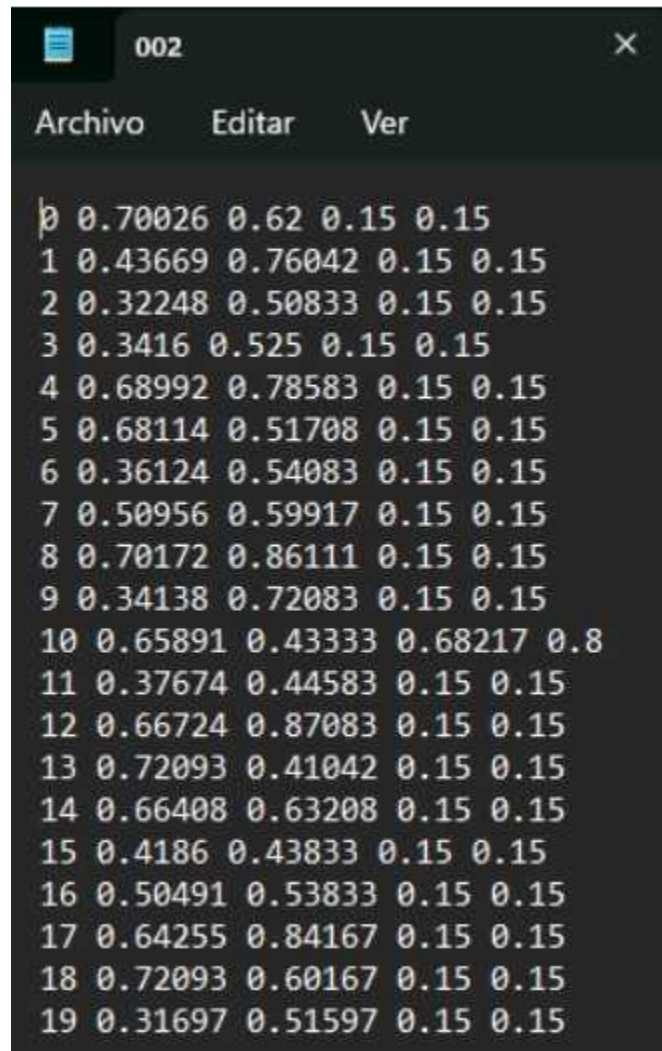
1.7. Resultados

Los resultados reportados por el proyecto muestran que el sistema automatizado logró identificar 19 de los 20 puntos cefalométricos de referencia y alcanzó una tasa de detección del 80 % para un umbral de 2 mm [1]. Este resultado sugiere que la mayor parte del trazado inicial puede automatizarse con un nivel de precisión suficiente para servir como apoyo al especialista.

En el análisis de errores radiales, el promedio de MRE fue de 1.487832 mm y la desviación estándar media fue de 0.974205 mm. La mayoría de los puntos se ubicaron por debajo de 2 mm de error, aunque el punto *Cm* presentó una dificultad significativamente mayor, con un MRE de 6.144785 mm y una SD de 4.209443 mm. El propio trabajo atribuye esta dificultad a que su ubicación depende no solo de una estructura anatómica puntual, sino también de la curva oclusal y del punto capitulare, lo que introduce complejidad geométrica adicional [1].

Además del desempeño cuantitativo, el resultado más importante desde el punto de vista aplicado fue la integración completa del flujo diagnóstico. El sistema no solo predice puntos: también permite revisarlos, corregirlos, almacenar el estado del proyecto, incorporar datos del paciente y emitir un reporte PDF. Esa integración reduce la fragmentación habitual del trabajo cefalométrico y facilita su uso en entornos académicos y clínicos [2, 3].

La Fig. 1.5 muestra una salida documental del sistema, correspondiente al reporte automatizado del análisis.



The image shows a screenshot of a PDF report window titled '002'. The window has a dark background and a menu bar with 'Archivo', 'Editar', and 'Ver' options. The main content is a table of numerical data representing cephalometric analysis results. The table has 20 rows, each starting with a point number from 0 to 19. Each row contains five numerical values. The values for point 10 are notably different from the others, with a final value of 0.8 instead of 0.15.

Point	Value 1	Value 2	Value 3	Value 4	Value 5
0	0.70026	0.62	0.15	0.15	
1	0.43669	0.76042	0.15	0.15	
2	0.32248	0.50833	0.15	0.15	
3	0.3416	0.525	0.15	0.15	
4	0.68992	0.78583	0.15	0.15	
5	0.68114	0.51708	0.15	0.15	
6	0.36124	0.54083	0.15	0.15	
7	0.50956	0.59917	0.15	0.15	
8	0.70172	0.86111	0.15	0.15	
9	0.34138	0.72083	0.15	0.15	
10	0.65891	0.43333	0.68217	0.8	
11	0.37674	0.44583	0.15	0.15	
12	0.66724	0.87083	0.15	0.15	
13	0.72093	0.41042	0.15	0.15	
14	0.66408	0.63208	0.15	0.15	
15	0.4186	0.43833	0.15	0.15	
16	0.50491	0.53833	0.15	0.15	
17	0.64255	0.84167	0.15	0.15	
18	0.72093	0.60167	0.15	0.15	
19	0.31697	0.51597	0.15	0.15	

Figura 1.5: Reporte automatizado del análisis cefalométrico exportado en formato PDF, como salida final estructurada del sistema.

1.8. Discusión

El principal mérito de esta propuesta es haber resuelto el problema de una forma técnicamente realista. En lugar de prometer una automatización cerrada, el desarrollo adopta una estrategia verificable: la detección automática se entiende como una hipótesis inicial de trabajo que el especialista puede validar. Esto vuelve al sistema más defendible desde el punto de vista clínico y más útil en la práctica [2].

También hay un valor metodológico importante. El proyecto instala una plataforma reutilizable para futuros estudios en cefalometría, imágenes biomédicas y apoyo diagnóstico.

La combinación entre modelo de inferencia, herramientas de edición, cálculo geométrico y generación documental deja una base tecnológica que puede crecer con nuevos conjuntos de datos, nuevas poblaciones y nuevas variantes del análisis [1, 5].

1.9. Conclusiones

Se desarrolló una herramienta de apoyo diagnóstico que automatiza de manera asistida el análisis cefalométrico de Bimler sobre radiografías laterales del cráneo. Su aporte central no es solamente localizar puntos de referencia, sino articular esa localización con validación experta, cálculo geométrico y generación de reportes dentro de una misma plataforma [2, 3, 5].

Los resultados muestran que el enfoque es viable y útil. La identificación de 19 de 20 puntos y la SDR del 80% a 2 mm indican que la inteligencia artificial puede asumir buena parte del trabajo repetitivo del trazado inicial, dejando al especialista la verificación fina y la interpretación clínica [1].

En términos más amplios, este desarrollo confirma que la incorporación de aprendizaje profundo en odontología debe orientarse a aumentar la capacidad del profesional, no a reemplazarla. Cuando la automatización está bien diseñada, se convierte en una asistencia técnica útil: reduce tiempos, mejora la organización del proceso y favorece decisiones más reproducibles.

Agradecimientos

Se agradece a la Universidad de Antioquia, a la Facultad de Odontología y a los actores académicos que participaron en el desarrollo, validación y documentación del sistema, así como al entorno institucional que permitió articular inteligencia artificial, procesamiento digital de imágenes y práctica clínica en un mismo proyecto.

Referencias

- [1] S. González Herrera, «DIACEFAL – Algoritmo para el diagnóstico cefalométrico utilizando radiografía lateral del cráneo,» Trabajo de grado, Universidad de Antioquia, Medellín, Colombia, 2024.

- [2] Universidad de Antioquia, *DIACEFAL: Diagnóstico cefalométrico de apoyo utilizando radiografías laterales del cráneo. Manual de Usuario*, Universidad de Antioquia, Medellín, Colombia, 2024.
- [3] Universidad de Antioquia, *DIACEFAL: Diagnóstico cefalométrico de apoyo utilizando radiografías laterales del cráneo. Manual Técnico*, Universidad de Antioquia, Medellín, Colombia, 2024.
- [4] H. P. Bimler, *La cefalometría de Bimler*. Barcelona: Marcombo, 1989.
- [5] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software. DIACEFAL: diagnóstico cefalométrico de apoyo utilizando radiografías laterales del cráneo*, Libro/Tomo/Partida 13-99-467, fecha de registro: 29-ago.-2024, 2024.
- [6] Ultralytics. «Ultralytics YOLO Documentation,» visitado 9 de abr. de 2026. dirección: <https://docs.ultralytics.com/>.
- [7] Roboflow. «Roboflow Documentation,» visitado 9 de abr. de 2026. dirección: <https://docs.roboflow.com/>.

Capítulo 2

Monitoreo remoto de variables ambientales en sistemas leníticos

Autores

Johan David Rojas Martínez¹; David Stephen Fernández Mc Cann²; Fabio de Jesús Vélez Macías³; Néstor Jaime Aguirre Ramírez⁴

Resumen

El monitoreo de variables ambientales en cuerpos de agua leníticos constituye una necesidad técnica y científica para comprender procesos ecológicos, vigilar la calidad del agua y construir series históricas útiles para análisis ambiental. Sin embargo, en escenarios como ciénagas, humedales, lagunas o sistemas de inundación estacional, la medición manual presenta restricciones operativas: depende de la presencia de personal en campo, introduce variabilidad asociada al operador y dificulta la consolidación de registros continuos. En este capítulo se presenta un sistema de monitoreo remoto orientado a registrar variables fisicoquímicas del agua mediante sensores, microcontroladores, comunicación serial, enlace por radiofrecuencia, conectividad WiFi, publicación en una base de datos en tiempo

¹Estudiante de pregrado, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Profesor titular, Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GeoLimna.

⁴Profesor titular, Escuela Ambiental, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GeoLimna.

real y consulta mediante plataforma web y aplicación móvil. La propuesta integra hardware, comunicaciones e infraestructura digital para transformar una medición de campo en un dato disponible para análisis remoto. El desarrollo se articula con actividades de investigación, docencia y extensión de los grupos GeoLimna y GEPAR, y constituye una base tecnológica reutilizable para el seguimiento de sistemas leníticos y otros entornos ambientales de difícil acceso [1], [2], [3].

Palabras clave: monitoreo ambiental; sistemas leníticos; sensores; IoT; LoRa; Firebase; GeoLimna; GEPAR; calidad del agua; variables fisicoquímicas.

2.1. Introducción

El monitoreo de variables ambientales en cuerpos de agua leníticos constituye una necesidad técnica y científica de alta relevancia para la comprensión de procesos ecológicos, la vigilancia de la calidad del agua y la construcción de series históricas útiles para análisis de comportamiento y predicción. Sin embargo, en escenarios reales como ciénagas, humedales, lagunas o sistemas de inundación estacional, la medición manual presenta restricciones operativas importantes: depende de la presencia del personal en campo, introduce variabilidad asociada al operador y dificulta la consolidación de registros continuos en el tiempo.

Frente a esta limitación, una alternativa de ingeniería consiste en desacoplar el punto de medición del punto de consulta. En lugar de entender el dato como un valor aislado obtenido en una visita puntual, el problema se aborda como una cadena completa de adquisición, transmisión, almacenamiento y visualización. En esa lógica, el software asociado al sistema no es el proceso en sí mismo, sino una herramienta que articula la electrónica de adquisición con la infraestructura digital de publicación y consulta remota.

El desarrollo descrito en este capítulo se orienta a resolver ese reto. La solución integra sensores para variables fisicoquímicas del agua, microcontroladores, comunicaciones seriales y por radiofrecuencia, conexión WiFi, publicación en una base de datos en tiempo real, estructuración de históricos en hojas de cálculo y visualización en plataforma web y aplicación móvil. La finalidad no es únicamente automatizar mediciones, sino crear una infraestructura de observación ambiental continua, útil para proyectos de investigación, extensión y formación académica.

La documentación institucional asociada al desarrollo muestra que el sistema fue concebido para registrar variables como pH, potencial redox, oxígeno disuelto, conductividad eléctrica y temperatura, con publicación en la nube y acceso desde herramientas web y

móviles [1], [2], [4]. A ello se suma un uso amplio en las actividades académicas y científicas de los grupos involucrados, lo cual confirma que el desarrollo trasciende la prueba de laboratorio y se proyecta como un recurso de trabajo para investigación aplicada [3].

Desde una perspectiva de ingeniería, el valor de esta propuesta reside en la integración sistémica. No se trata de un único dispositivo ni de una aplicación aislada, sino de una arquitectura modular en la que cada bloque cumple una función específica dentro del tránsito de la información: medir, codificar, transmitir, almacenar, estructurar y visualizar. Esa visión integral es la que se expone en las secciones siguientes.

2.2. Problema de ingeniería y propósito del sistema

La observación continua de variables ambientales en medios acuáticos remotos impone una dificultad fundamental: el lugar donde ocurre el fenómeno de interés no coincide con el lugar donde conviene procesar, consultar o resguardar la información. Esta separación obliga a diseñar una cadena de transformación del dato que garantice continuidad, confiabilidad y disponibilidad.

Si se denota por $x_i(t)$ la magnitud física de la variable ambiental i en el instante t , el objetivo del sistema puede expresarse como la implementación de una transformación general:

$$\mathcal{F} : x_i(t) \longrightarrow d_i(t),$$

donde $d_i(t)$ es el dato digital final disponible para consulta remota. Esta transformación no ocurre en un solo paso, sino a través de varios subsistemas: sensado, lectura electrónica, concatenación de tramas, transmisión local serial, transmisión remota por radiofrecuencia, publicación en red y almacenamiento estructurado.

Cuando el proceso se organiza bajo un periodo de muestreo Δt , la observación continua se convierte en una serie discreta:

$$x_i[k] = x_i(k\Delta t), \quad k = 0, 1, 2, \dots, N,$$

donde k es el índice de muestreo y N el número total de registros generados durante el intervalo de observación. Esta formulación pone en evidencia que el problema no se limita a adquirir un valor, sino a construir una secuencia temporal interpretable.

En el caso del desarrollo aquí descrito, el sistema se planteó para operar inicialmente en un contexto como la Ciénaga de Ayapel, aunque con posibilidad de extensión a otros entornos con características similares [5], [6]. La motivación de fondo es disponer de datos en tiempo real y, simultáneamente, consolidar una base histórica que permita estudiar tendencias, detectar cambios y, en el futuro, alimentar modelos predictivos.

2.3. Arquitectura funcional del monitoreo remoto

La arquitectura adoptada se organiza en dos puntos funcionalmente diferenciados. El primero corresponde al punto remoto de medición, ubicado cerca del cuerpo de agua y encargado de adquirir las señales de los sensores. El segundo corresponde al punto fijo o centro de datos, donde la información es recibida, reenrutada hacia la nube y puesta a disposición de plataformas de consulta.

En términos funcionales, el flujo completo puede representarse como:

$$\mathbf{x}(t) \rightarrow \mathbf{s}(t) \rightarrow \mathbf{r}(t) \rightarrow \mathbf{c}(t) \rightarrow \mathbf{h}(t),$$

donde $\mathbf{x}(t)$ es el vector de variables físicas medidas, $\mathbf{s}(t)$ representa la codificación digital local, $\mathbf{r}(t)$ la trama transmitida por radiofrecuencia, $\mathbf{c}(t)$ la información publicada en la nube y $\mathbf{h}(t)$ el histórico estructurado para consulta y análisis.

Para este sistema, el vector de variables ambientales puede escribirse como:

$$\mathbf{x}(t) = \begin{bmatrix} x_{\text{pH}}(t) \\ x_{\text{ORP}}(t) \\ x_{\text{OD}}(t) \\ x_{\text{EC}}(t) \\ x_{\text{T}}(t) \end{bmatrix},$$

donde $x_{\text{pH}}(t)$ corresponde al pH, $x_{\text{ORP}}(t)$ al potencial redox, $x_{\text{OD}}(t)$ al oxígeno disuelto, $x_{\text{EC}}(t)$ a la conductividad eléctrica y $x_{\text{T}}(t)$ a la temperatura.

El punto remoto integra sensores Atlas Scientific, placa Whitebox T2, Arduino Mega 2560 y un módulo Heltec LoRa WiFi configurado como transmisor. El punto fijo incluye un módulo Heltec LoRa receptor y un NodeMCU 1.0, responsable de conectarse a la red WiFi y publicar los datos en la infraestructura digital del sistema. A partir de allí, los valores son enviados a una base de datos en tiempo real, estructurados en una hoja de

cálculo y visualizados tanto en la plataforma web GEOVAM como en una aplicación móvil específica [6], [7].

La ventaja de esta disposición es que desacopla espacialmente la zona de medición de la zona de conectividad a Internet. Esto mejora la factibilidad operativa y permite que el sistema combine trabajo de campo con acceso remoto a la información.

La Fig. 2.1 presenta la arquitectura general del sistema, desde el punto de medición en campo hasta las plataformas de almacenamiento y visualización.

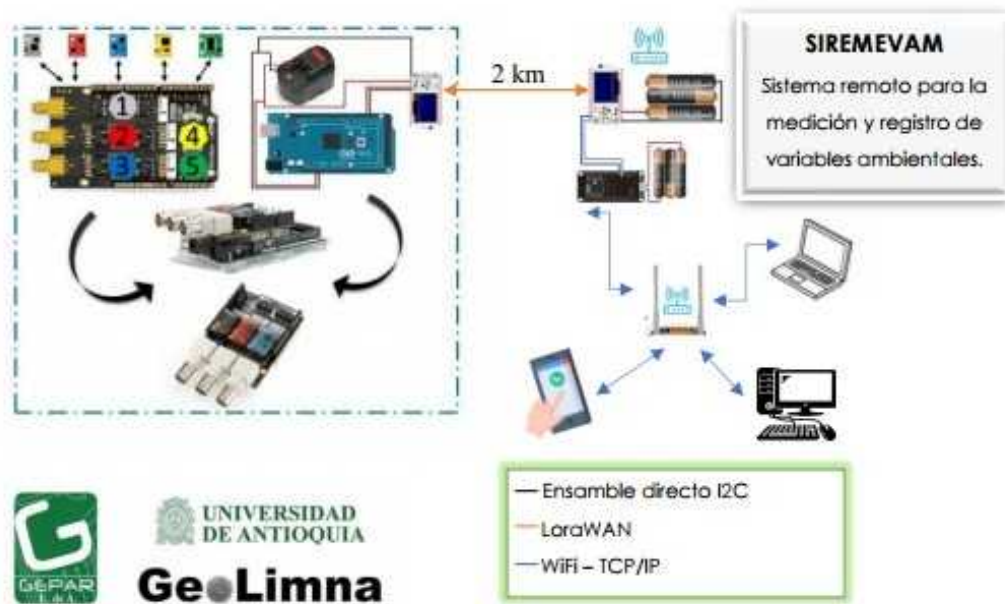


Figura 2.1: Arquitectura general del sistema de monitoreo remoto, con punto de medición en campo, enlace por radiofrecuencia, nodo con conectividad WiFi y plataformas finales de almacenamiento y visualización.

2.4. Adquisición de datos y construcción de la trama digital

La primera etapa operativa consiste en transformar magnitudes fisicoquímicas del entorno en datos digitales. Para ello, el sistema utiliza sensores especializados y un esquema de adquisición basado en el protocolo I2C. Esta decisión técnica simplifica la lectura de múltiples variables sobre una estructura común de comunicación y facilita la integración modular del sistema.

Si se representa por $\mathcal{S}_i(\cdot)$ la función de conversión del sensor asociado a la variable i , la lectura discreta puede expresarse como:

$$y_i[k] = \mathcal{S}_i(x_i(k\Delta t)),$$

donde $y_i[k]$ es el valor digital leído en el instante discreto k . Esta ecuación resume el paso desde la variable física continua hasta su representación numérica procesable por el microcontrolador.

Una vez realizadas las lecturas, el sistema organiza los datos en una cadena etiquetada. Esta decisión es importante porque evita que los valores viajen como números sin contexto. Cada medición queda asociada a una etiqueta que conserva su significado durante toda la cadena de transmisión. Conceptualmente, esta trama puede representarse como:

$$m[k] = \text{RTD}:y_{\text{T}}[k] \parallel \text{ORP}:y_{\text{ORP}}[k] \parallel \\ \text{EC}:y_{\text{EC}}[k] \parallel \text{OD}:y_{\text{OD}}[k] \parallel \text{PH}:y_{\text{PH}}[k],$$

donde \parallel representa la concatenación ordenada de cadenas. Esta formulación es coherente con la lógica implementada en el sistema: cada variable es leída, etiquetada y empacada para su tránsito hacia los módulos posteriores [6].

La rutina de adquisición también incorpora verificación de estado. Si una lectura es válida, se incorpora a la trama; si el sensor no responde adecuadamente, el sistema puede reportar estados equivalentes a fallo, espera o ausencia de dato. Esta capacidad diagnóstica fortalece la robustez del proceso y facilita tareas de mantenimiento y depuración.

La Fig. 2.2 muestra el entorno de adquisición y la construcción de la trama etiquetada antes de su transmisión.



Figura 2.2: Adquisición de datos en el microcontrolador y construcción de la trama etiquetada para su transmisión hacia los módulos de comunicación.

2.5. Transmisión serial local y enlace por radiofrecuencia

Después de construir la trama local, el sistema debe transferirla desde el microcontrolador de adquisición hacia el módulo transmisor y, posteriormente, enviarla hasta el punto fijo mediante radiofrecuencia. Esta etapa combina una comunicación serial local con un enlace inalámbrico de mayor alcance.

Si $m[k]$ es la trama generada por el subsistema de adquisición, el paquete de transmisión puede representarse como:

$$p[k] = \mathcal{P}(m[k]),$$

donde $\mathcal{P}(\cdot)$ es la función de empaquetamiento implementada en el módulo transmisor. El objetivo es convertir la cadena estructurada en un paquete que pueda ser enviado por el canal LoRa.

Desde un punto de vista conceptual, la señal recibida puede modelarse como:

$$\hat{p}[k] = p[k] + n[k],$$

donde $\hat{p}[k]$ es el paquete reconstruido en el receptor y $n[k]$ representa perturbaciones o alteraciones introducidas por el medio. Aunque el sistema real se implementa mediante bibliotecas y parámetros concretos de configuración, esta expresión permite comprender que la integridad del dato depende tanto del empaquetamiento como de la correcta configuración del enlace.

La documentación técnica muestra que el sistema utiliza parámetros compartidos de frecuencia, potencia, ancho de banda y eventos de transmisión y recepción para coordinar el intercambio entre módulos Heltec LoRa [6]. La principal virtud de esta solución es que permite mantener el punto de medición en el entorno acuático mientras el punto con conectividad a Internet se instala en un lugar más accesible y estable.

Esta separación no es un detalle menor. Desde la ingeniería de sistemas embebidos, constituye una forma eficiente de aislar funciones: el nodo remoto se concentra en medir y transmitir; el nodo fijo, en recibir, publicar y estructurar la información.

La Fig. 2.3 presenta la secuencia de transmisión y visualización de datos desde el nodo de adquisición hasta la plataforma de consulta.

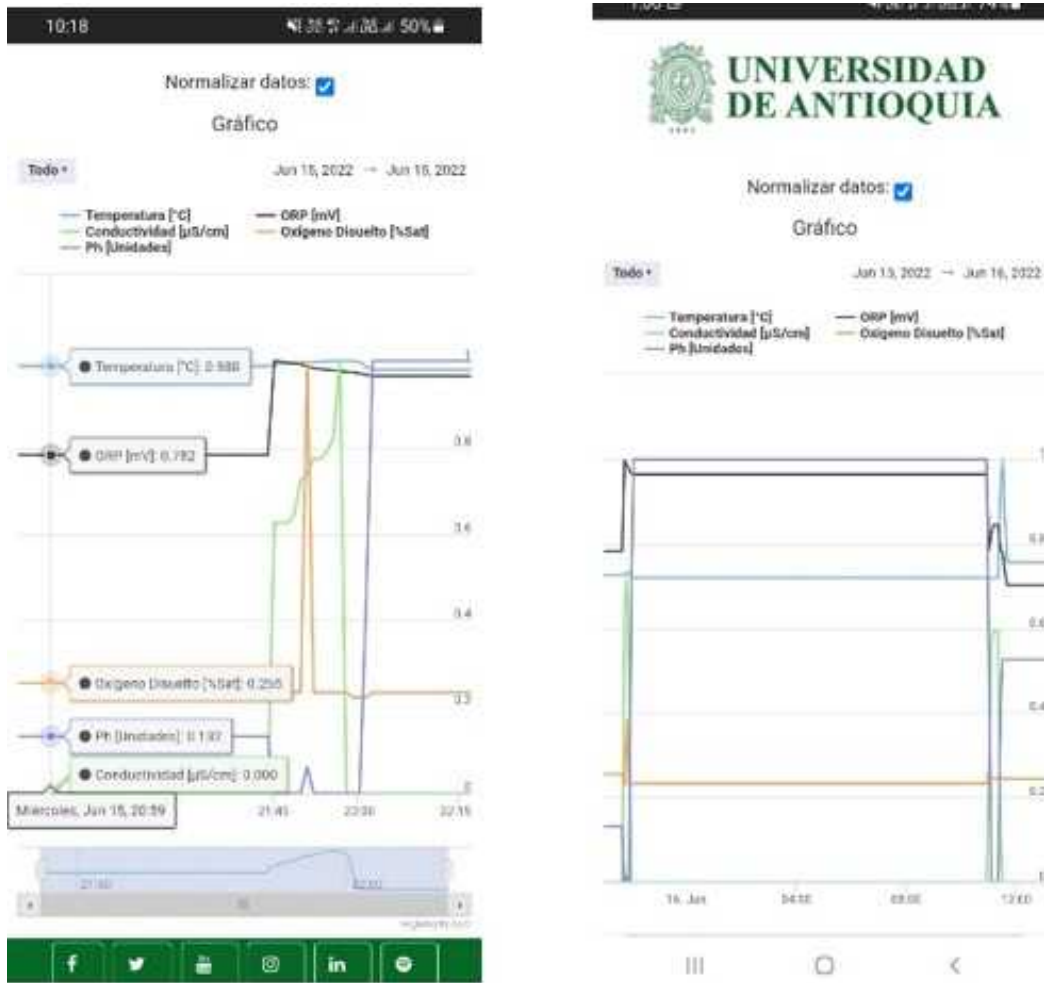


Figura 2.3: Secuencia de transmisión entre el punto de adquisición y el punto fijo mediante comunicación serial local y enlace LoRa, con publicación posterior en la plataforma GEOVAM.

2.6. Publicación en la nube, almacenamiento histórico y visualización

Una vez recibida la trama en el punto fijo, el NodeMCU realiza la conexión a una red WiFi local y publica los datos en una base de datos en tiempo real. Esta operación transforma el dato de campo en un recurso digital distribuido.

Si $\hat{m}[k]$ es la cadena recibida en el nodo con conectividad, el proceso de asignación a rutas de almacenamiento puede idealizarse como:

$$\mathcal{G} : \hat{m}[k] \longrightarrow \left\{ \begin{array}{l} z_{\text{RTD}}[k], z_{\text{ORP}}[k], z_{\text{EC}}[k], \\ z_{\text{OD}}[k], z_{\text{PH}}[k] \end{array} \right\},$$

donde cada $z_i[k]$ corresponde al valor publicado para una variable específica dentro de la estructura de la base de datos. En la implementación concreta, la lógica del sistema identifica las etiquetas dentro de la cadena y dirige cada valor hacia su ruta correspondiente en Firebase [2], [6].

Posteriormente, un script de Google Apps Script extrae esos valores y los organiza en un histórico. Si $h[k]$ representa el registro consolidado en el instante k , entonces:

$$h[k] = (t[k], z_T[k], z_{\text{ORP}}[k], z_{\text{EC}}[k], z_{\text{OD}}[k], z_{\text{PH}}[k]),$$

donde $t[k]$ es la marca temporal asociada al registro. Esta organización permite que la plataforma de visualización consulte tanto valores recientes como secuencias históricas completas.

La estructura histórica no solo cumple una función de almacenamiento. También convierte el sistema en una herramienta analítica. Gracias a ella es posible observar tendencias, detectar comportamientos anómalos y generar insumos para análisis posteriores. En este punto se vuelve clara la diferencia entre el proceso y la herramienta: el proceso es el monitoreo remoto de variables ambientales; el software es la pieza que permite automatizar, gestionar y poner en circulación la información producida por ese proceso.

La Fig. 2.4 muestra la publicación de datos y la consulta de históricos mediante las plataformas del sistema.

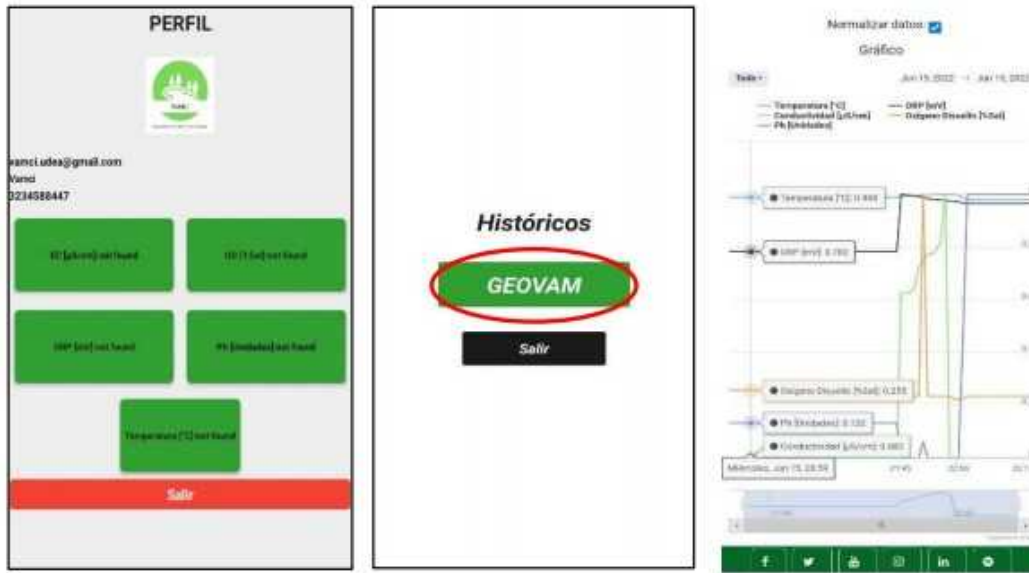


Figura 2.4: Publicación de datos en la infraestructura digital del sistema y organización de históricos para consulta en plataformas web y aplicación móvil.

2.7. Consulta distribuida mediante aplicación móvil y plataforma web

La etapa final del sistema corresponde a la interacción con el usuario. Para ello, el desarrollo integra una aplicación móvil construida en Kodular y una plataforma web de consulta histórica. La aplicación permite registro de usuarios, autenticación, visualización de perfil y acceso a datos publicados, mientras que la plataforma web facilita la exploración de históricos y su representación gráfica [6], [8].

Si se denota por \mathcal{U} el conjunto de usuarios registrados, por $\mathcal{A}(u)$ el proceso de autenticación del usuario u y por $\mathcal{Q}(u, t)$ la operación de consulta en el instante t , la lógica general de acceso puede escribirse como:

$$u \in \mathcal{U} \Rightarrow \mathcal{A}(u) \Rightarrow \mathcal{Q}(u, t).$$

Aunque esta expresión es abstracta, resume con claridad el comportamiento de la interfaz: el usuario se registra, accede al sistema y consulta variables o históricos según sus permisos y necesidades.

La existencia de esta capa de acceso distribuido tiene implicaciones importantes. Desde el

punto de vista de la apropiación tecnológica, el sistema deja de ser una solución encerrada en la electrónica y se convierte en un recurso utilizable por investigadores, estudiantes y otros actores del territorio. Desde el punto de vista del diseño, demuestra que el desarrollo fue pensado no solo para adquirir datos, sino también para hacerlos accesibles en condiciones reales de uso.

2.8. Aporte académico, institucional y proyección del desarrollo

La información de soporte asociada al desarrollo deja ver que este sistema ha aportado al fortalecimiento de capacidades científicas y tecnológicas en varios niveles. En primer lugar, ha contribuido a la formación de estudiantes mediante experiencias de trabajo con sensores, microcontroladores, conectividad IoT, almacenamiento en la nube y análisis de datos. En segundo lugar, ha fortalecido la infraestructura y la experiencia técnica de los grupos de investigación involucrados. Finalmente, ha abierto la posibilidad de ampliar servicios y productos relacionados con el monitoreo ambiental remoto [7].

La certificación de uso emitida desde GeoLimna indica que la herramienta se emplea en actividades de investigación, extensión y formación en medios acuáticos leníticos, y que la información producida tiene valor no solo para caracterización y diagnóstico, sino como futura fuente de Big Data ambiental [3]. Este aspecto es relevante porque muestra que el sistema ya no se justifica solo por su diseño técnico, sino también por su inserción en prácticas académicas concretas.

A ello se suma la formalización institucional del resultado mediante el registro ante la Dirección Nacional de Derecho de Autor, la cesión de derechos patrimoniales a la Universidad de Antioquia y la distribución explícita de participación entre los autores [2], [9], [10]. Estos elementos confirman que se trata de un resultado de investigación aplicado, reconocido y estructurado dentro de la institucionalidad universitaria.

En términos de proyección, el sistema deja una base sólida para incorporar nuevas variables, mejorar la autonomía energética, robustecer la infraestructura de almacenamiento y avanzar hacia modelos predictivos de mayor alcance. En consecuencia, su importancia no radica solo en lo ya logrado, sino en la plataforma tecnológica que deja disponible para desarrollos posteriores.

2.9. Conclusiones

El desarrollo presentado muestra que el monitoreo remoto de variables ambientales en sistemas leníticos puede abordarse eficazmente mediante una arquitectura modular que articule sensado, transmisión, almacenamiento y visualización. Su aporte principal no se limita a la automatización de una medición puntual, sino a la construcción de una cadena técnica completa que transforma una señal de campo en un recurso digital disponible para análisis continuo.

Desde el punto de vista de ingeniería, la solución integra sensores especializados, micro-controladores, radiofrecuencia, conectividad WiFi, bases de datos en tiempo real, almacenamiento estructurado y consulta distribuida. Esa integración permite reducir la dependencia de campañas manuales, mejorar la continuidad del registro y consolidar históricos útiles para procesos de investigación y análisis ambiental.

Asimismo, el sistema tiene un valor institucional claro. Ha fortalecido la formación de estudiantes, la capacidad de los grupos de investigación y la infraestructura tecnológica disponible para monitoreo ambiental. En ese sentido, la herramienta software asociada al desarrollo debe entenderse como parte de una estrategia más amplia de investigación aplicada y no como el eje temático exclusivo del capítulo.

Referencias

- [1] Universidad de Antioquia, *Solicitud de registro de soporte lógico - software: Sistema de registro y seguimiento de variables ambientales en ciénagas (VAMCI)*, Documento institucional interno, diligenciado el 05/06/2024, jun. de 2024.
- [2] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico - software: VAMCI: Variables Ambientales de Ciénagas*, Libro-Tomo-Partida 13-100-87, fecha de registro: 30-sept.-2024, 2024.
- [3] F. d. J. Vélez Macías, *Certificado de uso del software VAMCI*, Certificación emitida por el grupo GeoLimna, 2024.
- [4] D. S. F. McCann, F. de Jesús Vélez Macías, N. J. A. Ramírez y J. D. R. Martínez, *Solicitud para asignación de puntos por desarrollo de software: VAMCI*, Documento institucional, Universidad de Antioquia, oct. de 2024.
- [5] Grupo GEPAR y Grupo GeoLimna, *Sistema remoto para la medición y registro de variables ambientales: Manual de usuario*, Universidad de Antioquia, ago. de 2023.

- [6] *Variables Ambientales de Ciénagas: Manual Técnico*, Documento interno de los grupos GEPAR y GeoLimna, Universidad de Antioquia, abr. de 2024.
- [7] Universidad de Antioquia, *Formato de solicitud para asignación de puntos por desarrollo de software*, Documento institucional interno, Formato institucional empleado para documentar desarrollos de software registrados como productos tecnológicos., Medellín, Colombia, 2024.
- [8] *Variables Ambientales de Ciénagas: Manual de Usuario*, Documento interno de los grupos GEPAR y GeoLimna, Universidad de Antioquia, abr. de 2024.
- [9] D. S. F. M. Cann, F. de Jesús Vélez Macías, N. J. A. Ramírez y J. D. R. Martinez, *Contrato de cesión de derechos sobre la creación denominada Sistema de registro y seguimiento de variables ambientales en ciénagas (VAMCI)*, Minuta de cesión de derechos a la Universidad de Antioquia, 2024.
- [10] Universidad de Antioquia, *Acta definitiva de distribución de regalías entre los participantes relacionados con la Universidad de Antioquia para la creación VAMCI*, Acta interna, 05/06/2024, jun. de 2024.

Capítulo 3

Automatización de la alimentación y clasificación de esquejes en floricultura

Autores

Ricardo Moreno Sánchez¹; David Stephen Fernández Mc Cann²

Resumen

La propagación vegetativa por esquejes constituye una operación crítica del sector floricultor, pues la calidad del material sembrado incide directamente en el rendimiento agronómico, la uniformidad del cultivo y la eficiencia del proceso productivo. En este capítulo se presenta el desarrollo de un sistema integrado para la alimentación y clasificación automatizada de esquejes, concebido como una evolución de la máquina seleccionadora previamente desarrollada para material vegetal de crisantemo. El problema abordado surge de una restricción operativa concreta: aunque la clasificación por visión artificial había permitido mejorar la consistencia del proceso, la alimentación de los esquejes seguía dependiendo de la intervención manual. La solución integró diseño mecánico, automatización, visión por computador, comunicación con PLC y control de operación para transformar

¹Profesor, Grupo de Diseño Mecánico (GDM), Facultad de Ingeniería, Universidad de Antioquia.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Proyecto desarrollado en articulación con los grupos GDM y GEPAR, Universidad de Antioquia, con relacionamiento con Asocolflores.

la alimentación en un proceso continuo, sincronizado y articulado con la clasificación. El desarrollo incluyó un mecanismo alimentador basado en organización del material vegetal, toma neumática, posicionamiento y deposición sobre banda transportadora, así como una herramienta software para segmentación, medición, clasificación y generación de estadísticas. La experiencia muestra cómo una necesidad sectorial puede convertirse en un resultado tecnológico maduro, con documentación técnica, validación institucional, registro de software y proyección hacia sistemas de automatización agrícola más amplios [1], [2], [3], [4].

Palabras clave: floricultura; esquejes; visión artificial; automatización; alimentador automático; clasificación morfológica; PLC; GEPAR; GDM.

3.1. Introducción

La propagación vegetativa por esquejes constituye una de las operaciones críticas del sector floricultor, pues la calidad del material sembrado incide directamente sobre el rendimiento agronómico, la uniformidad del cultivo y la eficiencia del proceso productivo. En este contexto, la selección adecuada de esquejes no es una tarea menor, sino una etapa de control de calidad que condiciona la productividad posterior del sistema.

Durante años, esta selección se ha realizado de forma manual, apoyada en la experiencia del operario y en criterios visuales asociados a la longitud del tallo, la presencia de hojas en la base y otras características morfológicas. Sin embargo, este procedimiento presenta limitaciones estructurales: depende fuertemente del juicio humano, introduce variabilidad entre operarios, exige esfuerzos repetitivos y se convierte en un cuello de botella cuando se busca elevar la tasa de procesamiento. La experiencia previa de la máquina seleccionadora de esquejes mostró que la automatización de la clasificación podía mejorar la consistencia del proceso, alcanzando capacidades de entre 40 y 60 esquejes por minuto según el modo de alimentación, frente a aproximadamente 25 esquejes por minuto por operaria en la modalidad manual [1], [2], [5].

A partir de esa experiencia surgió un problema de segundo orden, pero de alta relevancia práctica: aun cuando la máquina seleccionadora disponía de una capacidad de análisis superior a la del proceso manual, seguía dependiendo de que los esquejes fueran organizados y suministrados por personas. En otras palabras, la clasificación había sido parcialmente automatizada, pero la alimentación continuaba siendo manual, lo que restringía el aprovechamiento pleno de la capacidad del sistema. Ese diagnóstico dio origen al proyecto de un mecanismo alimentador de esquejes, concebido para transformar la alimentación en un proceso continuo, sincronizado y articulado con la etapa de clasificación [1], [2], [6].

El desarrollo posterior integró dos dimensiones complementarias. Por una parte, se diseñó y construyó un mecanismo físico de alimentación automática basado en organización del material vegetal, toma neumática, posicionamiento y deposición sincronizada sobre la banda transportadora. Por otra, se consolidó una herramienta software de visión por computador y comunicación con PLC que permitió clasificar automáticamente los esquejes, registrar variables de interés y generar información estadística para el análisis del proceso [3], [4], [7], [8].

Este capítulo presenta esa evolución como una arquitectura de ingeniería integrada. No se trata únicamente de describir una máquina o un programa, sino de mostrar cómo la combinación de diseño mecánico, automatización, visión artificial y control permitió intervenir una operación concreta del sector floricultor y convertirla en un sistema tecnológico más productivo, repetible y escalable.

3.2. Planteamiento del problema

El proceso de selección de esquejes parte de una condición fundamental: el material vegetal que llega a siembra no debe tratarse como una colección indiferenciada de tallos, sino como un conjunto de especímenes con atributos morfológicos que afectan el comportamiento posterior del cultivo. Por ello, las floricultoras requieren separar esquejes ideales, cortos, largos o con hoja en base, de manera que la siembra ocurra con mayor homogeneidad y calidad. Esta necesidad fue reconocida por Asocolflores y dio origen, en una primera etapa, al desarrollo de la máquina seleccionadora de esquejes, orientada a clasificar automáticamente material de crisantemo mediante visión artificial [5], [9].

Sin embargo, una vez validada la capacidad de clasificación, apareció una limitación operativa evidente. Los esquejes seguían siendo desempacados, organizados y alimentados manualmente sobre la banda. De este modo, la operación global continuaba dependiendo de tareas repetitivas que reducían la productividad total del sistema, aun cuando la etapa de clasificación ya estuviera parcialmente optimizada. La propuesta técnica del proyecto fue construir un mecanismo alimentador que permitiera pasar de la alimentación manual a una alimentación automática y continua, articulada con la máquina clasificadora existente [1], [2], [6].

Desde el punto de vista ingenieril, este problema puede formularse como una relación entre la tasa de alimentación del sistema, denotada por q_a , y la tasa de clasificación de la máquina, denotada por q_c . Si la capacidad efectiva del proceso global se representa por q_g , entonces dicha capacidad no puede superar al subsistema más lento. En consecuencia,

$$q_g = \min(q_a, q_c),$$

donde q_g es la tasa global de procesamiento en esquejes por unidad de tiempo, q_a es la tasa de suministro del alimentador y q_c es la tasa de clasificación de la máquina. Esta ecuación resume el problema central del proyecto: aunque q_c fuera elevada, si q_a continuaba limitada por la intervención manual, el sistema completo seguiría operando por debajo de su potencial.

En ese marco, la automatización de la alimentación deja de ser un accesorio y pasa a convertirse en una condición estructural para incrementar la productividad. El objetivo no era simplemente mover esquejes hacia la banda, sino hacerlo con orientación suficiente, ritmo estable, sincronización dinámica y mínima afectación mecánica del material vegetal. Tal exigencia explica por qué el desarrollo no puede entenderse como una adecuación menor, sino como una intervención tecnológica que modifica sustancialmente la arquitectura del proceso original [10].

3.3. Antecedentes y evolución del desarrollo

Los antecedentes del sistema muestran una evolución progresiva desde una solución centrada en la clasificación hacia una solución integral de clasificación y alimentación. La primera etapa, documentada en reportes institucionales y en divulgación universitaria, estuvo asociada al desarrollo de una máquina seleccionadora de esquejes para crisantemo, capaz de capturar imágenes del material vegetal, identificar rasgos morfológicos y depositar los esquejes en compartimientos distintos según su clasificación [5], [9].

Esa máquina demostró que la visión por computador podía aportar valor directo al sector floricultor. Según la información divulgada en la Universidad de Antioquia, la solución permitió clasificar esquejes a partir de imágenes procesadas por software y depositarlos en compartimientos correspondientes a categorías como largo, corto, hoja en base e ideal [5]. Además, la validación empresarial y el relacionamiento con Asocolflores confirmaron que el desarrollo respondía a una necesidad real del sector y que su desempeño justificaba avanzar hacia una segunda fase [11].

La segunda fase consistió en intervenir la restricción que permanecía abierta: la alimentación. La ficha técnica del proyecto y la propuesta técnico-económica plantearon con claridad que el nuevo objetivo era diseñar y construir un mecanismo alimentador de esquejes para una máquina seleccionadora, buscando aumentar por lo menos en un 20% la capacidad de selección y automatizar la etapa que seguía siendo manual [1], [2]. A

partir de ello, se definieron cuatro objetivos específicos: estudiar las variables del prototipo actual, realizar el diseño conceptual, establecer el diseño experimental y construir el prototipo con sus pruebas y puesta a punto [1], [6].

En paralelo, el desarrollo software evolucionó hacia una herramienta más completa. Aunque el nombre del software no estructura este capítulo, sí es importante señalar que dicha herramienta consolidó la interfaz de operación, el procesamiento de imágenes, la selección de variedades, la comunicación con PLC y la generación de estadísticas, ampliando el alcance funcional del sistema físico [3], [7], [8]. La convergencia entre mecanismo alimentador y software de clasificación permitió pasar de una máquina aislada a una célula tecnológica más cercana a una solución industrial.

La Fig. 3.1 resume los antecedentes del desarrollo, desde la máquina seleccionadora inicial hasta la formulación del sistema de alimentación automática.



Figura 3.1: Antecedentes del desarrollo tecnológico, desde la máquina seleccionadora de esquejes hasta la formulación del sistema de alimentación automática articulado con la etapa de clasificación.

3.4. Arquitectura funcional del sistema integrado

La solución puede entenderse como un sistema mecatrónico compuesto por cuatro bloques funcionales: adquisición del material vegetal, alimentación y posicionamiento, clasificación

por visión y actuación sobre la separación física del esqueje. Estos bloques no operan de manera independiente, sino articulados en una secuencia de eventos que exige coordinación temporal y consistencia geométrica.

Si se representa por u_k el esqueje número k que ingresa al sistema, el flujo funcional puede expresarse conceptualmente como

$$u_k \longrightarrow a_k \longrightarrow i_k \longrightarrow c_k \longrightarrow d_k,$$

donde a_k representa la etapa de alimentación y posicionamiento, i_k la captura y procesamiento de la imagen, c_k la clase asignada al esqueje y d_k la deposición final en el compartimiento correspondiente. Esta formulación muestra que el sistema completo no se limita a observar un objeto, sino que lo transforma en una decisión de clasificación y en una acción física posterior.

Desde la perspectiva mecánica, la arquitectura propuesta contempló una mesa de depósito vibratorio para organizar el material vegetal, un sistema de captura mediante ventosas neumáticas, una estructura de traslado y un mecanismo de deposición coordinado con la banda transportadora de la máquina clasificadora. En la propuesta inicial, las ventosas toman los esquejes desde una canasta con vibración y flujo de aire y los llevan hasta los compartimientos de la banda alimentadora de la clasificadora [2].

Desde la perspectiva de control, la coordinación entre alimentador y banda era indispensable. La propuesta técnico-económica identificó explícitamente un subsistema de sincronización entre la banda transportadora y el alimentador, necesario para depositar los esquejes teniendo en cuenta cambios de velocidad en la banda [2]. Posteriormente, el software documentó la interacción con PLC y el control del variador como parte del funcionamiento real del sistema [3].

Desde la perspectiva computacional, la herramienta de visión por computador permite capturar una imagen del esqueje, segmentarlo, extraer medidas de interés y asignarle una categoría. Esta clasificación es luego utilizada por el sistema para gobernar la actuación física asociada a la separación.

La Fig. 3.2 presenta la arquitectura funcional del sistema integrado.

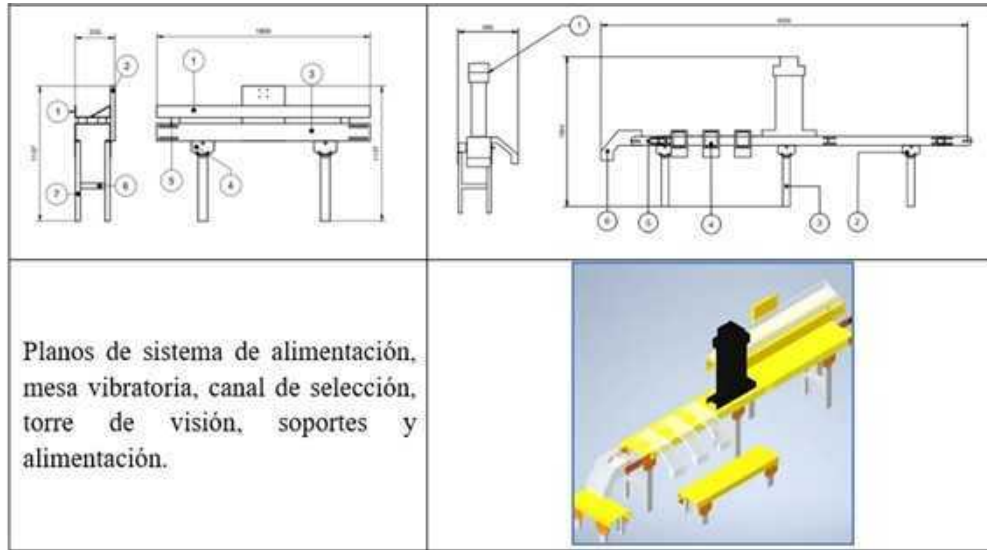


Figura 3.2: Arquitectura funcional del sistema integrado de alimentación, captura, clasificación y deposición de esquejes.

3.5. Modelo de productividad y condición de sincronización

Una forma útil de analizar el impacto del alimentador consiste en comparar el tiempo total de procesamiento antes y después de la automatización. Si N denota el número total de esquejes procesados en una jornada, T_m representa el tiempo requerido en operación predominantemente manual y T_a el tiempo requerido con alimentación automatizada, entonces las productividades respectivas pueden escribirse como

$$P_m = \frac{N}{T_m}, \quad P_a = \frac{N}{T_a},$$

donde P_m y P_a se expresan en esquejes por unidad de tiempo. El incremento relativo de productividad, denotado por η_P , puede definirse como

$$\eta_P = \frac{P_a - P_m}{P_m} \times 100\%.$$

Aquí, η_P expresa el porcentaje de mejora atribuible a la automatización. Esta formulación es coherente con el planteamiento del proyecto, en el cual se esperaba que la incorporación del mecanismo alimentador incrementara al menos en un 20% la capacidad de selección

[2]. Además, documentos de uso y evaluación técnica resaltan que el desarrollo impacta directamente indicadores medibles como esquejes procesados por minuto, reducción de tiempos muertos y mayor aprovechamiento de la capacidad de la máquina seleccionadora [10].

Ahora bien, el incremento de productividad no depende solamente de mover más rápido el material, sino de sincronizar correctamente la deposición con el movimiento de la banda. Si v_b denota la velocidad lineal de la banda transportadora, L_e la separación entre eslabones o compartimientos útiles y τ_d el tiempo disponible para depositar un esqueje en el sitio correcto, entonces debe cumplirse la relación

$$\tau_d \leq \frac{L_e}{v_b}.$$

En esta ecuación, $\frac{L_e}{v_b}$ representa el intervalo temporal entre dos posiciones consecutivas disponibles sobre la banda. La desigualdad indica que el mecanismo alimentador debe completar la toma, el traslado y la deposición del esqueje dentro de ese intervalo; de lo contrario, se pierde sincronía y disminuye la calidad operativa del sistema.

Esta condición resume uno de los principales retos de incertidumbre tecnológica del proyecto: no bastaba con diseñar un alimentador capaz de capturar esquejes, sino uno que lo hiciera con precisión temporal suficiente para acoplarse a una banda en movimiento y adaptarse a cambios de velocidad. Por ello, la sincronización dinámica entre banda y alimentador se convirtió en un componente esencial del diseño experimental y del desarrollo tecnológico logrado [2], [10].

3.6. Procesamiento de imágenes y decisión de clasificación

La clasificación automática del esqueje se apoya en una cadena de procesamiento digital de imágenes implementada en un entorno de software de escritorio. La documentación técnica del sistema indica que la aplicación fue desarrollada en Python 3.9, con soporte de bibliotecas para interfaz gráfica, procesamiento de imágenes, análisis numérico, manejo de archivos y comunicación de red [3]. A nivel funcional, la herramienta permite operar tanto con cámara como con directorio de imágenes, seleccionar variedades, establecer conexión con PLC y generar estadísticas del proceso [3], [8].

Si la imagen capturada del esqueje se denota por $I(x, y)$, donde x e y representan las coordenadas del píxel, el proceso de segmentación puede entenderse como la obtención de

una máscara binaria $M(x, y)$, definida como

$$M(x, y) = \begin{cases} 1, & \text{si el píxel pertenece al esqueje,} \\ 0, & \text{en otro caso.} \end{cases}$$

Una vez obtenida la máscara, es posible calcular atributos morfológicos relevantes. Si A_f denota el área foliar estimada y L_t la longitud del tallo en centímetros, la decisión de clasificación puede modelarse como una función

$$c = f(L_t, A_f, \theta),$$

donde c es la clase asignada al esqueje y θ representa el conjunto de parámetros umbral asociados a una variedad específica. Esta representación es consistente con la lógica documentada en los manuales, según la cual el software utiliza medidas suministradas por la empresa floricultora para clasificar los esquejes en categorías como corto, largo, hoja en base e ideal, y además calcula medidas de interés como área foliar y tallo promedio para análisis estadístico [3], [8].

La documentación técnica identifica la función `segmentacion()` como el núcleo del software, encargada de ejecutar preprocesado de imagen, operaciones de morfología y filtros para detección de hoja en base, retornando la imagen procesada, la longitud del esqueje y su clasificación [3]. Esa modularidad resulta importante porque demuestra que la clasificación no es una simple inspección visual, sino una secuencia computacional reproducible y parametrizable.

Desde el punto de vista del sistema global, la visión por computador cumple una doble función: decide la categoría del material vegetal y provee variables cuantitativas para el análisis del proceso y la trazabilidad operativa.

La Fig. 3.3 presenta la interfaz de operación y un ejemplo de procesamiento de esqueje mediante segmentación y extracción de variables.

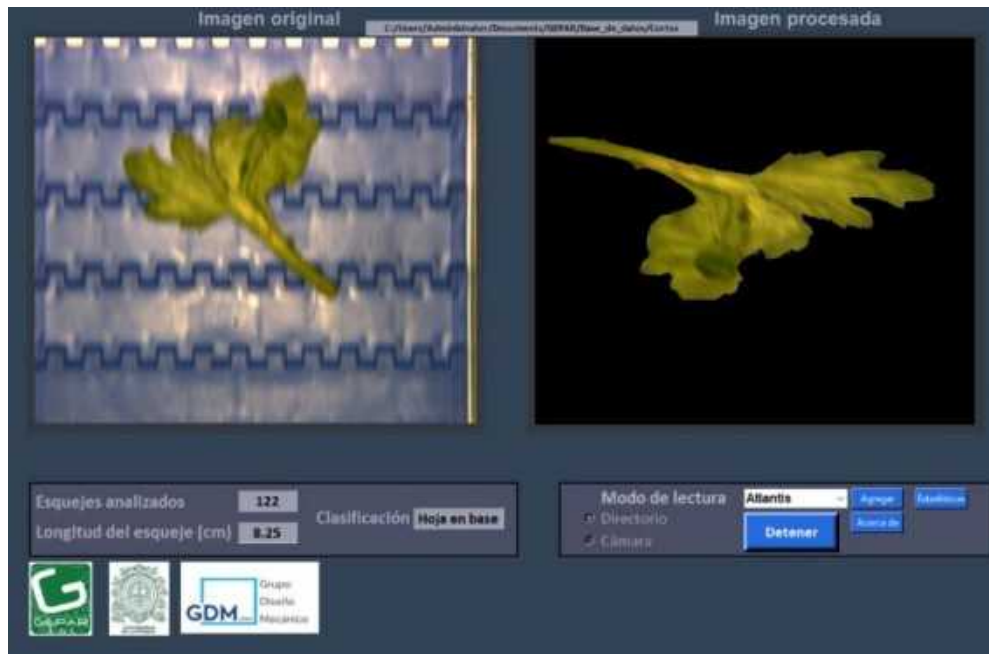


Figura 3.3: Interfaz de operación del sistema de visión y ejemplo de procesamiento de un esqueje, incluyendo visualización original, imagen segmentada y variables de clasificación.

3.7. Diseño experimental y desarrollo del prototipo

El proyecto fue estructurado bajo una lógica explícita de investigación aplicada y desarrollo experimental. La documentación institucional consigna objetivos, compromisos, cronograma, financiación y resultados verificables, entre ellos la construcción del prototipo de alimentadora de esquejes y la entrega de documentación necesaria para su operación [6]. Esta formalización permite interpretar el desarrollo no como una adaptación empírica aislada, sino como una ruta metodológica con fases claramente definidas.

En la propuesta del proyecto se plantearon actividades de caracterización del transporte y posicionamiento, mediciones con mecanismos de ventosas, tabla morfológica, evaluación de alternativas, cálculos de neumática y resistencia de materiales, elaboración de planos, construcción, montaje electrónico y control, acople hardware-software, pruebas y puesta a punto [2]. El cronograma, organizado para nueve meses, evidencia una secuencia lógica de diseño conceptual, diseño detallado, fabricación, montaje e integración.

Si se denota por \mathbf{r} el vector de requerimientos funcionales del sistema, por \mathbf{d} el vector de variables de diseño y por \mathbf{y} el vector de respuestas experimentales observadas en pruebas, la lógica general de desarrollo puede escribirse como

$$\mathbf{d}^* = \arg \min_{\mathbf{d}} \mathcal{J}(\mathbf{r}, \mathbf{y}(\mathbf{d})),$$

donde \mathcal{J} es una función de costo o discrepancia entre los requerimientos funcionales y el comportamiento observado del sistema. Aunque esta expresión resume el problema de manera abstracta, es útil porque representa el carácter iterativo del proyecto: diseñar, probar, medir, ajustar y volver a probar hasta obtener una configuración funcionalmente aceptable.

Los informes y respuestas técnicas posteriores destacan que el proyecto implicó pruebas con múltiples configuraciones de boquillas, presiones de aire, mecanismos de posicionamiento y sincronización con la banda, precisamente para superar incertidumbres asociadas a la manipulación automática de un material vegetal variable [10], [12]. Esta evidencia confirma que el prototipo final fue el resultado de una exploración sistemática y no de una decisión de diseño tomada de antemano sin validación.

La Fig. 3.4 muestra elementos del desarrollo experimental y de la revisión técnica del prototipo.



Figura 3.4: Elementos del diseño experimental del sistema, incluyendo pilotos, prototipos y revisión técnica del prototipo elegido.

3.8. Integración hardware-software y operación con PLC

Uno de los aspectos más robustos del desarrollo es la integración entre sistema físico y sistema lógico. El manual de usuario muestra que la aplicación dispone de modos de

lectura por cámara y por directorio, selección de tipo de esqueje, incorporación de nuevas variedades, conexión con PLC, funcionamiento en línea y módulo de estadísticas [8]. El manual técnico, por su parte, documenta funciones asociadas a la clase principal, segmentación de imagen, ejecución por cámara y directorio, lectura y escritura hacia el PLC, y actualización de velocidad del variador [3].

Si se representa por s_k el estado del sistema en el instante discreto k , por m_k la medición obtenida del procesamiento de imágenes y por u_k la acción enviada al PLC, la integración de control puede describirse mediante la ecuación de actualización

$$s_{k+1} = g(s_k, m_k, u_k),$$

donde $g(\cdot)$ resume la dinámica de acoplamiento entre observación, decisión y actuación. En este contexto, la medición m_k contiene información como la clase asignada al esqueje y variables de interés, mientras que u_k contiene instrucciones relacionadas con velocidad, movimiento de banda, confirmación de conexión y sincronización operativa.

La documentación técnica indica que la función `conectar_PLC()` verifica la comunicación con el PLC mediante IP y puerto; `leer_PLC()` obtiene estados de la máquina usando protocolo Modbus; `escribir_PLC()` envía datos para controlar la máquina en sincronía con el software; y `actualizar_velocidad()` verifica valores válidos del variador [3]. Esto confirma que el sistema no actúa solo como visor de imágenes, sino como una capa de supervisión y decisión integrada al control de la máquina.

La conexión con PLC tiene además importancia práctica. El manual de usuario advierte que, si la comunicación no está correctamente establecida, el sistema no debe iniciar, y detalla mensajes de error para fallos de conexión, tiempos de espera agotados y problemas de lectura de cámara o CSV [8]. Esta atención a los errores operativos es indicio de madurez tecnológica, porque traslada el desarrollo desde el prototipo de laboratorio hacia un escenario de uso más cercano a la operación real.

La Fig. 3.5 presenta el montaje general del sistema y la articulación entre el alimentador, la máquina seleccionadora y los elementos de operación.



Figura 3.5: Montaje e integración del sistema, mostrando la articulación física entre alimentador, máquina seleccionadora, computador de operación y elementos de actuación.

3.9. Resultados, uso y validación institucional

La evidencia documental disponible permite afirmar que el sistema alcanzó un nivel de validación importante tanto en el entorno universitario como en el contexto de aplicación sectorial. El acta de registro y propiedad intelectual del proyecto ALIES deja consignados el objetivo, los compromisos verificables, la duración, la cofinanciación entre Universidad de Antioquia y Asocolflore, y la construcción del prototipo alimentador como producto del proyecto [6]. Por su parte, el certificado de uso del software asociado indica que este es utilizado por los grupos GEPAR y GDM en proyectos de automatización del alimentador de esquejes y en estudios asociados a procesos automatizados de selección [13].

El registro ante la Dirección Nacional de Derecho de Autor formalizó además la existencia del software asociado, indicando que automatiza la selección y clasificación de esquejes, emplea visión por computador y PLC, mide longitud del tallo y área foliar, clasifica en cuatro categorías y genera reportes para optimizar calidad y eficiencia del proceso [4]. Esa formalización aporta trazabilidad sobre la madurez del resultado y su carácter de producto tecnológico.

La validación empresarial también resulta significativa. Una constancia de uso de Aso-

colflores señala que la máquina seleccionadora de esquejes fue validada para uso en la clasificación de esquejes según sus cualidades y que, debido a su buen desempeño, se inició un nuevo desarrollo conjunto complementario para pasar de alimentación manual a automática, buscando mayor velocidad útil en los procesos de clasificación [11]. Esta afirmación vincula directamente la segunda fase del proyecto con una necesidad surgida del desempeño real del prototipo anterior.

A nivel de impacto institucional y sectorial, documentos de evaluación docente subrayan que la alimentadora introduce elementos técnicos y metodológicos nuevos en el entorno, modifica sustancialmente la estructura del proceso original, supera barreras técnicas sin solución fácil en el sector y produce impactos medibles en productividad, reducción de tiempos muertos y estandarización del proceso [10]. En la misma línea, una publicación informativa de la Facultad de Ingeniería destacó el aporte del profesor David Fernández y del grupo GEPAR al mejoramiento de procesos del sector floricultor, incluyendo el desarrollo de la seleccionadora y de la alimentadora como evolución reciente del sistema [9].

3.10. Discusión

La principal contribución del desarrollo presentado radica en haber intervenido una operación aparentemente simple, la alimentación de esquejes, y haberla reformulado como un problema de sincronización mecatrónica, visión por computador y control industrial. Esta reformulación tiene consecuencias importantes, porque desplaza la discusión desde la mecanización aislada hacia la integración de subsistemas heterogéneos en una solución productiva real.

En términos de ingeniería, la solución muestra al menos cuatro niveles de valor. El primero es mecánico, pues resuelve la manipulación de un material vegetal variable mediante ventosas, vibración y posicionamiento. El segundo es electrónico y de control, al integrar sensores, PLC, variador y lógica de sincronización. El tercero es computacional, al clasificar automáticamente el material mediante procesamiento de imágenes. El cuarto es organizacional, al producir estadísticas, documentación y condiciones de uso para la operación continua.

Otro aspecto destacable es que el desarrollo no reemplaza simplemente una labor humana, sino que reconfigura el rol del operario dentro del sistema. La propia documentación empresarial sugiere que la intención no es desplazar mano de obra, sino reubicarla en tareas de mayor valor dentro del proceso productivo [11]. Desde esta perspectiva, la automatización no se presenta como sustitución ciega, sino como reorganización del trabajo

apoyada por tecnología.

Finalmente, el proyecto ofrece una lección metodológica de interés para la ingeniería aplicada: cuando un sistema parcial muestra buenos resultados, las restricciones suelen desplazarse hacia otra etapa del proceso. La máquina seleccionadora resolvió la clasificación, pero dejó abierta la alimentación; la alimentadora resolvió parte del acople físico, pero exigió una integración más sólida con software y control. Así, la innovación aparece menos como un artefacto aislado y más como una cadena de cierres progresivos de cuellos de botella.

3.11. Conclusiones

El desarrollo descrito en este capítulo demuestra que la mejora de procesos en floricultura puede beneficiarse significativamente de soluciones integradas de ingeniería que combinen diseño mecánico, visión por computador, automatización y control industrial. La transición desde una clasificación parcialmente automatizada con alimentación manual hacia un sistema de alimentación y clasificación articulado responde a una necesidad concreta del entorno productivo y muestra una ruta clara de maduración tecnológica.

Desde el punto de vista funcional, el proyecto resolvió el problema de que la productividad global del sistema estaba limitada por la alimentación manual. La formulación del mecanismo alimentador, junto con su sincronización con la banda transportadora y su integración con el software de clasificación, permitió concebir una solución continua, repetible y más eficiente. En este sentido, la alimentadora no fue un accesorio añadido, sino una modificación estructural del proceso.

Desde el punto de vista metodológico, el proyecto evidencia una etapa comprobable de investigación aplicada y desarrollo experimental. La existencia de objetivos definidos, cronograma, prototipo, manuales, documentación técnica, registro de software, uso institucional y validación sectorial confirma que se trata de un resultado tecnológico maduro, sustentado en iteración, prueba y ajuste.

Finalmente, el caso muestra cómo una necesidad sectorial concreta puede convertirse en una oportunidad de articulación universidad-empresa. La experiencia acumulada en torno a la selección y alimentación de esquejes deja una plataforma tecnológica y conceptual para futuros desarrollos en automatización agrícola, análisis morfológico de material vegetal y sistemas de clasificación industrial asistidos por visión.

Referencias

- [1] Universidad de Antioquia, *Ficha técnica: Mecanismo Alimentador de Esquejes – Prototipo*, Vicerrectoría de Investigación, documento interno de proyecto, feb. de 2021.
- [2] R. M. Sánchez y D. F. M. Cann, *Propuesta técnico-económica: Mecanismo Alimentador de Esquejes*, Grupo de Diseño Mecánico y Grupo GEPAR, Universidad de Antioquia, dic. de 2020.
- [3] Grupos GEPAR y Diseño Mecánico, *ALSOFT: Manual técnico*, Universidad de Antioquia, 2024.
- [4] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: ALSOFT*, Libro-Tomo-Partida 13-101-238, fecha de registro 06-Dec-2024, 2024.
- [5] L. J. Q. Martínez, *Investigación en Ingeniería aporta al desarrollo empresarial*, Portal institucional Universidad de Antioquia, abr. de 2018.
- [6] Universidad de Antioquia, *Acta de registro y de propiedad intelectual del proyecto ALIES Mecanismo Alimentador de Esquejes Prototipo*, Acta No. 2021-41210, Comité para el Desarrollo de la Investigación - CODI, 2022.
- [7] D. S. F. McCann y R. M. Sanchez, *Solicitud para asignación de puntos por desarrollo de software ALSOFT*, Documento institucional, Universidad de Antioquia, oct. de 2024.
- [8] Grupos GEPAR y Diseño Mecánico, *ALSOFT: Manual de usuario*, Universidad de Antioquia, 2024.
- [9] L. J. Q. Martínez, «El profesor David Fernández fue galardonado por Ceniflores,» *Ingeniemos*, dic. de 2024, Edición 54, Facultad de Ingeniería, Universidad de Antioquia.
- [10] Universidad de Antioquia, *Alimentadora de esquejes (ALIES): respuestas a los criterios de jerarquía docente para sustentar la innovación tecnológica*, Documento de soporte técnico, 2024.
- [11] M. A. O. Ramírez, *Constancia de uso de la máquina seleccionadora de esquejes y desarrollo complementario de alimentación automática*, Asocolflores, feb. de 2022.
- [12] R. M. Sanchez y D. F. M., *ALIES: Alimentadora de Esquejes – Prototipo – Informe de avance*, Convocatoria de Investigación Aplicada y Desarrollo Experimental, mayo de 2024.
- [13] R. M. Sanchez y D. S. F. M. Cann, *Certificación de uso del software ALSOFT*, Grupos GDM y GEPAR, Universidad de Antioquia, dic. de 2024.

Capítulo 4

Sistema móvil de visión artificial para el control de calidad de ramos de flores

Autores

David Stephen Fernández Mc Cann¹; Miguel Ángel Restrepo Rueda²

Resumen

El control de calidad de ramos de flores es una etapa crítica en la floricultura de exportación, pues permite verificar el cumplimiento de parámetros dimensionales, morfológicos y de presentación antes del despacho. Tradicionalmente, esta inspección se realiza de forma manual, con criterios visuales, mediciones parciales y registros que luego deben ser consolidados. Este procedimiento introduce variabilidad entre operarios, fatiga, demoras y dificultades de trazabilidad. En este capítulo se presenta un sistema móvil de visión artificial orientado a automatizar la inspección de ramos recién cosechados mediante una plataforma que integra estructura mecánica móvil, cámaras RGB y de profundidad, sensor de peso, procesamiento digital de imágenes, modelos de aprendizaje automático, interfaz gráfica y almacenamiento estructurado de resultados. El sistema permite estimar variables como número de tallos, diámetro, área superior de floración, longitud del

¹Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Ingeniero electrónico, Facultad de Ingeniería, Universidad de Antioquia. Desarrollo asociado al proyecto en articulación con la empresa Flores El Capiro S.A. y con apoyo de Tecondor S.A.S.

ramo, posición del elástico, longitud de pelada foliar y peso. Además, registra imágenes y datos asociados a cada inspección, generando una base útil para trazabilidad, análisis productivo y mejora continua. La validación en condiciones reales de operación permitió avanzar hacia un nivel de madurez tecnológica alto, con potencial de transferencia al sector floricultor y proyección hacia otras aplicaciones agroindustriales [1], [2], [3], [4].

Palabras clave: visión artificial; floricultura; control de calidad; ramos de flores; procesamiento de imágenes; aprendizaje automático; trazabilidad; poscosecha; GEPAR.

4.1. Introducción

La floricultura de exportación descansa sobre una paradoja técnica notable: se trata de una industria altamente sofisticada en sus exigencias comerciales, pero todavía dependiente, en múltiples etapas críticas, de operaciones manuales intensivas. Una de ellas es el control de calidad de los ramos de flores, proceso en el que se verifica si cada ramo cumple criterios dimensionales, estructurales y de presentación antes de su despacho. En la práctica, esta labor suele recaer en inspectores que observan visualmente el ramo, toman algunas medidas, estiman otras y registran la información en formatos que luego deben ser consolidados. Esta dinámica introduce subjetividad, variabilidad entre operarios, fatiga y demoras en la disponibilidad del dato [1], [2].

En el caso de Flores El Capiro S.A., el control de calidad de los ramos está asociado al cumplimiento de parámetros como número de tallos, diámetro de tallos, longitud del ramo, área superior ocupada por las flores, posición de la liga o elástico y peso, entre otros [1], [5]. Tales variables no solo permiten decidir si un ramo satisface una especificación comercial, sino que también ofrecen información valiosa para trazabilidad, logística y mejora del proceso. El problema no consiste únicamente en medir, sino en medir de forma confiable, oportuna, estandarizada y en condiciones reales de operación.

A partir de esta necesidad surgió el desarrollo de un sistema móvil para control automático de calidad de ramos, concebido como una plataforma que integra estructura mecánica móvil, cámaras, sensores de peso, cómputo local, interfaz gráfica y algoritmos de visión por computador para automatizar la inspección de ramos recién cosechados y reportar la información a los sistemas de la empresa [1], [4], [6]. En esta perspectiva, el sistema no debe entenderse como un accesorio instrumental, sino como una reconfiguración tecnológica del proceso de inspección.

El presente capítulo expone el fundamento ingenieril, el diseño metodológico y la lógica de operación de este desarrollo. El interés no está en presentar únicamente un dispositivo

o un software, sino en mostrar cómo la integración entre procesamiento digital de imágenes, aprendizaje automático, visión artificial, sensado físico y diseño mecatrónico puede traducirse en una solución aplicada para un problema concreto del sector floricultor.

4.2. Contexto productivo y formulación del problema

En el proceso productivo de flores de corte, los tallos son cosechados, agrupados y organizados en ramos según especificaciones definidas por el comprador. Una vez armado, cada ramo es sometido a una verificación de calidad para establecer si cumple con la longitud requerida, el número de tallos definido, el diámetro esperado, la apertura superior del ramo y las restricciones de peso y presentación [1], [2]. En el procedimiento manual observado en campo, el inspector revisa algunos ramos por lote, registra en cuadernos o formularios las no conformidades y posteriormente transcribe esos resultados a medios digitales. Este mecanismo es funcional, pero presenta limitaciones estructurales de velocidad, repetibilidad y trazabilidad [2].

Desde un punto de vista formal, el problema puede modelarse como la necesidad de estimar automáticamente un vector de variables de calidad del ramo. Si se denota por \mathbf{q} el vector de control de calidad, entonces

$$\mathbf{q} = [n_t \quad d_t \quad A_f \quad L_r \quad L_e \quad L_p \quad w],$$

donde n_t representa el número de tallos del ramo, d_t el diámetro característico de los tallos, A_f el área superior asociada a la floración visible, L_r la longitud del ramo, L_e la distancia desde el centro del elástico hasta la base del tallo, L_p la longitud de pelada foliar y w el peso del ramo. Estas magnitudes resumen la información esencial que el sistema debe capturar o inferir automáticamente [2], [5].

Una vez estimado este vector, la conformidad del ramo puede expresarse mediante una función de verificación respecto a un vector de especificaciones \mathbf{q}^* y una tolerancia admisible ε . De manera general,

$$C(\mathbf{q}) = \begin{cases} 1, & \text{si } |q_i - q_i^*| \leq \varepsilon_i \quad \forall i, \\ 0, & \text{en otro caso,} \end{cases}$$

donde $C(\mathbf{q})$ es la decisión binaria de cumplimiento y q_i representa cada parámetro medido. Esta formulación expresa el desafío del sistema: trasladar un procedimiento manual y parcialmente subjetivo hacia una evaluación cuantitativa reproducible.

En consecuencia, el problema de ingeniería no era simplemente construir un carro o instalar cámaras, sino diseñar una plataforma capaz de desplazarse en ambientes agrícolas, adquirir información multimodal, transformar imágenes y señales en mediciones útiles, y presentar resultados en tiempo real mediante una interfaz operable por personal de planta.

4.3. Antecedentes y génesis del desarrollo

El desarrollo del carro de calidad no apareció de forma aislada. Se inscribe en una trayectoria previa del grupo GEPAR y de la Universidad de Antioquia en el sector floricultor, con antecedentes relevantes en automatización de selección de esquejes, estudios morfológicos y sistemas de conteo e identificación basados en visión artificial [1]. Entre estos antecedentes se encuentran la máquina seleccionadora de esquejes, los estudios comparativos de medidas de esquejes, SICOP para conteo de plántulas, ALIES como alimentadora de la máquina seleccionadora, e iniciativas de clasificación de esquejes a partir de rasgos visibles [1].

Ese recorrido fue importante por dos razones. La primera es metodológica: permitió consolidar experiencia en adquisición de imágenes, segmentación, extracción de rasgos, entrenamiento de modelos y articulación universidad-empresa. La segunda es conceptual: mostró que muchos procesos de inspección agrícola que parecían inevitablemente manuales podían transformarse si se abstraían como problemas de medición automatizada.

En el caso específico del control de calidad de ramos, se realizaron primero estudios previos y montajes de prueba conceptual para examinar si era posible capturar imágenes pertinentes y traducirlas en medidas útiles. En esos estudios se construyó un *mock-up* del sistema, con imágenes de entrada y resultados de salida, y se concluyó que era viable obtener medidas automáticas precisas, almacenarlas, generar estadísticas y transmitir las a un sistema central para fines de control, logística e inventario [1]. Este punto fue clave, porque la utilidad del sistema ya no se limitaba a replicar la inspección humana, sino que abría la posibilidad de realizar una inspección más exhaustiva, ramo por ramo, en lugar de depender únicamente de muestreos estadísticos.

La Fig. 4.1 sintetiza los principales aspectos de calidad que motivaron el desarrollo del sistema.



Figura 4.1: Aspectos técnicos evaluados en el control de calidad de ramos de flores, tradicionalmente medidos o estimados mediante inspección manual.

4.4. Arquitectura funcional del sistema

El sistema desarrollado puede entenderse como una arquitectura integrada de cinco bloques: plataforma física móvil, adquisición multimodal, procesamiento de imágenes, inferencia algorítmica y gestión de resultados. Esta estructura fue concebida para operar tanto en camas de flores como en bouquetera y almacén o poscosecha, es decir, en escenarios con variabilidad ambiental, lumínica y operativa [2].

Si se denota por R_k el ramo número k , el flujo funcional del sistema puede resumirse mediante la transformación

$$R_k \longrightarrow \mathcal{A}_k \longrightarrow \mathcal{P}_k \longrightarrow \hat{\mathbf{q}}_k \longrightarrow \mathcal{G}_k,$$

donde \mathcal{A}_k representa la adquisición de datos del ramo, \mathcal{P}_k el preprocesamiento y la organización de la información visual, $\hat{\mathbf{q}}_k$ el vector estimado de parámetros de calidad y \mathcal{G}_k la generación de salidas en forma de interfaz, registro e integración con el sistema de información. Esta secuencia expresa que el sistema no es solo un equipo de captura, sino una cadena de producción de dato de calidad.

La etapa de adquisición se apoya en tres vistas principales del ramo: una vista de flores, una vista de tallos y una vista superior. Según los manuales y el informe de práctica, estas vistas se presentan en la interfaz de captura como tres marcos independientes, permitiendo al usuario verificar visualmente que la escena esté bien posicionada antes de activar la captura [2], [5]. La etapa posterior de control de calidad despliega cuatro marcos: tallos segmentados, flores delimitadas, zona foliar-elástico y longitud del ramo, además de una columna con los parámetros medidos y datos complementarios de trazabilidad [5].

Desde la perspectiva de diseño, esta arquitectura refleja una decisión adecuada: distri-

buir el problema de medición en vistas especializadas, en lugar de exigir que una sola imagen resuelva simultáneamente todas las variables. Tal partición mejora la estabilidad del sistema, simplifica la inferencia y hace más transparente la relación entre medición y evidencia visual.

La Fig. 4.2 presenta la arquitectura funcional del sistema, desde la captura multimodal hasta la gestión de resultados.

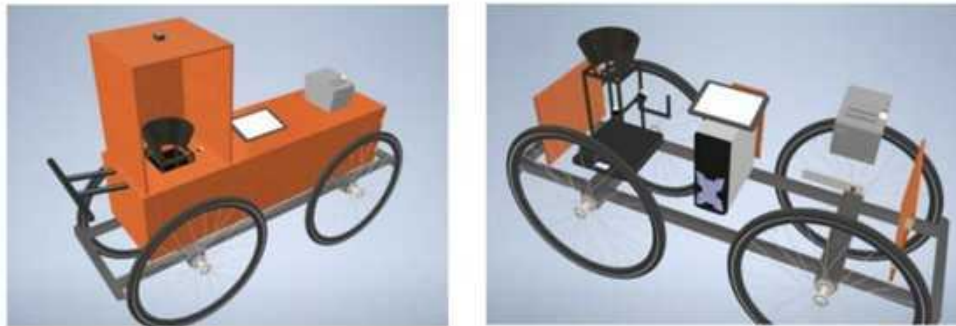


Figura 4.2: Arquitectura funcional del sistema, mostrando la captura multimodal del ramo, el procesamiento visual, la estimación de parámetros de calidad y la gestión de resultados.

4.5. Diseño del sistema y selección tecnológica

El diseño del sistema atravesó una etapa exploratoria en la cual se definieron los alcances, los requerimientos funcionales y las restricciones de operación. En esta fase se estudió el procedimiento manual vigente, se identificaron las variables críticas del control de calidad y se estableció que el dispositivo debía ser móvil, resistente, fácil de usar y apto para desplazarse en terreno irregular e incluso para ser suspendido por garrucha en determinados escenarios [1], [2].

Las primeras pruebas de concepto exploraron configuraciones con el ramo en posición vertical y horizontal. La comparación entre prototipos mostró que la orientación horizontal facilitaba la captura de imágenes y la medición de los parámetros relevantes, por lo que el prototipo final adoptó esta disposición [2]. A partir de ello, se redefinió la estructura física, se ajustó la distribución de dispositivos y se consolidó el montaje actual.

En cuanto a hardware, el sistema integró pantalla táctil, báscula digital, sistema de iluminación, cámara a color para vista superior, cámaras estereoscópicas para vistas laterales, teclado con ratón integrado, impresora y elementos de adquisición de peso y control [2],

[5]. La selección de cámaras Intel RealSense D405 resultó relevante debido a su capacidad para medir profundidad y filtrar escenas por distancia, reduciendo ruido de fondo y facilitando tareas como segmentación de tallos y delimitación de floración [2].

Para explicitar el criterio de medición geométrica, conviene señalar que el sistema trabaja sobre imágenes RGB y sobre mapas de profundidad. Si se denota por $z(x, y)$ la profundidad medida en un píxel ubicado en la coordenada (x, y) , y por Ω la región de interés correspondiente a un objeto, entonces la profundidad media de esa región puede expresarse como

$$\bar{z}_\Omega = \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} z(x, y),$$

donde $|\Omega|$ representa el número de píxeles pertenecientes a la región. Esta operación permite aislar objetos por cercanía y estabilizar la medición posterior sobre regiones específicas del ramo.

La Fig. 4.3 muestra la evolución física del prototipo y la configuración móvil adoptada para condiciones reales de uso.



Figura 4.3: Diseño conceptual y evolución física del prototipo, desde las pruebas iniciales hasta la configuración móvil orientada al uso en condiciones reales de campo.

4.6. Procesamiento de imágenes y modelos de inferencia

El sistema fue implementado en Python y articuló bibliotecas para adquisición de imágenes, procesamiento visual, sincronización de dispositivos, manejo de interfaz y ejecución de modelos de aprendizaje automático [2], [5]. En términos metodológicos, el procesamiento se apoya en un principio fundamental: cada parámetro del control de calidad debe asociarse a una vista adecuada y a una cadena de tratamiento específica.

En el caso de tallos y flores, las cámaras estereoscópicas permiten generar mapas de profundidad con los cuales se filtran objetos fuera del rango de interés. Posteriormente, las imágenes se convierten a arreglos numéricos y se aplican operaciones de umbralado, apertura y cierre morfológico, suavizado y determinación de regiones de interés [2]. En la vista superior, la cámara RGB de alta resolución captura la escena completa del ramo

y dicha imagen se somete a recortes, redimensionamientos y acondicionamientos para generar entradas consistentes hacia los modelos.

Si la imagen original se denota por $I(x, y)$, el proceso de segmentación puede entenderse como la construcción de una máscara binaria $M(x, y)$, definida como

$$M(x, y) = \begin{cases} 1, & \text{si el píxel pertenece al objeto de interés,} \\ 0, & \text{en otro caso.} \end{cases}$$

A partir de esta máscara, es posible estimar áreas, contornos, diámetros y distancias. Por ejemplo, si A representa el área proyectada del objeto en la imagen segmentada, entonces

$$A = \sum_{(x,y)} M(x, y),$$

que en términos discretos equivale a la suma de píxeles pertenecientes a la región de interés. En el sistema real, esta magnitud puede escalarse a unidades físicas a partir de calibración geométrica o del uso de profundidad.

El proyecto utilizó modelos YOLO tanto para detección como para segmentación. Según la documentación disponible, se entrenaron modelos específicos para longitud del ramo, follaje-elástico y tallos, con el fin de inferir automáticamente las variables de interés a partir de imágenes previamente etiquetadas en Roboflow y luego entrenadas en Google Colab [2]. La modularidad de los modelos fue una decisión adecuada, pues evitó exigir a una sola red la resolución simultánea de tareas heterogéneas.

El informe de práctica reporta, para el modelo de tallos, métricas de desempeño elevadas: precisión de detección de 0.992, *recall* de 0.972, mAP50 de 0.992 y precisión de segmentación de 0.994, con mAP50 de 0.993 [2]. Sin sobrecargar este capítulo con todas las tablas, estos resultados permiten afirmar que el sistema alcanzó un desempeño robusto para uso aplicado en campo.

La Fig. 4.4 presenta ejemplos del procesamiento visual y de las salidas de inferencia empleadas por el sistema.

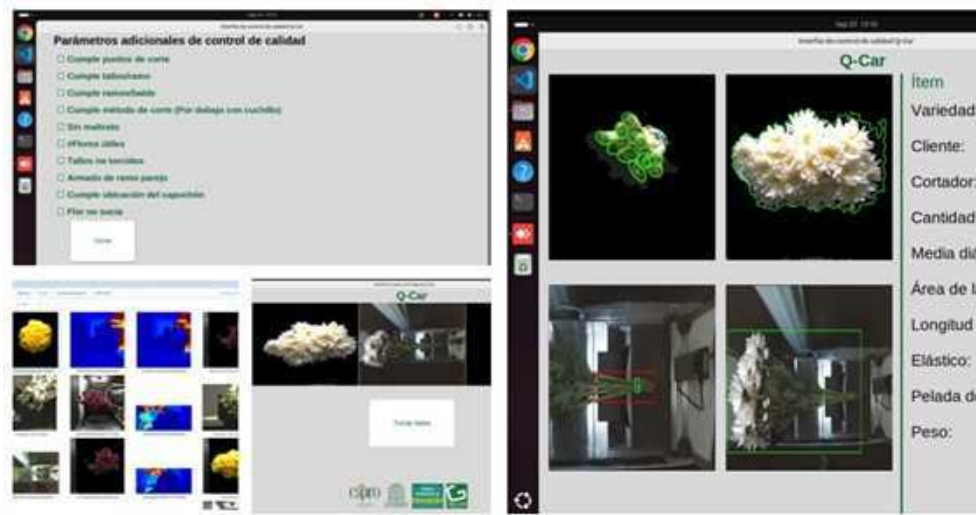


Figura 4.4: Ejemplos de captura, preprocesamiento, etiquetado e inferencia visual en las diferentes vistas empleadas por el sistema de control de calidad.

4.7. Interfaz de usuario, trazabilidad y almacenamiento

Una contribución importante del sistema no reside solamente en la medición automática, sino en la forma en que el dato es puesto a disposición del operario y del sistema productivo. El software asociado organiza el proceso en dos módulos de interacción: una interfaz de captura y una interfaz de control de calidad [5]. En la primera, el usuario verifica las tres vistas del ramo y activa la toma de datos; en la segunda, observa las imágenes resultantes del procesamiento y los parámetros medidos, además de registrar o seleccionar datos complementarios como variedad, cliente y cortador [2], [5].

Desde el punto de vista funcional, esta organización cumple dos propósitos. El primero es reducir errores de operación, al permitir que el usuario confirme visualmente la posición del ramo antes de la captura. El segundo es consolidar la trazabilidad, pues cada inspección queda asociada a imágenes y a datos estructurados almacenados en carpetas organizadas por fecha y en un archivo Excel con enlaces a las imágenes procesadas [5], [6].

Si se denota por \mathcal{D}_k el registro digital asociado al ramo k , entonces dicho registro puede expresarse como

$$\mathcal{D}_k = \{\hat{\mathbf{q}}_k, \mathbf{m}_k, \mathbf{i}_k, t_k\},$$

donde $\hat{\mathbf{q}}_k$ representa el vector estimado de parámetros de calidad, \mathbf{m}_k los metadatos de trazabilidad, \mathbf{i}_k el conjunto de imágenes asociadas a la inspección y t_k la marca temporal del registro.

Tal enfoque convierte el control de calidad en una fuente de datos productivos. En lugar de agotar su función en la decisión inmediata de conformidad, el proceso genera una base histórica útil para identificar patrones, entrenar nuevos modelos y fortalecer la toma de decisiones a nivel operativo y gerencial [4], [6].

La Fig. 4.5 resume la relación entre interfaz de captura, resultados y trazabilidad de las inspecciones.

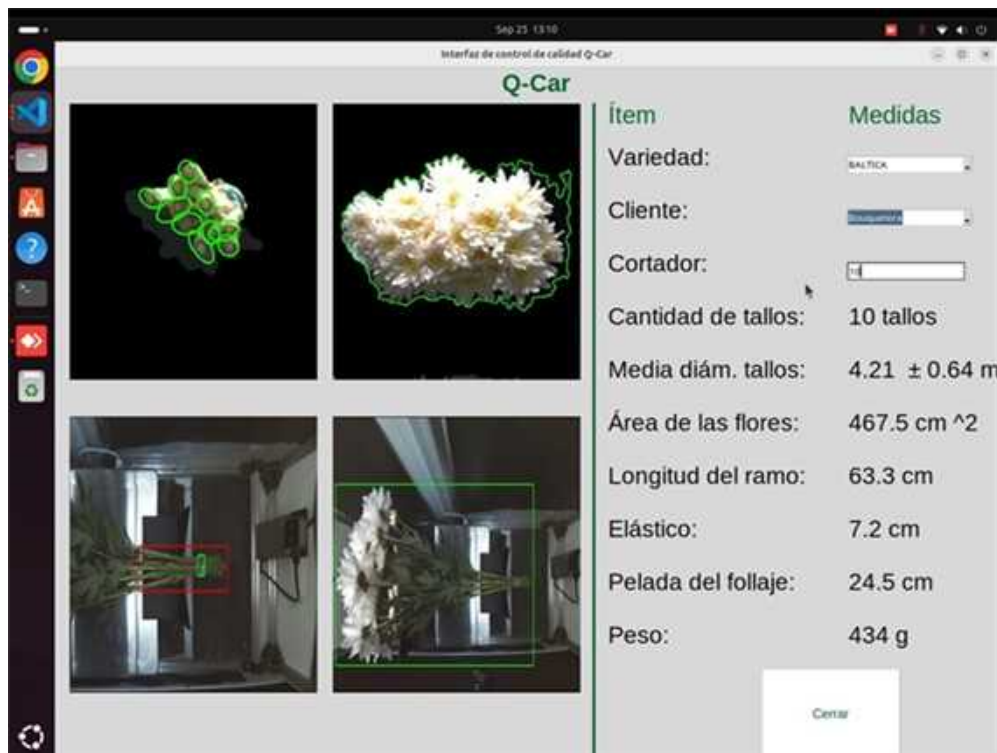


Figura 4.5: Interfaz de captura, interfaz de resultados y esquema general de almacenamiento y trazabilidad de las inspecciones realizadas por el sistema.

4.8. Validación en campo e innovación tecnológica

El valor del desarrollo no puede juzgarse únicamente por la sofisticación del algoritmo o por la calidad del prototipo en laboratorio. Su relevancia depende de la validación en condiciones reales de operación. En ese sentido, la evidencia documental disponible muestra que el sistema fue diseñado, ajustado y probado en escenarios de uso de Flores

El Capiro S.A., incluyendo campo, bouquetera y almacén, y que a partir de esas pruebas se incorporaron mejoras funcionales y operativas [2], [3].

El informe final del proyecto señala que el prototipo alcanzó un nivel de madurez tecnológica TRL 7, correspondiente a demostración del sistema en entorno operativo real [3]. Allí también se consigna que las pruebas permitieron verificar precisión en medición, ajuste de movilidad en terrenos irregulares, mejora en algoritmos y validación del uso por parte de operarios [3]. Del mismo modo, documentos posteriores de innovación tecnológica subrayan que el sistema sustituye un esquema manual, subjetivo y no trazable por uno automatizado, estandarizado y trazable, con reducción de tiempos de medición y generación de bases de datos estructuradas [4], [6].

Este punto permite sostener que el desarrollo corresponde a una innovación tecnológica y no a una mera adaptación incremental. La novedad no está en usar cámaras, sensores o modelos conocidos de forma aislada, sino en integrarlos sistémicamente en una plataforma móvil específica para el control de calidad de ramos, validada bajo condiciones reales del sector floricultor [4]. En otras palabras, la innovación emerge de la articulación funcional de subsistemas heterogéneos para transformar una práctica productiva.

Para expresar esa mejora de forma cuantitativa general, si se denota por T_m el tiempo promedio requerido para una inspección manual y por T_a el tiempo promedio de inspección automatizada, entonces la ganancia relativa de eficiencia puede definirse como

$$\eta_T = \frac{T_m - T_a}{T_m} \times 100 \%$$

Aquí, η_T representa el porcentaje de reducción del tiempo de inspección. Aunque el efecto del sistema no se agota en esta métrica, la expresión captura una de sus contribuciones más importantes: hacer que el dato llegue más rápido y con menor variabilidad al sistema productivo.

4.9. Discusión

El desarrollo del carro de calidad muestra cómo la ingeniería aplicada puede intervenir una operación aparentemente rutinaria y convertirla en un sistema de captura, medición y decisión basado en datos. La inspección manual de ramos puede parecer una tarea simple, pero en realidad concentra dificultades de geometría variable, condiciones lumínicas cambiantes, diversidad de variedades, necesidad de movilidad y exigencia de rapidez en un contexto productivo real. Resolver ese conjunto de restricciones exigió una mirada

sistémica.

Uno de los méritos del proyecto es haber evitado la tentación de resolver todo con una sola técnica. El sistema combina cámaras RGB, profundidad, preprocesamiento geométrico, modelos especializados, sensado de peso, estructura móvil y gestión de interfaz. Esta combinación parece más compleja que una solución monolítica, pero es precisamente lo que hace viable la robustez operativa. Cada subsistema asume una parte bien delimitada del problema.

Otro aspecto relevante es el valor técnico del dato generado. En el esquema manual, el control de calidad produce decisiones puntuales y registros sujetos a error de transcripción. En el esquema propuesto, la inspección genera un rastro digital completo, reutilizable y auditable. Esta diferencia cambia la naturaleza del proceso: ya no se trata solo de aprobar o rechazar un ramo, sino de convertir cada inspección en una unidad de información productiva.

Además, la plataforma deja abierta una ruta clara de crecimiento. Los documentos del proyecto señalan la posibilidad de incorporar variables adicionales, extender el uso a otras especies, adaptar el sistema a otros contextos agroindustriales y reutilizar módulos de software y de inferencia para nuevas aplicaciones [4], [6]. En esa medida, el desarrollo no se cierra sobre sí mismo, sino que se constituye en plataforma.

4.10. Conclusiones

El desarrollo del sistema móvil de visión artificial para control de calidad de ramos demuestra que es posible automatizar una etapa crítica de la poscosecha floricultora mediante una integración efectiva de diseño mecánico, sensado físico, procesamiento digital de imágenes, aprendizaje automático e interfaz de usuario.

El problema abordado fue formulado como un problema de medición estandarizada en condiciones reales, y no simplemente como una tarea de captura fotográfica. Esta diferencia de enfoque permitió diseñar una solución que no solo observa ramos, sino que estima variables relevantes, apoya decisiones de conformidad y genera trazabilidad digital para la empresa.

Desde la perspectiva metodológica, el proyecto evidencia una transición ordenada desde pruebas de concepto, exploración de configuraciones y selección tecnológica, hasta implementación operativa y validación en entorno productivo. El hecho de que el sistema haya alcanzado validación de campo y un nivel TRL 7 refuerza su valor como desarrollo tecnológico aplicado.

Finalmente, el caso muestra que la automatización en floricultura no debe entenderse únicamente como reemplazo de trabajo manual, sino como creación de infraestructura de datos y capacidad analítica para el proceso. En este sentido, el desarrollo constituye una contribución relevante tanto para la modernización del sector como para la consolidación de una línea de investigación y transferencia tecnológica en visión artificial aplicada.

Referencias

- [1] Universidad de Antioquia, *Ficha técnica del proyecto Carro de Calidad para ramos de flores – Prototipo*, Vicerrectoría de Investigación, Centro de Investigación Facultad de Ingeniería, feb. de 2022.
- [2] M. Á. R. Rueda, «Desarrollo de Q-Car (Quality Car), carro para el control de calidad automático de ramos de flores,» Universidad de Antioquia, Facultad de Ingeniería, Informe de práctica, sep. de 2025.
- [3] D. Fernández, *Q-CAR: Carro de calidad para ramos de flores – Avance #3 e informe final*, Informe final del proyecto, Universidad de Antioquia, dic. de 2024.
- [4] Universidad de Antioquia, *Formulario de desarrollo tecnológico: Carrito de calidad para ramos de flores*, Documento de soporte para innovación tecnológica, 2025.
- [5] Universidad de Antioquia, *Q-Car: Software para carro de calidad para ramos de flores. Manual de usuario*, Manual técnico y de usuario del sistema, 2024.
- [6] D. S. F. McCann, *Solicitud para asignación de puntos por desarrollo de software Q-Car*, Documento institucional, Universidad de Antioquia, feb. de 2025.

Capítulo 5

Aplicación móvil para la gestión y trazabilidad de videos en el monitoreo de plántulas

Autores

David Stephen Fernández Mc Cann¹; Mateo Alejandro Bravo Revelo²; Jhon Fredy García Hincapié³

Resumen

La gestión ordenada de videos capturados en campo constituye una etapa crítica en sistemas de visión artificial aplicados al monitoreo agrícola. En el caso de SICOP, sistema orientado al conteo de plántulas mediante captura de video, generación de panorámicas e individualización analítica de plantas, la calidad del dato de entrada depende de que cada video sea obtenido, editado, almacenado, etiquetado y sincronizado de manera trazable. En este capítulo se presenta una aplicación móvil Android desarrollada para gestionar videos de monitoreo de plántulas en campo. La herramienta permite seleccionar finca, bloque y cama; administrar videos locales; editar atributos técnicos; conectarse por WiFi

¹Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Ingeniero en desarrollo de software asociado al proyecto, Universidad de Antioquia.

³Director de Innovación Tecnológica y Nuevos Negocios, Flores El Capiro S.A.

³Desarrollo realizado en el marco del convenio FIT2023-59235 entre la Universidad de Antioquia y Flores El Capiro S.A., con participación del grupo GEPAR.

a una cámara Insta360 X3; descargar archivos al dispositivo móvil; y sincronizar la información con servidores de Flores El Capiro S.A. La aplicación fue implementada en Kotlin sobre Android Studio y emplea librerías para persistencia local, comunicación con APIs, edición de medios y conectividad con cámara externa. Su aporte principal no consiste en realizar directamente el conteo de plántulas, sino en cerrar la brecha entre la captura audiovisual en campo y la disponibilidad organizada del dato para los módulos posteriores de SICOP. El desarrollo cuenta con documentación técnica, certificado de uso y registro ante la Dirección Nacional de Derecho de Autor, lo que respalda su condición de resultado tecnológico aplicado [1], [2], [3], [4], [5].

Palabras clave: SICOP; aplicación móvil; trazabilidad; videos agrícolas; Android; Kotlin; visión artificial; monitoreo de plántulas; agricultura de precisión; GEPAR.

5.1. Introducción

La transformación digital de los procesos agrícolas no depende únicamente de algoritmos avanzados de visión artificial o de plataformas analíticas de alto desempeño. En muchos casos, el cuello de botella se encuentra en la etapa anterior: la obtención ordenada, estandarizada y trazable del dato en campo. Esta observación resulta particularmente relevante en sistemas como SICOP, cuyo objetivo es contar plántulas en camas de cultivo mediante una cadena compuesta por captura de video, construcción de panorámicas e individualización analítica de plántulas [4], [6].

En la segunda fase del proyecto SICOP se reconoció que la utilidad del sistema completo dependía no solo de los módulos panorámico y analítico, sino también de la calidad operativa del proceso de captura y del manejo posterior de los videos. En efecto, el acta de inicio y propiedad intelectual del proyecto estableció como objetivo general el diseño y construcción de una segunda etapa de SICOP que incorporara nuevos equipos y protocolos para agilizar y generalizar su uso en un ambiente productivo [7]. Dentro de esa lógica, surgió la necesidad de desarrollar una aplicación móvil que permitiera gestionar de forma estructurada los videos capturados en campo, asociarles metadatos de ubicación, editarlos, almacenarlos localmente, sincronizarlos con servidores de la empresa e integrarse con una cámara Insta360 X3 para captura remota [1], [5].

El desarrollo resultante fue una aplicación Android orientada a la gestión integral de videos de monitoreo de plántulas. Su papel dentro de la arquitectura de SICOP no es accesorio. Por el contrario, constituye la interfaz operativa que convierte un procedimiento de captura potencialmente disperso y manual en una secuencia digital estructurada, repetible y trazable. Desde el punto de vista funcional, la aplicación permite seleccionar finca,

bloque y cama; ajustar propiedades técnicas del video; gestionar listas de grabaciones; conectarse por WiFi a una cámara Insta360 X3; descargar videos al dispositivo móvil; y sincronizar la información con los servidores de Flores El Capiro S.A. [4], [5].

El presente capítulo expone este desarrollo como una solución de ingeniería de software aplicada a agricultura de precisión. El interés principal no consiste en describir una interfaz aislada, sino en mostrar cómo una aplicación móvil puede formularse como un componente estructural de una cadena de visión artificial, garantizando integridad del dato, consistencia de metadatos, trazabilidad y sincronización operativa entre campo y servidor.

5.2. Contexto del problema y motivación del desarrollo

SICOP es un sistema de conteo de plántulas compuesto por tres etapas articuladas: captura, panorámica y analítica. La primera tiene como salida un video útil de una cama de plántulas; la segunda produce una imagen panorámica segmentada de alta resolución; y la tercera individualiza plántulas y genera estadísticas morfológicas y de conteo [6]. Esta estructura, aunque conceptualmente robusta, depende críticamente de que los videos de entrada sean obtenidos con calidad suficiente, nombrados de forma coherente, asociados a la cama correcta y puestos a disposición del sistema central sin ambigüedad.

Antes del desarrollo de la aplicación móvil, esta fase podía verse afectada por una combinación de dificultades prácticas: almacenamiento desestructurado de archivos, asociación incompleta de metadatos, dispersión de formatos, dependencia de acciones manuales repetitivas y ausencia de una ruta uniforme para transferir el video capturado hacia el flujo de procesamiento posterior. En términos de sistemas de información, el problema puede formularse como una falta de acoplamiento entre el evento físico de captura y el registro digital formal requerido por la analítica posterior.

Sea v_k el video correspondiente a la cama número k , y sea \mathbf{m}_k el vector de metadatos asociados a dicho video. En el caso de SICOP, este vector puede representarse como

$$\mathbf{m}_k = [f_k \quad b_k \quad c_k \quad t_k \quad r_k \quad h_k \quad \phi_k]^\top,$$

donde f_k representa la finca, b_k el bloque, c_k la cama, t_k la marca temporal de captura, r_k la resolución del video, h_k la velocidad de fotogramas y ϕ_k un conjunto de atributos de edición como orientación o recorte temporal. La función de la aplicación es garantizar

que el par (v_k, \mathbf{m}_k) se genere y preserve de manera coherente.

Desde esta perspectiva, el problema no se reduce a grabar y subir videos. Se trata de construir un sistema móvil capaz de transformar una secuencia audiovisual en un objeto digital estructurado, apto para ser incorporado al flujo técnico de SICOP. Si se define por \mathcal{R}_k el registro digital de la captura k , entonces

$$\mathcal{R}_k = (v_k, \mathbf{m}_k, s_k),$$

donde s_k representa el estado del registro dentro del flujo operativo, por ejemplo: pendiente, listo para cargar, cargado exitosamente o con error. Esta estructura aparece reflejada en la interfaz de “Mis Videos” de la aplicación, donde cada video se presenta con su metadato esencial y con un estado de gestión visible para el usuario [4].

5.3. Inserción de la aplicación dentro de la arquitectura SICOP

La aplicación móvil debe entenderse como un componente funcional de la etapa de captura de SICOP. Según el manual del método de captura, el sistema completo requiere un procedimiento estandarizado basado en un arnés lateral, una barra de soporte de altura, un estabilizador tipo gimbal, un celular configurado a 1080p FHD y 60 fps, y testigos visuales azules para marcar inicio y fin de la cama [6], [8]. El objetivo de este procedimiento es producir un video útil para el módulo Panorámica.

Sin embargo, una vez obtenido el video, se requiere una capa adicional que permita incorporarlo al flujo digital de manera organizada. Allí se ubica la aplicación desarrollada. Conceptualmente, la arquitectura completa puede expresarse como la composición de tres niveles:

$$\mathcal{S}_{\text{SICOP}} = \mathcal{C} \circ \mathcal{G} \circ \mathcal{A},$$

donde \mathcal{A} representa la etapa de adquisición y gestión móvil de videos, \mathcal{G} la generación de panorámicas y \mathcal{C} la etapa de conteo y analítica. En esta notación, la aplicación móvil forma parte esencial de \mathcal{A} , porque no solo participa en la captura, sino también en el enriquecimiento del video mediante metadatos, edición y sincronización.

La documentación del método de captura muestra que el último paso del procedimiento consiste en descargar el video del celular y cargarlo al módulo de software SICOP [8].

La aplicación móvil materializa ese puente operativo entre el dispositivo de captura y el ecosistema software del sistema completo. Esto justifica que, aunque el capítulo se centre en una aplicación, su relevancia deba entenderse a la escala del sistema de conteo de plántulas.

La Fig. 5.1 ubica funcionalmente la aplicación dentro de la arquitectura general de SICOP.

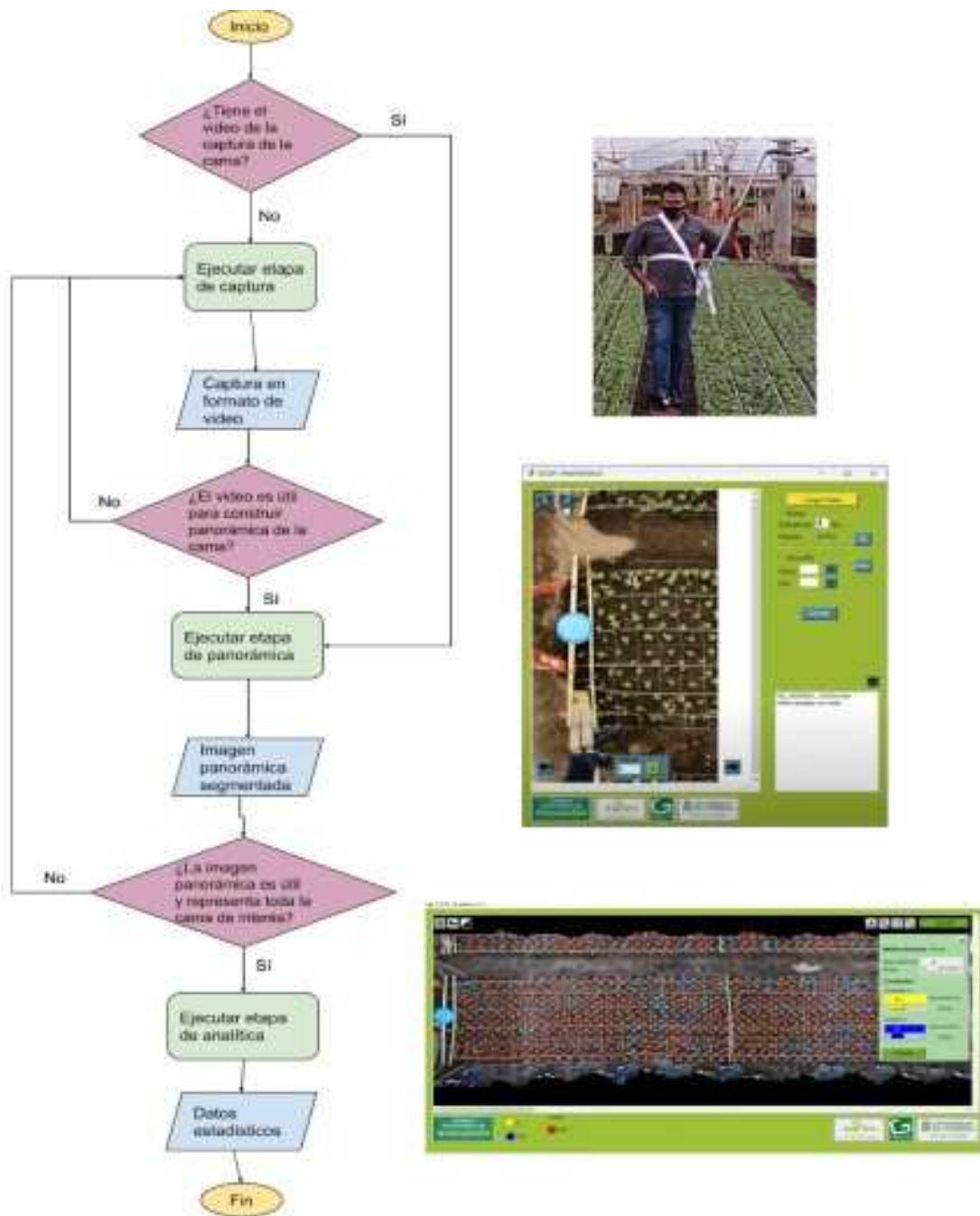


Figura 5.1: Ubicación funcional de la aplicación móvil dentro de la arquitectura general de SICOP, articulando la etapa de captura con la generación de panorámicas y la analítica de plántulas.

5.4. Arquitectura funcional de la aplicación

La documentación de usuario y el manual técnico coinciden en que la aplicación fue diseñada bajo una arquitectura modular compuesta por cuatro bloques principales: módulo de ingreso, módulo de gestión de videos, módulo de edición y módulo de conexión con cámara Insta360 X3 [4], [5]. Esta descomposición funcional reduce la complejidad de la interacción humana y permite distribuir responsabilidades dentro del sistema.

Si se denota por u el usuario, por V el conjunto de videos locales y por D el conjunto de datos sincronizados desde la empresa, el flujo principal de la aplicación puede describirse como

$$u \longrightarrow D_f \longrightarrow D_b \longrightarrow V \longrightarrow V' \longrightarrow S,$$

donde D_f corresponde a la sincronización de fincas, D_b a la sincronización de bloques, V al conjunto de videos capturados o seleccionados, V' al conjunto de videos editados y enriquecidos con metadatos, y S al estado final de sincronización con servidor.

En el módulo de ingreso, el usuario selecciona la finca y sincroniza la información disponible en servidores de Flores El Capiro S.A. [4]. En el módulo de gestión de videos, la aplicación muestra los videos registrados y permite conocer su estado de carga. En el módulo de edición, el usuario puede asignar bloque, cama, resolución, ancho, alto y FPS, además de recortar y rotar el video. Finalmente, el módulo de conexión con Insta360 permite establecer comunicación por WiFi, iniciar y detener la grabación, y descargar el archivo generado al almacenamiento del dispositivo [4], [5].

Desde el punto de vista de diseño de software, esta organización modular puede modelarse como una partición del conjunto de funcionalidades \mathcal{F} :

$$\mathcal{F} = \mathcal{F}_{\text{ingreso}} \cup \mathcal{F}_{\text{gestion}} \cup \mathcal{F}_{\text{edicion}} \cup \mathcal{F}_{\text{captura}},$$

con intersecciones controladas entre módulos. Esta estrategia reduce el acoplamiento, mejora el mantenimiento y facilita la escalabilidad funcional del sistema.

La Fig. 5.2 resume la estructura modular de la aplicación.

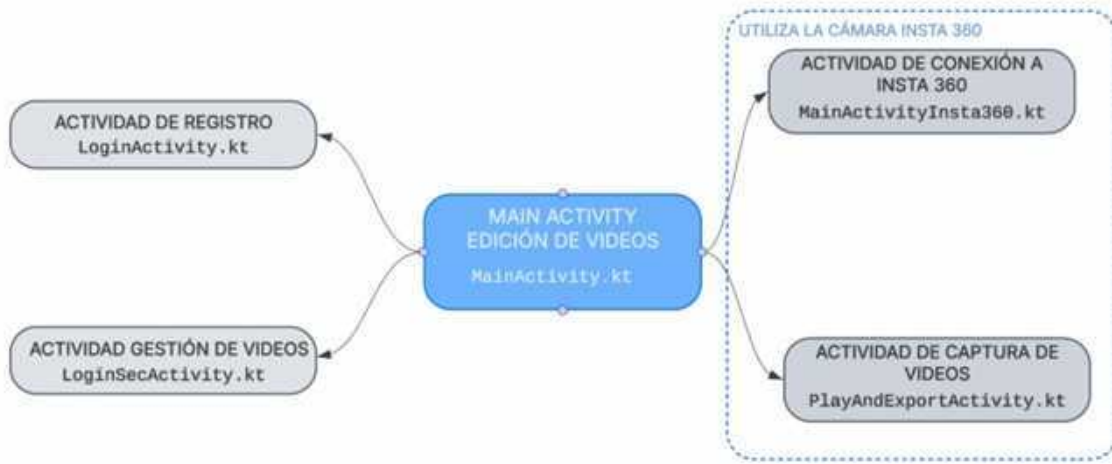


Figura 5.2: Estructura modular de la aplicación, incluyendo ingreso, gestión de videos, edición y conexión con cámara Insta360 X3.

5.5. Modelo de datos y trazabilidad operativa

Uno de los aportes más relevantes de la aplicación es la formalización del dato audiovisual como un objeto estructurado. En escenarios manuales, un video puede existir como archivo aislado, sin vínculo explícito con el lugar de captura ni con las condiciones técnicas bajo las que fue generado. La aplicación resuelve este problema imponiendo una estructura de registro.

Sea r_i el registro correspondiente al video i . Este registro puede modelarse como la 5-upla

$$r_i = (p_i, \mathbf{m}_i, e_i, \tau_i, \sigma_i),$$

donde p_i es la referencia al archivo de video, \mathbf{m}_i es el vector de metadatos de ubicación y características técnicas, e_i representa los parámetros de edición aplicados, τ_i es la marca temporal del registro y σ_i es su estado de sincronización. Esta formulación refleja la intención del diseño: cada video no es solo un medio digital, sino una entidad informacional trazable.

El vector de metadatos puede reescribirse como

$$\mathbf{m}_i = [\text{finca}_i \quad \text{bloque}_i \quad \text{cama}_i \quad \text{fecha}_i \quad \text{hora}_i]^\top,$$

mientras que el vector de características técnicas puede expresarse como

$$\mathbf{e}_i = [W_i \ H_i \ F_i \ \theta_i \ \Delta t_i]^\top,$$

donde W_i y H_i corresponden a ancho y alto del video, F_i a los cuadros por segundo, θ_i a la orientación aplicada y Δt_i al intervalo temporal conservado tras el recorte. Este nivel de estructuración mejora significativamente la integridad del flujo de datos.

En el módulo “Mis Videos”, el usuario no solo visualiza el archivo, sino también su estado dentro del sistema: listo para cargar, cargado exitosamente o con error [4]. Por tanto, el estado σ_i pertenece a un conjunto discreto

$$\sigma_i \in \{\text{pendiente, listo, cargado, error}\}.$$

Este mecanismo de estados es particularmente valioso para trazabilidad operativa, ya que convierte un proceso potencialmente opaco en uno auditable.

5.6. Modelo de interacción y flujo de uso

El manual técnico presenta los requerimientos funcionales del sistema mediante cuatro grupos principales: gestión de videos, gestión de ubicación, edición de videos y conexión con cámara externa [5]. Esta clasificación puede traducirse a una descripción formal del flujo de uso.

Sea \mathcal{U} el espacio de acciones posibles del usuario. Entonces el flujo de interacción elemental puede expresarse como

$$a_1 \rightarrow a_2 \rightarrow a_3 \rightarrow a_4 \rightarrow a_5,$$

donde

$$\begin{aligned} a_1 &= \text{seleccionar finca,} & a_2 &= \text{sincronizar bloques,} \\ a_3 &= \text{capturar o seleccionar video,} & a_4 &= \text{editar y etiquetar,} \\ a_5 &= \text{subir al servidor.} \end{aligned}$$

En términos más generales, si x_k representa el estado del sistema en el paso k y u_k la acción del usuario, la evolución operativa puede modelarse como

$$x_{k+1} = f(x_k, u_k),$$

donde $f(\cdot)$ describe la lógica de transición entre pantallas, validaciones y actualizaciones de base de datos. Esta ecuación permite entender que la aplicación implementa una máquina de estados finitos con transiciones sujetas a condiciones de validez, por ejemplo: no se puede ingresar si no hay finca seleccionada, ni agregar videos si no hay bloques sincronizados [4], [5].

La interfaz de usuario fue diseñada con base en un criterio de simplicidad operativa. El módulo de ingreso contiene un menú desplegable de fincas y dos botones; el módulo “Mis Videos” concentra el estado de los registros y las acciones de sincronización; el módulo de edición organiza los metadatos técnicos y de ubicación; y el módulo de Insta360 expone solo las acciones esenciales de conexión, grabación y descarga [4]. Esta economía de interfaz es coherente con escenarios de trabajo en campo, donde el usuario requiere un flujo robusto y directo más que una interfaz sobrecargada.

La Fig. 5.3 presenta la secuencia de interacción de la aplicación.

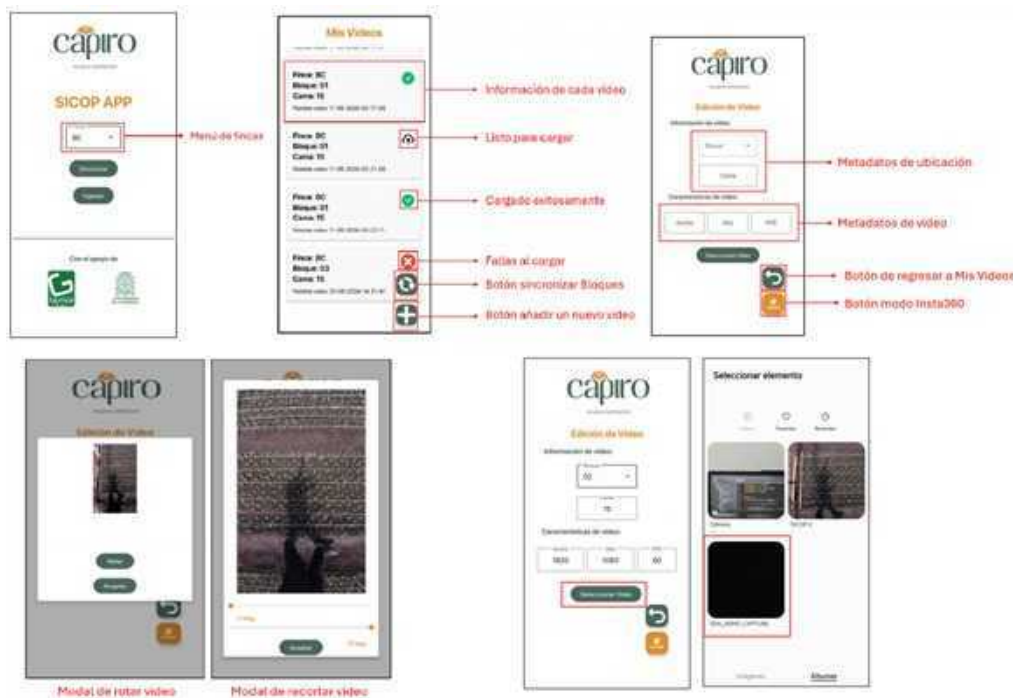


Figura 5.3: Secuencia de interacción de la aplicación, desde el ingreso y sincronización de finca hasta la edición, gestión y envío de videos al servidor.

5.7. Integración con captura remota mediante Insta360 X3

Uno de los elementos de novedad más importantes del desarrollo es la integración directa con una cámara Insta360 X3 mediante WiFi [1], [4]. Esta funcionalidad permite desacoplar parcialmente la adquisición audiovisual del control físico inmediato del dispositivo móvil, habilitando captura remota, descarga posterior y continuidad del flujo dentro de la misma aplicación.

Desde el punto de vista del sistema, esta integración puede interpretarse como la incorporación de una fuente externa de video \mathcal{V}_{ext} al conjunto de entradas del sistema móvil. Si $v_i^{(m)}$ es un video proveniente de la galería del dispositivo y $v_i^{(c)}$ uno proveniente de la cámara externa, entonces el conjunto de videos gestionables por la aplicación puede escribirse como

$$\mathcal{V} = \mathcal{V}_m \cup \mathcal{V}_c.$$

La conexión con la cámara requiere establecer un enlace WiFi, verificar el estado de la conexión y luego habilitar las acciones de inicio y finalización de captura [4], [5]. Si se denota por $\gamma(t) \in \{0, 1\}$ el estado de la conexión en el tiempo t , donde $\gamma = 1$ indica conexión activa, entonces la función de captura solo es admisible si

$$\gamma(t) = 1.$$

Una vez detenida la captura, el sistema habilita la descarga del archivo y lo almacena localmente en un directorio específico [4]. Esto significa que el video capturado por la cámara se incorpora al flujo normal de gestión y edición. En términos operativos, la aplicación resuelve la composición de dos tareas que frecuentemente aparecen separadas en entornos reales: capturar y administrar.

5.8. Arquitectura técnica de implementación

El manual técnico documenta que la aplicación fue desarrollada en Kotlin sobre Android Studio, y que emplea un conjunto de librerías orientadas a interfaz, persistencia local, comunicación con APIs, procesamiento de medios y conectividad con cámara externa [5]. Entre ellas se mencionan Retrofit, Gson Converter, Room, Mobile FFmpeg, OkGo,

Permission X y los SDK de Insta360 [5].

La arquitectura interna está organizada alrededor de actividades principales como `LoginActivity.kt`, `LoginSecActivity.kt`, `MainActivity.kt`, `MainActivityInsta360.kt` y `CaptureActivity.kt` [5]. Esta organización responde a una separación por responsabilidades: autenticación contextual, gestión de videos, edición, conexión con cámara y captura.

Desde una perspectiva abstracta, la arquitectura puede representarse como un grafo dirigido

$$\mathcal{G} = (\mathcal{N}, \mathcal{E}),$$

donde \mathcal{N} es el conjunto de actividades y \mathcal{E} el conjunto de transiciones entre ellas. Si se define

$$\mathcal{N} = \left\{ \begin{array}{l} \text{LoginActivity,} \\ \text{LoginSecActivity,} \\ \text{MainActivity,} \\ \text{MainActivityInsta360,} \\ \text{CaptureActivity} \end{array} \right\},$$

entonces las aristas de \mathcal{E} representan navegación permitida, transferencia de parámetros y continuidad del contexto de trabajo. Este modelo permite entender la aplicación como una máquina de navegación dirigida, no como una colección dispersa de pantallas.

El almacenamiento local usa Room como capa ORM sobre SQLite, lo que permite persistencia segura y organizada de los videos y sus metadatos [5]. Si se denota por \mathcal{B} la base de datos local y por `VideoDao` la capa de acceso, entonces la operación de recuperación del conjunto de videos puede representarse simbólicamente como

$$V = \text{VideoDao.getAll}()(\mathcal{B}).$$

Este tipo de estructura resulta apropiado para operación móvil, pues permite trabajar aun cuando la sincronización con servidor no se haya ejecutado todavía.

La Fig. 5.4 resume la arquitectura técnica y los principales casos de uso de la aplicación.

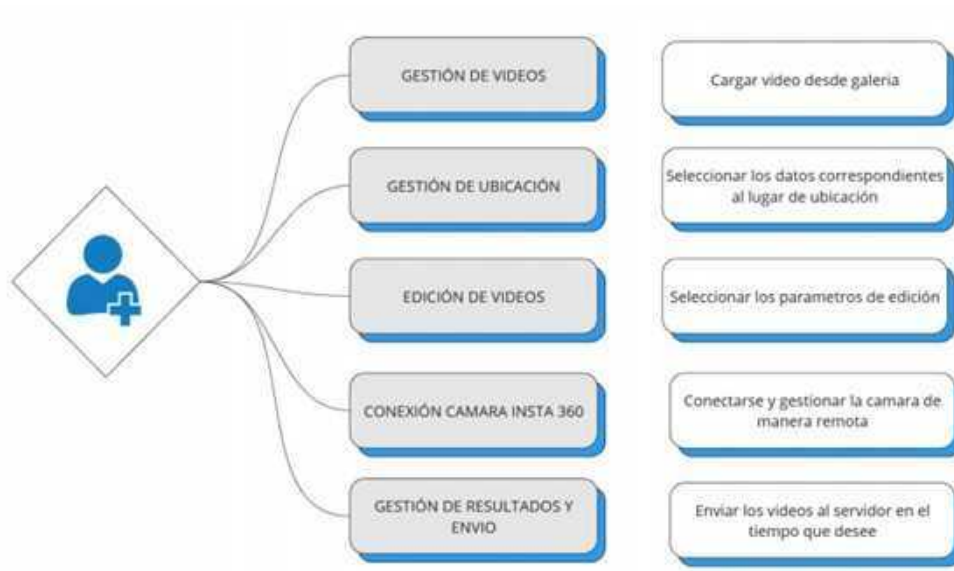


Figura 5.4: Arquitectura técnica y casos de uso de implementación, mostrando las actividades principales de la aplicación y su relación funcional dentro del sistema móvil.

5.9. Requerimientos técnicos y restricciones operativas

La documentación establece requerimientos mínimos para el funcionamiento correcto de la aplicación: procesador de 2 GHz o superior, al menos 4 GB de RAM, 100 MB de almacenamiento libre, conectividad a internet y, preferiblemente, Android 13 o superior [4], [5]. Adicionalmente, para la sincronización de fincas, bloques y camas, se requiere acceso a la red interna de la empresa o a la VPN correspondiente [4].

Estas especificaciones pueden modelarse como un conjunto de restricciones de operación \mathcal{R} :

$$\mathcal{R} = \{\omega \geq 2 \text{ GHz}, M \geq 4 \text{ GB}, A \geq 100 \text{ MB}, \chi = 1, \psi = 1\},$$

donde ω es la frecuencia efectiva de procesamiento, M la memoria RAM disponible, A el almacenamiento libre, χ la disponibilidad de conectividad y ψ el acceso a la red corporativa o VPN. Solo si estas condiciones se satisfacen, el sistema puede desplegar toda su funcionalidad con estabilidad razonable.

En un contexto agrícola real, estas restricciones no son menores. La conectividad, la disponibilidad de dispositivos y la estabilidad energética suelen ser variables más críticas que en

entornos de oficina. Por tanto, el diseño de la aplicación tuvo que equilibrar funcionalidad, ligereza y tolerancia a condiciones operativas variables.

5.10. Validación, uso institucional y valor tecnológico

La solicitud institucional de puntos por desarrollo de software describe a la aplicación como una herramienta móvil para dispositivos Android, desarrollada en el marco del convenio FIT2023-59235 entre la Universidad de Antioquia y Flores El Capiro S.A., orientada a gestionar y editar videos capturados en campo durante procesos de monitoreo agrícola [1]. En ese mismo documento se resalta su arquitectura modular, su implementación en Android Studio con Kotlin y su capacidad de sincronización automática con servidores locales [1].

El certificado de uso emitido en junio de 2025 señala que el software ha sido utilizado tanto por el grupo GEPAR como por la empresa Flores El Capiro S.A. en investigación aplicada y en procesos operativos reales de registro y organización visual de camas de siembra [2]. Allí se destaca además que el uso del software ha fortalecido competencias en desarrollo móvil, sincronización de datos y visión artificial, y ha contribuido a sistematizar la captura y transferencia de información audiovisual para análisis posterior [2].

A nivel de propiedad intelectual, la Dirección Nacional de Derecho de Autor registró la obra como software inédito, originario y colectivo, con el título oficial *SICOP APP: Sistema de Gestión y Edición de Videos para SICOP*, fecha de registro 03 de junio de 2025, e identificó como autores a David Fernández Mc Cann, Mateo Alejandro Bravo Revelo y Jhon Fredy García Hincapié, con titularidad patrimonial compartida entre la Universidad de Antioquia y Flores El Capiro S.A. [3]. Este aspecto muestra que el desarrollo no se limitó a una herramienta interna de uso informal, sino que alcanzó un nivel de formalización jurídica y documental consistente con un producto tecnológico.

5.11. Discusión

La aplicación desarrollada para SICOP pone en evidencia una idea que con frecuencia pasa desapercibida en proyectos de visión artificial aplicada: la inteligencia del sistema no reside solo en los módulos analíticos finales, sino también en la calidad estructural de los insumos y en la forma como esos insumos son organizados. En este caso, la aplicación móvil no hace conteo directo de plántulas, pero cumple una función estratégica al garantizar que el video de entrada llegue con contexto, formato y trazabilidad suficientes al resto de

la cadena.

Desde el punto de vista ingenieril, el desarrollo integra al menos cuatro dimensiones de valor. La primera es operativa, porque simplifica y estandariza la captura y gestión de videos en campo. La segunda es informacional, porque transforma archivos audiovisuales en registros estructurados y auditables. La tercera es tecnológica, porque integra dispositivos móviles, bases de datos locales, APIs corporativas y cámaras externas. La cuarta es formativa, pues la documentación institucional subraya el papel del desarrollo en la formación de estudiantes y en el fortalecimiento de capacidades del grupo GEPAR [1], [2].

Otro aspecto destacable es su carácter modular y extensible. Aunque hoy la aplicación está orientada al flujo de SICOP, la misma estructura conceptual permite proyectarla hacia otros escenarios de agricultura de precisión en los que la captura y trazabilidad de videos en campo sean fundamentales. En este sentido, el desarrollo no debe interpretarse como una solución aislada, sino como una plataforma base para una familia de aplicaciones móviles de monitoreo agrícola.

5.12. Conclusiones

El desarrollo presentado en este capítulo demuestra que la gestión móvil de videos puede constituirse en un componente estructural de sistemas más amplios de visión artificial aplicada a agricultura de precisión. En el caso de SICOP, la aplicación permite cerrar la brecha entre el acto físico de captura y la disponibilidad organizada del dato para etapas posteriores de panorámica y análisis.

La aplicación fue concebida e implementada como una solución Android modular, orientada a la selección contextual de ubicación, edición técnica de videos, gestión local de registros, sincronización con servidores empresariales e integración con captura remota mediante Insta360 X3. Esta combinación funcional le confiere valor como herramienta operativa y como producto tecnológico formalizable.

Desde el punto de vista metodológico, el desarrollo responde a una necesidad real surgida dentro de la segunda fase de SICOP, asociada a agilizar y generalizar el uso del sistema en ambientes productivos. La documentación técnica, la validación institucional, el certificado de uso y el registro ante la Dirección Nacional de Derecho de Autor respaldan su madurez como resultado de investigación aplicada.

Finalmente, la aplicación muestra que la digitalización efectiva de procesos agrícolas requiere no solo modelos analíticos, sino también herramientas de software capaces de es-

estructurar el dato desde su origen. En este sentido, el desarrollo constituye una contribución relevante para la infraestructura digital de la agricultura de precisión y para la consolidación de una línea de transferencia tecnológica en software móvil aplicado.

Referencias

- [1] D. S. F. McCann, R. D. Liñán y P. M. de Santa Lucía Jaramillo, *Solicitud para asignación de puntos por desarrollo de software: SICOP APP*, Documento institucional, Universidad de Antioquia, mayo de 2025.
- [2] D. S. F. McCann y J. F. G. Hincapié, *Certificación de uso del software SICOP APP por parte del grupo GEPAR y la empresa Flores El Capiro*, Documento institucional de certificación de uso, jun. de 2025.
- [3] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: SICOP APP: Sistema de Gestión y Edición de Videos para SICOP*, Libro-Tomo-Partida 13-104-155, fecha de registro 03-Jun-2025, jun. de 2025.
- [4] Universidad de Antioquia, *SICOP APP: Sistema de Gestión y Edición de Videos para SICOP. Manual de usuario*, Universidad de Antioquia, 2024.
- [5] Universidad de Antioquia, *SICOP APP: Sistema de Gestión y Edición de Videos para SICOP. Manual técnico*, Universidad de Antioquia, 2024.
- [6] Grupo GEPAR, *SICOP: Método de captura para software SICOP. Manual de usuario*, Universidad de Antioquia, 2024.
- [7] Universidad de Antioquia, *Acta de registro y de propiedad intelectual del proyecto Sistema de Conteo de Plántulas en Cama segunda fase SICOP 2 – Prototipo*, Acta N.º 2023-59235, CODI, Vicerrectoría de Investigación, oct. de 2023.
- [8] Grupo GEPAR, *Método de captura de video SICOP: guía rápida*, Universidad de Antioquia, 2024.

Capítulo 6

Conteo automatizado de plántulas en cama mediante visión por computador

Autores

Johan David Rojas Martínez¹; Jhon Fredy García Hincapié²; David Stephen Fernández Mc Cann³

Resumen

El conteo de plántulas en cama constituye una operación crítica en sistemas de producción agrícola intensiva, pues permite estimar densidad de siembra, eficiencia de establecimiento, comportamiento del cultivo y proyección de cosecha. Tradicionalmente, esta tarea se realiza mediante conteo manual o muestreos parciales, procedimientos expuestos a fatiga, subjetividad, baja escalabilidad y limitada trazabilidad digital. En este capítulo se presenta un sistema de conteo automatizado de plántulas basado en visión por computador, desarrollado como una aplicación de escritorio en Python con interfaz gráfica en Tkinter, modelos de detección basados en YOLO y rutinas de almacenamiento estructurado

¹Ingeniero asociado al desarrollo del software, Universidad de Antioquia. Grupo de investigación GEPAR.

²Director de Innovación Tecnológica y Nuevos Negocios, Flores El Capiro S.A.

³Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Desarrollo realizado en el marco del convenio FIT2023-59235 entre la Universidad de Antioquia y Flores El Capiro S.A., con participación del grupo GEPAR.

de resultados. El sistema permite cargar uno o varios videos capturados en campo, procesarlos automáticamente, visualizar el avance del conteo, generar fotogramas anotados, almacenar recortes individuales de plántulas, producir un video final con anotaciones y guardar un archivo de resultados con el total detectado. Aunque el software cuenta con nombre propio y registro ante la Dirección Nacional de Derecho de Autor, el énfasis de este capítulo se sitúa en la solución ingenieril: transformar secuencias de video agrícola en información cuantitativa, reproducible y trazable para apoyar procesos de monitoreo en agricultura de precisión [1], [2], [3].

Palabras clave: conteo de plántulas; visión por computador; agricultura de precisión; YOLO; procesamiento de video; trazabilidad; Python; GEPAR; Flores El Capiro.

6.1. Introducción

El conteo de plántulas en cama constituye una operación crítica en los sistemas de producción agrícola intensiva, pues de su correcta ejecución dependen estimaciones posteriores relacionadas con densidad de siembra, eficiencia de establecimiento, pronóstico de cosecha y evaluación del comportamiento del cultivo. En la práctica tradicional, esta tarea se ha realizado de forma manual o mediante revisiones parciales por muestreo, lo que introduce errores asociados a fatiga, subjetividad del operario y baja escalabilidad cuando se trabaja sobre grandes superficies o múltiples camas de cultivo [2], [4].

En el contexto del proyecto SICOP 2, orientado al conteo de plántulas en cama, se consolidó una cadena tecnológica compuesta por captura de video, construcción panorámica, analítica y, en esta nueva etapa, una herramienta software orientada al procesamiento automatizado del material capturado. Este desarrollo responde a la necesidad de transformar el video agrícola en una fuente cuantitativa de información, reduciendo la dependencia del conteo visual manual y generando resultados estructurados, reproducibles y almacenables digitalmente [1], [3].

El software desarrollado fue concebido como una aplicación de escritorio en Python, con interfaz gráfica en Tkinter, capaz de cargar uno o varios videos, procesarlos mediante modelos de detección basados en YOLO, visualizar el avance del conteo en tiempo real y guardar automáticamente las evidencias generadas durante el análisis. Entre dichas evidencias se incluyen fotogramas anotados, recortes individuales de plántulas, un video final con anotaciones y un archivo de texto con el total de plántulas detectadas [1], [2]. Aunque el nombre del software se emplea en la documentación técnica y en el registro ante la Dirección Nacional de Derecho de Autor, en este capítulo el interés principal no reside en la denominación de la herramienta, sino en la solución ingenieril que representa

para el problema de conteo automatizado de plántulas.

El objetivo de este capítulo es presentar dicho desarrollo como un sistema formal de visión por computador aplicado a agricultura de precisión. Para ello, se expone el problema de conteo como una tarea de estimación automática sobre secuencias de video, se describe la arquitectura funcional del software, se introducen formulaciones matemáticas del proceso de detección y conteo, y se discute su valor como producto tecnológico y como componente de transferencia universidad–empresa.

6.2. Planteamiento del problema

El conteo de plántulas puede formularse como un problema de estimación del cardinal de un conjunto de objetos visibles sobre una secuencia de imágenes adquiridas en condiciones agrícolas reales. Sea \mathcal{P} el conjunto real de plántulas presentes en una cama de cultivo y sea $N = |\mathcal{P}|$ el número total de plántulas que se desea estimar. El reto consiste en aproximar N a partir de una secuencia de fotogramas extraídos de un video, donde las plántulas pueden presentar variaciones de tamaño aparente, oclusión parcial, iluminación no uniforme, presencia de mallas o mangueras y cambios de perspectiva.

Si el video de entrada se representa como una secuencia ordenada de fotogramas

$$\mathcal{V} = \{I_1, I_2, \dots, I_T\},$$

donde I_t denota el fotograma correspondiente al instante discreto t y T es el número total de fotogramas, el sistema debe construir una estimación \widehat{N} del número de plántulas a partir de la información visual contenida en \mathcal{V} .

En una formulación general, puede considerarse que cada fotograma I_t produce un conjunto de detecciones

$$\mathcal{D}_t = \{d_{t,1}, d_{t,2}, \dots, d_{t,n_t}\},$$

donde cada detección $d_{t,j}$ corresponde a una hipótesis de presencia de plántula en el fotograma t , y n_t es el número de detecciones en dicho fotograma. Sin embargo, el problema no se resuelve simplemente acumulando todos los n_t , pues una misma plántula puede aparecer en varios fotogramas consecutivos. Por ello, el conteo exige una estrategia de seguimiento o un criterio geométrico que permita evitar el sobreconteo.

Desde el punto de vista del sistema, el problema puede resumirse mediante la construcción de una función de estimación

$$\widehat{N} = \mathcal{F}(\mathcal{V}, \Theta),$$

donde Θ representa el conjunto de parámetros del modelo de detección, del criterio de conteo y de los umbrales de procesamiento. En este caso, la función \mathcal{F} se implementa mediante una cadena que combina lectura de video, inferencia visual basada en YOLO, anotación, seguimiento sobre una línea de conteo y almacenamiento de resultados [2].

6.3. Contexto técnico y necesidad de automatización

La documentación técnica del desarrollo subraya que uno de los principales problemas del conteo manual radica en el tiempo requerido para evaluar grandes extensiones de cultivo, así como en la exposición del proceso a errores humanos derivados de cansancio, variabilidad metodológica y baja trazabilidad digital [2]. A ello se suman las condiciones variables del entorno agrícola, que dificultan la inspección consistente: iluminación cambiante, distribución heterogénea de plántulas, elementos estructurales del cultivo y necesidad de comparar resultados entre jornadas o camas distintas.

En este marco, el conteo automatizado no debe entenderse únicamente como un reemplazo del conteo manual, sino como una transformación del dato agrícola en información cuantitativa sistemática. Si se denota por N_m el conteo manual y por N_a el conteo automatizado, puede definirse el error absoluto del sistema automático respecto a una referencia de validación como

$$E = |N_a - N_m|.$$

Análogamente, si se desea una medida relativa, puede emplearse el error porcentual

$$E_r = \frac{|N_a - N_m|}{N_m} \times 100\%.$$

Estas expresiones sitúan la discusión del software en el terreno de la medición cuantitativa. La meta no es solo contar rápido, sino hacerlo con un error suficientemente bajo y con una consistencia superior a la disponible en procedimientos no automatizados.

El certificado de uso emitido para el desarrollo indica que el software ha sido empleado

tanto por el grupo GEPAR como por Flores El Capiro S.A. en actividades de validación experimental y operativa, incluyendo procesamiento de videos tomados en campo, generación de registros cuantitativos y evaluación del estado de siembra [4]. Esto confirma que el problema abordado es real y que la solución fue llevada más allá de una prueba académica aislada.

6.4. Arquitectura funcional del software

De acuerdo con los manuales de usuario y técnico, el sistema está organizado en cuatro módulos principales: carga de video, conteo del video, almacenamiento de resultados y conteo múltiple con guardado automático [1], [2]. Esta organización modular permite separar la interacción con el usuario de la lógica de procesamiento y facilita la expansión futura del software.

Si se denota por \mathcal{V}_{in} el conjunto de videos de entrada y por \mathcal{R} el conjunto de resultados generados, el flujo principal del sistema puede describirse como

$$\mathcal{V}_{in} \xrightarrow{\mathcal{L}} \mathcal{V}_{val} \xrightarrow{\mathcal{C}} \mathcal{A} \xrightarrow{\mathcal{G}} \mathcal{R},$$

donde \mathcal{L} representa la validación y carga del video, \mathcal{C} la etapa de conteo e inferencia visual, \mathcal{A} el conjunto de anotaciones y evidencias intermedias, y \mathcal{G} el módulo de generación y almacenamiento de resultados.

En la interfaz gráfica, el usuario interactúa principalmente con tres botones: **Cargar**, **Contar** y **Guardar**. Esta simplicidad operativa no implica trivialidad técnica. Detrás de cada acción se articula una serie de procesos relacionados con selección de archivos, lectura de atributos del video, inferencia del modelo, construcción de carpetas, generación de fotogramas anotados y escritura de archivos de salida [1]. El módulo adicional de conteo y guardado automático permite extender esa lógica al procesamiento por lotes de múltiples videos.

Desde una perspectiva de diseño, el sistema puede representarse como una partición del conjunto de funcionalidades \mathcal{F} :

$$\mathcal{F} = \mathcal{F}_{carga} \cup \mathcal{F}_{conteo} \cup \mathcal{F}_{guardado} \cup \mathcal{F}_{lote},$$

donde cada subconjunto agrupa responsabilidades funcionales bien definidas. Esta modularidad constituye una decisión acertada de ingeniería de software, pues reduce acopla-

miento, mejora mantenibilidad y favorece la trazabilidad del flujo de uso.

La Fig. 6.1 presenta la arquitectura general del software y su relación entre interfaz gráfica y módulo de conteo.

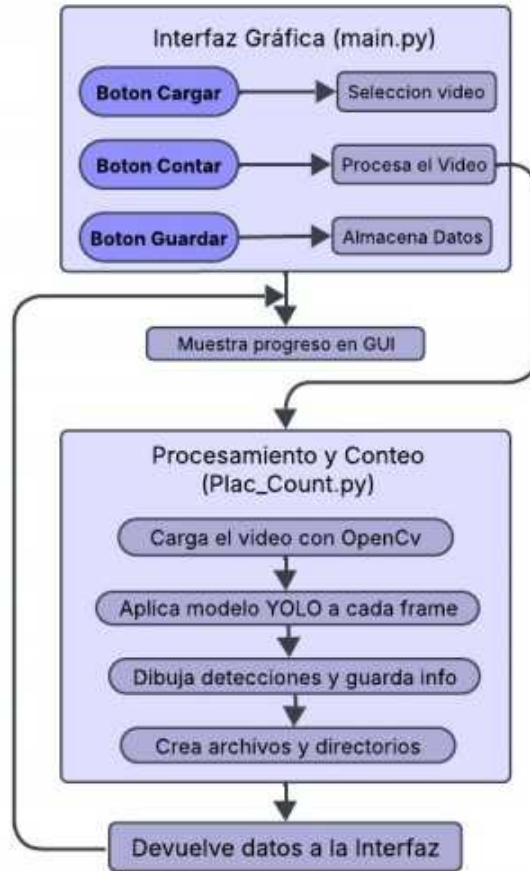


Figura 6.1: Arquitectura general del software, mostrando la relación entre la interfaz gráfica principal y el módulo de procesamiento y conteo de plántulas.

6.5. Modelo matemático del proceso de detección y conteo

El núcleo del sistema consiste en transformar una secuencia de fotogramas en una estimación robusta del número de plántulas. Sea $I_t : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R}^3$ el fotograma RGB en el instante t , con dominio espacial Ω . El modelo de detección genera sobre cada fotograma un conjunto de cajas delimitadoras

$$\mathcal{B}_t = \{B_{t,1}, B_{t,2}, \dots, B_{t,n_t}\},$$

donde cada $B_{t,j}$ es una región rectangular asociada a una plántula detectada. Cada detección viene acompañada por una puntuación de confianza $s_{t,j} \in [0, 1]$, y el sistema conserva únicamente aquellas detecciones que satisfacen

$$s_{t,j} \geq \tau,$$

donde τ es un umbral de confianza previamente definido en el modelo.

Para evitar contar múltiples veces una misma plántula a lo largo del video, la documentación técnica indica el uso de herramientas de anotación y seguimiento, así como la definición de una línea de conteo [2]. Sea dicha línea un segmento orientado determinado por dos puntos del plano imagen,

$$L = \overline{p_s p_e},$$

donde $p_s = (x_s, y_s)$ y $p_e = (x_e, y_e)$ representan los puntos de inicio y fin de la línea de conteo. Para cada objeto detectado y seguido en el tiempo, se asocia un centroide $c_{t,j} = (x_{t,j}, y_{t,j})$. El conteo se incrementa cuando la trayectoria del objeto cruza la línea L en el sentido establecido.

Si $g(c_{t,j})$ representa una función de signo asociada a la posición del centroide respecto a la línea, entonces una condición de cruce puede modelarse como

$$g(c_{t,j}) g(c_{t+1,j}) < 0,$$

lo que indica que el objeto pasó de un lado de la línea al otro entre dos instantes consecutivos. De esta forma, el conteo acumulado puede expresarse como

$$\widehat{N} = \sum_{j=1}^M \delta_j,$$

donde M es el número de trayectorias válidas seguidas durante el video y δ_j es una variable indicadora definida como

$$\delta_j = \begin{cases} 1, & \text{si la trayectoria } j \text{ cruza } L \text{ según el criterio de conteo,} \\ 0, & \text{en otro caso.} \end{cases}$$

Esta formulación muestra que el sistema no solo detecta plántulas por fotograma, sino que integra detección y lógica de conteo sobre secuencias temporales.

6.6. Arquitectura de implementación

El manual técnico señala que el software se organiza alrededor de dos archivos principales: `Main.py`, encargado de la interfaz de usuario y del control del flujo, y `Plac_Count.py`, responsable del procesamiento del video y del conteo de plántulas [2]. Esta decisión de diseño es congruente con una arquitectura de separación entre capa de presentación y capa de procesamiento.

Sea $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ un grafo dirigido que representa la interacción entre módulos. En este caso puede definirse el conjunto de nodos como

$$\mathcal{N} = \left\{ \begin{array}{l} \text{Main.py,} \\ \text{Plac_Count.py,} \\ \text{/frames,} \\ \text{/plantulas,} \\ \text{resultado.mp4,} \\ \text{resultados.txt} \end{array} \right\},$$

donde los nodos iniciales representan módulos de software y los demás representan estructuras de salida persistente. Las aristas de \mathcal{E} describen dependencias de ejecución, paso de datos y generación de resultados.

La documentación del código muestra que `Plac_Count.py` importa bibliotecas como `os`, `cv2`, `PIL`, `numpy`, `supervision`, `ultralytics.YOLO` y `time`, mientras que `Main.py` administra la interfaz gráfica, la selección de videos, el arranque del proceso, la actualización del conteo y el guardado de resultados [2]. Esta estructura permite modelar el software como una composición funcional

$$\mathcal{S} = \mathcal{I} \circ \mathcal{P},$$

donde \mathcal{P} es el módulo de procesamiento visual y \mathcal{I} el módulo de interacción y control. En términos prácticos, la interfaz gráfica orquesta el proceso, pero la lógica de detección y conteo reside en el motor de procesamiento.

La Fig. 6.2 presenta la relación entre interfaz gráfica y módulo interno de procesamiento.



Figura 6.2: Relación entre la interfaz gráfica principal y el módulo interno de procesamiento de video y conteo.

6.7. Almacenamiento estructurado y trazabilidad de resultados

Uno de los rasgos más valiosos del desarrollo es que no se limita a mostrar un número final en pantalla. El sistema construye una estructura persistente de resultados para cada video procesado. Según el manual de usuario, por cada video de entrada se crea automáticamente una carpeta madre con el nombre del archivo, dentro de la cual se generan una subcarpeta `frames`, una subcarpeta `plantulas`, un archivo `resultado.mp4` y un archivo `resultados.txt` [1].

Si el video procesado se denota por v_i , entonces el resultado estructurado asociado puede modelarse como

$$\mathcal{R}_i = (\mathcal{F}_i, \mathcal{P}_i, m_i, r_i),$$

donde \mathcal{F}_i es el conjunto de fotogramas anotados, \mathcal{P}_i el conjunto de imágenes individuales

de plántulas detectadas, m_i el video final anotado y r_i el archivo textual con el resumen del conteo. Esta estructura convierte el resultado del conteo en un objeto digital trazable y auditable.

El archivo de resultados puede representarse como una función resumen

$$r_i = \Psi(v_i, \widehat{N}_i, t_i),$$

donde \widehat{N}_i es el conteo total estimado para el video v_i y t_i el tiempo de procesamiento registrado. Esta formalización resalta que el sistema produce no solo evidencia visual sino también información cuantitativa compacta.

En aplicaciones de investigación y monitoreo agrícola, esta decisión de diseño es crucial, porque permite revisar posteriormente el comportamiento del algoritmo, contrastar resultados con inspecciones humanas y construir bases de datos históricas para validación o entrenamiento adicional.

La Fig. 6.3 muestra la estructura de almacenamiento producida por el sistema.



Figura 6.3: Estructura de almacenamiento generada por el software para cada video analizado, incluyendo fotogramas procesados, recortes de plántulas, video anotado y archivo de resultados.

6.8. Procesamiento por lotes y escalabilidad operativa

Además del procesamiento individual, el sistema incorpora un módulo de conteo y guardado automático para múltiples videos [1]. Esta funcionalidad es especialmente importante cuando se requiere analizar varias camas de cultivo, ya que evita la necesidad de repetir manualmente la secuencia cargar–contar–guardar sobre cada archivo.

Sea

$$\mathcal{V}_B = \{v_1, v_2, \dots, v_K\}$$

un lote de K videos seleccionados por el usuario. Entonces el procesamiento por lotes puede expresarse como

$$\mathcal{R}_B = \{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_K\},$$

donde cada \mathcal{R}_i es el resultado estructurado del video v_i . En términos de implementación, el sistema recorre secuencialmente cada elemento del conjunto y ejecuta la misma función de procesamiento:

$$\mathcal{R}_i = \mathcal{F}(v_i, \Theta), \quad i = 1, \dots, K.$$

Esta capacidad mejora la escalabilidad operativa del software. Si el tiempo medio de procesamiento por video se denota por \bar{t} , el tiempo total de análisis de un lote puede aproximarse como

$$T_B \approx K \bar{t},$$

suponiendo procesamiento secuencial homogéneo. Aunque esta relación es simple, resulta útil para dimensionar el comportamiento del sistema y justificar la necesidad de requerimientos computacionales más robustos cuando se procesan lotes grandes de videos.

6.9. Requerimientos técnicos y restricciones de uso

La documentación técnica establece requisitos claros de hardware y software. En particular, se requiere:

- procesador Intel Core i5 o equivalente AMD;
- al menos 12 GB de RAM;
- un mínimo de 5 GB de espacio libre en disco;
- resolución de pantalla de 1366×800 ;
- sistema operativo Windows 11;
- Python 3.11 y las dependencias listadas en `requirements.txt`.

Estas especificaciones pueden resumirse mediante el conjunto

$$\mathcal{R} = \{\omega \geq \omega_{\min}, M \geq 12 \text{ GB}, A \geq 5 \text{ GB}, \Pi = \text{Windows 11}\},$$

donde ω es la capacidad de procesamiento efectiva del equipo, M la memoria RAM disponible, A el almacenamiento libre y Π la plataforma de sistema operativo. La necesidad de estos recursos no es arbitraria: deriva del uso de modelos de inteligencia artificial, lectura intensiva de video y generación simultánea de archivos de salida [2].

A nivel funcional, el sistema también impone condiciones lógicas de operación: el video debe estar previamente cargado antes de iniciar el conteo, el archivo debe estar en un formato compatible y el proceso debe finalizar correctamente antes de que el guardado manual tenga sentido [1], [2]. Esto equivale a decir que el sistema implementa un flujo de estados con precondiciones bien definidas.

6.10. Validación institucional, uso y protección intelectual

El certificado de uso del software indica que la herramienta ha sido desarrollada y se encuentra en uso dentro del marco de actividades de investigación, desarrollo e innovación del grupo GEPAR, así como en procesos de validación experimental y operativa de Flores El Capiro S.A. [4]. El documento también resalta que el sistema se ha empleado como insumo para formación de estudiantes de ingeniería y como base tecnológica para nuevas líneas de trabajo en agricultura de precisión [4].

Por su parte, el certificado de registro expedido por la Dirección Nacional de Derecho

de Autor identifica el software como obra inédita, originaria y colectiva, con título oficial *PLANT-IQ: Software para el conteo de plántulas en cama*, fecha de registro 06 de junio de 2025 y autores David Fernández Mc Cann, Jhon Fredy García Hincapié y Johan David Rojas Martínez, con titularidad patrimonial compartida entre la Universidad de Antioquia y Flores El Capiro S.A. [3]. La descripción oficial de la obra destaca que el sistema detecta y cuenta plántulas en videos capturados en entornos agrícolas, organiza fotogramas, almacena resultados estructurados, genera reportes y produce un video final anotado [3].

Estos documentos permiten afirmar que el desarrollo no corresponde solo a un ejercicio experimental de laboratorio, sino a un producto tecnológico con formalización técnica, jurídica y operativa.

6.11. Discusión

El software presentado en este capítulo resuelve un problema agrícola concreto mediante una estrategia de visión por computador que articula detección automática, seguimiento temporal, estructura de almacenamiento y operación mediante interfaz gráfica. Su valor no reside únicamente en el uso de YOLO como modelo detector, sino en la manera como dicho modelo se integra dentro de un flujo reproducible de trabajo para el usuario final.

Desde una perspectiva de ingeniería aplicada, el sistema muestra al menos cuatro fortalezas. La primera es funcional, pues automatiza una labor de alto costo operativo cuando se ejecuta manualmente. La segunda es informacional, ya que transforma el conteo en una fuente trazable de datos y evidencias visuales. La tercera es computacional, al integrar procesamiento de video, detección en tiempo real y procesamiento por lotes. La cuarta es institucional, dado que se inserta en una ruta efectiva de transferencia tecnológica entre universidad y empresa.

Otro aspecto relevante es su potencial de escalamiento. La misma arquitectura podría adaptarse a otros tipos de cultivos, a variaciones del modelo detector o a procesos más complejos de caracterización morfológica. En este sentido, el desarrollo constituye una plataforma base más que una solución cerrada.

6.12. Conclusiones

Se presentó un sistema de conteo automatizado de plántulas en cama basado en visión por computador, concebido como una herramienta software de escritorio para el procesamiento

de videos agrícolas y la generación estructurada de resultados.

El problema fue formulado como una tarea de estimación del número de plántulas a partir de secuencias de video, y la solución implementada integró carga de archivos, inferencia con modelos YOLO, lógica geométrica de conteo, almacenamiento organizado y procesamiento por lotes. Esta integración permitió trasladar una labor tradicionalmente manual hacia un entorno reproducible, trazable y escalable.

Desde el punto de vista metodológico, el desarrollo se articuló con el proyecto SICOP 2 y con las necesidades reales de Flores El Capiro S.A., alcanzando uso activo en investigación y validación operativa, así como registro formal ante la Dirección Nacional de Derecho de Autor. Esto respalda su condición de resultado tecnológico maduro dentro de una línea de agricultura de precisión apoyada en visión artificial.

Finalmente, el sistema evidencia que la automatización del monitoreo agrícola requiere tanto modelos de detección robustos como diseños de software capaces de organizar el dato, comunicarlo al usuario y preservarlo de forma útil para decisiones posteriores. En este sentido, constituye una contribución relevante a la digitalización del seguimiento de cultivos.

Referencias

- [1] Universidad de Antioquia, *PLANT-IQ: Software para el conteo de plántulas en cama. Manual de usuario*, Universidad de Antioquia, 2025.
- [2] Universidad de Antioquia, *PLANT-IQ: Software para el conteo de plántulas en cama. Manual técnico*, Universidad de Antioquia, 2025.
- [3] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: PLANT-IQ: Software para el conteo de plántulas en cama*, Libro-Tomo-Partida 13-104-204, fecha de registro 06-Jun-2025, jun. de 2025.
- [4] D. S. F. McCann y J. F. G. Hincapié, *Certificación de uso del software PLANT-IQ por parte del grupo GEPAR y la empresa Flores El Capiro*, Documento institucional de certificación de uso, jun. de 2025.
- [5] D. S. F. McCann, *Solicitud para asignación de puntos por desarrollo de software: PLANT-IQ*, Documento institucional, Universidad de Antioquia, mayo de 2025.

Capítulo 7

Reconocimiento automático de placas vehiculares en video mediante visión artificial y OCR profundo

Autores

Johan David Rojas Martínez¹; David Stephen Fernández Mc Cann²; Claudia Cristina Botero Suárez³

Resumen

El reconocimiento automático de placas vehiculares constituye una tarea relevante para la supervisión vial, el control de acceso, la trazabilidad de vehículos y la generación de registros estructurados a partir de video. La revisión manual de grabaciones presenta limitaciones operativas asociadas al tiempo de inspección, la variabilidad del observador, la calidad del video, los cambios de iluminación y el movimiento de los vehículos. En este capítulo se presenta un desarrollo de visión artificial orientado a detectar y reconocer placas vehiculares en videos capturados por cámaras estáticas. El sistema integra un

¹Estudiante de Ingeniería Electrónica, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Coautora vinculada a InterTelco S.A.S.

³Desarrollo realizado en el marco del proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, articulado entre la Universidad de Antioquia, el grupo GEPAR e InterTelco S.A.S.

detector basado en YOLOv8, un módulo de reconocimiento óptico de caracteres mediante TrOCR, procesamiento de video en Python, manipulación de imágenes con OpenCV y PIL, ejecución de modelos con Torch e interfaz gráfica en Tkinter [1], [2]. La herramienta permite cargar videos, iniciar, pausar y reanudar el procesamiento, visualizar detecciones, recortar placas, reconocer caracteres y exportar resultados en carpetas organizadas. El desarrollo fue construido en el marco del proyecto VISIONARY y se encuentra en uso por el grupo GEPAR e InterTelco S.A.S. para validación experimental, integración industrial y formación en visión artificial y aprendizaje profundo [3], [4]. La formalización del producto se respalda con el certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor, bajo Libro–Tomo–Partida 13-106-30, con fecha de registro 29 de septiembre de 2025 [5].

Palabras clave: reconocimiento automático de placas; visión artificial; YOLOv8; TrOCR; OCR; procesamiento de video; seguridad vial; inteligencia vial; GEPAR; VISIONARY.

7.1. Introducción

El análisis de video aplicado a la movilidad permite transformar grabaciones viales en información estructurada sobre circulación, acceso, permanencia y trazabilidad vehicular. Entre las variables de interés, la placa del vehículo ocupa un lugar central porque conecta la evidencia visual con un identificador alfanumérico verificable. Esta relación hace posible construir registros de tránsito, apoyar controles de acceso, fortalecer sistemas de seguridad, estudiar flujos vehiculares y alimentar plataformas de supervisión urbana o logística.

La revisión manual de videos presenta dificultades operativas claras. El operador debe examinar secuencias extensas, identificar vehículos, ubicar placas visibles, interpretar caracteres y registrar la información resultante. Este procedimiento demanda tiempo, introduce variabilidad humana y pierde eficiencia cuando aumenta el número de cámaras, la duración de los videos o la densidad vehicular. Además, las condiciones de captura rara vez son ideales: iluminación cambiante, movimiento, oclusiones parciales, ángulos desfavorables, baja resolución o compresión del video pueden afectar la lectura directa de los caracteres [1].

El desarrollo presentado en este capítulo responde a ese problema mediante una arquitectura de visión artificial que combina detección de objetos y reconocimiento óptico de caracteres. El sistema automatiza el flujo desde la carga del video hasta la exportación de evidencias visuales y textos reconocidos. La herramienta fue documentada en manual técnico y manual de usuario bajo el estándar IEEE 1063-2001, lo cual facilita su operación, mantenimiento, validación y transferencia [1], [2].

Dentro del proyecto VISIONARY, este desarrollo se vincula con una línea de trabajo orientada a sistemas avanzados de visión artificial para seguridad vial y supervisión de infraestructura. El acta institucional del proyecto establece como objetivos investigar algoritmos de visión artificial, desarrollar modelos aplicados y registrar conjuntamente productos de software derivados de la cooperación entre la Universidad de Antioquia e InterTelco S.A.S. [4]. En ese marco, el software registrado ante la Dirección Nacional de Derecho de Autor constituye una pieza tecnológica concreta para la identificación automática de placas vehiculares en video.

7.2. Planteamiento del problema

Sea un video digital capturado por una cámara estática y representado como una secuencia ordenada de fotogramas

$$V = \{I_1, I_2, \dots, I_T\},$$

donde I_t corresponde al fotograma en el instante discreto t , y T representa el número total de fotogramas disponibles. El problema consiste en detectar, para cada fotograma, las regiones que contienen placas vehiculares y asignar a cada región una cadena alfanumérica reconocida.

Formalmente, para cada imagen I_t , se busca construir una función de detección

$$D_\theta(I_t) = \{(b_{t,i}, s_{t,i})\}_{i=1}^{N_t},$$

donde $b_{t,i}$ representa el recuadro delimitador de la placa i detectada en el fotograma t , $s_{t,i}$ corresponde a la confianza de la detección, N_t es el número de placas detectadas en dicho fotograma y θ agrupa los parámetros del modelo detector. En este desarrollo, la detección se implementa mediante YOLOv8, empleado para localizar placas en imágenes y videos [1].

Una vez localizada la placa, se define una operación de recorte

$$R_{t,i} = C(I_t, b_{t,i}),$$

donde $C(\cdot)$ extrae la región de interés asociada al recuadro $b_{t,i}$. Sobre esta región se aplica un modelo OCR que estima una cadena de caracteres:

$$O_\phi(R_{t,i}) = (\hat{y}_{t,i}, q_{t,i}),$$

donde $\hat{y}_{t,i}$ es el texto reconocido, $q_{t,i}$ representa una medida de confianza o consistencia de la lectura y ϕ agrupa los parámetros del modelo de reconocimiento. En el sistema desarrollado, esta etapa se realiza mediante TrOCR, un modelo de reconocimiento óptico de caracteres basado en transformadores [1], [6].

El resultado final del procesamiento debe producir un conjunto estructurado de registros:

$$\mathcal{R} = \{(\hat{y}_j, R_j, t_j, q_j)\}_{j=1}^M,$$

donde cada elemento conserva la lectura reconocida, la imagen recortada de la placa, el instante o fotograma asociado y una medida de confianza. Esta estructura permite que el resultado no sea solo una lectura aislada, sino una evidencia trazable, revisable y exportable.

7.3. Contexto institucional y alcance del desarrollo

El sistema fue desarrollado dentro del proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*. La documentación institucional del proyecto lo clasifica como investigación aplicada y plantea como objetivos la investigación de algoritmos de visión artificial, el desarrollo de modelos para cumplir los objetivos técnicos, el fortalecimiento de la cooperación entre la Universidad de Antioquia e InterTelco S.A.S. y el registro conjunto de productos de software [4].

El certificado de uso indica que el software fue construido en el marco de la cooperación entre la Universidad de Antioquia e InterTelco S.A.S. y que se encuentra en uso activo por el grupo GEPAR y por la empresa para labores de validación experimental e integración industrial. El mismo documento señala aplicaciones como el procesamiento de videos capturados en entornos viales, la generación de registros cuantitativos de tránsito, la evaluación de cámaras y algoritmos en condiciones reales, y su uso como herramienta de formación en visión artificial y aprendizaje profundo [3].

Esta condición sitúa el desarrollo en una zona propia de la investigación aplicada: no corresponde únicamente a una demostración algorítmica, sino a una herramienta documentada, operable, validable y formalizada como soporte lógico. Su utilidad se ubica tanto en escenarios industriales e institucionales como en procesos académicos de formación y experimentación.

7.4. Arquitectura funcional del sistema

La arquitectura funcional integra cuatro bloques principales: entrada de video, detección de placas, reconocimiento de caracteres y exportación estructurada de resultados. Estos bloques se articulan mediante una interfaz gráfica desarrollada en Tkinter, que permite operar el sistema sin requerir intervención directa sobre el código fuente [1], [2].

La Fig. 7.1 presenta la interfaz principal del sistema, desde la cual se cargan los videos y se controlan las operaciones de procesamiento.

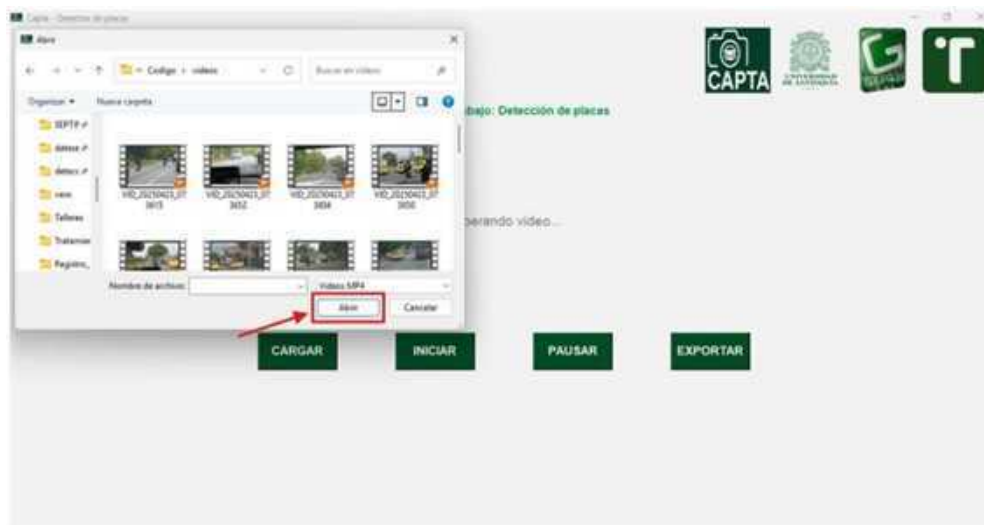


Figura 7.1: Interfaz principal del sistema para carga de video, inicio de procesamiento, pausa y exportación de resultados. La operación se organiza en botones de acceso directo, lo que permite ejecutar el flujo completo desde una misma ventana.

El primer bloque corresponde a la carga y verificación del video. El usuario selecciona un archivo desde la interfaz, el sistema verifica que el formato sea compatible y extrae información básica como duración, resolución y velocidad de fotogramas. El manual de usuario indica compatibilidad con formatos como MP4, AVI y MOV, y descarta archivos que no correspondan a video [2].

El segundo bloque ejecuta el procesamiento cuadro a cuadro. Cada fotograma se analiza mediante el detector de placas. Las detecciones válidas se recortan y se envían al módulo OCR. Esta separación entre detección y lectura permite tratar la placa como una región de interés, reduciendo el espacio visual que debe interpretar el reconocedor de caracteres.

El tercer bloque corresponde al reconocimiento óptico de caracteres. En esta etapa se toma cada recorte $R_{t,i}$ y se estima una cadena alfanumérica $\hat{y}_{t,i}$. La arquitectura documentada

incorpora TrOCR como modelo principal y contempla OCR alternativo dentro de las bibliotecas especializadas incluidas en el sistema [1].

El cuarto bloque organiza los resultados. La exportación crea una carpeta con el nombre del video procesado, almacena imágenes de las placas detectadas y genera una subcarpeta `placas_finales/` para conservar una imagen por placa única, seleccionada por reglas de confianza o coherencia entre detecciones [2].

La Fig. 7.2 resume el flujo completo del sistema, desde la carga del video hasta la exportación de evidencias.

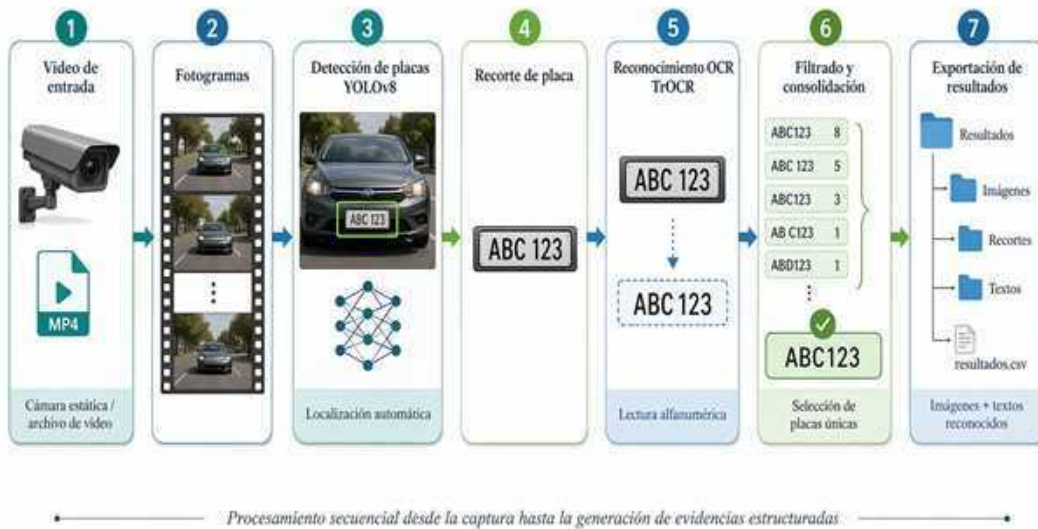


Figura 7.2: Flujo funcional del sistema: carga del video, lectura secuencial de fotogramas, detección de placas, recorte de regiones de interés, reconocimiento OCR, visualización y exportación estructurada de evidencias.

7.5. Modelo computacional de detección y reconocimiento

La operación del sistema puede expresarse como una composición de funciones:

$$S(V) = E \circ A \circ O_{\phi} \circ C \circ D_{\theta}(V),$$

donde D_{θ} detecta las placas, C recorta las regiones de interés, O_{ϕ} reconoce los caracteres, A agrupa o filtra resultados equivalentes y E exporta las evidencias en una estructura organizada de almacenamiento.

La detección se realiza sobre cada fotograma. Para un recuadro delimitador

$$b_{t,i} = (x_{t,i}, y_{t,i}, w_{t,i}, h_{t,i}),$$

las variables $x_{t,i}$ y $y_{t,i}$ representan la posición de referencia del recuadro, mientras que $w_{t,i}$ y $h_{t,i}$ corresponden a su ancho y alto. A partir de estas coordenadas, se extrae la región de placa que será procesada por el OCR.

Cuando una misma placa aparece en varios fotogramas, el sistema puede obtener lecturas ligeramente diferentes debido a desenfoque, movimiento, ángulo o iluminación. Para representar la consolidación de lecturas repetidas, puede definirse una agrupación de cadenas reconocidas:

$$G_k = \{\hat{y}_j : d(\hat{y}_j, \hat{y}_k) \leq \tau\},$$

donde $d(\cdot, \cdot)$ es una distancia entre cadenas, por ejemplo distancia de edición, y τ es un umbral de similitud. Una lectura final puede seleccionarse mediante

$$\hat{y}_k^* = \arg \max_{\hat{y}_j \in G_k} Q(\hat{y}_j),$$

donde $Q(\cdot)$ representa una función de calidad asociada a la confianza del detector, la confianza del OCR, la frecuencia de aparición o la coherencia de la lectura. Esta formulación resume la lógica documentada de filtrado inteligente y conservación de placas finales no duplicadas [2].

7.6. Interfaz de operación y módulos de usuario

El manual de usuario organiza la operación en cuatro módulos: carga de video, inicio de procesamiento, pausa o reanudación del análisis, y exportación de resultados [2]. Esta estructura reduce la complejidad del flujo para el usuario final y permite separar las acciones críticas del proceso.

7.6.1. Carga del video

La carga del video constituye el punto de entrada al sistema. El usuario selecciona un archivo grabado con cámara fija, y el sistema verifica su compatibilidad antes de habilitar

el procesamiento. Esta validación previene errores asociados a formatos no admitidos, archivos corruptos o selección accidental de documentos que no corresponden a video.

7.6.2. Procesamiento y visualización

El procesamiento se activa mediante el botón de inicio. A partir de ese momento, el sistema analiza el video por fotogramas, detecta regiones que contienen placas y aplica reconocimiento óptico de caracteres. Durante la ejecución, las detecciones y lecturas se visualizan en la interfaz, lo que permite al usuario verificar el comportamiento del sistema durante el análisis.

La Fig. 7.3 muestra el módulo de procesamiento en operación.



Figura 7.3: Módulo de inicio de procesamiento. El sistema activa el análisis automático del video y aplica los modelos de detección y reconocimiento sobre los fotogramas de entrada.

7.6.3. Pausa y reanudación

El sistema permite pausar y reanudar el procesamiento. Esta función conserva el estado del análisis, mantiene visible el último fotograma procesado y evita que el usuario deba reiniciar el video desde el comienzo. En aplicaciones prácticas, esta capacidad es útil cuando se requiere revisar una detección, verificar una lectura o controlar la ejecución sobre videos largos.

7.6.4. Exportación de resultados

La exportación genera una estructura de almacenamiento con imágenes recortadas de placas y textos reconocidos. El sistema crea una carpeta principal asociada al video procesado y una subcarpeta `placas_finales/`, donde se almacenan las placas filtradas, únicas y seleccionadas por confianza o reglas definidas [2].

La Fig. 7.4 presenta la estructura de exportación.

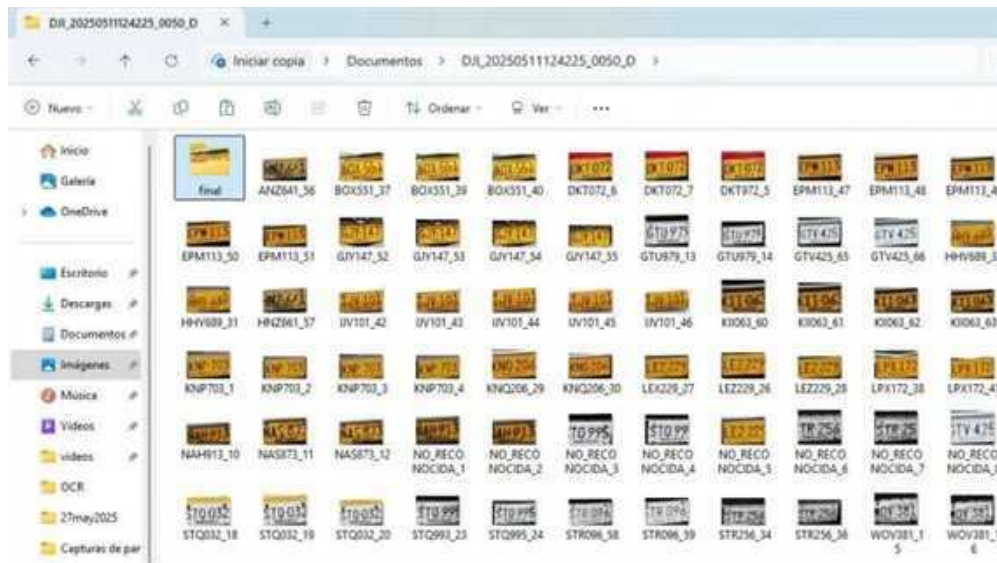


Figura 7.4: Estructura de exportación de resultados. El sistema organiza las placas procesadas y conserva una subcarpeta de placas finales con lecturas filtradas y no duplicadas.

7.7. Requerimientos técnicos

El sistema está diseñado para ejecutarse en computadores capaces de procesar video y ejecutar modelos de aprendizaje profundo. La documentación recomienda procesador Intel Core i5 como mínimo, Intel Core i7/i9 o AMD Ryzen 7 como configuración recomendada, 8 GB de RAM como mínimo y 16 GB o más como configuración recomendada. Para aceleración de inferencia, se recomienda GPU NVIDIA con soporte CUDA, preferiblemente RTX 3060 o superior [1], [2].

En software, la documentación indica compatibilidad con Windows 10/11 y Linux Ubuntu 20.04 o superior, Python 3.8 o superior y dependencias administradas mediante un archivo de requerimientos. Las bibliotecas principales incluyen Ultralytics YOLO, Transformers para TrOCR, EasyOCR como alternativa, OpenCV, PIL, Tkinter y Torch [1].

Tabla 7.1: Requerimientos técnicos generales del sistema.

Componente	Mínimo	Recomendado
Procesador	Intel Core i5, cuatro núcleos	Intel Core i7/i9 o AMD Ryzen 7 o superior
Memoria RAM	8 GB	16 GB o más
GPU	NVIDIA GTX 1050 o superior con CUDA	NVIDIA RTX 3060 o superior
Almacenamiento	2 GB disponibles	10 GB o más para resultados
Pantalla	1366 × 768	1920 × 1080 o superior
Sistema operativo	Windows 10/11 o Ubuntu 20.04	Sistema con Python y controladores CUDA configurados

El video de entrada debe provenir preferiblemente de cámara estática, con iluminación adecuada y resolución mínima de 640×480 píxeles. Estas condiciones no eliminan la variabilidad propia de escenarios reales, pero favorecen la estabilidad de la detección y la legibilidad de las regiones recortadas.

7.8. Validación funcional, uso y protección intelectual

La validación disponible en los documentos anexos es funcional, operativa e institucional. El certificado de uso señala que el software se encuentra activo dentro de las actividades de investigación, desarrollo e innovación del grupo GEPAR y de InterTelco S.A.S. También documenta su uso en procesamiento de videos viales, generación de registros cuantitativos de tránsito, evaluación de desempeño de cámaras y algoritmos, y formación de estudiantes en visión artificial y aprendizaje profundo [3].

La protección intelectual está respaldada por el certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor. El certificado identifica la obra con el título oficial *CAPTA - Captura Automática de Placas en Tránsito Avanzado*, registra como autores a David Fernández Mc Cann, Claudia Cristina Botero Suárez y Johan David Rojas Martínez, y establece como titulares patrimoniales y productores a la Universidad de Antioquia e InterTelco S.A.S. [5].

El registro corresponde al Libro–Tomo–Partida 13-106-30, con fecha 29 de septiembre de 2025. La descripción oficial de la obra indica que el software detecta y reconoce automáticamente placas vehiculares en videos capturados por cámaras estáticas, está implementado

en Python con YOLOv8 y TrOCR, utiliza OpenCV, PIL y Torch, y cuenta con una interfaz gráfica en Tkinter para cargar videos, procesarlos en tiempo real, pausar y reanudar el análisis, y exportar imágenes de placas detectadas con sus textos reconocidos [5].

Estos elementos permiten situar el desarrollo como producto tecnológico formalizado. No se trata únicamente de un prototipo algorítmico, sino de una herramienta documentada, registrada, usada y articulada a un convenio universidad–empresa.

7.9. Aplicaciones y transferencia

El desarrollo puede aplicarse en sistemas de control vehicular, gestión de tráfico, seguridad vial, accesos empresariales, peajes, zonas logísticas, seguimiento de flotas y analítica urbana. Su arquitectura modular permite integrarlo con cámaras IP, plataformas IoT o sistemas de monitoreo por video, siempre que se garantice un flujo de captura compatible con el procesamiento de fotogramas y la lectura de placas [7].

En el ámbito académico, el sistema funciona como caso de estudio para cursos y semilleros relacionados con visión artificial, aprendizaje profundo, procesamiento de video, detección de objetos y OCR. Su documentación facilita que estudiantes y docentes comprendan la cadena completa: captura, inferencia, recorte, reconocimiento, filtrado, visualización y exportación. Esta condición fortalece la capacidad del grupo GEPAR para convertir problemas industriales en ejercicios formativos y productos de investigación aplicada [7].

La articulación con InterTelco S.A.S. amplía el alcance del desarrollo hacia escenarios de validación industrial. El sistema puede evolucionar hacia soluciones de análisis distribuido, integración con redes de cámaras, tableros de supervisión, bases de datos de tránsito o servicios de inteligencia vial. Esta proyección depende de nuevas etapas de validación, robustecimiento, despliegue y evaluación cuantitativa en condiciones operativas diversas.

7.10. Discusión

El desarrollo resuelve una necesidad práctica: convertir videos viales en registros estructurados de placas vehiculares mediante una cadena automatizada de detección, OCR y exportación. La contribución técnica no reside únicamente en usar modelos de aprendizaje profundo, sino en integrarlos dentro de un flujo operable por usuario, documentado y formalizable.

Desde el punto de vista de ingeniería, el sistema articula cinco decisiones relevantes. La

primera es la separación entre detección de placa y reconocimiento de caracteres, lo cual permite especializar cada etapa del procesamiento. La segunda es el uso de una interfaz gráfica que reduce la barrera de operación. La tercera es la exportación de evidencias visuales junto con textos reconocidos, lo que favorece trazabilidad. La cuarta es el filtrado de placas finales, que evita conservar únicamente detecciones repetidas sin consolidación. La quinta es su documentación técnica y de usuario, necesaria para mantenimiento, adopción y transferencia.

El desarrollo también plantea retos. La precisión del sistema puede depender de la resolución del video, el ángulo de captura, la velocidad del vehículo, la iluminación, el estado físico de la placa y la calidad del enfoque. Además, los documentos disponibles respaldan el uso activo y la estructura funcional del sistema, pero no presentan una evaluación estadística completa con métricas como precisión, exhaustividad, tasa de caracteres correctos o desempeño por condición de captura. Una etapa posterior debería incorporar protocolos de validación cuantitativa con bases de video etiquetadas, separación por escenarios y comparación contra lectura humana o sistemas de referencia.

A pesar de estas limitaciones, el sistema deja una base tecnológica reutilizable. La arquitectura puede adaptarse a otras configuraciones de cámara, nuevos modelos OCR, integración con bases de datos o módulos de seguimiento multiobjeto. En ese sentido, el desarrollo funciona como plataforma de trabajo para sistemas de inteligencia vial más amplios.

7.11. Conclusiones

El sistema presentado automatiza la detección y reconocimiento de placas vehiculares en video mediante una arquitectura que integra YOLOv8, TrOCR, procesamiento de imágenes, manejo de video e interfaz gráfica de usuario. Su flujo permite cargar videos, procesarlos cuadro a cuadro, detectar placas, reconocer caracteres, visualizar resultados y exportar evidencias en carpetas estructuradas.

El desarrollo responde a una necesidad concreta de supervisión vial y control vehicular: reducir la dependencia de revisión manual, organizar resultados y generar registros reutilizables. La herramienta conserva valor tanto en aplicaciones industriales como en entornos académicos, donde permite formar estudiantes en visión artificial aplicada y fortalecer capacidades del grupo GEPAR.

La documentación técnica, el certificado de uso y el registro ante la Dirección Nacional de Derecho de Autor respaldan su condición de producto tecnológico. El certificado DNDA, con Libro–Tomo–Partida 13-106-30 y fecha de registro 29 de septiembre de 2025, formaliza

el soporte lógico asociado a la detección y reconocimiento automático de placas vehiculares en videos capturados por cámaras estáticas.

El trabajo consolida una ruta de cooperación universidad–empresa en visión artificial aplicada a seguridad vial. Su continuidad natural incluye validación cuantitativa en escenarios reales, integración con sistemas de cámaras, conexión con bases de datos, despliegue en arquitecturas distribuidas y evaluación de desempeño bajo condiciones variables de captura.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia, al grupo de investigación GEPAR y a InterTelco S.A.S. por el entorno técnico, académico e institucional que hizo posible el desarrollo del sistema. Se reconoce además el marco del proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, orientado al desarrollo de sistemas de visión artificial para seguridad vial y supervisión de infraestructura.

Referencias

- [1] Universidad de Antioquia, *CAPTA: Captura Automática de Placas en Tránsito Avanzado. Manual técnico*, Documento técnico elaborado conforme al estándar IEEE 1063-2001., Universidad de Antioquia, Medellín, Colombia, 2025.
- [2] Universidad de Antioquia, *CAPTA: Captura Automática de Placas en Tránsito Avanzado. Manual de usuario*, Documento elaborado en el marco del proyecto VISIONARY entre la Universidad de Antioquia e InterTelco S.A.S., Universidad de Antioquia, Medellín, Colombia, 2025.
- [3] Grupo de Investigación GEPAR e InterTelco S.A.S., *Certificación de uso del software CAPTA por parte del grupo de investigación GEPAR y la empresa InterTelco S.A.S.* Certificación institucional, Certificado expedido el 15 de octubre de 2025., Medellín, Colombia, oct. de 2025.
- [4] Universidad de Antioquia, *Acta de registro y de propiedad intelectual del proyecto VISIONARY: Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, Acta institucional CODI, Vicerrectoría de Investigación, Acta N.º 2024-74783, fecha 13 de noviembre de 2024., Medellín, Colombia, nov. de 2024.

- [5] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: CAPTA - Captura Automática de Placas en Tránsito Avanzado*, Certificado de registro de soporte lógico, Libro–Tomo–Partida 13-106-30. Fecha de registro: 29 de septiembre de 2025. Autores: David Fernández Mc Cann, Claudia Cristina Botero Suárez y Johan David Rojas Martínez., Bogotá, Colombia, sep. de 2025.
- [6] M. Li et al., «TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models,» en *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023, págs. 13 094-13 102.
- [7] Universidad de Antioquia, *Solicitud para asignación de puntos por desarrollo de software: CAPTA, Captura Automática de Placas en Tránsito Avanzado*, Documento institucional interno, Formato de solicitud por software diligenciado el 30 de mayo de 2025., Medellín, Colombia, mayo de 2025.

Capítulo 8

Estimación automática de edad ósea a partir de radiografías de mano mediante visión artificial y aprendizaje profundo

Autores

Raúl Daza Liñán¹; David Stephen Fernández Mc Cann²; Pedro María de Santa Lucía Jaramillo Vallejo³

Resumen

La estimación de la edad ósea a partir de radiografías de mano constituye una tarea relevante en pediatría, odontología, ortopedia y ciencias forenses, debido a su utilidad para valorar la maduración esquelética, contrastar el desarrollo biológico con la edad cronológica y apoyar decisiones clínicas relacionadas con crecimiento y tratamiento. En este capítulo se presenta un desarrollo de software orientado a automatizar esta tarea mediante procesamiento digital de imágenes, visión artificial y aprendizaje profundo. El

¹Estudiante de Ingeniería Electrónica, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Universidad de Antioquia. Grupo de investigación POPCAD.

³Desarrollo interdisciplinario asociado a la Universidad de Antioquia, con articulación entre los grupos de investigación GEPAR y POPCAD.

sistema utiliza un modelo YOLOv8 para localizar regiones anatómicas inspiradas en el método Tanner–Whitehouse y una red neuronal convolucional para estimar la edad ósea a partir de las regiones detectadas. La solución incorpora una interfaz gráfica de usuario, permite cargar radiografías individuales o carpetas de imágenes, visualizar regiones de interés, ejecutar la predicción de edad ósea, ajustar manualmente el resultado cuando el criterio clínico lo requiera y guardar los resultados en formato DICOM. El desarrollo fue registrado ante la Dirección Nacional de Derecho de Autor bajo Libro–Tomo–Partida 13-104-111, con fecha de registro 30 de mayo de 2025. Su uso ha sido certificado por los grupos GEPAR y POPCAD, en el marco de actividades de investigación interdisciplinaria, formación estudiantil y transferencia de conocimiento en ingeniería aplicada a la salud [1], [2], [3], [4].

Palabras clave: edad ósea; carpogramas; radiografía de mano; Tanner–Whitehouse; YOLOv8; CNN; ResNet50; DICOM; visión artificial; aprendizaje profundo.

8.1. Introducción

La edad ósea permite estimar el grado de maduración esquelética de un paciente a partir de estructuras anatómicas visibles en radiografías de mano y muñeca. Esta información es útil cuando se requiere evaluar retrasos o adelantos en el crecimiento, planear intervenciones odontológicas u ortopédicas, acompañar estudios pediátricos o apoyar análisis forenses. A diferencia de la edad cronológica, que se mide desde el nacimiento, la edad ósea expresa el estado de desarrollo de los tejidos esqueléticos y puede diferir de manera significativa entre individuos de la misma edad.

Los métodos clásicos de evaluación de edad ósea se basan en la comparación visual con atlas de referencia o en la valoración sistemática de regiones anatómicas específicas. Entre estos enfoques, el método Tanner–Whitehouse ha tenido amplia relevancia porque propone evaluar huesos seleccionados de la mano y la muñeca mediante criterios de maduración esquelética definidos [5]. Esta aproximación ofrece una estructura clara para el análisis, pero exige conocimiento experto, tiempo de revisión y consistencia en la interpretación radiográfica.

El desarrollo presentado en este capítulo responde a la necesidad de apoyar esta evaluación mediante herramientas computacionales. El sistema implementado integra técnicas de procesamiento de imágenes, detección automática de regiones de interés y estimación basada en aprendizaje profundo. La herramienta fue diseñada para facilitar su uso por parte de profesionales de la salud y odontología, sin exigir que el usuario interactúe directamente con el código o con los modelos de inteligencia artificial.

La interfaz inicial del sistema permite cargar radiografías o carpetas de imágenes, y constituye el punto de entrada del flujo de trabajo. La Fig. 8.1 muestra esta vista, desde la cual el usuario inicia la operación del sistema.

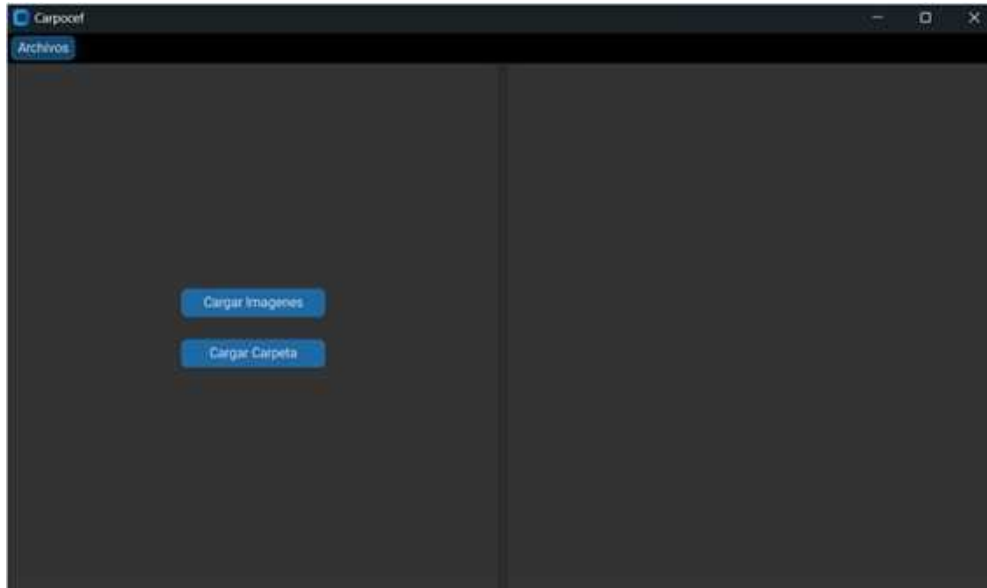


Figura 8.1: Vista inicial del sistema para la carga de radiografías individuales o carpetas de imágenes. La interfaz organiza el ingreso de los estudios antes de avanzar hacia las etapas de análisis y predicción.

El software fue registrado ante la Dirección Nacional de Derecho de Autor con el título *CARPOSOFT - Software para la estimación de edad ósea a partir de carpogramas*. El certificado DNDA lo describe como un sistema que estima la edad ósea a partir de radiografías de mano mediante procesamiento de imágenes, visión artificial y aprendizaje profundo, con detección de regiones clave basadas en el método Tanner–Whitehouse y cálculo de edad ósea mediante redes neuronales convolucionales [1]. En este capítulo se conserva el nombre del software dentro del texto y en las referencias, pero el título del capítulo se formula de acuerdo con la función técnica y científica del desarrollo.

8.2. Problema abordado

La evaluación manual de la edad ósea presenta tres dificultades operativas. La primera está asociada al tiempo requerido para revisar la radiografía, localizar estructuras relevantes y estimar la maduración. La segunda corresponde a la variabilidad entre observadores, especialmente cuando el análisis depende de criterios visuales finos. La tercera se relaciona

con la trazabilidad del proceso, pues no siempre se conservan de manera estructurada las regiones evaluadas, el resultado final, la información del paciente y los ajustes realizados por el especialista.

Desde el punto de vista computacional, el problema puede expresarse como la estimación de una variable continua o discretizada de maduración esquelética a partir de una imagen radiográfica de entrada. Sea

$$I : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R}$$

una radiografía digital de mano, definida sobre un dominio discreto de píxeles Ω . El objetivo es construir una función

$$F(I) = \hat{a},$$

donde \hat{a} representa la edad ósea estimada. Sin embargo, en lugar de aplicar directamente un modelo sobre la radiografía completa, el desarrollo implementado descompone el problema en una secuencia de operaciones: preprocesamiento, detección de regiones anatómicas, recorte, normalización, composición de regiones y predicción final.

Esta estrategia busca que el sistema use información anatómica localizada, coherente con la lógica Tanner–Whitehouse. De esta manera, la estimación no depende únicamente de una representación global de la imagen, sino de regiones de interés que concentran información de maduración esquelética.

8.3. Arquitectura funcional del sistema

La arquitectura funcional del sistema se organiza como una cadena de procesamiento en la que cada etapa transforma la información de entrada antes de entregarla al siguiente bloque. La radiografía se preprocesa, luego se detectan regiones de interés mediante YOLOv8, se recortan y normalizan las regiones Tanner–Whitehouse, se construye una imagen compuesta y finalmente se estima la edad ósea mediante una red neuronal convolucional. El flujo incluye además una etapa de revisión clínica opcional y una salida estructurada en formato DICOM.

La Fig. 8.2 presenta el flujograma general del sistema. Esta figura resume la lógica completa del procesamiento, desde la radiografía de entrada hasta la exportación de resultados.

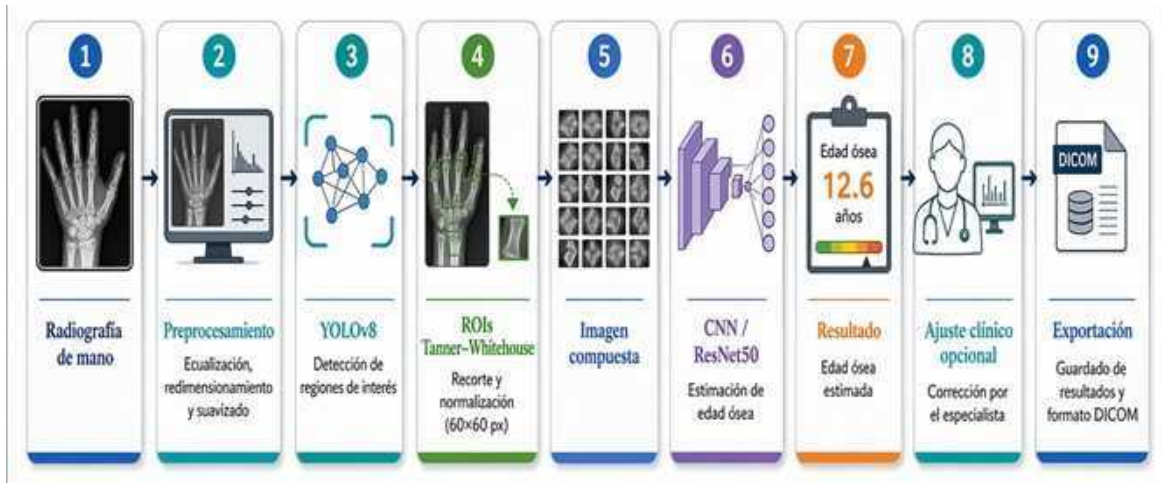


Figura 8.2: Flujo funcional del sistema para estimación automática de edad ósea. El procesamiento integra preprocesamiento radiográfico, detección de regiones de interés mediante YOLOv8, recorte y normalización de regiones Tanner–Whitehouse, construcción de una imagen compuesta, predicción mediante CNN/ResNet50, ajuste clínico opcional y exportación de resultados en formato DICOM.

Formalmente, el sistema puede representarse como una composición de funciones:

$$S(I) = E \circ A \circ f_\phi \circ \Gamma \circ \mathcal{N} \circ C \circ D_\theta \circ P(I),$$

donde P corresponde al preprocesamiento de la radiografía, D_θ es el detector de regiones de interés, C realiza el recorte de cada región, \mathcal{N} normaliza geométrica y radiométricamente los recortes, Γ organiza las regiones en una imagen compuesta, f_ϕ estima la edad ósea, A representa el ajuste clínico opcional y E exporta los resultados.

El manual técnico describe que el preprocesamiento aplica ecualización de histograma, redimensionamiento del eje mayor a 640 píxeles y un difuminado ligero que mejora el desempeño del detector [4]. Esta operación puede expresarse como

$$I_p = \mathcal{G}_\sigma (\mathcal{R}_{640} (\mathcal{H}(I))),$$

donde \mathcal{H} representa la ecualización de histograma, \mathcal{R}_{640} el redimensionamiento proporcional de la imagen y \mathcal{G}_σ el suavizado aplicado con parámetro σ . La imagen resultante I_p se entrega al modelo de detección.

8.4. Detección de regiones de interés

La detección automática de regiones anatómicas constituye el primer bloque de inteligencia artificial del sistema. Sobre la radiografía preprocesada se aplica un modelo YOLOv8 entrenado para identificar regiones clave inspiradas en el método Tanner–Whitehouse. El resultado de esta etapa es un conjunto de cajas delimitadoras:

$$\mathcal{B} = D_{\theta}(I_p) = \{b_1, b_2, \dots, b_n\},$$

donde cada b_k representa una región candidata y θ agrupa los parámetros del detector. Cada caja puede escribirse como

$$b_k = (x_k, y_k, w_k, h_k, s_k),$$

donde x_k y y_k indican la posición de referencia, w_k y h_k corresponden al ancho y alto de la región, y s_k representa la confianza de la detección.

La interfaz permite visualizar las regiones detectadas, lo cual aporta trazabilidad al proceso y facilita que el usuario comprenda qué partes de la radiografía están siendo utilizadas por el sistema. La Fig. 8.3 muestra la vista correspondiente a las regiones de interés.



Figura 8.3: Vista de regiones de interés Tanner–Whitehouse detectadas sobre la radiografía de mano. El panel derecho muestra la región seleccionada y permite guardar los recortes para documentación o análisis posterior.

A partir de cada caja b_k , el sistema extrae una región de interés:

$$R_k = C(I_p, b_k),$$

donde $C(\cdot)$ representa la operación de recorte. El manual técnico señala que estas regiones se transforman en imágenes cuadradas, se ecualizan, se redimensionan a 60×60 píxeles y luego se organizan para formar una imagen compuesta [4]. Este proceso puede representarse como

$$\tilde{R}_k = \mathcal{N}(R_k), \quad \tilde{R}_k \in \mathbb{R}^{60 \times 60},$$

y posteriormente como

$$M = \Gamma(\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_n),$$

donde M es la imagen compuesta que resume las regiones anatómicas detectadas.

8.5. Predicción de edad ósea

La etapa de predicción recibe la imagen compuesta M y la procesa mediante una red neuronal convolucional. La documentación técnica indica que el sistema utiliza una CNN para la estimación de edad ósea y refiere el uso de ResNet50 como modelo preentrenado asociado a esta tarea [4]. La predicción puede expresarse como

$$\hat{a} = f_\phi(M),$$

donde f_ϕ representa el modelo de aprendizaje profundo y \hat{a} la edad ósea estimada.

La vista de predicción BAA presenta al usuario la radiografía, la información del paciente y el resultado estimado. También permite modificar el valor final cuando el profesional considera que el resultado automático debe ajustarse. La Fig. 8.4 muestra esta etapa de interacción.

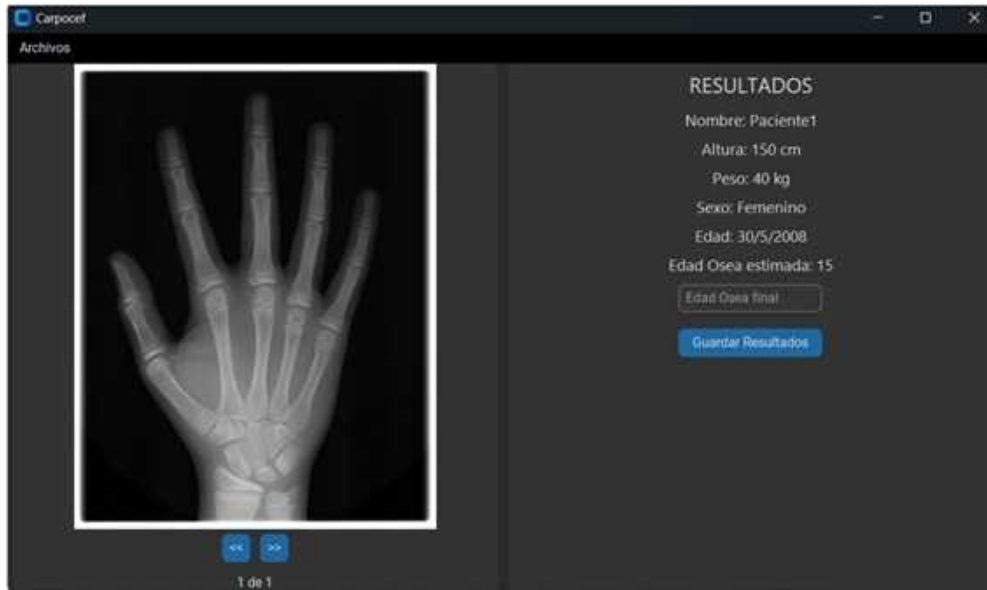


Figura 8.4: Vista de predicción BAA. El sistema presenta la edad ósea estimada, conserva la información del paciente y habilita un campo de corrección para que el especialista ajuste el resultado antes de guardarlo.

La posibilidad de edición clínica es relevante porque mantiene el carácter asistivo del sistema. El resultado automático no reemplaza el juicio experto, sino que lo apoya. Si se denota por a^* el valor ajustado por el especialista, la edad ósea final almacenada se define como

$$a_f = \begin{cases} a^*, & \text{si el especialista modifica la predicción,} \\ \hat{a}, & \text{si se conserva la predicción automática.} \end{cases}$$

Esta estructura permite integrar inferencia automática y criterio profesional dentro de un mismo flujo operativo.

8.6. Interfaz de usuario y organización del trabajo clínico

La interfaz gráfica fue diseñada para reducir la complejidad operativa del sistema. El usuario no interactúa directamente con los modelos ni con los archivos internos de pesos, sino con vistas organizadas que guían el flujo de trabajo. La vista de imágenes cargadas

permite revisar las radiografías incorporadas, navegar entre ellas e ingresar información del paciente antes de ejecutar el análisis.

La Fig. 8.5 muestra la vista de imágenes cargadas, en la que el usuario organiza el caso antes de acceder a las regiones de interés o a la predicción.



Figura 8.5: Vista de imágenes cargadas. El sistema muestra la radiografía seleccionada, permite navegar entre estudios y habilita el ingreso de información del paciente antes de la detección de regiones y la predicción de edad ósea.

El manual de usuario describe cuatro vistas principales: carga de carpetas o imágenes, imágenes cargadas, regiones TW ROIs y predicción BAA [3]. Esta organización separa las acciones del usuario según el estado del análisis: primero se ingresan las imágenes, luego se revisa la información del paciente, después se visualizan las regiones anatómicas y finalmente se obtiene la predicción.

El sistema también incorpora un menú que permite añadir imágenes, añadir carpetas, eliminar imágenes, regresar a la vista de carga, acceder a la vista de regiones, acceder a la vista de predicción y salir de la aplicación. La Fig. 8.6 documenta esta forma de navegación.

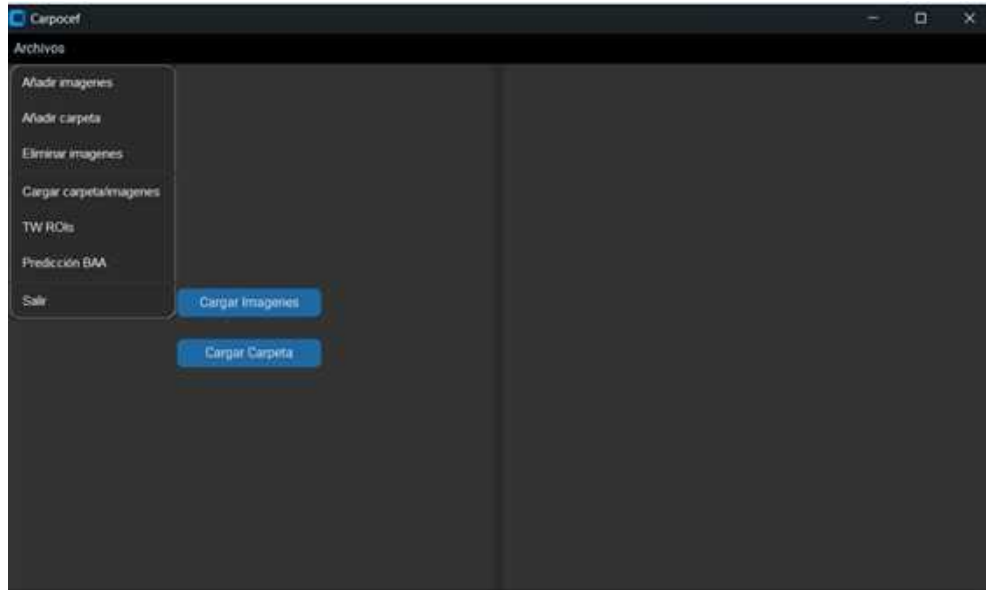


Figura 8.6: Menú de navegación del sistema. Las opciones permiten gestionar radiografías, cambiar entre vistas de análisis y cerrar la aplicación de forma controlada.

La interfaz cumple una función técnica y formativa. Técnica, porque encapsula modelos de visión artificial en un entorno accesible para usuarios clínicos. Formativa, porque permite mostrar de manera secuencial cómo una radiografía se transforma en regiones, predicción y resultado exportable.

8.7. Formato DICOM y trazabilidad

El sistema guarda resultados en formato DICOM, estándar usado para transmisión, almacenamiento, recuperación, procesamiento y visualización de información médica. La documentación técnica indica que el archivo generado conserva cabecera DICOM, datos del paciente, información específica de la evaluación y datos de imagen [4].

El conjunto de metadatos incluye identificadores de estudio, serie e instancia, modalidad, nombre del paciente, fecha de nacimiento, sexo, peso, estatura, comentarios adicionales y edad ósea. Esta organización permite que el resultado no quede reducido a un número aislado, sino que se conserve asociado a la radiografía y a la información clínica del caso.

En términos de trazabilidad, cada análisis puede representarse como una tupla:

$$T = (I, \mathcal{B}, M, \hat{a}, a_f, \mathcal{D}),$$

donde I es la radiografía original, \mathcal{B} el conjunto de regiones detectadas, M la imagen compuesta, \hat{a} la edad ósea estimada automáticamente, a_f la edad final aceptada o ajustada, y \mathcal{D} el conjunto de metadatos DICOM asociados al paciente y al estudio.

Esta estructura favorece la revisión posterior del análisis, la conservación de evidencia, la interoperabilidad y la integración con entornos clínicos o investigativos.

8.8. Uso institucional y formación

El certificado de uso indica que el software es utilizado por los grupos GEPAR y POPCAD en proyectos interdisciplinarios con participación de docentes y estudiantes de ingeniería y odontología [2]. El mismo documento señala que la herramienta apoya investigación sobre maduración esquelética en población colombiana y se proyecta como recurso útil en odontología, medicina y ciencias forenses.

En la línea de GEPAR, el desarrollo fortalece capacidades en visión por computador, aprendizaje profundo, procesamiento de imágenes médicas, diseño de interfaces y estructuración de productos tecnológicos. En la línea de POPCAD, el sistema se inserta en estudios de crecimiento y desarrollo craneofacial, prácticas clínicas, semilleros de investigación y proyectos colaborativos con áreas de salud.

La participación de estudiantes es un componente importante del producto. La solicitud de asignación de puntos por software indica que el desarrollo involucró tareas de entrenamiento de modelos, diseño de interfaz gráfica, procesamiento de imágenes médicas y estructuración de la arquitectura del sistema [6]. En este sentido, el software opera también como plataforma de aprendizaje aplicado, donde convergen programación, inteligencia artificial, imagen médica y necesidades clínicas reales.

8.9. Protección intelectual

El certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor identifica la obra como *CARPOSOFT - Software para la estimación de edad ósea a partir de carpogramas*, con fecha de registro 30 de mayo de 2025 y Libro–Tomo–Partida 13-104-111 [1]. El documento reconoce como autores a David Fernández Mc Cann, Pedro María de Santa Lucía Jaramillo Vallejo y Raúl Daza Liñán, y como titular patrimonial y productor a la Universidad de Antioquia.

La descripción oficial establece que el software estima la edad ósea a partir de radiografías

de mano mediante procesamiento de imágenes, visión artificial y aprendizaje profundo. También indica que detecta regiones clave basadas en el método Tanner–Whitehouse, calcula la edad ósea mediante redes neuronales convolucionales, organiza registros, automatiza la evaluación, almacena resultados y genera reportes en formato DICOM para integración en entornos clínicos.

Este registro formaliza el desarrollo como producto tecnológico de soporte lógico. Su valor institucional no reside solo en la existencia del modelo computacional, sino en la integración de algoritmo, interfaz, documentación, uso certificado y protección intelectual.

8.10. Discusión

El desarrollo se ubica dentro de una línea internacional de automatización de la evaluación de edad ósea. La literatura reciente muestra enfoques basados en carpogramas, modelos globales sobre la mano completa, estrategias global-locales, métodos multimodales y arquitecturas orientadas a reducir error en meses sobre bases como RSNA [7], [8]. En ese panorama, el sistema presentado se aproxima a los enfoques por regiones, al usar detección de zonas anatómicas relacionadas con Tanner–Whitehouse antes de ejecutar la predicción final.

La comparación con sistemas modernos debe hacerse con cautela. Algunos trabajos reportan métricas como MAE o MAD en meses, mientras otros reportan clasificación por umbrales de edad, exactitud acumulada o medidas derivadas de tareas binarias. El documento comparativo revisado para este proyecto señala que las métricas no son directamente equivalentes y que la eficiencia debe evaluarse considerando precisión, costo computacional, necesidad de anotaciones adicionales y generalización a nuevas cohortes [9].

En este capítulo no se reporta una validación estadística externa del desempeño del sistema, porque los documentos disponibles se concentran en arquitectura, uso, registro y operación. Una etapa posterior debería incluir evaluación cuantitativa con una base de radiografías anotadas, separación entre entrenamiento y prueba, análisis por sexo y edad, comparación contra estimaciones expertas y reporte de métricas como MAE, desviación estándar del error y porcentaje de casos dentro de márgenes clínicamente aceptables.

Aun así, el desarrollo representa una contribución relevante como producto tecnológico universitario. Integra modelos de detección y predicción, interfaz gráfica, trazabilidad DICOM, ajuste clínico opcional y uso interdisciplinario. Esa combinación lo diferencia de un experimento aislado de aprendizaje profundo, porque lo acerca a una herramienta de apoyo clínico, formación e investigación aplicada.

8.11. Conclusiones

Se presentó un sistema para la estimación automática de edad ósea a partir de radiografías de mano, basado en procesamiento digital de imágenes, detección de regiones de interés mediante YOLOv8 y predicción con redes neuronales convolucionales. La herramienta permite cargar radiografías, ingresar información del paciente, visualizar regiones Tanner–Whitehouse, estimar la edad ósea, ajustar el resultado bajo criterio profesional y exportar la información en formato DICOM.

El desarrollo conserva una relación explícita entre ingeniería y salud. Desde la ingeniería, aporta una arquitectura funcional de visión artificial aplicada a imágenes médicas. Desde el ámbito clínico y odontológico, ofrece una herramienta de apoyo para valorar maduración esquelética. Desde la perspectiva institucional, se consolida como producto de investigación aplicada con uso certificado por los grupos GEPAR y POPCAD y registro de soporte lógico ante la Dirección Nacional de Derecho de Autor.

El sistema constituye una base tecnológica para trabajos posteriores en validación clínica, adaptación poblacional, mejoramiento de modelos, integración con sistemas hospitalarios y ampliación hacia otras tareas de análisis de maduración ósea. Su desarrollo muestra cómo una necesidad clínica puede transformarse en un producto computacional documentado, operable y protegido intelectualmente.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia y a los grupos de investigación GEPAR y POPCAD por el entorno académico e interdisciplinario que permitió el desarrollo del sistema. Se reconoce además la participación de estudiantes y docentes vinculados a las actividades de diseño, entrenamiento, documentación y aplicación del software en el marco de la investigación en ingeniería aplicada a la salud.

Referencias

- [1] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: CARPOSOFIT - Software para la estimación de edad ósea a partir de carpogramas*, Certificado de registro de soporte lógico, Libro–Tomo–Partida 13-104-111. Fecha de registro: 30 de mayo de 2025. Autores: David Fernández Mc Cann,

- Pedro María de Santa Lucía Jaramillo Vallejo y Raúl Daza Liñán., Bogotá, Colombia, mayo de 2025.
- [2] Grupo de Investigación GEPAR y Grupo POPCAD, *Certificación de uso del software CARPOSOF*T por parte de los grupos de investigación GEPAR y POPCAD, Certificación institucional, Certificado expedido el 12 de junio de 2025., Medellín, Colombia, jun. de 2025.
- [3] Universidad de Antioquia, *Carposoft: Manual de usuario. Software para la estimación de edad ósea a partir de carpogramas*, Universidad de Antioquia, Medellín, Colombia, 2025.
- [4] Universidad de Antioquia, *Carposoft: Manual técnico. Software para la estimación de edad ósea a partir de carpogramas*, Universidad de Antioquia, Medellín, Colombia, 2025.
- [5] J. M. Tanner, M. J. R. Healy, H. Goldstein y N. Cameron, *Assessment of Skeletal Maturity and Prediction of Adult Height (TW3 Method)*, 3.^a ed. London: Saunders, 2001.
- [6] Universidad de Antioquia, *Solicitud para asignación de puntos por desarrollo de software: CARPOSOF*T – *Software para la estimación de edad ósea a partir de carpogramas*, Documento institucional interno, Formato de solicitud por software diligenciado el 30 de mayo de 2025., Medellín, Colombia, mayo de 2025.
- [7] F. Torres, M. A. Bravo, E. Salinas, G. Triana y P. Arbeláez, «Bone Age Detection via Carpogram Analysis Using Convolutional Neural Networks,» en *Proceedings of the 13th International Symposium on Medical Information Processing and Analysis*, 2017, pág. 1 057 217. DOI: 10.1117/12.2310208.
- [8] Radiological Society of North America, *RSNA Pediatric Bone Age Challenge 2017*, AI Image Challenge, Competencia de referencia para evaluación automática de edad ósea pediátrica., 2017.
- [9] Fernández Mc Cann, David Stephen, *Comparación de sistemas para evaluación de edad ósea: síntesis comparativa basada en la línea CINFONIA/Uniandes y reportes recientes de la literatura*, Documento de trabajo, Documento preparado el 22 de abril de 2026., Medellín, Colombia, abr. de 2026.

Capítulo 9

Detección, clasificación y conteo automatizado de microalgas verdes en imágenes microscópicas

Autores

Juan Guillermo Quevedo¹; Yarin Tatiana Puerta Quintana²; David Stephen Fernández Mc Cann³

Resumen

La identificación y cuantificación de microalgas verdes de agua dulce constituye una tarea relevante en ecotoxicología, monitoreo ambiental y evaluación de calidad del agua. Sin embargo, el conteo manual en imágenes microscópicas demanda tiempo, entrenamiento taxonómico, control visual sostenido y criterios consistentes de clasificación. En este capítulo se presenta un desarrollo de software orientado a automatizar la detección, clasificación y conteo de microalgas verdes en imágenes microscópicas. El sistema se enfoca en tres especies de interés ecotoxicológico: *Pseudopediastrum boryanum*, *Desmodesmus communis* y *Raphidocelis subcapitata*. La herramienta integra una interfaz gráfica de usuario, carga

¹Universidad de Antioquia. Grupo de investigación GeoLimna.

²Universidad de Antioquia. Grupo de investigación GeoLimna.

³Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Desarrollo realizado en el marco de la articulación entre los grupos de investigación GEPAR y GeoLimna, con soporte en proyectos asociados a SIMEVAM y a la automatización de análisis de microalgas verdes de agua dulce mediante visión artificial.

de imágenes microscópicas, recorte automático del campo útil, inferencia con un modelo YOLO previamente entrenado, conteo por clase y exportación de resultados en archivos estructurados. El software fue registrado ante la Dirección Nacional de Derecho de Autor bajo Libro–Tomo–Partida 13-105-375, con fecha de registro 05 de septiembre de 2025. Su uso ha sido certificado por los grupos GEPAR y GeoLimna en actividades de investigación, formación estudiantil, monitoreo de calidad del agua y validación experimental en laboratorio [1], [2], [3], [4].

Palabras clave: microalgas; visión artificial; ecotoxicología; imágenes microscópicas; YOLO; conteo automático; bioindicadores; calidad del agua; GeoLimna; GEPAR.

9.1. Introducción

Las microalgas ocupan una posición central en los ecosistemas acuáticos. Su papel como productores primarios, su sensibilidad frente a cambios ambientales y su uso en ensayos de toxicidad las convierten en organismos de alto valor para la ecotoxicología y el monitoreo de calidad del agua. En particular, especies verdes de agua dulce como *Pseudopediastrum boryanum*, *Desmodesmus communis* y *Raphidocelis subcapitata* son relevantes como modelos biológicos en ensayos ambientales y estudios de respuesta frente a contaminantes [3], [5].

La identificación y cuantificación de microalgas en muestras microscópicas se realiza tradicionalmente mediante observación directa. Este procedimiento exige personal entrenado, tiempo de laboratorio, criterios taxonómicos estables y una alta concentración visual. Cuando el volumen de imágenes aumenta o cuando se requiere comparar resultados entre ensayos, la dependencia del observador introduce variabilidad y limita la reproducibilidad del proceso.

El avance del procesamiento digital de imágenes, la visión por computador y los modelos de aprendizaje profundo permite abordar esta tarea desde una perspectiva automatizada. La detección y clasificación de organismos microscópicos puede formularse como un problema de reconocimiento de patrones: a partir de una imagen microscópica, el sistema debe localizar individuos, asignarles una clase biológica y producir un conteo verificable por especie.

El software registrado ante la Dirección Nacional de Derecho de Autor con el título *ALGID Identificación de microalgas utilizando inteligencia artificial* fue concebido para automatizar la detección, clasificación y conteo de tres especies de microalgas verdes en imágenes de microscopio. Su interfaz principal, mostrada en la Fig. 9.1, organiza el flujo de traba-

jo en torno a la carga de imágenes, la ejecución del modelo, el conteo por especie y el guardado de resultados.

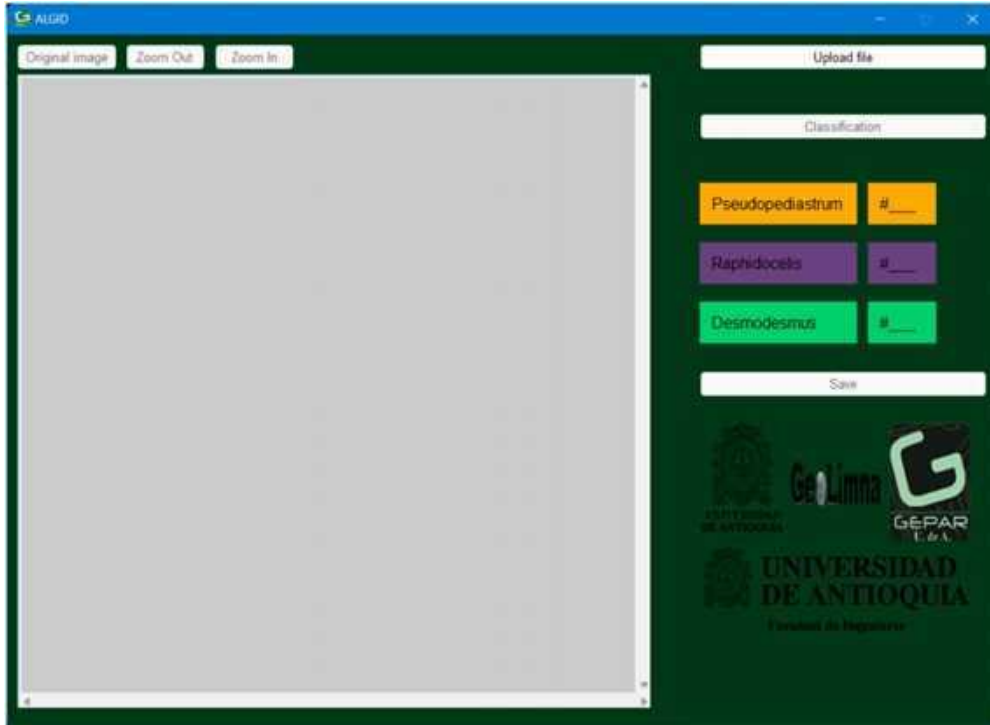


Figura 9.1: Interfaz principal del sistema. El panel izquierdo contiene el área de visualización de la imagen microscópica, mientras que el panel derecho concentra las acciones de carga, clasificación, conteo por especie y guardado de resultados.

El certificado de registro describe la obra como un software para detección, clasificación y conteo automatizado de microalgas en imágenes microscópicas mediante inteligencia artificial y visión por computador, con generación de reportes estructurados de las especies detectadas [1]. En este capítulo se conserva el nombre del software dentro del texto y en las referencias, pero el título se formula de acuerdo con la función técnica y científica del desarrollo.

9.2. Problema abordado

El conteo manual de microalgas puede entenderse como una operación secuencial en la que un observador localiza organismos, reconoce su morfología, los asigna a una especie y registra el número de individuos encontrados. Este flujo presenta tres dificultades principales: dependencia del experto, consumo de tiempo y variabilidad en la interpretación de imágenes con diferente densidad, iluminación, enfoque o presencia de residuos.

Desde el punto de vista computacional, el problema consiste en construir una función que transforme una imagen microscópica de entrada en una imagen anotada y un conjunto de conteos por especie. Sea I_0 la imagen microscópica original. El sistema busca implementar una función

$$F : I_0 \longrightarrow (I_{\text{ann}}, \mathbf{N}, \mathcal{R}),$$

donde I_{ann} es la imagen anotada con los individuos detectados, \mathbf{N} es el vector de conteos por clase y \mathcal{R} es el conjunto de archivos generados como salida del procesamiento.

El vector de clases se define como

$$\mathcal{K} = \left\{ \begin{array}{l} Pseudopediastrum boryanum, \\ Desmodesmus communis, \\ Raphidocelis subcapitata \end{array} \right\}.$$

Para cada clase $k \in \mathcal{K}$, el conteo automático se expresa como

$$N_k = \sum_{i=1}^M \mathbf{1}(\hat{y}_i = k),$$

donde M representa el número total de individuos detectados, \hat{y}_i es la clase asignada al individuo i , y $\mathbf{1}(\cdot)$ es la función indicadora. El resultado final del análisis queda dado por

$$\mathbf{N} = (N_{\text{Pseudopediastrum}}, N_{\text{Desmodesmus}}, N_{\text{Raphidocelis}}).$$

Esta formulación permite entender el desarrollo no solo como una interfaz gráfica, sino como una cadena de medición computacional. La imagen se convierte en evidencia anotada, la evidencia en conteos y los conteos en información utilizable para estudios ambientales o ecotoxicológicos.

9.3. Descripción funcional del sistema

El sistema opera como una cadena funcional de siete etapas: carga de imagen, recorte del campo microscópico útil, visualización de la imagen, inferencia con YOLO, clasificación por especie, conteo automático y exportación de resultados. Esta arquitectura convierte

una imagen capturada en microscopio en un conjunto de evidencias visuales y datos cuantitativos organizados.

La Fig. 9.2 resume esta lógica. La imagen microscópica ingresa al sistema, se recorta para conservar el campo útil, se entrega al modelo de detección y clasificación, se generan recuadros de identificación por individuo, se agrupan las detecciones por especie y se exportan los resultados.



Figura 9.2: Flujo funcional del sistema para detección, clasificación y conteo de microalgas. El procesamiento inicia con la imagen microscópica, continúa con el recorte del campo útil, la inferencia con YOLO, el conteo por especie y finaliza con la generación de archivos de salida.

La operación puede representarse como una composición de funciones:

$$S(I_0) = E \circ C_{\text{count}} \circ D_{\theta} \circ R \circ L(I_0),$$

donde L representa la carga de imagen, R el recorte del campo útil del microscopio, D_{θ} la detección y clasificación mediante el modelo YOLO parametrizado por θ , C_{count} el conteo por clase, y E la exportación de resultados.

El manual de usuario señala que el programa recorta la imagen del microscopio para conservar el foco e ignorar la zona negra de la captura. Esta decisión reduce la región de búsqueda y concentra el análisis en el área donde se encuentran los organismos [3]. Si se denota por Ω_0 el dominio total de la imagen original y por Ω_f el dominio útil de observación, el recorte puede expresarse como

$$I_f(x, y) = I_0(x, y), \quad (x, y) \in \Omega_f \subset \Omega_0.$$

El detector se aplica entonces sobre I_f , no sobre toda la imagen original. Esto mejora

la relación entre información útil y fondo no informativo, y hace más estable el flujo de inferencia.

9.4. Contexto ambiental y biológico

El proyecto original de identificación de algas plantea que el incremento en la producción de compuestos sintéticos y su entrada a ecosistemas acuáticos hacen necesario disponer de herramientas más eficientes para evaluar respuestas biológicas. Los cuerpos de agua reciben contaminantes de diferente naturaleza y las microalgas, por su sensibilidad y posición ecológica, pueden servir como indicadores de estado ambiental y como modelos de ensayo en laboratorio [5].

Las tres especies consideradas tienen interés particular. *Raphidocelis subcapitata* es ampliamente usada en ensayos de toxicidad acuática; *Desmodesmus communis* y *Pseudopediastrum boryanum* presentan morfologías coloniales que exigen criterios visuales distintos a los de especies unicelulares aisladas. Esta diversidad morfológica hace que el problema sea adecuado para modelos de visión por computador, porque combina detección de objetos pequeños, diferenciación de clases y conteo de individuos o agrupaciones visibles.

La utilidad de una herramienta como la desarrollada está en reducir la fricción entre la imagen microscópica y el dato cuantitativo. En un laboratorio de ecotoxicología, el resultado relevante no es únicamente la imagen, sino la cantidad de organismos observados, su distribución por especie y la posibilidad de conservar evidencia visual del proceso.

9.5. Carga y preparación de imágenes

El flujo de trabajo inicia con la selección de una imagen microscópica. El usuario carga el archivo desde la interfaz, y el sistema prepara la imagen para el análisis. La documentación indica que el software fue diseñado para imágenes capturadas desde microscopio con aumento de 400×, y que después de cargar la imagen el programa realiza el recorte automático del campo visual [3].

La Fig. 9.3 muestra un ejemplo de imagen ya incorporada al sistema. Esta etapa fija el punto de entrada del análisis y habilita la ejecución del modelo de clasificación.



Figura 9.3: Imagen microscópica cargada en el sistema después del recorte del campo útil. La interfaz habilita la clasificación automática y conserva la imagen como evidencia visual del análisis.

La preparación de la imagen cumple una función metodológica: estabiliza el área de análisis y evita que el modelo procese regiones no informativas. En términos de flujo de datos, el sistema transforma la imagen original I_0 en una imagen de análisis I_f , más adecuada para la inferencia.

9.6. Detección y clasificación mediante YOLO

El núcleo de inteligencia artificial del sistema se activa mediante el botón de clasificación. La documentación técnica indica que el software utiliza un modelo YOLO previamente entrenado, almacenado en el archivo `best.pt`, ubicado en el directorio `log`. Este archivo es fundamental para la ejecución del programa y puede reemplazarse si se dispone de un modelo mejorado, siempre que se conserve el nombre esperado por el sistema [4].

El modelo produce un conjunto de detecciones sobre la imagen recortada:

$$D_{\theta}(I_f) = \{(b_i, \hat{y}_i, p_i)\}_{i=1}^M,$$

donde b_i es la caja delimitadora del individuo detectado, \hat{y}_i es la clase asignada, p_i es la confianza de la predicción y M es el número total de detecciones. Cada caja puede escribirse como

$$b_i = (x_i, y_i, w_i, h_i),$$

donde x_i y y_i representan la posición de referencia, mientras w_i y h_i corresponden al ancho y alto de la región detectada.

La Fig. 9.4 presenta la salida de clasificación. Cada recuadro encierra un individuo detectado, muestra una confianza asociada y utiliza un color vinculado a la especie correspondiente. A la derecha de la imagen, la interfaz despliega el conteo por clase.

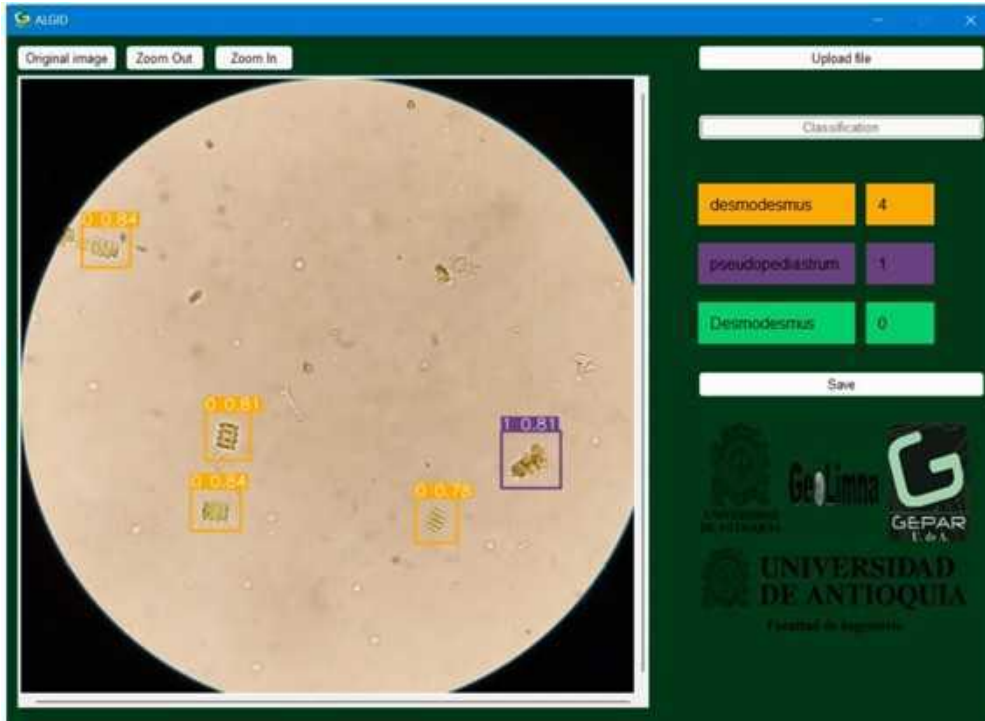


Figura 9.4: Resultado de clasificación automática. Cada recuadro delimita un individuo detectado y se acompaña de la clase asignada y de la confianza del modelo. El panel derecho resume el conteo por especie.

El conteo por clase se obtiene agrupando las detecciones según la clase predicha:

$$N_k = |\{i : \hat{y}_i = k, p_i \geq \tau\}|,$$

donde τ es un umbral de confianza. Esta expresión resume la lógica operativa del sistema:

solo las detecciones aceptadas por el modelo contribuyen al conteo final de la especie correspondiente.

9.7. Exportación de resultados

Al terminar el proceso de clasificación, el usuario puede guardar los resultados. El sistema crea una carpeta de salida con la imagen anotada y archivos de texto que registran tanto la localización de los individuos como el número de organismos detectados por especie. Según el manual de usuario, el directorio de resultados contiene la imagen clasificada, un archivo con coordenadas y clases de los individuos detectados, y un archivo con el número de individuos por especie [3].

La Fig. 9.5 muestra la salida generada por el sistema. Esta organización es importante porque separa la evidencia visual, las coordenadas de detección y los conteos por especie.

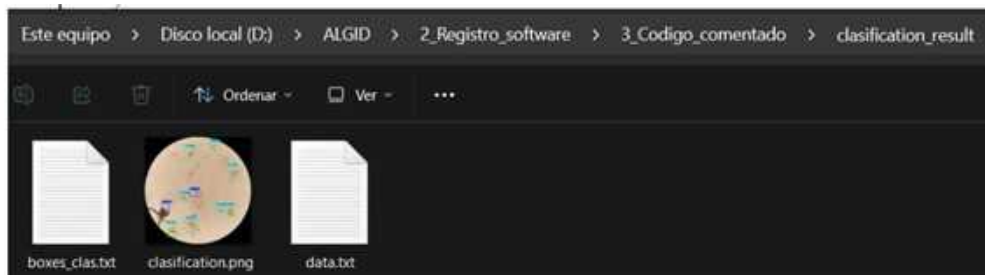


Figura 9.5: Estructura de resultados generada por el sistema. La salida conserva la imagen clasificada, el archivo con las coordenadas de las detecciones y el archivo con conteos por especie.

El conjunto de salida puede representarse como

$$\mathcal{R} = \{I_{\text{ann}}, B, \mathbf{N}\},$$

donde I_{ann} corresponde a la imagen anotada, B contiene las coordenadas y clases de los individuos detectados, y \mathbf{N} resume los conteos por especie. Esta estructura permite revisar el análisis, reconstruir la ubicación de cada detección y conservar un registro cuantitativo del resultado.

9.8. Arquitectura técnica del software

El manual técnico describe que el sistema fue desarrollado en un entorno Windows 11, programado en Spyder y ejecutado con Anaconda de 64 bits y Python 3.11.5. Entre las librerías principales se encuentran NumPy, Tkinter, OpenCV, Ultralytics, PIL, pathlib y collections [4]. Estas librerías cubren las funciones de interfaz gráfica, procesamiento de imágenes, inferencia con YOLO, manipulación de rutas y conteo de objetos.

El sistema requiere como mínimo un procesador AMD Ryzen 2200G de 3.5 GHz, 8 GB de memoria RAM, 4 GB de espacio disponible y tarjeta gráfica AMD Radeon RX Vega 8. Estos requisitos muestran que el desarrollo fue orientado a una operación local, viable en equipos de laboratorio sin depender necesariamente de infraestructura de alto desempeño.

Tabla 9.1: Componentes técnicos principales del sistema.

Componente	Función dentro del sistema
Tkinter	Construcción de la interfaz gráfica de usuario.
OpenCV	Procesamiento de imágenes y operaciones de visión por computador.
NumPy	Operaciones numéricas y manejo de arreglos.
Ultralytics	Ejecución del modelo YOLO para detección y clasificación.
PIL	Manipulación y visualización de imágenes.
pathlib	Gestión de rutas de archivos y directorios.
collections	Conteo y manejo de estructuras de datos.

El manual técnico documenta funciones principales como `askfile`, `cut_image`, `clasif_img`, `zoomIn`, `zoomOut`, `original` y `save_data`. Esta organización permite entender el software como una máquina de estados operativa:

$$s_0 \xrightarrow{\text{askfile}} s_1 \xrightarrow{\text{cut_image}} s_2 \xrightarrow{\text{clasif_img}} s_3 \xrightarrow{\text{save_data}} s_4,$$

donde

- s_0 : sin imagen cargada,
- s_1 : imagen seleccionada,
- s_2 : campo microscópico recortado,
- s_3 : detecciones y conteos generados,
- s_4 : resultados exportados.

Esta representación muestra que el sistema no opera como una ejecución monolítica, sino como una secuencia controlada por el usuario. Cada estado habilita una acción posterior y mantiene trazabilidad sobre la etapa del análisis.

9.9. Uso institucional, formación y transferencia

El certificado de uso indica que el software es empleado por los grupos GEPAR y GeoLimna en proyectos de innovación tecnológica, formación estudiantil, transferencia de conocimiento, ecotoxicología y monitoreo de calidad de agua. También señala que ha sido utilizado en prácticas de laboratorio, semilleros de investigación y proyectos colaborativos con entidades relacionadas con gestión de recursos hídricos [2].

El desarrollo se vincula con los proyectos SIMEVAM I y II, asociados al sistema de registro y seguimiento de variables ambientales en la ciénaga de Ayapel, y con SIMEVAM III, orientado a la medición de variables ambientales en sistemas leníticos [2]. Esta relación ubica el software dentro de una línea más amplia de instrumentación, monitoreo ambiental y automatización de análisis biológicos.

Desde la perspectiva de GEPAR, el desarrollo fortalece capacidades en visión artificial, procesamiento de imágenes microscópicas, inteligencia artificial aplicada y construcción de interfaces. Desde GeoLimna, aporta una herramienta para apoyar análisis de organismos bioindicadores, experimentos ecotoxicológicos y procesos de formación en limnología experimental. La integración de ambas capacidades expresa la lógica del *Seminario GeoLimna–GEPAR*: convertir problemas ambientales específicos en herramientas técnicas documentadas y transferibles.

9.10. Protección intelectual

El certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor identifica la obra con el título *ALGID Identificación de microalgas utilizando inteligencia artificial*. La fecha de registro es 05 de septiembre de 2025 y el Libro–Tomo–Partida corresponde a 13-105-375 [1]. El documento reconoce como autores a David Fernández Mc Cann, Yarin Tatiana Puerta Quintana y Juan Guillermo Quevedo, y como titular patrimonial y productor a la Universidad de Antioquia.

La descripción oficial del registro establece que el software detecta, clasifica y cuenta microalgas en imágenes microscópicas mediante inteligencia artificial y visión por computador. También indica que automatiza el proceso de identificación de especies, acelera el

procesamiento mediante redes neuronales convolucionales y genera reportes estructurados con información cuantitativa de las especies detectadas [1].

La formalización ante la Dirección Nacional de Derecho de Autor respalda el carácter tecnológico del desarrollo y lo diferencia de una rutina experimental aislada. El producto integra modelo, interfaz, documentación, uso institucional y registro de propiedad intelectual.

9.11. Discusión

La literatura reciente sobre automatización del conteo de microalgas muestra que el problema no se limita a detectar objetos. También exige definir qué se cuenta, cómo se conserva la evidencia, cómo se exportan los resultados y cómo se mantiene la robustez frente a variaciones de densidad, iluminación, enfoque, residuos y agrupamientos. El estado del arte revisado para el proyecto destaca que la cuantificación manual presenta limitaciones de tiempo, variabilidad entre operadores y sensibilidad a las condiciones de captura [6].

En ese contexto, el sistema aporta una solución integrada para un conjunto específico de especies verdes de agua dulce. Su contribución principal no está únicamente en usar un modelo YOLO, sino en cerrar un flujo operativo: cargar imagen, recortar campo útil, clasificar, contar y exportar. Esta integración es relevante para laboratorios, porque reduce la distancia entre el modelo computacional y el dato útil para experimentación.

El desarrollo también presenta retos. La precisión del conteo puede depender de la calidad de la imagen, el aumento utilizado, la densidad de organismos, la presencia de colonias superpuestas, la limpieza de la muestra y la representatividad del conjunto de entrenamiento. Por ello, una etapa posterior debería incluir validación cuantitativa con imágenes anotadas por expertos, métricas de precisión por especie, análisis de error de conteo, evaluación bajo diferentes densidades y comparación con conteo manual.

La posibilidad de reemplazar el archivo `best.pt` permite actualizar el modelo cuando se disponga de nuevas bases de datos o de arquitecturas mejoradas. Esto convierte el software en una plataforma extensible, no en un producto cerrado. Con nuevos datos, el sistema podría ampliarse a otras especies de interés ecológico o ecotoxicológico.

9.12. Conclusiones

Se presentó un sistema para la detección, clasificación y conteo automatizado de microalgas verdes en imágenes microscópicas. La herramienta integra interfaz gráfica, recorte del campo útil, inferencia mediante YOLO, conteo por especie y exportación estructurada de resultados. Su flujo permite convertir una imagen de microscopio en una imagen anotada y en archivos con coordenadas y conteos por clase.

El desarrollo responde a una necesidad concreta de los laboratorios ambientales: reducir la dependencia del conteo manual, mejorar la trazabilidad del análisis y generar datos reproducibles para estudios ecotoxicológicos y de calidad del agua. Su enfoque sobre *Pseudopediastrum boryanum*, *Desmodesmus communis* y *Raphidocelis subcapitata* lo sitúa en un campo de aplicación relevante para el monitoreo de ecosistemas de agua dulce.

El sistema constituye un producto tecnológico formalizado, con uso certificado por GEPAR y GeoLimna y registro ante la Dirección Nacional de Derecho de Autor. Su aporte institucional reside en articular investigación ambiental, visión artificial, formación estudiantil y transferencia tecnológica dentro de una herramienta operable y documentada.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia y a los grupos de investigación GEPAR y GeoLimna por el entorno académico e interdisciplinario que permitió el desarrollo del sistema. Se reconoce además el marco de los proyectos SIMEVAM I, II y III, orientados al monitoreo de variables ambientales y al fortalecimiento de herramientas de análisis para ecosistemas acuáticos.

Referencias

- [1] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: ALGID Identificación de microalgas utilizando inteligencia artificial*, Certificado de registro de soporte lógico, Libro–Tomo–Partida 13-105-375. Fecha de registro: 05 de septiembre de 2025. Autores: David Fernández Mc Cann, Yarin Tatiana Puerta Quintana y Juan Guillermo Quevedo., Bogotá, Colombia, sep. de 2025.

- [2] Grupo de Investigación GEPAR y Grupo de Investigación GeoLimna, *Certificación de uso del software ALGID por parte de los grupos de investigación GEPAR y GeoLimna*, Certificación institucional, Certificado expedido el 25 de septiembre de 2025., Medellín, Colombia, sep. de 2025.
- [3] Universidad de Antioquia, *ALGID: Identificación de microalgas utilizando inteligencia artificial. Manual de usuario*, Universidad de Antioquia, Medellín, Colombia, 2025.
- [4] Universidad de Antioquia, *ALGID: Identificación de microalgas utilizando inteligencia artificial. Manual técnico*, Universidad de Antioquia, Medellín, Colombia, 2025.
- [5] Universidad de Antioquia y Universidade Federal do Rio Grande, *Algoritmo para la identificación de algas ALGID*, Documento de formulación de proyecto, Proyecto de identificación de microalgas verdes mediante machine learning y deep learning., 2026.
- [6] Universidad de Antioquia, *Estado del arte de algoritmos utilizados para automatizar la detección, clasificación y conteo de microalgas verdes de agua dulce a partir de imágenes microscópicas*, Documento de trabajo del proyecto Algas-FURG, Medellín, Colombia, mar. de 2026.

Capítulo 10

Estimación forense de edad a partir de radiografías panorámicas mediante visión artificial y aprendizaje profundo

Autores

Isabella Bermón Rojas¹; Liliana Tamara Patiño²; David Stephen Fernández Mc Cann³; Ruby Amparo Vásquez Escobar⁴; Elizabeth Llano Sánchez⁵; Laura Cristina Vargas Álvarez⁶; Ricardo Ortiz Ruiz⁷; Ana Milena Doria Martínez⁸; Tania Camila Niño Sandoval⁹

¹Estudiante de Ingeniería Electrónica, Universidad de Antioquia. Facultad de Ingeniería. Grupo de investigación GEPAR.

²Instituto Nacional de Medicina Legal y Ciencias Forenses.

³Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

⁴Instituto Nacional de Medicina Legal y Ciencias Forenses.

⁵Profesora, Facultad de Odontología, Universidad de Antioquia. Grupo de investigación GIONORTO.

⁶Universidad de Antioquia. Facultad de Odontología. Grupo de investigación GIONORTO.

⁷Instituto Nacional de Medicina Legal y Ciencias Forenses.

⁸Instituto Nacional de Medicina Legal y Ciencias Forenses.

⁹Instituto Nacional de Medicina Legal y Ciencias Forenses.

⁹Desarrollo realizado en el marco del convenio de investigación entre la Universidad de Antioquia y el Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, código SIIU 2022-57265 / PPY-2022-005, con articulación entre los grupos GIONORTO y GEPAR.

Resumen

La estimación de edad en contextos forenses constituye una tarea técnica de alta responsabilidad, especialmente cuando se requiere apoyar decisiones judiciales, administrativas o periciales relacionadas con personas sin documentación confiable, verificación de mayoría o minoría de edad, identificación humana o caracterización biológica. En este capítulo se presenta un desarrollo de software orientado a apoyar la determinación de edad a partir de radiografías panorámicas dentales mediante visión artificial y aprendizaje profundo. El sistema integra modelos de segmentación dental basados en YOLOv8, mecanismos de etiquetamiento manual bajo nomenclatura FDI, corrección supervisada de detecciones, estimación automática de edad en años y meses mediante redes neuronales convolucionales, y generación de informes de validación que comparan resultados manuales y automáticos. El desarrollo fue realizado en el marco de un convenio entre la Universidad de Antioquia y el Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, con aprobación del Comité de Bioética de la Facultad de Medicina de la Universidad de Antioquia. El software se encuentra registrado ante la Dirección Nacional de Derecho de Autor bajo Libro–Tomó–Partida 13-107-337, con fecha de registro 20 de enero de 2026. Su uso ha sido certificado por los grupos GIONORTO y GEPAR como herramienta de investigación aplicada, formación estudiantil y validación tecnológica en la intersección entre odontología forense e ingeniería [1], [2], [3], [4].

Palabras clave: odontología forense; estimación de edad; radiografía panorámica; inteligencia artificial; YOLOv8; segmentación dental; aprendizaje profundo; nomenclatura FDI; validación; GIONORTO; GEPAR.

10.1. Introducción

La estimación de edad es una actividad fundamental en odontología forense, medicina legal e identificación humana. En escenarios judiciales y administrativos, la edad cronológica estimada puede incidir en decisiones relacionadas con responsabilidad penal, protección de menores, migración, identificación de personas no documentadas y caracterización biológica de individuos. En estos contextos, las estructuras dentales son de especial interés porque su desarrollo, mineralización, erupción, cierre apical y cambios morfológicos ofrecen información útil para inferir rangos de edad.

Las radiografías panorámicas constituyen una fuente ampliamente utilizada para observar simultáneamente maxilar, mandíbula y piezas dentales. Sin embargo, el análisis manual de estas imágenes exige conocimiento experto, tiempo de evaluación, criterios de segmen-

tación consistentes y documentación trazable del proceso. La interpretación individual puede variar entre observadores, especialmente cuando se trabaja con dentición mixta, piezas superpuestas, dientes intraóseos o estructuras de difícil delimitación.

El desarrollo presentado en este capítulo responde a esa necesidad mediante una plataforma que integra odontología forense, visión artificial e inteligencia artificial. El sistema registrado ante la Dirección Nacional de Derecho de Autor con el título *SICED* fue concebido como una herramienta de apoyo al experto, no como sustituto del criterio pericial. Su propósito es organizar el flujo de trabajo, automatizar tareas repetitivas, facilitar la segmentación dental, estimar la edad en años y meses, y generar evidencias comparables entre etiquetamientos manuales y detecciones automáticas.

El software fue desarrollado en el marco de un convenio de investigación entre la Universidad de Antioquia, bajo el código SIIU 2022-57265, y el Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, bajo el proyecto PPY-2022-005. La documentación del sistema indica además que el proyecto contó con aprobación del Comité de Bioética de la Facultad de Medicina de la Universidad de Antioquia, código F-017-00 [1], [2]. La ejecución del proyecto fue aprobada como proyecto inscrito el 22 de febrero de 2023, y su informe final fue aprobado por el Comité Técnico mediante Acta No. 030 del 30 de agosto de 2024 [5].

10.2. Problema abordado

La estimación forense de edad a partir de radiografías panorámicas puede entenderse como un problema de interpretación estructurada de imagen médica. Una radiografía contiene información anatómica relevante, pero esta debe convertirse en datos analizables: piezas dentales identificadas, regiones segmentadas, etiquetas normalizadas, rasgos derivados y predicción final de edad.

Sea P_i una radiografía panorámica correspondiente al individuo i . El objetivo general del sistema puede expresarse como la construcción de una función computacional

$$F(P_i) = (\mathcal{D}_i, \hat{a}_i, \mathcal{V}_i),$$

donde \mathcal{D}_i representa el conjunto de piezas dentales detectadas o etiquetadas, \hat{a}_i es la edad estimada por el sistema y \mathcal{V}_i corresponde al conjunto de salidas de validación y trazabilidad.

Cada panorámica P_i puede estar asociada a un conjunto de estructuras dentales:

$$\mathcal{D}_i = \{d_{i1}, d_{i2}, \dots, d_{in_i}\},$$

donde d_{ij} corresponde a la pieza dental j del individuo i , y n_i es el número de piezas observables o segmentables en la radiografía. Para que estas piezas sean comparables entre observadores y entre ejecuciones del algoritmo, el sistema emplea la nomenclatura internacional FDI, en la cual cada pieza dental se identifica mediante un código numérico estandarizado [6].

El problema técnico combina cuatro tareas: etiquetamiento manual, segmentación automática, estimación de edad y validación. En el etiquetamiento manual, el experto delimita piezas dentales y asigna etiquetas FDI. En la segmentación automática, un modelo de aprendizaje profundo detecta y segmenta piezas dentales. En la estimación de edad, una red neuronal convolucional produce una predicción en años y meses. En la validación, el sistema compara resultados manuales y automáticos para evaluar consistencia, precisión operativa y necesidad de corrección.

10.3. Descripción funcional del sistema

El sistema está organizado en cuatro módulos: etiquetamiento manual, detección automática dental con corrección manual, estimación automática de edad e informe de validación. El manual de usuario señala que los tres primeros módulos pueden ejecutarse de forma independiente; sin embargo, para generar el informe de validación se requiere haber realizado tanto el etiquetamiento manual como la detección automática [1]. Esta organización permite que el sistema funcione como herramienta de análisis, entrenamiento, corrección y comparación.

La Fig. 10.1 resume el flujo funcional general. La radiografía panorámica se carga en el sistema, luego puede pasar por etiquetamiento manual o por detección automática, se habilita la corrección supervisada de etiquetas, se ejecuta la estimación de edad y finalmente se genera un informe de validación cuando existen resultados manuales y automáticos comparables.



Figura 10.1: Flujo funcional del sistema para estimación forense de edad a partir de radiografías panorámicas. El proceso integra carga de imagen, etiquetamiento manual, segmentación automática mediante YOLOv8, corrección experta, estimación de edad con CNN e informe de validación entre resultados manuales y automáticos.

Desde el punto de vista funcional, la cadena de procesamiento puede representarse como

$$S(P_i) = G \circ E_\phi \circ C \circ D_\theta(P_i),$$

donde D_θ representa el módulo de detección y segmentación dental, C la corrección supervisada por el usuario, E_ϕ el modelo de estimación de edad y G la generación de salidas, reportes o evidencias. Esta formulación resume el flujo automático. Cuando interviene el etiquetamiento manual, se incorpora además un conjunto de polígonos definidos por el experto:

$$\mathcal{M}_i = \{(m_{ij}, y_{ij})\}_{j=1}^{n_i},$$

donde m_{ij} es el polígono o máscara manual asociado a la pieza dental j , y y_{ij} corresponde a la etiqueta FDI asignada por el usuario. La detección automática genera un conjunto análogo:

$$\mathcal{A}_i = \{(\hat{m}_{ik}, \hat{y}_{ik}, s_{ik})\}_{k=1}^{r_i},$$

donde \hat{m}_{ik} es la máscara automática, \hat{y}_{ik} la etiqueta predicha o asignada y s_{ik} la confianza de la detección. La comparación entre \mathcal{M}_i y \mathcal{A}_i permite construir el informe de validación.

10.4. Módulo de etiquetamiento manual

El módulo de etiquetamiento manual permite que el usuario delimite piezas dentales en una radiografía panorámica y asigne a cada una su respectiva nomenclatura FDI. Este módulo es fundamental porque produce una referencia experta que puede utilizarse tanto para documentación clínica como para validación de resultados automáticos.

La Fig. 10.2 muestra un ejemplo de etiquetamiento sobre una radiografía panorámica. En esta etapa, el usuario selecciona puntos sobre el contorno de la pieza dental, escribe la etiqueta correspondiente y guarda la información. El manual de usuario precisa que el etiquetamiento debe conservar el espacio del ligamento periodontal y aproximadamente un milímetro de hueso alveolar; también ofrece indicaciones para dientes multirradiculares, dientes intraóseos, superpuestos y supernumerarios [1].



Figura 10.2: Ejemplo de etiquetamiento manual de una pieza dental en radiografía panorámica. El usuario delimita el contorno de la estructura y asigna la nomenclatura FDI correspondiente, generando una referencia supervisada para análisis o validación.

Matemáticamente, una etiqueta manual puede representarse mediante un polígono cerrado:

$$m_{ij} = \{(x_{ij1}, y_{ij1}), (x_{ij2}, y_{ij2}), \dots, (x_{ijq}, y_{ijq})\},$$

donde cada par (x_{ijl}, y_{ijl}) corresponde a un punto seleccionado por el usuario sobre la imagen. La pieza etiquetada queda entonces definida como

$$d_{ij}^M = (m_{ij}, y_{ij}),$$

donde y_{ij} es la etiqueta FDI. El conjunto de todas las piezas etiquetadas manualmente constituye la anotación experta de la radiografía.

Este módulo aporta trazabilidad porque conserva los segmentos, coordenadas y etiquetas. Además, permite construir conjuntos de referencia para comparar el comportamiento del modelo automático. En un contexto forense, esta trazabilidad es importante porque el resultado no puede reducirse a una predicción numérica; debe estar acompañado de evidencia visual y de decisiones verificables.

10.5. Módulo de detección automática dental

El módulo de detección automática utiliza modelos de aprendizaje profundo para localizar y segmentar piezas dentales en radiografías panorámicas. El manual técnico indica que el sistema emplea YOLOv8 para la segmentación automática de piezas dentales y que esta salida puede ser corregida manualmente por el usuario [2]. Esta combinación permite que la inteligencia artificial acelere el proceso, sin eliminar la capacidad del experto para revisar y ajustar el resultado.

La Fig. 10.3 muestra la visualización de detecciones automáticas. Cada pieza dental aparece con su delimitación y etiqueta, lo que permite revisar el resultado global de la inferencia.



Figura 10.3: Visualización de detección automática de piezas dentales en una radiografía panorámica. El sistema segmenta las estructuras dentales, asigna etiquetas y permite su revisión posterior por parte del usuario.

La detección automática puede representarse como

$$D_{\theta}(P_i) = \{(\hat{b}_{ik}, \hat{m}_{ik}, \hat{y}_{ik}, s_{ik})\}_{k=1}^{r_i},$$

donde \hat{b}_{ik} es la caja delimitadora, \hat{m}_{ik} la máscara segmentada, \hat{y}_{ik} la etiqueta asociada y s_{ik} la confianza de la predicción. El parámetro θ agrupa los pesos del modelo de segmentación.

Para cada detección, la caja delimitadora puede expresarse como

$$\hat{b}_{ik} = (x_{ik}, y_{ik}, w_{ik}, h_{ik}),$$

donde x_{ik} y y_{ik} indican la posición de referencia, y w_{ik} y h_{ik} corresponden al ancho y alto de la región. La máscara \hat{m}_{ik} conserva la forma de la pieza dental de manera más precisa que la caja, lo cual resulta útil para segmentación, recorte y validación.

El módulo incorpora además edición manual de detecciones. El usuario puede borrar piezas repetidas, renombrar dientes o corregir etiquetas. Esta funcionalidad es coherente

con la naturaleza pericial del sistema: el algoritmo propone una interpretación, pero el experto conserva control sobre la decisión final.

10.6. Módulo de estimación automática de edad

El módulo de estimación automática de edad toma como entrada la información disponible de la radiografía y genera una predicción expresada en años y meses. El manual técnico indica que el sistema utiliza redes neuronales convolucionales para estimar la edad a partir de radiografías dentales [2]. Esta tarea se integra con el análisis de piezas dentales, pero puede ejecutarse como módulo independiente dentro de la interfaz.

La Fig. 10.4 muestra la vista de estimación de edad. El sistema presenta en pantalla el resultado estimado para la radiografía procesada.

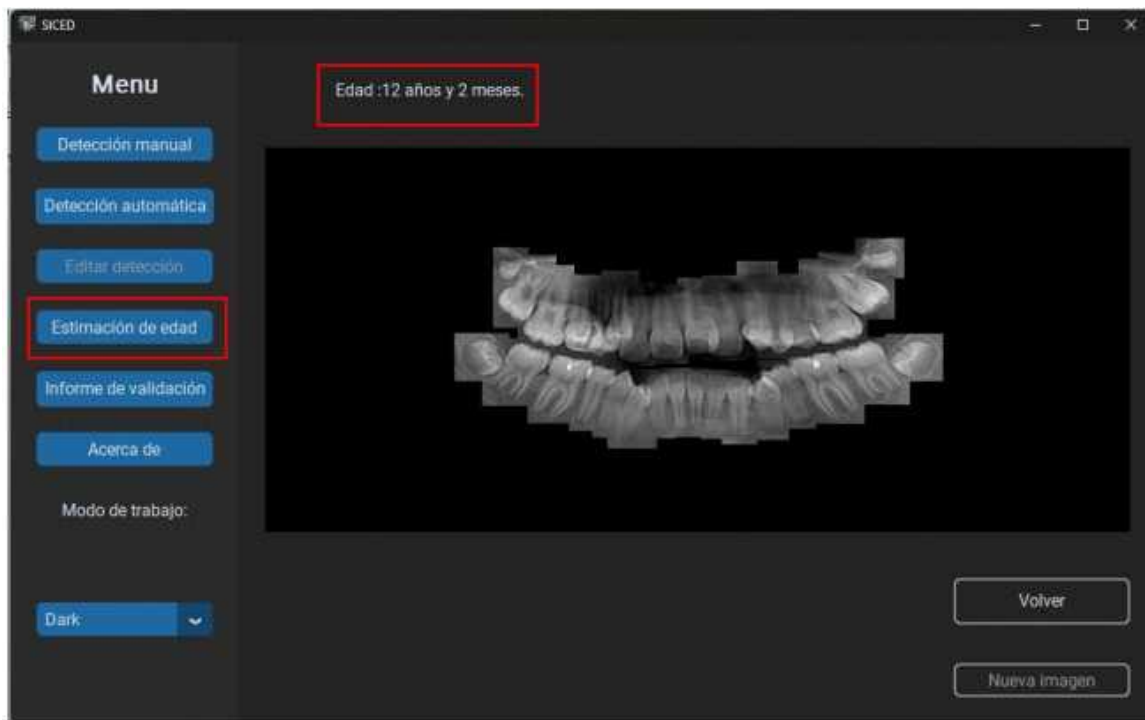


Figura 10.4: Módulo de estimación automática de edad. El sistema procesa la radiografía panorámica y entrega una predicción expresada en años y meses, como apoyo al análisis forense.

La estimación puede formalizarse como

$$\hat{a}_i = E_{\phi}(X_i),$$

donde E_ϕ es el modelo de estimación de edad parametrizado por ϕ , y X_i representa la información visual o derivada de la radiografía P_i . La salida \hat{a}_i puede descomponerse en años y meses:

$$\hat{a}_i = (\hat{a}_i^{(y)}, \hat{a}_i^{(m)}),$$

donde $\hat{a}_i^{(y)}$ corresponde a los años estimados y $\hat{a}_i^{(m)}$ a los meses adicionales.

En un escenario de uso pericial, esta predicción no reemplaza el concepto experto. Su función es ofrecer una estimación objetiva y reproducible que puede contrastarse con otros métodos, con el juicio profesional y con la documentación del caso. Por esta razón, el valor del sistema reside tanto en el resultado como en la trazabilidad del proceso que lo produce.

10.7. Módulo de validación

El módulo de validación compara resultados de etiquetamiento manual y detección automática. Para ello, el usuario carga los datos producidos por ambos módulos y el sistema genera salidas comparativas. Según el manual de usuario, este módulo requiere que se hayan ejecutado previamente el etiquetamiento manual y la detección automática sobre la radiografía correspondiente [1].

La Fig. 10.5 presenta un ejemplo de resultados de validación. Esta etapa permite revisar la correspondencia entre la segmentación realizada por el experto y la producida por el algoritmo, lo cual aporta evidencia para evaluar el comportamiento del sistema.

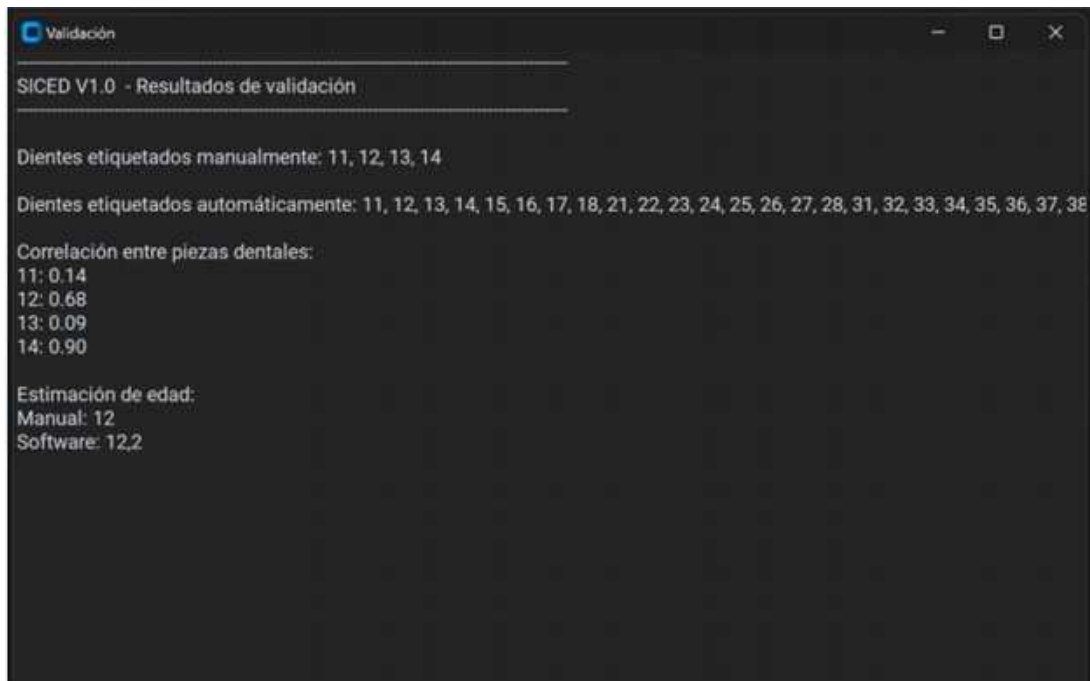


Figura 10.5: Resultados de validación entre etiquetamiento manual y detección automática. El módulo permite comparar la salida experta con la salida algorítmica y generar soportes visuales para el análisis de consistencia.

Una forma natural de comparar una máscara manual m_{ij} con una máscara automática \hat{m}_{ik} es la intersección sobre la unión:

$$\text{IoU}(m_{ij}, \hat{m}_{ik}) = \frac{|m_{ij} \cap \hat{m}_{ik}|}{|m_{ij} \cup \hat{m}_{ik}|}.$$

También puede utilizarse un coeficiente de solapamiento tipo Dice:

$$\text{Dice}(m_{ij}, \hat{m}_{ik}) = \frac{2|m_{ij} \cap \hat{m}_{ik}|}{|m_{ij}| + |\hat{m}_{ik}|}.$$

Estas expresiones permiten cuantificar la concordancia espacial entre el etiquetamiento manual y la segmentación automática. Aunque el manual describe el módulo desde la perspectiva operativa, estas medidas expresan matemáticamente el tipo de comparación que sustenta la validación entre regiones dentales.

El informe de validación cumple una función técnica e institucional. Técnica, porque permite evaluar el desempeño del sistema frente a referencias manuales. Institucional, porque documenta el proceso de comparación y permite conservar evidencia para auditoría, me-

jora de modelos y formación de usuarios.

10.8. Arquitectura técnica y requerimientos

El manual técnico describe que el sistema fue desarrollado para ejecutarse en Windows 11, con Python 3.7 o superior sin superar la versión 3.11, y con librerías especificadas en el archivo `requirements.txt`. En hardware, recomienda procesador Intel Core i5 o equivalente AMD, 16 GB de RAM, 518 MB de espacio libre y pantalla con resolución mínima de 1366×800 [2].

Tabla 10.1: Requerimientos técnicos generales del sistema.

Componente	Requerimiento reportado
Sistema operativo	Windows 11.
Lenguaje y entorno	Python 3.7 o superior, sin pasar de Python 3.11.
Procesador	Intel Core i5 o superior, o equivalente AMD.
Memoria RAM	16 GB o más para rendimiento óptimo con modelos de aprendizaje profundo.
Almacenamiento	518 MB de espacio libre para el programa.
Pantalla	Resolución mínima de 1366×800 .
Modelos principales	YOLOv8 para segmentación dental y redes neuronales convolucionales para estimación de edad.

Desde el punto de vista de software, la arquitectura se apoya en módulos entrenados para tareas específicas: segmentación dental, etiquetamiento, estimación de edad y validación. Esta separación modular permite ejecutar procesos independientes y facilita el mantenimiento, la revisión y la mejora progresiva de los modelos.

El sistema puede representarse como una máquina de estados:

$$s_0 \xrightarrow{\text{cargar radiografía}} s_1 \xrightarrow{\text{etiquetar o detectar}} s_2 \xrightarrow{\text{corregir}} s_3 \xrightarrow{\text{estimar edad}} s_4 \xrightarrow{\text{validar}} s_5,$$

donde

- s_0 : inicio sin radiografía cargada,
- s_1 : radiografía disponible,
- s_2 : piezas dentales segmentadas o etiquetadas,
- s_3 : detecciones revisadas por el usuario,
- s_4 : edad estimada,
- s_5 : informe de validación generado.

Esta representación refleja el carácter asistivo del sistema. El usuario conserva puntos de intervención, especialmente en el etiquetamiento, corrección y validación.

10.9. Uso institucional, formación y transferencia

El certificado de uso indica que el software registrado es utilizado por los grupos GIONORTO y GEPAR en actividades de investigación aplicada, formación de estudiantes, semilleros y procesos de validación tecnológica. También señala que el sistema fortalece la articulación interdisciplinaria entre odontología forense e ingeniería, y contribuye a generar capacidades institucionales para el desarrollo y uso de herramientas computacionales en contextos académicos y científicos [4].

La presentación institucional del proyecto sintetiza una línea de evolución: en 2024 se desarrolló el software, se entrenaron modelos y se realizó validación técnica; en 2025 se finalizó el proyecto, se consolidó la documentación y se fortaleció el uso institucional; en 2026 se obtuvo el registro de software y se proyectó la transferencia [7]. Esta trayectoria sitúa el desarrollo como un producto que avanza desde la investigación hacia la innovación aplicada.

El acta de finalización del proyecto indica que el trabajo corresponde al proyecto *Sistema Colombiano de determinación de la edad usando odontología forense e inteligencia artificial*, aprobado como proyecto inscrito el 22 de febrero de 2023, con informe final aprobado el 30 de agosto de 2024. También documenta la participación de coinvestigadores y estudiantes de pregrado vinculadas al proceso [5].

Desde la perspectiva de GIONORTO, el sistema contribuye a la investigación en odontología del niño, ortodoncia y estimación de edad. Desde GEPAR, fortalece capacidades en visión artificial, aprendizaje profundo, procesamiento de imágenes médicas, desarrollo de interfaces y transferencia tecnológica. Desde el Instituto Nacional de Medicina Legal y Ciencias Forenses, se vincula con necesidades operativas del ámbito pericial y de identificación humana.

10.10. Protección intelectual

El certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor identifica la obra con el título *SICED*. La fecha de registro es 20 de enero de 2026 y el Libro–Tomo–Partida corresponde a 13-107-337 [3]. El documento reconoce como autores a Liliana Tamara, David Fernández Mc Cann, Isabella Bermón Rojas, Ruby Amparo Vásquez Escobar, Elizabeth Llano Sánchez, Laura Cristina Vargas Álvarez, Ricardo Ortiz Ruiz, Ana Milena Doria Martínez y Tania Camila Niño Sandoval. También identifica como titulares patrimoniales y productores a la Universidad de Antioquia y al Instituto Nacional de Medicina Legal y Ciencias Forenses.

La descripción oficial del registro establece que el software procesa imágenes, detecta dientes en radiografías y estima la edad con inteligencia artificial. También indica que permite cargar imágenes, segmentar piezas dentales automáticamente, editar posibles errores, etiquetar manualmente, estimar edad y validar el etiquetamiento manual para casos de estudio [3].

El registro formaliza el desarrollo como producto tecnológico de soporte lógico. Su importancia institucional no se limita al modelo de inteligencia artificial, sino que comprende la integración de módulos, interfaz, documentación, validación, uso certificado y protección de propiedad intelectual.

10.11. Discusión

El sistema descrito se ubica en una línea de desarrollo donde la inteligencia artificial no actúa como sustituto del perito, sino como sistema de apoyo a la decisión. Esta condición es esencial en odontología forense, porque los resultados deben ser interpretables, trazables y revisables. La plataforma permite que el experto intervenga en el etiquetamiento, revise la detección automática, corrija errores y compare salidas manuales y algorítmicas.

La principal contribución del sistema reside en integrar tareas que usualmente se encuentran separadas: segmentación dental, etiquetamiento FDI, estimación de edad e informe de validación. Esta integración favorece la estandarización del flujo de trabajo, reduce tiempos operativos y conserva evidencia visual y documental del proceso.

El desarrollo también plantea retos. La estimación de edad depende de la representatividad de las radiografías utilizadas para entrenamiento, de la calidad de imagen, de la variabilidad poblacional, de la condición de dentición y de la presencia de piezas ausentes, superpuestas, restauradas o alteradas. Por esta razón, el sistema debe entenderse como

una plataforma en evolución, susceptible de mejorar con nuevas bases de datos validadas, nuevos protocolos de evaluación y comparación con métodos forenses tradicionales.

Una etapa posterior debería fortalecer la validación cuantitativa mediante métricas de segmentación, error absoluto medio en edad, análisis por rangos etarios, comparación entre observadores, desempeño en dentición mixta y permanente, y evaluación de robustez frente a imágenes con diferente calidad. Esta información permitiría establecer con mayor precisión el alcance pericial del sistema y sus condiciones de uso.

10.12. Conclusiones

Se presentó un sistema para la estimación forense de edad a partir de radiografías panorámicas dentales, basado en visión artificial, segmentación con YOLOv8 y redes neuronales convolucionales para predicción de edad. La herramienta integra etiquetamiento manual, detección automática, corrección supervisada, estimación en años y meses, y generación de informes de validación.

El desarrollo responde a una necesidad concreta del ámbito forense: disponer de herramientas que apoyen la estimación de edad con mayor trazabilidad, estandarización y reproducibilidad. Su diseño conserva el papel del experto, permitiendo que la automatización complemente el juicio profesional y no lo reemplace.

El sistema constituye un producto tecnológico formalizado, con registro ante la Dirección Nacional de Derecho de Autor, uso certificado por GIONORTO y GEPAR, y desarrollo en convenio entre la Universidad de Antioquia y el Instituto Nacional de Medicina Legal y Ciencias Forenses. Su valor institucional reside en articular investigación aplicada, formación estudiantil, odontología forense, visión artificial e innovación tecnológica en una plataforma funcional.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia, al Instituto Nacional de Medicina Legal y Ciencias Forenses, al grupo GIONORTO y al grupo GEPAR por el entorno académico, técnico e institucional que permitió el desarrollo del sistema. Se reconoce además el marco del proyecto SIIU 2022-57265 / PPY-2022-005 y la participación de estudiantes, docentes e investigadores vinculados a la construcción, validación y documentación del software.

Referencias

- [1] Universidad de Antioquia e Instituto Nacional de Medicina Legal y Ciencias Forenses, *SICED: Sistema Colombiano de Determinación de la Edad. Manual de usuario*, Universidad de Antioquia, Medellín, Colombia, 2026.
- [2] Universidad de Antioquia e Instituto Nacional de Medicina Legal y Ciencias Forenses, *SICED: Sistema Colombiano de Determinación de la Edad. Manual técnico*, Universidad de Antioquia, Medellín, Colombia, 2026.
- [3] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: SICED*, Certificado de registro de soporte lógico, Libro–Tomo–Partida 13-107-337. Fecha de registro: 20 de enero de 2026., Bogotá, Colombia, ene. de 2026.
- [4] Grupo GIONORTO y Grupo GEPAR, *Certificación de uso del software SICED por parte de los grupos de investigación GIONORTO y GEPAR*, Certificación institucional, Certificado expedido el 12 de febrero de 2026., Medellín, Colombia, feb. de 2026.
- [5] Universidad de Antioquia, *Acta de finalización del proyecto Sistema Colombiano de determinación de la edad usando odontología forense e inteligencia artificial*, Acta institucional de finalización de proyecto, Proyecto aprobado como Proyecto Inscrito el 22 de febrero de 2023; informe final aprobado mediante Acta No. 030 del 30 de agosto de 2024., Medellín, Colombia, ago. de 2025.
- [6] Fédération Dentaire Internationale, *Two-digit system of designating teeth*, International dental nomenclature system, 1971.
- [7] Universidad de Antioquia, *SICED – Sistema Colombiano de Determinación de la Edad: de la investigación a la innovación*, Presentación institucional, Medellín, Colombia, 2026.

Capítulo 11

Reconocimiento móvil de placas vehiculares en video mediante visión artificial y OCR profundo

Autores

Johan David Rojas Martínez¹; David Stephen Fernández Mc Cann²; Claudia Cristina Botero Suárez³

Resumen

El reconocimiento automático de placas vehiculares en video constituye una tarea relevante para la supervisión vial, el patrullaje inteligente, el control vehicular y la generación de registros estructurados de movilidad. A diferencia de los sistemas basados en cámaras fijas, el análisis de video capturado desde cámaras en movimiento introduce retos adicionales asociados a vibración, variación de ángulo, cambios rápidos de iluminación, desenfoque, oclusiones y variabilidad en la escala aparente de las placas. En este capítulo se presenta un desarrollo de software orientado a detectar y reconocer placas vehiculares

¹Estudiante de Ingeniería Electrónica, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Coautora vinculada a InterTelco S.A.S.

³Desarrollo realizado en el marco del proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, articulado entre la Universidad de Antioquia, el grupo GEPAR e InterTelco S.A.S.

en videos obtenidos desde plataformas móviles. El sistema integra un detector YOLOv8 para localizar placas y un modelo TrOCR para reconocimiento óptico de caracteres. La herramienta incorpora una interfaz gráfica en Tkinter que permite cargar videos, iniciar el procesamiento, pausar y reanudar el análisis, visualizar resultados OCR y exportar reportes estructurados. El desarrollo fue realizado dentro del proyecto VISIONARY, en cooperación entre la Universidad de Antioquia e InterTelco S.A.S., y se encuentra registrado ante la Dirección Nacional de Derecho de Autor bajo Libro–Tomo–Partida 13-106-183, con fecha de registro 23 de octubre de 2025. Su uso ha sido certificado por el grupo GEPAR e InterTelco S.A.S. en procesos de validación experimental, integración industrial, evaluación de cámaras y formación de estudiantes en visión artificial y aprendizaje profundo [1], [2], [3], [4].

Palabras clave: reconocimiento móvil de placas; visión artificial; YOLOv8; TrOCR; OCR; procesamiento de video; patrullaje inteligente; seguridad vial; inteligencia vial; GEPAR; VISIONARY.

11.1. Introducción

El análisis automatizado de placas vehiculares permite transformar registros de video en información estructurada para supervisión vial, control vehicular, patrullaje inteligente, monitoreo de corredores, gestión de flotas y apoyo a sistemas urbanos de seguridad. Cuando la captura se realiza desde una cámara fija, el problema se concentra en detectar placas dentro de una escena relativamente estable. En cambio, cuando la cámara está en movimiento, el sistema debe enfrentar variaciones continuas de perspectiva, vibración, cambios de distancia, desenfoque por movimiento y transiciones rápidas de iluminación.

Estas condiciones convierten el reconocimiento móvil de placas en una tarea más exigente que la lectura de placas en escenarios controlados. El sistema debe localizar placas en secuencias de video, recortar regiones de interés, reconocer caracteres, consolidar resultados y exportar evidencias que puedan revisarse posteriormente. El procesamiento manual de este tipo de videos demanda tiempo, concentración del operador y no escala bien cuando se incrementa el número de recorridos, cámaras o fuentes de grabación.

El desarrollo presentado en este capítulo responde a ese problema mediante una arquitectura de visión artificial y reconocimiento óptico de caracteres. El sistema registrado ante la Dirección Nacional de Derecho de Autor con el título *MOVILPR (Reconocimiento Móvil de Placas)* fue diseñado para analizar videos capturados desde cámaras móviles, como dispositivos montados en vehículos, drones o plataformas de vigilancia en movimiento. Su propósito es automatizar el reconocimiento de placas en entornos dinámicos, con una

interfaz operativa accesible y una salida organizada para análisis posterior.

La Fig. 11.1 muestra la interfaz inicial del sistema. Desde esta vista, el usuario puede cargar el video, iniciar el procesamiento, pausar la ejecución y guardar los resultados. La estructura de operación conserva una lógica sencilla, aunque internamente articula modelos de aprendizaje profundo y procesamiento de video.



Figura 11.1: Interfaz principal para la carga de video en el sistema. El usuario selecciona una secuencia capturada desde una cámara en movimiento y habilita el flujo de procesamiento automático.

El desarrollo se inscribe en el proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, ejecutado conjuntamente por la Universidad de Antioquia e InterTelco S.A.S. En la documentación institucional del proyecto se establecen como objetivos investigar algoritmos de visión artificial, desarrollar modelos aplicados y registrar productos de software asociados a la supervisión de seguridad vial e infraestructura [5], [6].

11.2. Problema abordado

Sea un video digital capturado desde una cámara móvil y representado como una secuencia de fotogramas:

$$V = \{I_1, I_2, \dots, I_T\},$$

donde I_t corresponde al fotograma en el instante discreto t , y T es el número total de fotogramas disponibles. El objetivo del sistema consiste en detectar las regiones que contienen placas vehiculares y estimar, para cada región válida, una cadena alfanumérica asociada al texto de la placa.

La detección de placas en cada fotograma puede expresarse como

$$D_{\theta}(I_t) = \{(b_{t,i}, s_{t,i})\}_{i=1}^{N_t},$$

donde $b_{t,i}$ representa la caja delimitadora de la placa i detectada en el fotograma t , $s_{t,i}$ es la confianza de la detección y N_t es el número de placas detectadas en dicho fotograma. El parámetro θ agrupa los pesos del modelo de detección, implementado mediante YOLOv8.

Cada caja delimitadora puede escribirse como

$$b_{t,i} = (x_{t,i}, y_{t,i}, w_{t,i}, h_{t,i}),$$

donde $x_{t,i}$ y $y_{t,i}$ indican la posición de referencia de la caja, mientras que $w_{t,i}$ y $h_{t,i}$ corresponden a su ancho y alto. A partir de esta región se obtiene un recorte de placa:

$$R_{t,i} = C(I_t, b_{t,i}),$$

donde $C(\cdot)$ representa la operación de recorte de la región de interés. Sobre esta imagen recortada se aplica un modelo OCR:

$$O_{\phi}(R_{t,i}) = (\hat{y}_{t,i}, q_{t,i}),$$

donde $\hat{y}_{t,i}$ corresponde al texto reconocido, $q_{t,i}$ representa una medida de confianza o consistencia del reconocimiento y ϕ agrupa los parámetros del modelo TrOCR.

En video móvil, una misma placa puede aparecer en varios fotogramas con lecturas diferentes. Por esta razón, el problema no se limita a detectar una placa aislada, sino a consolidar resultados repetidos o parcialmente ambiguos. El sistema debe agrupar lecturas similares y seleccionar una representación final confiable.

11.3. Descripción funcional del sistema

El sistema opera como una cadena funcional de seis etapas: carga del video, lectura de fotogramas, detección de placas, recorte de regiones, reconocimiento OCR y exportación estructurada de resultados. Esta arquitectura permite transformar una secuencia de video capturada desde una plataforma móvil en un conjunto de imágenes de placas y textos reconocidos.

La Fig. 11.2 resume el flujo general del sistema. La figura debe corresponder a un diagrama propio, elaborado con el mismo estilo gráfico de los capítulos anteriores, y no a una captura del manual.



Figura 11.2: Flujo funcional del sistema para reconocimiento móvil de placas vehiculares. El procesamiento inicia con un video capturado desde una cámara en movimiento, continúa con lectura de fotogramas, detección mediante YOLOv8, recorte de placas, reconocimiento OCR con TrOCR, filtrado inteligente y exportación estructurada de evidencias.

La operación completa puede representarse como una composición de funciones:

$$S(V) = E \circ F_{\text{ocr}} \circ O_{\phi} \circ C \circ D_{\theta} \circ L(V),$$

donde L representa la lectura del video en fotogramas, D_{θ} la detección de placas, C el recorte de las regiones detectadas, O_{ϕ} el reconocimiento OCR, F_{ocr} el filtrado y consolidación de lecturas, y E la exportación de resultados.

El manual de usuario describe cuatro módulos principales: carga de video, procesamiento, pausa o reanudación, y exportación. Además, incluye una lógica de filtrado OCR inteligente que compara resultados en caso de ambigüedad, agrupa placas similares y selecciona automáticamente la versión más confiable [1]. Esta lógica es relevante porque en un vi-

deo móvil una placa puede ser leída varias veces con pequeñas variaciones debidas al movimiento, desenfoque o cambios de perspectiva.

11.4. Carga y validación del video

El primer módulo permite seleccionar un archivo de video desde el equipo del usuario. La documentación de usuario indica que el sistema admite formatos como `.mp4`, `.avi` y `.mov`, y que no permite imágenes ni documentos de otro tipo [1]. La carga constituye el punto de entrada del proceso porque determina la fuente visual sobre la cual se ejecutarán las etapas de detección y reconocimiento.

Desde el punto de vista funcional, la carga de video puede representarse como

$$L(V) = \{I_t\}_{t=1}^T,$$

donde L corresponde al proceso de lectura del archivo y extracción secuencial de fotogramas. En esta etapa también se puede obtener información técnica del video, como resolución, duración y tasa de cuadros por segundo, útil para ajustar parámetros internos de procesamiento.

La validación inicial del archivo evita errores operativos. Si el archivo no corresponde a un formato compatible o si está dañado, el sistema debe notificarlo al usuario sin interrumpir de manera abrupta la ejecución. Esta condición es importante en entornos de campo, donde los videos pueden provenir de diferentes dispositivos y configuraciones de grabación.

11.5. Procesamiento de fotogramas y reconocimiento OCR

Una vez cargado el video, el usuario inicia el análisis. El sistema procesa la secuencia cuadro a cuadro, detectando placas, recortando las regiones correspondientes y aplicando reconocimiento óptico de caracteres. La documentación técnica indica que el sistema emplea YOLOv8 para la detección y TrOCR para el reconocimiento de texto en placas vehiculares [2].

La Fig. 11.3 muestra el panel de resultados durante el procesamiento. En esta etapa, la interfaz presenta una miniatura de la placa detectada junto con el texto reconocido.



Figura 11.3: Panel de procesamiento con detección de placa y resultado OCR. El sistema muestra la región recortada de la placa y el texto reconocido durante el análisis del video.

El procesamiento de cada fotograma puede expresarse como

$$\mathcal{P}_t = \{(R_{t,i}, \hat{y}_{t,i}, q_{t,i})\}_{i=1}^{N_t},$$

donde $R_{t,i}$ es el recorte de placa, $\hat{y}_{t,i}$ la cadena reconocida y $q_{t,i}$ la confianza o calidad asociada al reconocimiento. El conjunto de resultados del video completo queda dado por

$$\mathcal{P} = \bigcup_{t=1}^T \mathcal{P}_t.$$

En condiciones móviles, el mismo vehículo puede permanecer visible durante varios fotogramas. Por tanto, el conjunto \mathcal{P} puede contener lecturas repetidas o variantes de una misma placa. Para consolidarlas, se define una agrupación de cadenas reconocidas:

$$G_k = \{\hat{y}_j : d(\hat{y}_j, \hat{y}_k) \leq \tau\},$$

donde $d(\cdot, \cdot)$ es una distancia entre cadenas, como la distancia de edición, y τ es un umbral de similitud. La lectura final de un grupo puede seleccionarse mediante

$$\hat{y}_k^* = \arg \max_{\hat{y}_j \in G_k} Q(\hat{y}_j),$$

donde $Q(\cdot)$ puede considerar confianza OCR, frecuencia de aparición, consistencia entre fotogramas y calidad visual del recorte. Esta formulación resume la lógica de filtrado inteligente descrita en la documentación del sistema.

11.6. Pausa, reanudación y control operativo

El sistema permite pausar y reanudar el análisis durante la ejecución. Esta función es relevante en videos largos o en pruebas de campo, donde el usuario puede requerir revisar una detección, detener temporalmente el procesamiento o controlar el flujo de inferencia sin perder el estado de la ejecución.

La Fig. 11.4 muestra la interfaz asociada al control de pausa y reanudación. Esta funcionalidad no modifica los modelos de detección u OCR, pero mejora la operabilidad del sistema y evita reiniciar el análisis desde cero ante interrupciones o revisiones intermedias.



Figura 11.4: Control de pausa y reanudación del procesamiento. Esta función permite detener temporalmente el análisis del video y continuar posteriormente sin reiniciar la secuencia completa.

Desde una perspectiva de estados, el procesamiento puede describirse como

$$s_0 \xrightarrow{\text{cargar video}} s_1 \xrightarrow{\text{procesar}} s_2 \xrightarrow{\text{pausar}} s_3 \xrightarrow{\text{reanudar}} s_2 \xrightarrow{\text{finalizar}} s_4,$$

donde

- s_0 : sin video cargado,
- s_1 : video cargado y validado,
- s_2 : procesamiento activo,
- s_3 : procesamiento pausado,
- s_4 : procesamiento finalizado.

Esta representación muestra que la interfaz no es solo un elemento visual, sino un componente de control del flujo computacional.

11.7. Exportación y trazabilidad de resultados

Al finalizar el procesamiento, el sistema permite guardar los resultados. La documentación técnica establece que el sistema genera un directorio con el nombre del video procesado y organiza allí las imágenes recortadas de las placas, los resultados OCR y los archivos de resumen. También contempla un subdirectorio `placas_finales/`, donde se almacena una imagen representativa por cada placa única, seleccionada según criterios de confianza o reglas definidas [2].

La Fig. 11.5 muestra la salida exportada por el sistema.

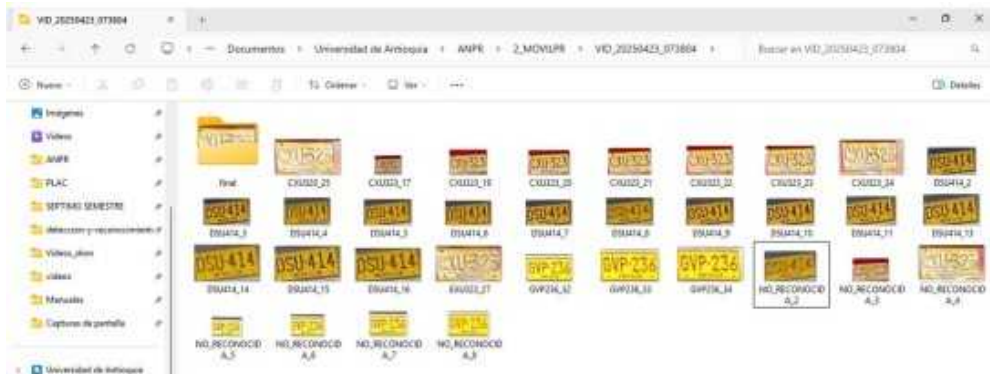


Figura 11.5: Exportación estructurada de resultados. El sistema organiza las imágenes recortadas, las lecturas OCR y las placas finales consolidadas en carpetas asociadas al video procesado.

El resultado final del procesamiento puede representarse como

$$\mathcal{R} = \{I_{\text{ann}}, \mathcal{C}, \mathcal{Y}, \mathcal{Y}^*\},$$

donde I_{ann} representa evidencias visuales o fotogramas anotados, \mathcal{C} el conjunto de recortes de placas, \mathcal{Y} las lecturas OCR obtenidas durante el video y \mathcal{Y}^* las lecturas consolidadas finales. Esta estructura permite conservar evidencia visual, resultados intermedios y salidas depuradas para análisis posterior.

La trazabilidad es un componente central del desarrollo. Un sistema de reconocimiento móvil no debe limitarse a entregar una cadena de texto; debe conservar el recorte que la originó, el video de procedencia, el momento de detección y la estructura de archivos que permita revisar o auditar el resultado.

11.8. Arquitectura técnica y requerimientos

El manual técnico describe el sistema como una aplicación de escritorio desarrollada en Python, con interfaz gráfica en Tkinter, motor de detección basado en YOLOv8 mediante la biblioteca Ultralytics y reconocimiento OCR mediante TrOCR. La arquitectura se organiza en módulos especializados para detección, reconocimiento, visualización, exportación y control de flujo [2].

Tabla 11.1: Requerimientos técnicos generales del sistema.

Componente	Mínimo	Recomendado
Procesador	Intel Core i5, cuatro núcleos	Intel Core i7/i9 o AMD Ryzen 7 o superior
Memoria RAM	8 GB	16 GB o más
GPU	NVIDIA GTX 1050 o equivalente con CUDA	NVIDIA RTX 3060 o superior
Almacenamiento	8 GB disponibles	10 GB o más para resultados
Pantalla	1366 × 768	1920 × 1080 o superior
Sistema operativo	Windows 10/11 o Ubuntu 20.04	Sistema con entorno Python y soporte CUDA configurado
Python	Versión 3.8 o superior	Entorno virtual con dependencias controladas

El sistema admite videos en formatos estándar como `.mp4` y `.avi`; el manual de usuario también refiere compatibilidad con `.mov`. Para mejorar los resultados, se recomienda contar con buena iluminación, resolución suficiente y condiciones de captura que permitan observar las placas con nitidez [1], [2].

La estructura modular facilita futuras adaptaciones. El detector YOLOv8 puede ser re-entrenado con nuevas placas, ángulos o condiciones viales; el OCR puede ajustarse con

datos adicionales; y la interfaz puede ampliarse para conectar el sistema con bases de datos, tableros de monitoreo o plataformas de supervisión vial.

11.9. Uso institucional, formación y transferencia

El certificado de uso indica que el software registrado se encuentra activo dentro de las actividades de investigación, desarrollo tecnológico e innovación del grupo GEPAR, en el marco del convenio con InterTelco S.A.S. El mismo documento señala que el sistema es utilizado para validación experimental, integración industrial, procesamiento de videos capturados en entornos viales, generación de registros cuantitativos de tránsito y evaluación de desempeño de cámaras y algoritmos en condiciones reales [4].

La articulación con InterTelco S.A.S. se enmarca en el Convenio Específico de Investigación No. 21330002-0002-2024, cuyo objeto es desarrollar el proyecto *Uso de técnicas de visión artificial para sistemas de seguridad y supervisión de vías*. El convenio establece como alcance investigar algoritmos para identificación de placas de vehículos a partir de imágenes o videos tomados por drones en vuelo, desarrollar modelos basados en visión artificial, estrechar la cooperación científica entre UdeA e InterTelco y registrar software derivado del proyecto [6].

El acta de propiedad intelectual del proyecto VISIONARY documenta su carácter de investigación aplicada, la participación de GEPAR, la duración de 24 meses desde el 29 de julio de 2024 hasta el 29 de julio de 2026, y los compromisos de registro de software relacionados con identificación de placas vehiculares e irregularidades en pavimento [5].

En el ámbito académico, el sistema ha sido empleado como herramienta de formación para estudiantes en visión artificial y aprendizaje profundo. Su arquitectura permite comprender el ciclo completo de una solución de inteligencia artificial aplicada: adquisición de video, detección, OCR, depuración, interfaz y exportación de resultados. Esta condición fortalece las capacidades del grupo GEPAR en transferencia tecnológica hacia el sector productivo.

11.10. Protección intelectual

El certificado de registro de soporte lógico expedido por la Dirección Nacional de Derecho de Autor identifica la obra con el título *MOVILPR (Reconocimiento Móvil de Placas)*. La fecha de registro es 23 de octubre de 2025 y el Libro–Tomo–Partida corresponde a 13-106-183 [3]. El certificado reconoce como autores a David Fernández Mc Cann, Claudia

Cristina Botero Suárez y Johan David Rojas Martínez, y como titulares patrimoniales y productores a la Universidad de Antioquia e InterTelco S.A.S.

La descripción oficial establece que el software detecta y reconoce automáticamente placas vehiculares en videos capturados desde cámaras en movimiento. También indica que está implementado en Python con modelos YOLOv8 y TrOCR, incorpora una interfaz gráfica en Tkinter y permite procesar, visualizar y exportar resultados de forma estructurada, con aplicabilidad en patrullaje inteligente, control vehicular y supervisión de seguridad vial [3].

Este registro formaliza el desarrollo como producto tecnológico conjunto universidad–empresa. Su valor institucional no se limita al modelo computacional, sino que integra desarrollo algorítmico, interfaz, documentación, validación, uso certificado e inscripción de propiedad intelectual.

11.11. Discusión

El sistema descrito extiende el reconocimiento automático de placas hacia un escenario más exigente: cámaras en movimiento. Esta condición implica que el sistema debe tolerar variaciones de ángulo, escala, iluminación, vibración y desenfoque. En comparación con sistemas de cámara fija, el reconocimiento móvil exige mayor robustez temporal y una estrategia de consolidación de lecturas, porque una placa puede aparecer varias veces con resultados OCR ligeramente distintos.

La principal contribución del desarrollo está en integrar detección, OCR, interfaz y exportación dentro de una plataforma operable. El sistema no se limita a aplicar un detector sobre imágenes aisladas; procesa video, muestra resultados durante la ejecución, permite controlar el análisis y conserva evidencias organizadas. Esta integración es clave para aplicaciones de patrullaje, supervisión vial y validación industrial.

El desarrollo también presenta retos. La precisión puede degradarse ante videos con baja resolución, placas lejanas, movimiento brusco, exposición variable, reflejos, oclusiones o caracteres deteriorados. Por ello, una etapa posterior debería incorporar validación cuantitativa con videos representativos de diferentes condiciones viales, métricas de detección, tasa de caracteres correctos, precisión por placa completa, análisis por velocidad de captura y comparación con revisión humana.

A pesar de estas limitaciones, el sistema constituye una base tecnológica reutilizable. Su arquitectura puede ampliarse hacia detección de otros elementos viales, integración con GPS, generación de trayectorias, conexión con bases de datos autorizadas o despliegue en

plataformas móviles de patrullaje y supervisión.

11.12. Conclusiones

Se presentó un sistema para reconocimiento móvil de placas vehiculares en video, basado en visión artificial y OCR profundo. La herramienta integra YOLOv8 para detección de placas, TrOCR para reconocimiento óptico de caracteres, procesamiento de video, interfaz gráfica en Tkinter y exportación estructurada de resultados.

El desarrollo responde a una necesidad práctica de la supervisión vial: automatizar la lectura de placas en videos capturados desde cámaras en movimiento, reduciendo la carga de revisión manual y generando evidencias reutilizables. Su diseño permite procesar secuencias de video, visualizar resultados durante la ejecución, consolidar lecturas y conservar imágenes de placas detectadas.

El sistema constituye un producto tecnológico formalizado, con registro ante la Dirección Nacional de Derecho de Autor, uso certificado por GEPAR e InterTelco S.A.S., y desarrollo dentro del proyecto VISIONARY. Su aporte institucional reside en articular investigación aplicada, formación estudiantil, visión artificial, inteligencia vial y transferencia tecnológica hacia el sector productivo.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia, al grupo de investigación GEPAR y a InterTelco S.A.S. por el entorno técnico, académico e institucional que hizo posible el desarrollo del sistema. Se reconoce además el marco del proyecto VISIONARY: *Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, orientado al desarrollo de sistemas de visión artificial para seguridad vial y supervisión de infraestructura.

Referencias

- [1] Universidad de Antioquia, *MOVILPR: Reconocimiento Móvil de Placas. Manual de usuario*, Documento elaborado siguiendo el estándar IEEE 1063-2001., Universidad de Antioquia, Medellín, Colombia, 2025.

- [2] Universidad de Antioquia, *MOVILPR: Reconocimiento Móvil de Placas. Manual técnico*, Documento técnico del sistema desarrollado en el marco del proyecto VISIONARY., Universidad de Antioquia, Medellín, Colombia, 2025.
- [3] Dirección Nacional de Derecho de Autor, *Certificado de registro de soporte lógico – software: MOVILPR (Reconocimiento Móvil de Placas)*, Certificado de registro de soporte lógico, Libro–Tomo–Partida 13-106-183. Fecha de registro: 23 de octubre de 2025. Autores: David Fernández Mc Cann, Claudia Cristina Botero Suárez y Johan David Rojas Martínez., Bogotá, Colombia, oct. de 2025.
- [4] Grupo de Investigación GEPAR e InterTelco S.A.S., *Certificación de uso del software MOVILPR por parte del grupo de investigación GEPAR y la empresa InterTelco S.A.S.* Certificación institucional, Certificado expedido el 4 de noviembre de 2025., Medellín, Colombia, nov. de 2025.
- [5] Universidad de Antioquia, *Acta de registro y de propiedad intelectual del proyecto VISIONARY: Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, Acta institucional CODI, Vicerrectoría de Investigación, Acta N.º 2024-74783, fecha 13 de noviembre de 2024., Medellín, Colombia, nov. de 2024.
- [6] Universidad de Antioquia e InterTelco S.A.S., *Convenio específico de investigación entre la Universidad de Antioquia e InterTelco S.A.S. para el uso de técnicas de visión artificial en sistemas de seguridad y supervisión de vías*, Convenio específico de investigación, Documento suscrito el 29 de julio de 2024., Medellín, Colombia, jul. de 2024.

Capítulo 12

Análisis multitemporal de coberturas terrestres en zonas de minería aluvial mediante imágenes Landsat

Autores

Maura Melissa Herrera¹; David Stephen Fernández Mc Cann²; Fabio de Jesús Vélez Macías³; Néstor Jaime Aguirre Ramírez⁴

Resumen

La minería de oro de aluvión ha transformado de manera significativa distintos sectores del municipio de Ayapel, Córdoba, especialmente en zonas asociadas a microcuencas que drenan hacia la Ciénaga de Ayapel. Este humedal, reconocido como sitio RAMSAR desde 2018, constituye un ecosistema estratégico por su biodiversidad, su función hidrológica y su importancia socioeconómica para las comunidades locales. Este capítulo presenta una síntesis técnica, reorganizada para el *Seminario GeoLimna–GEPAR*, del artículo publicado por Herrera et al. (2025), en el cual se analizó la dinámica multitemporal de coberturas

¹Universidad de Antioquia. Grupo de investigación GeoLimna.

²Profesor titular, Departamento de Ingeniería Electrónica y Telecomunicaciones, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

³Universidad de Antioquia. Grupo de investigación GeoLimna.

⁴Universidad de Antioquia. Grupo de investigación GeoLimna.

⁴Capítulo elaborado como síntesis técnica y editorial del artículo publicado: Herrera, M. M., Fernández, D. S., Vélez, F. J., & Aguirre, N. J. (2025). *Temporal analysis of land cover change in Ayapel's alluvial gold mining zones, Colombia. Earth Sciences Research Journal*, 29(1), 1–13 [1].

terrestres en la microcuenca Quebradona entre 1985 y 2020 mediante imágenes Landsat, índices espectrales, clasificación no supervisada y minería de patrones secuenciales [1]. El flujo metodológico integró imágenes Landsat 4, 5 y 8 con corrección radiométrica TOA, cálculo del índice SAVI para apoyar la discriminación entre coberturas vegetales y suelo desnudo, clasificación K-means en el espacio de color L^*a^*b , postprocesamiento mediante NDWI, máscaras de nubes y red hídrica vectorizada, y análisis secuencial mediante el método SPADE. Los resultados muestran un incremento sustancial de las áreas asociadas a minería activa, con un aumento de 503 ha entre 1985 y 2020 y un pico máximo en 2010. También se identificaron cambios relevantes en suelo desnudo, vegetación baja, vegetación media y vegetación alta. El estudio demuestra la utilidad de la teledetección y del análisis de patrones temporales para documentar la transformación del paisaje en zonas de minería aluvial y para apoyar procesos de monitoreo ambiental en territorios de alta sensibilidad ecológica.

Palabras clave: teledetección; minería aluvial; imágenes Landsat; Ciénaga de Ayapel; SAVI; NDWI; K-means; minería de patrones secuenciales; coberturas terrestres; GeoLimna; GEPAR.

12.1. Introducción

La minería de oro de aluvión modifica de manera directa la estructura del paisaje. La apertura de frentes de extracción, la remoción de cobertura vegetal, la excavación de piscinas mineras, la exposición del suelo y la alteración de cauces generan señales espaciales observables desde sensores remotos. En territorios donde la actividad minera ocurre de manera informal o por fuera de marcos regulatorios, el monitoreo satelital se convierte en una herramienta fundamental para documentar cambios históricos, identificar tendencias y aportar evidencia técnica a la gestión ambiental [1], [2], [3].

La Ciénaga de Ayapel constituye uno de los sistemas cenagosos más importantes del Caribe colombiano. Su condición de humedal RAMSAR reconoce su valor ecológico, hidrológico y socioeconómico; sin embargo, las microcuencas que la alimentan han estado sometidas a presiones derivadas de la minería aluvial. En particular, las quebradas Quebradona, Escobillas y Trejos han sido reportadas como zonas con afectación por deforestación, alteración del cauce y transformación de coberturas asociadas a la actividad minera [1].

El trabajo sintetizado en este capítulo corresponde a un artículo ya publicado en *Earth Sciences Research Journal*, en el cual se analizaron los cambios de cobertura terrestre en la microcuenca Quebradona, localizada al sureste de la Ciénaga de Ayapel. Para su inclu-

sión en este fascículo, el contenido se reorganiza como reporte académico del *Seminario GeoLimna–GEPAR*, manteniendo el fundamento metodológico y los resultados principales del artículo original, pero enfatizando la arquitectura de procesamiento, la lectura matemática del método y su valor como experiencia de monitoreo ambiental apoyada en procesamiento digital de imágenes [1].

El estudio empleó imágenes Landsat correspondientes al periodo 1985–2020, combinando procesamiento digital de imágenes, índices espectrales, clasificación no supervisada y minería de patrones secuenciales [1], [4]. Esta integración permitió no solo estimar áreas por clase, sino también reconstruir trayectorias de cambio en el tiempo.

La Fig. 12.1 presenta el área de estudio. La delimitación espacial corresponde al polígono de análisis usado para la microcuenca Quebradona, en el municipio de Ayapel, Córdoba. La imagen combina mapas de localización regional y una escena satelital en falso color que permite reconocer el polígono de trabajo dentro del entorno de la Ciénaga de Ayapel.

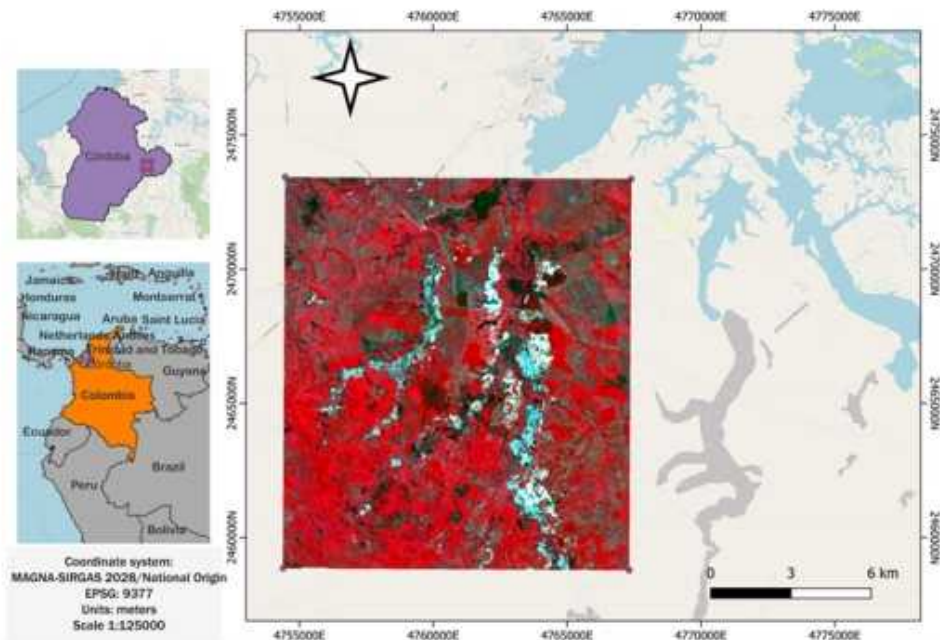


Figura 12.1: Área de estudio correspondiente a la microcuenca Quebradona, localizada al sureste de la Ciénaga de Ayapel, Córdoba. El polígono delimita la zona usada para el análisis multitemporal de coberturas terrestres.

El capítulo conserva el énfasis técnico del artículo original, pero lo presenta como una memoria estructurada dentro del seriado *Seminario GeoLimna–GEPAR*. Se privilegia la descripción del problema, la formalización del flujo metodológico, la interpretación de resultados y el valor institucional del trabajo como articulación entre teledetección,

procesamiento digital de imágenes y monitoreo ambiental.

12.2. Problema abordado

La minería aluvial produce cambios físicos y espectrales que pueden ser observados en series de imágenes satelitales. La desaparición de coberturas vegetales, la aparición de suelo desnudo, la formación de cuerpos de agua aislados y la alteración de redes de drenaje son expresiones del proceso extractivo. Sin embargo, distinguir entre suelo desnudo, cuerpos de agua, vegetación de distinta densidad y áreas mineras activas requiere un flujo de procesamiento capaz de integrar información espectral, espacial y temporal [1], [5].

Desde el punto de vista computacional, el problema consiste en transformar una colección de imágenes multiespectrales en una serie ordenada de mapas clasificados. Sea

$$\mathcal{I} = \{I_{t_1}, I_{t_2}, \dots, I_{t_T}\}$$

el conjunto de imágenes Landsat disponibles para las fechas t_1, t_2, \dots, t_T . Cada imagen I_t está compuesta por bandas espectrales

$$I_t = \{B_{t,NIR}, B_{t,R}, B_{t,G}, B_{t,QA}\},$$

donde $B_{t,NIR}$ corresponde a la banda de infrarrojo cercano, $B_{t,R}$ a la banda roja, $B_{t,G}$ a la banda verde y $B_{t,QA}$ a la banda de calidad empleada para el manejo de nubes. Esta nomenclatura coincide con la organización funcional mostrada en la Fig. 12.2, donde las entradas principales corresponden a imágenes Landsat 4, 5 y 8, el área de estudio de la microcuenca Quebradona y las bandas NIR, rojo, verde y QA.

El objetivo es construir, para cada fecha t , una imagen clasificada

$$C_t : \Omega \rightarrow \mathcal{Y},$$

donde Ω es el dominio espacial de píxeles y \mathcal{Y} representa el conjunto de clases de cobertura consideradas en el análisis:

$$\mathcal{Y} = \left\{ \begin{array}{l} \text{agua, suelo desnudo, vegetación baja,} \\ \text{vegetación media, vegetación alta,} \\ \text{suelo minero activo, nubes} \end{array} \right\}.$$

A partir de la serie $\{C_t\}_{t=1}^T$, el análisis multitemporal busca identificar trayectorias de cambio por píxel. Para cada ubicación (i, j) , se define una secuencia:

$$S_{i,j} = (C_{t_1}(i, j), C_{t_2}(i, j), \dots, C_{t_T}(i, j)).$$

Estas secuencias permiten estudiar transiciones como vegetación alta hacia vegetación media, vegetación baja hacia suelo desnudo, o suelo minero activo hacia vegetación baja. El interés no se limita a saber cuánta área ocupa cada clase en una fecha, sino a comprender cómo evoluciona el territorio a través del tiempo [6], [7].

12.3. Descripción funcional del flujo metodológico

El análisis se organizó en cuatro bloques funcionales: entradas, clasificación, postprocesamiento y análisis multitemporal. El primer bloque corresponde a las imágenes Landsat 4, 5 y 8 adquiridas entre 1985 y 2020, la delimitación del área de estudio en la microcuenca Quebradona y las bandas espectrales NIR, rojo, verde y QA. El segundo bloque corresponde a la clasificación inicial, donde se realiza corrección y recorte, cálculo del índice SAVI, conversión al espacio L^*a^*b , agrupamiento K-means y obtención de coberturas preliminares. El tercer bloque corresponde al postprocesamiento, donde se calcula NDWI, se detectan cuerpos de agua, se incorporan máscaras de nubes mediante la banda QA, se usa la red hídrica vectorizada y se identifica suelo minero activo. El cuarto bloque corresponde al análisis multitemporal, donde se trabaja con la serie de mapas clasificados, se construyen secuencias por píxel y se aplica minería de patrones secuenciales mediante SPADE.

La Fig. 12.2 presenta el flujo general de procesamiento en una versión sintética, diseñada para mostrar la lógica completa del método desde las entradas hasta la salida final de monitoreo multitemporal de coberturas y minería aluvial en Ayapel.



Figura 12.2: Flujo general de procesamiento. El método parte de imágenes Landsat 4, 5 y 8, del área de estudio en la microcuenca Quebradona y de las bandas NIR, rojo, verde y QA; luego ejecuta clasificación mediante SAVI y K-means, postprocesamiento mediante NDWI, máscaras de nubes y red hídrica vectorizada, y finalmente análisis multitemporal mediante secuencias por píxel y SPADE.

De manera compacta, el flujo puede representarse como una composición de funciones:

$$F(\mathcal{I}) = M \circ P \circ K \circ V(\mathcal{I}),$$

donde V representa el cálculo de índices y transformaciones espectrales, K la clasificación no supervisada, P el postprocesamiento espacial y espectral, y M el análisis multitemporal mediante patrones secuenciales.

El resultado final no es una única imagen, sino un conjunto de mapas clasificados, áreas por cobertura, trayectorias de cambio y tendencias temporales:

$$F(\mathcal{I}) = \left(\{C_t\}_{t=1}^T, \{A_{y,t}\}, \mathcal{S} \right),$$

donde $A_{y,t}$ representa el área de la clase y en la fecha t , y \mathcal{S} es el conjunto de secuencias relevantes encontradas en el análisis temporal.

12.4. Adquisición y preparación de imágenes Landsat

El estudio empleó imágenes Landsat T1 obtenidas desde Google Earth Engine, correspondientes a las misiones Landsat 4, Landsat 5 y Landsat 8 [1], [4]. Las imágenes utilizadas cuentan con preprocesamiento radiométrico de reflectancia TOA, lo que permite trabajar con valores normalizados entre 0 y 1 para las bandas espectrales consideradas.

Las bandas usadas fueron infrarrojo cercano, rojo visible, verde visible y la banda de calidad. Esta selección responde a dos necesidades metodológicas. Por un lado, las bandas NIR y roja permiten calcular índices de vegetación, como SAVI [8]. Por otro lado, las bandas verde e infrarroja permiten calcular NDWI para detección de cuerpos de agua [9]. La banda QA se utilizó para identificar nubes y reducir su efecto sobre la clasificación.

La zona de estudio fue delimitada mediante un polígono que cubre la microcuenca Quebradona. Esta delimitación permitió aplicar el mismo dominio espacial a todas las imágenes de la serie, condición necesaria para comparar áreas y calcular secuencias temporales por píxel.

12.5. Clasificación inicial mediante SAVI y K-means

La primera fase del procesamiento consistió en calcular el índice SAVI, diseñado para reducir la influencia del brillo del suelo en zonas con vegetación dispersa [8]. Este índice se ajusta mejor que NDVI en paisajes donde coexisten vegetación, suelo desnudo y áreas degradadas, condición frecuente en zonas afectadas por minería aluvial [1].

Para cada píxel (i, j) , el índice SAVI se calculó como

$$B_{i,j,SAVI} = \frac{B_{i,j,NIR} - B_{i,j,R}}{B_{i,j,NIR} + B_{i,j,R} + L}(1 + L),$$

donde $B_{i,j,NIR}$ es la reflectancia en infrarrojo cercano, $B_{i,j,R}$ es la reflectancia en rojo visible y L es el factor de corrección por suelo. En el estudio se empleó $L = 0,5$, valor recomendado para superficies con exposición moderada de suelo. Se usa la notación $B_{i,j,SAVI}$ para mantener consistencia con la Fig. 12.3, donde el índice SAVI se representa como una imagen derivada de las bandas espectrales.

La Fig. 12.3 presenta la fase de clasificación, donde el cálculo de SAVI se articula con la conversión al espacio $L \times a \times b$ y la clasificación no supervisada mediante K-means sobre los canales a y b .

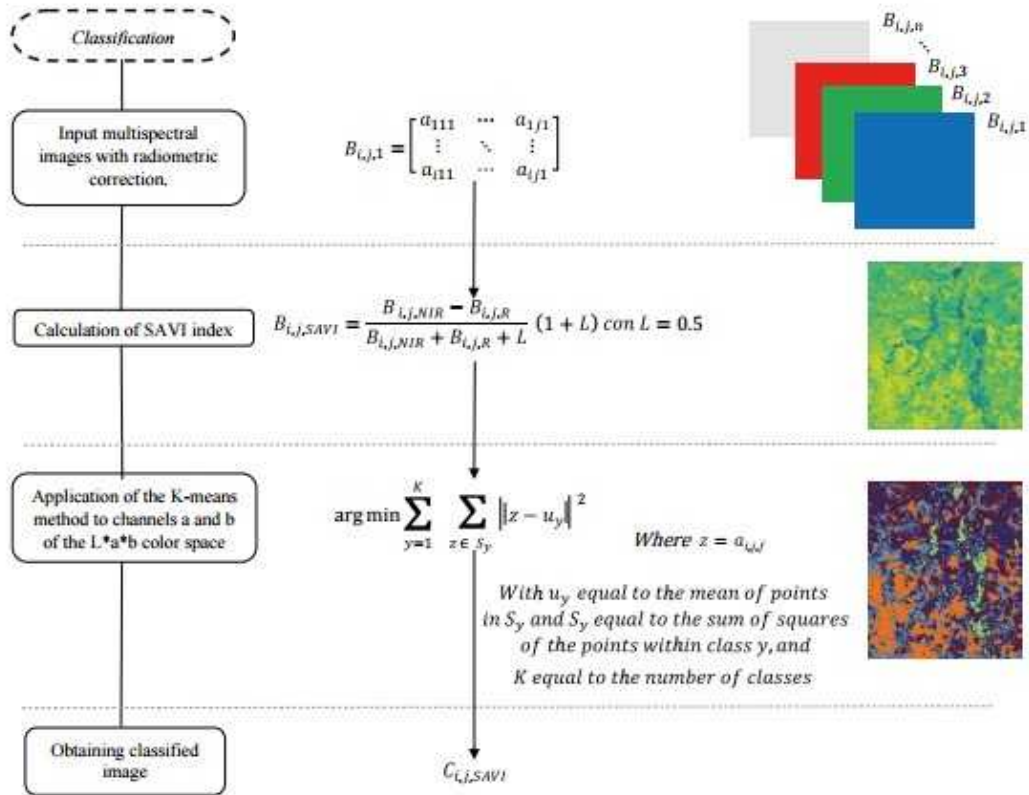


Figura 12.3: Fase de clasificación de imágenes satelitales. El proceso recibe imágenes multispectrales con corrección radiométrica, calcula el índice SAVI, convierte la imagen al espacio L^*a^*b y aplica K-means sobre los canales a y b para obtener una imagen clasificada preliminar.

Posteriormente, la imagen derivada del índice SAVI fue convertida al espacio de color L^*a^*b . La clasificación se aplicó sobre los canales a y b , descartando el canal L , asociado a luminosidad. Esta decisión buscó reducir el efecto de variaciones de brillo y enfatizar contrastes cromáticos relacionados con las coberturas.

El algoritmo K-means asigna cada observación a uno de K grupos, minimizando la suma de distancias cuadráticas dentro de cada clúster [10]. Formalmente, dado un conjunto de observaciones $X = \{x_1, x_2, \dots, x_n\}$, el objetivo es minimizar:

$$\arg \min_{\{S_y\}_{y=1}^K} \sum_{y=1}^K \sum_{x \in S_y} \|x - \mu_y\|^2,$$

donde S_y es el conjunto de píxeles asignados a la clase y , y μ_y es el centroide corres-

pondiente. En el estudio, esta etapa produjo una imagen preliminar $C_{i,j,SAVI}$ con clases asociadas a suelo desnudo, vegetación baja, vegetación media y vegetación alta.

12.6. Postprocesamiento: agua, nubes y suelo minero activo

La clasificación inicial fue complementada mediante una fase de postprocesamiento. Esta fase permitió incorporar clases que no quedaban completamente resueltas por el agrupamiento inicial: agua, nubes y suelo minero activo. Para la identificación de cuerpos de agua se empleó el índice NDWI, ampliamente usado para resaltar cuerpos de agua en imágenes de sensores remotos [9], calculado como

$$B_{i,j,NDWI} = \frac{B_{i,j,G} - B_{i,j,NIR}}{B_{i,j,G} + B_{i,j,NIR}},$$

donde $B_{i,j,G}$ es la reflectancia en la banda verde. A partir del NDWI se construyó una máscara de agua:

$$C_{i,j,W} = \begin{cases} 1, & \text{si } B_{i,j,NDWI} > 0,3, \\ 0, & \text{en otro caso.} \end{cases}$$

Esta máscara permitió identificar cuerpos de agua aislados, que en el contexto minero pueden corresponder a piscinas de extracción. Para distinguir estos cuerpos de agua de ríos y quebradas, se utilizó una red hídrica vectorizada, denotada como $V_{i,j,W}$. La operación de comparación entre la máscara de agua y la red hídrica permitió conservar cuerpos de agua aislados y diferenciar cauces naturales de posibles piscinas mineras [1], [5].

La Fig. 12.4 resume la fase de postprocesamiento y usa la misma nomenclatura: $C_{i,j,SAVI}$ como imagen clasificada preliminar, $B_{i,j,NDWI}$ como índice de agua, $C_{i,j,W}$ como máscara de cuerpos de agua, $V_{i,j,W}$ como red hídrica vectorizada y $B_{i,j,QA}$ como banda de calidad para detección de nubes.

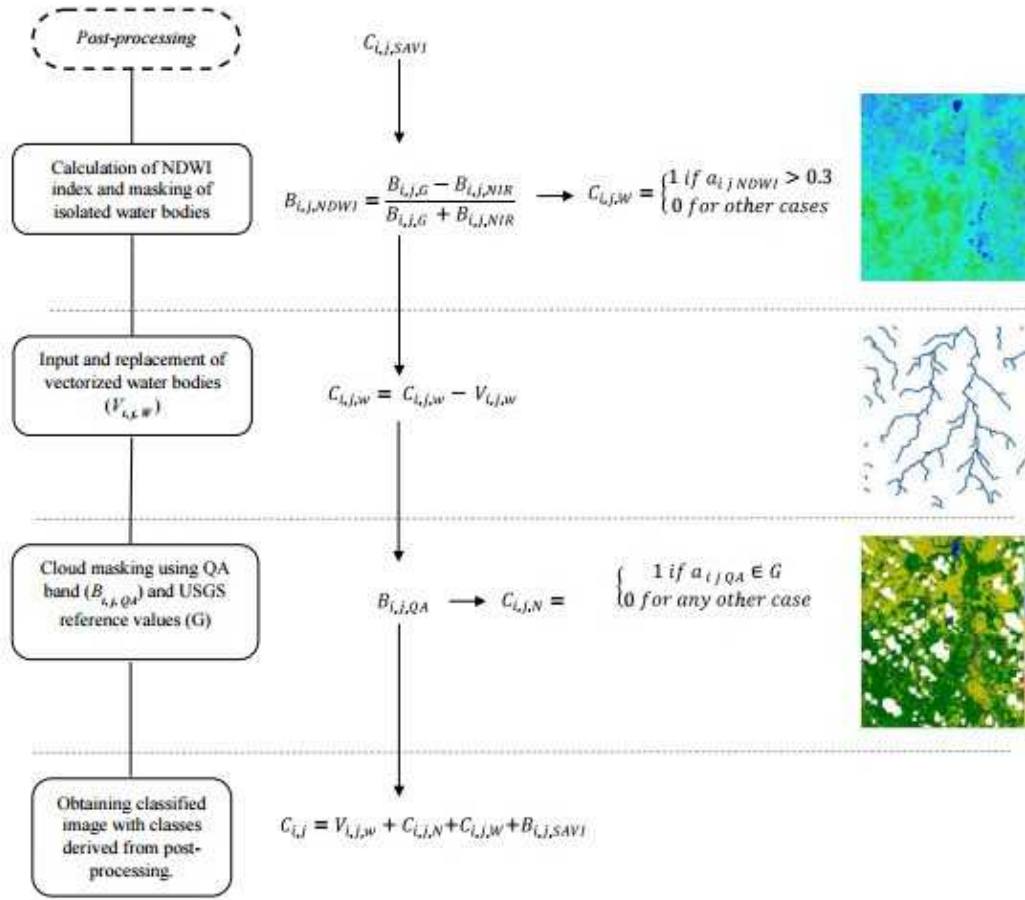


Figura 12.4: Fase de postprocesamiento. El método calcula NDWI para detectar cuerpos de agua aislados, incorpora la red hídrica vectorizada $V_{i,j,W}$, aplica una máscara de nubes a partir de la banda $B_{i,j,QA}$ y obtiene una imagen clasificada final con clases adicionales derivadas del postprocesamiento.

La identificación de nubes se realizó mediante la banda QA de Landsat. Sea G el conjunto de valores de referencia asociados a nubes. La máscara de nubes puede expresarse como

$$C_{i,j,N} = \begin{cases} 1, & \text{si } B_{i,j,QA} \in G, \\ 0, & \text{en otro caso.} \end{cases}$$

La imagen final clasificada se obtuvo integrando las clases derivadas de la clasificación inicial, las máscaras de agua, la red hídrica vectorizada y la máscara de nubes:

$$C_{i,j} = V_{i,j,W} + C_{i,j,N} + C_{i,j,W} + C_{i,j,SAVI}.$$

Esta integración permitió obtener mapas de cobertura más completos y adecuados para el análisis multitemporal. La identificación de suelo minero activo se apoyó en la detección de cuerpos de agua aislados y en la presencia de suelo desnudo alrededor de dichos cuerpos, de acuerdo con la lógica metodológica de las piscinas mineras asociadas a minería aluvial.

12.7. Análisis multitemporal mediante patrones secuenciales

La fase multitemporal se basó en la construcción de secuencias de clases para cada píxel. Dado un conjunto de imágenes clasificadas

$$\{C_{i,j,1}, C_{i,j,2}, \dots, C_{i,j,t}\},$$

cada píxel (i, j) genera una secuencia ordenada:

$$S_{i,j} = (y_{i,j,1}, y_{i,j,2}, \dots, y_{i,j,t}),$$

donde $y_{i,j,t}$ es la clase asignada al píxel (i, j) en la fecha t . Las nubes fueron excluidas del análisis porque representan ausencia de información útil y pueden distorsionar la interpretación de cambios.

El estudio utilizó el método SPADE, *Sequential Pattern Discovery using Equivalence Classes*, para identificar secuencias frecuentes [6]. Este tipo de minería de patrones secuenciales ha sido empleado en análisis de series espaciales y temporales derivadas de sensores remotos [7]. La Fig. 12.5 resume esta fase.

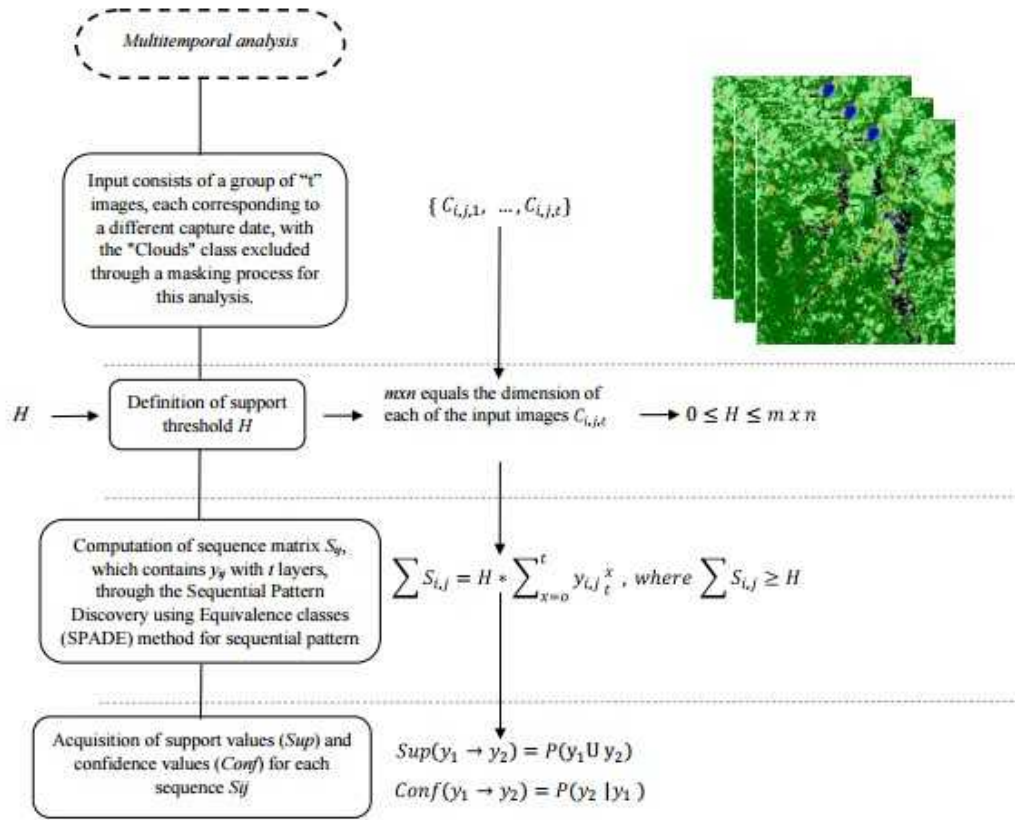


Figura 12.5: Fase de análisis multitemporal. La serie de imágenes clasificadas se transforma en secuencias por píxel $S_{i,j}$; posteriormente se define un umbral de soporte H , se aplica SPADE y se obtienen valores de soporte Sup y confianza $Conf$ para cada secuencia.

El soporte de una secuencia s se define como la proporción de ocurrencias de dicha secuencia respecto al número total de píxeles válidos:

$$Sup(s) = \frac{N(s)}{N_{\Omega}},$$

donde $N(s)$ es el número de píxeles que presentan la secuencia s , y N_{Ω} es el número total de píxeles considerados. La confianza de una transición $y_1 \rightarrow y_2$ se expresa como

$$Conf(y_1 \rightarrow y_2) = P(y_2|y_1) = \frac{P(y_1 \cup y_2)}{P(y_1)}.$$

Estas medidas permiten diferenciar entre secuencias frecuentes y transiciones condicionalmente probables. En términos ambientales, esto resulta útil para identificar persistencia

de coberturas, degradación, revegetación o alternancia entre suelo desnudo, minería y vegetación.

12.8. Resultados principales

La clasificación basada en SAVI, procesamiento de imágenes y K-means alcanzó una exactitud global de $74,37\% \pm 0,10$ y un coeficiente Kappa de 0.658 [1]. La evaluación de exactitud se apoyó en prácticas habituales de matrices de confusión, interpretación visual y estimación de calidad temática en clasificación de coberturas [11], [12], [13]. De acuerdo con la escala interpretativa de Landis y Koch, este valor corresponde a un acuerdo sustancial. Aunque existen algoritmos supervisados que pueden alcanzar valores superiores, la combinación SAVI-K-means ofreció resultados adecuados para el objetivo del estudio y permitió procesar de manera consistente una serie histórica de imágenes Landsat.

El resultado espacial de las imágenes clasificadas muestra la evolución de las coberturas entre 1985 y 2020. En la Fig. 12.6 se presenta la gráfica de cambios de cobertura, expresada como número de hectáreas por clase en función del año de análisis.

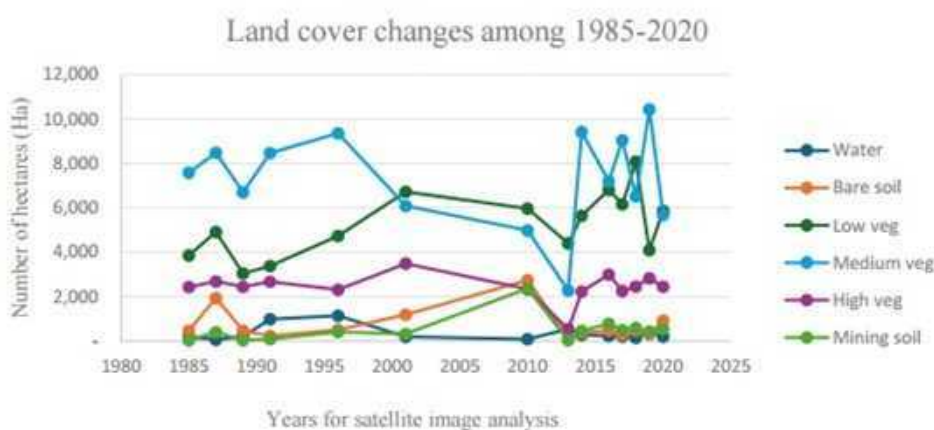


Figura 12.6: Cambios de coberturas terrestres entre 1985 y 2020. La gráfica resume la evolución temporal de agua, suelo desnudo, vegetación baja, vegetación media, vegetación alta y suelo minero activo en la microcuenca Quebradona.

Entre 1985 y 2020, la minería activa aumentó 503 ha, lo que equivale a 12 veces el área registrada en 1985. El pico máximo se observó en 2010, cuando el área minera fue 54 veces mayor que en 1985. El suelo desnudo aumentó 96 % para 2020. La vegetación alta se mantuvo relativamente constante, pasando de 2434 ha en 1985 a 2446 ha en 2020.

La vegetación baja aumentó de 3855 ha a 5859 ha, mientras que la vegetación media disminuyó de 7588 ha a 5674 ha, equivalente a una reducción del 25 % [1].

La interpretación temporal también muestra que los años 1989 y 2013 presentan una alta presencia de nubosidad, lo cual generó subestimación de las áreas clasificadas. Esta limitación es importante porque en estudios multitemporales con sensores ópticos la nubosidad puede alterar las series y producir falsas disminuciones de cobertura.

El comportamiento del suelo minero activo muestra un incremento inicial entre 1985 y 1987, asociado por entrevistas locales a la incorporación de maquinaria en la extracción aurífera. Posteriormente, el estudio identifica un periodo de menor actividad relativa entre la década de 1990 y 2004, seguido por una expansión importante entre 2001 y 2010. El aumento de 2039 ha entre esos años coincide con reportes de intensificación de la actividad minera después de 2005 [1].

12.9. Interpretación de los patrones secuenciales

El análisis de patrones secuenciales permitió identificar trayectorias de persistencia y cambio. En el análisis estacional, la subsecuencia *suelo minero activo* \rightarrow *suelo minero activo* presentó el mayor número de ocurrencias, con 6081 casos y una confianza de 58.71 %. Esto indica que una proporción importante de los píxeles clasificados como minería activa tendió a permanecer en esa condición en la siguiente observación [1].

También se identificaron subsecuencias como *suelo minero activo* \rightarrow *vegetación baja*, con una confianza cercana al 47.97 %, y *suelo minero activo* \rightarrow *vegetación baja* \rightarrow *vegetación baja*, con confianza de 59.42 %. Estos patrones sugieren que parte de las áreas mineras o intervenidas pueden transitar hacia coberturas de vegetación baja, lo cual puede interpretarse como revegetación inicial, abandono de zonas mineras o cambios estacionales en el paisaje.

En el análisis de vegetación, la secuencia *vegetación alta* \rightarrow *vegetación alta* fue una de las más recurrentes, lo que indica estabilidad relativa de algunas zonas boscosas. Sin embargo, también se observaron trayectorias de degradación, como *vegetación alta* \rightarrow *vegetación media* \rightarrow *vegetación media*, con una confianza de 41.61 % en el periodo 1985–1989. Esta secuencia puede interpretarse como pérdida de densidad o transformación progresiva de cobertura forestal.

El valor del análisis secuencial reside en que permite ir más allá de la comparación entre mapas aislados. Una misma área puede cambiar varias veces a lo largo de la serie; por ello, las trayectorias por píxel aportan una lectura más fina de persistencia, degradación,

revegetación o transición entre estados.

12.10. Discusión

El estudio original demuestra que la combinación de teledetección, índices espectrales, clasificación no supervisada y minería de patrones secuenciales permite documentar de forma consistente la transformación de coberturas en zonas afectadas por minería aluvial [1]. En territorios como Ayapel, donde el acceso de campo puede estar limitado por condiciones logísticas o de seguridad, las imágenes satelitales ofrecen una vía objetiva para reconstruir procesos históricos y generar evidencia espacial.

La relación entre suelo desnudo y minería activa aparece como una señal relevante. La remoción de cobertura vegetal suele preceder o acompañar la actividad extractiva, de modo que el aumento de suelo desnudo puede funcionar como indicador indirecto de presión minera. No obstante, esta relación debe interpretarse con cautela, porque el suelo desnudo también puede estar asociado a actividades agropecuarias, caminos, sequías, erosión u otras formas de intervención.

La exactitud global obtenida, aunque aceptable, evidencia las limitaciones de una metodología basada en Landsat y clasificación no supervisada. La resolución espacial de 30 m, la presencia de nubes, la variación entre sensores y la complejidad del mosaico de coberturas pueden afectar la precisión. Por ello, una línea futura de trabajo debería incorporar imágenes de mayor resolución, métodos supervisados como Random Forest o Support Vector Machine, algoritmos de aprendizaje profundo y mayor validación en campo [14].

El aporte metodológico más significativo está en la lectura temporal. La minería de patrones secuenciales permite interpretar no solo estados, sino trayectorias. Esto abre posibilidades para estudiar recuperación vegetal, persistencia de minería activa, expansión de suelo desnudo y transiciones asociadas a periodos sociales o económicos del territorio. En el caso de Ayapel, las entrevistas locales permitieron contextualizar cambios detectados satelitalmente, especialmente los relacionados con incorporación de maquinaria, presencia de actores armados y expansión posterior de la minería.

12.11. Aporte al monitoreo ambiental de la Ciénaga de Ayapel

El análisis de coberturas en la microcuenca Quebradona tiene relevancia directa para la protección de la Ciénaga de Ayapel. Las microcuencas que drenan hacia el sistema cenagoso funcionan como corredores de transporte de sedimentos, contaminantes y alteraciones hidrológicas. Por tanto, la transformación de coberturas aguas arriba puede afectar procesos ecológicos dentro del humedal.

El trabajo se articula con una línea de investigación más amplia orientada al monitoreo ambiental de sistemas leníticos y humedales estratégicos. Desde GeoLimna, aporta conocimiento sobre dinámica territorial, coberturas y presión minera. Desde GEPAR, consolida el uso de procesamiento digital de imágenes, análisis computacional y herramientas reproducibles para extraer información ambiental a partir de datos remotos.

La metodología descrita puede ser reutilizada en otros sistemas cenagosos o microcuencas afectadas por minería, deforestación o expansión agropecuaria. Su implementación en Python permite automatizar etapas de procesamiento, reducir variabilidad operativa y facilitar actualizaciones periódicas cuando se disponga de nuevas imágenes [1], [4].

12.12. Conclusiones

Este capítulo presentó una síntesis técnica y editorial de un artículo ya publicado sobre análisis multitemporal de coberturas terrestres en la microcuenca Quebradona, municipio de Ayapel, Córdoba [1]. El estudio permitió documentar transformaciones relevantes asociadas a minería de oro de aluvión entre 1985 y 2020. La minería activa aumentó de manera marcada durante el periodo de estudio, con un máximo en 2010, mientras que el suelo desnudo presentó un incremento significativo para 2020. La vegetación alta se mantuvo relativamente estable, la vegetación baja aumentó y la vegetación media disminuyó respecto al inicio de la serie.

La metodología basada en imágenes Landsat, SAVI, NDWI, K-means, máscaras de nubes, red hídrica vectorizada y minería de patrones secuenciales permitió construir una lectura integrada del territorio. El enfoque resultó útil para estimar áreas por cobertura, observar tendencias y analizar trayectorias de cambio por píxel. Aunque la resolución espacial y la nubosidad introducen limitaciones, el procedimiento ofrece una base reproducible para monitoreo histórico y seguimiento ambiental.

El estudio evidencia que la teledetección constituye una herramienta clave para territorios

donde el acceso de campo es limitado o donde se requiere reconstruir procesos de largo plazo. En el caso de Ayapel, el monitoreo de las microcuencas asociadas a la Ciénaga resulta fundamental para orientar acciones de gestión, restauración y control ambiental frente a la expansión de la minería aluvial.

Agradecimientos

Los autores agradecen a la Universidad de Antioquia y a los grupos de investigación GeoLimna y GEPAR por el soporte académico, técnico e institucional que permitió el desarrollo del estudio. Este trabajo se desarrolló en el marco del proyecto SIIU 2020-37870: *Sistema de registro y seguimiento de variables ambientales en la Ciénaga de Ayapel*, orientado al fortalecimiento de capacidades de monitoreo ambiental en sistemas cenagosos estratégicos.

Referencias

- [1] M. M. Herrera, D. S. Fernández, F. d. J. Vélez y N. J. Aguirre, «Temporal analysis of land cover change in Ayapel's alluvial gold mining zones, Colombia,» *Earth Sciences Research Journal*, vol. 29, n.º 1, págs. 1-13, 2025. DOI: 10.15446/esrj.v29n1.111372.
- [2] T. Lillesand, R. W. Kiefer y J. Chipman, *Remote Sensing and Image Interpretation*, 7.^a ed. John Wiley & Sons, 2015.
- [3] Z. Zhu y C. E. Woodcock, «Continuous monitoring of forest disturbance using all available Landsat imagery,» *Remote Sensing of Environment*, vol. 144, págs. 152-171, 2014.
- [4] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau y R. Moore, «Google Earth Engine: Planetary-scale geospatial analysis for everyone,» *Remote Sensing of Environment*, vol. 202, págs. 18-27, 2017.
- [5] E. Ibrahim, L. Lema, P. Barnabé, P. Lacroix y E. Pirard, «Small-scale surface mining of gold placers: Detection, mapping, and temporal analysis through the use of free satellite imagery,» *International Journal of Applied Earth Observation and Geoinformation*, vol. 93, pág. 102194, 2020.
- [6] M. J. Zaki, «SPADE: An efficient algorithm for mining frequent sequences,» *Machine Learning*, vol. 42, n.º 1-2, págs. 31-60, 2001.

- [7] X. Wu y X. Zhang, «An efficient pixel clustering-based method for mining spatial sequential patterns from serial remote sensing images,» *Computers & Geosciences*, vol. 126, págs. 12-23, 2019.
- [8] A. R. Huete, «A soil-adjusted vegetation index (SAVI),» *Remote Sensing of Environment*, vol. 25, n.º 3, págs. 295-309, 1988.
- [9] S. K. McFeeters, «The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features,» *International Journal of Remote Sensing*, vol. 17, n.º 7, págs. 1425-1432, 1996.
- [10] J. A. Hartigan y M. A. Wong, «A K-Means Clustering Algorithm,» *Journal of the Royal Statistical Society: Series C*, vol. 28, n.º 1, págs. 100-108, 1979.
- [11] R. G. Congalton y K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 3.ª ed. CRC Press, 2019.
- [12] G. M. Foody, «Status of land cover classification accuracy assessment,» *Remote Sensing of Environment*, vol. 80, n.º 1, págs. 185-201, 2002.
- [13] P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock y M. A. Wulder, «Good practices for estimating area and assessing accuracy of land change,» *Remote Sensing of Environment*, vol. 148, págs. 42-57, 2014.
- [14] S. Talukdar et al., «Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review,» *Remote Sensing*, vol. 12, n.º 7, pág. 1135, 2020.

Capítulo 13

Detección automatizada de irregularidades en pavimento mediante visión artificial y aprendizaje profundo

Jose Andres Henao Alzate¹
Claudia Cristina Botero Suárez²
David Stephen Fernández Mc Cann³

13.1. Introducción

La evaluación del estado superficial de los pavimentos constituye una tarea relevante para la gestión de infraestructura vial, especialmente cuando se requiere identificar de manera oportuna la presencia de grietas, huecos y patrones de deterioro asociados con la denominada piel de cocodrilo. Estos defectos no solo afectan la comodidad y seguridad de los usuarios de la vía, sino que también aportan información técnica para planificar labores de mantenimiento, priorizar intervenciones y construir registros históricos sobre la evolución del deterioro.

En este contexto se desarrolló CRACKFINDER, un software orientado a la detección au-

¹Estudiante, Facultad de Ingeniería, Universidad de Antioquia. Grupo de investigación GEPAR.

²InterTelco S.A.S. Coinvestigadora industrial del proyecto.

³Profesor, Facultad de Ingeniería, Universidad de Antioquia. Grupos de investigación GEPAR y Geolimna.

tomatizada de irregularidades en pavimentos mediante técnicas de procesamiento digital de imágenes, visión por computador y aprendizaje profundo. El sistema fue concebido en el marco del convenio específico de investigación entre la Universidad de Antioquia e InterTelco S.A.S., asociado al proyecto *VISIONARY: Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, cuyo alcance incluye el desarrollo de modelos basados en visión artificial para seguridad vial y supervisión de vías [1], [2].

El desarrollo se inscribe en una línea de trabajo aplicada en la que la Universidad de Antioquia, a través del grupo GEPAR y en articulación con capacidades de GeoLimna, aporta conocimiento en procesamiento de imágenes, visión artificial y modelado computacional, mientras que InterTelco S.A.S. contribuye con necesidades de validación tecnológica, criterios de aplicación y condiciones de uso asociadas al sector de la videoanalítica. Esta relación permitió orientar el desarrollo hacia un producto de software con trazabilidad técnica, documentación de usuario, manual técnico y una estructura funcional susceptible de integración futura en sistemas de gestión vial.

CRACKFINDER se encuentra en proceso de registro como soporte lógico ante la Dirección Nacional de Derecho de Autor, con radicado de solicitud No. 1-2026-39681 del 2 de marzo de 2026 [3]. En consecuencia, el presente capítulo describe el desarrollo tecnológico y metodológico del sistema sin afirmar que el registro haya sido concedido. La obra se presenta como un resultado de investigación aplicada derivado del convenio UdeA–InterTelco, con una arquitectura implementada en Python 3.12, una interfaz gráfica en Tkinter, módulos de preprocesamiento, inferencia con YOLOv8 y generación automática de reportes en video y hoja de cálculo [4], [5].

13.2. Planteamiento del problema

La inspección visual de pavimentos suele depender de recorridos en campo, observación directa, registro fotográfico y posterior clasificación manual de deterioros. Aunque esta práctica conserva valor técnico, presenta limitaciones cuando se requiere analizar grandes trayectos, repetir mediciones de manera consistente o generar información estructurada en tiempos reducidos. La variabilidad de criterio entre observadores, las condiciones de iluminación, el movimiento del vehículo de captura y la diversidad de texturas del pavimento introducen incertidumbre en la detección y clasificación de imperfecciones.

Desde la perspectiva de visión artificial, el problema puede formularse como una tarea de detección de objetos en secuencias de video. Dado un conjunto de fotogramas obtenidos sobre la superficie de una vía, el sistema debe localizar regiones de interés y asignarles una

clase asociada al tipo de deterioro observado. En CRACKFINDER, las clases de interés se definieron como grietas, huecos y piel de cocodrilo. Esta formulación permite pasar de una inspección cualitativa a un flujo de análisis que produce evidencias visuales y registros temporales cuantificables.

Sea V un video de entrada compuesto por una secuencia ordenada de fotogramas:

$$V = \{I_1, I_2, \dots, I_T\}, \quad (13.1)$$

donde I_t representa el fotograma en el instante discreto t y T corresponde al número total de fotogramas analizados. El objetivo del sistema es obtener, para cada fotograma, un conjunto de detecciones:

$$\mathcal{D}_t = \{(b_i, c_i, s_i)\}_{i=1}^{N_t}, \quad (13.2)$$

donde b_i es la caja delimitadora de la región detectada, c_i es la clase asignada, s_i es el puntaje de confianza del modelo y N_t es el número de detecciones presentes en el fotograma I_t . Para este caso, el conjunto de clases se define como:

$$\mathcal{C} = \{\text{grieta, hueco, piel de cocodrilo}\}. \quad (13.3)$$

La dificultad técnica está asociada a la variabilidad de escala, contraste, textura, iluminación y perspectiva. Las grietas pueden aparecer como estructuras delgadas y de bajo contraste; los huecos pueden confundirse con sombras o manchas superficiales; y la piel de cocodrilo puede manifestarse como un patrón distribuido, con bordes irregulares y geometría no uniforme. Por esta razón, el sistema incorpora una etapa de homogenización de iluminación antes de la inferencia, con el fin de reducir la influencia de cambios locales de brillo y mejorar la estabilidad visual del análisis.

13.3. Adquisición de datos y condiciones de captura

La adquisición de datos se realizó mediante videos e imágenes de pavimentos reales capturados con una cámara instalada sobre un vehículo en movimiento. La cámara se ubicó a una altura constante, buscando mantener una perspectiva estable y una distancia de observación relativamente uniforme respecto a la superficie de la vía. Esta condición permite reducir variaciones geométricas innecesarias durante el entrenamiento y posterior inferencia del modelo.



Figura 13.1: Montaje empleado para la captura de datos sobre pavimento. La cámara se instala sobre un vehículo en movimiento para registrar la superficie vial bajo una perspectiva relativamente constante.

La Figura 13.1 ilustra el montaje de captura utilizado como referencia para la generación del material visual. La estabilidad relativa de la cámara, junto con la selección de horarios de iluminación natural homogénea, contribuye a obtener imágenes con menor ruido visual. Se evitó, en lo posible, la captura en condiciones de oscuridad, lluvia reciente, humedad superficial o zonas de sombra intensa, debido a que estas condiciones pueden aumentar la aparición de falsos positivos o reducir la sensibilidad del modelo ante irregularidades pequeñas.

El conjunto de datos fue construido a partir de imágenes representativas de diferentes texturas y estados del pavimento. Las muestras fueron etiquetadas manualmente para identificar las clases de deterioro consideradas. Esta anotación permitió estructurar un conjunto de entrenamiento y validación orientado a la detección supervisada, de modo que el modelo pudiera aprender patrones visuales asociados con grietas, huecos y piel de cocodrilo.

13.4. Preprocesamiento de imágenes

El primer componente del flujo computacional corresponde al preprocesamiento de los fotogramas. Su función es mejorar la calidad visual de la información antes de la inferencia, particularmente en presencia de variaciones de iluminación. En CRACKFINDER se

implementó una estrategia de homogenización basada en filtrado gaussiano, que permite estimar una componente suavizada de la imagen y compensar cambios lentos de brillo.

Sea $I_t(x, y)$ el valor de intensidad del fotograma t en la posición espacial (x, y) . La componente suavizada de iluminación puede expresarse como:

$$L_t(x, y) = G_\sigma(x, y) * I_t(x, y), \quad (13.4)$$

donde G_σ es un filtro gaussiano con desviación estándar σ y $*$ denota la operación de convolución. A partir de esta estimación, se obtiene una imagen corregida:

$$I'_t(x, y) = I_t(x, y) - L_t(x, y) + \bar{I}_t, \quad (13.5)$$

donde \bar{I}_t corresponde al valor promedio del fotograma original. Esta operación reduce la influencia de variaciones suaves de iluminación y conserva estructuras locales relevantes, como bordes, discontinuidades y cambios de textura asociados con deterioros del pavimento.

El resultado se normaliza antes de ser entregado al modelo de detección. En términos generales, el preprocesamiento puede representarse como una transformación:

$$\tilde{I}_t = \mathcal{P}(I_t), \quad (13.6)$$

donde $\mathcal{P}(\cdot)$ agrupa las operaciones de suavizado, compensación de iluminación y normalización. Esta etapa es especialmente importante para la identificación de grietas de bajo contraste, que pueden confundirse con variaciones naturales de textura cuando el fotograma presenta iluminación irregular.

13.5. Modelo de detección

La detección automática de irregularidades se implementó mediante YOLOv8, un modelo de detección de objetos basado en redes neuronales convolucionales, utilizado en este desarrollo para localizar y clasificar regiones defectuosas en fotogramas de pavimento [6]. La elección de un detector de una sola etapa responde a la necesidad de procesar secuencias de video y generar resultados visuales de forma eficiente, manteniendo una relación adecuada entre capacidad de detección y tiempo de inferencia.

El modelo recibe como entrada el fotograma preprocesado \tilde{I}_t y produce un conjunto de detecciones:

$$\mathcal{D}_t = f_\theta(\tilde{I}_t), \quad (13.7)$$

donde f_θ representa la red neuronal entrenada y θ agrupa sus parámetros. Cada detección contiene la localización espacial, la clase asignada y una medida de confianza. En términos operativos, el modelo marca sobre el fotograma las regiones en las que identifica una irregularidad y genera un conteo por clase.

Para cada fotograma analizado se calcula un vector de conteo:

$$\mathbf{n}_t = \begin{bmatrix} n_t^{grieta} \\ n_t^{hueco} \\ n_t^{piel} \end{bmatrix}, \quad (13.8)$$

donde n_t^{grieta} , n_t^{hueco} y n_t^{piel} representan, respectivamente, el número de grietas, huecos y zonas de piel de cocodrilo detectadas en el fotograma t . Esta representación permite convertir las salidas visuales del modelo en registros tabulares útiles para análisis posteriores.

La evaluación del desempeño del modelo puede realizarse mediante métricas clásicas de detección y clasificación, entre ellas precisión, exhaustividad y medida F_1 . Para una clase dada, estas métricas se definen como:

$$\text{Precisión} = \frac{VP}{VP + FP}, \quad (13.9)$$

$$\text{Exhaustividad} = \frac{VP}{VP + FN}, \quad (13.10)$$

$$F_1 = 2 \cdot \frac{\text{Precisión} \cdot \text{Exhaustividad}}{\text{Precisión} + \text{Exhaustividad}}, \quad (13.11)$$

donde VP corresponde a verdaderos positivos, FP a falsos positivos y FN a falsos negativos. Estas métricas permiten analizar si el modelo tiende a sobredetectar imperfecciones, omitir defectos reales o mantener un equilibrio aceptable entre sensibilidad y confiabilidad.

13.6. Arquitectura funcional del software

CRACKFINDER fue desarrollado como una aplicación de escritorio en Python 3.12, integrando bibliotecas de procesamiento de imágenes, cálculo numérico, manejo de datos, visualización e inferencia con aprendizaje profundo. La arquitectura se organiza en dos componentes principales: una interfaz gráfica de usuario y un backend de procesamiento. Esta separación permite mantener independencia entre la interacción del usuario y la lógica computacional encargada de procesar los fotogramas.

La interfaz fue implementada en Tkinter con apoyo de Pillow para la visualización de imágenes. Desde esta capa, el usuario puede cargar el video de entrada, definir la ruta de salida, seleccionar intervalos temporales de análisis, iniciar la inferencia y observar el avance del procesamiento. El backend, estructurado alrededor de la clase `PavementProcessor`, administra la lectura del video, el preprocesamiento, la inferencia con YOLOv8, la actualización de contadores, la generación del video anotado y la escritura del archivo de resultados.



Figura 13.2: Interfaz principal de CRACKFINDER. La aplicación integra visualización del fotograma original, visualización del resultado procesado, controles de ejecución, barra de progreso y contadores de imperfecciones.

La Figura 13.2 muestra la organización general de la interfaz. El diseño permite comparar el fotograma original con el resultado inferido, lo que facilita la revisión visual de las

detecciones. Además, los controles de tiempo permiten ejecutar el análisis sobre todo el video o sobre un tramo específico, una función útil cuando solo una parte del recorrido contiene información relevante.

El flujo de procesamiento puede expresarse de manera sintética como:

$$V \rightarrow \{I_t\}_{t=1}^T \rightarrow \{\tilde{I}_t\}_{t=1}^T \rightarrow \{\mathcal{D}_t\}_{t=1}^T \rightarrow (V', R), \quad (13.12)$$

donde V es el video de entrada, I_t son los fotogramas originales, \tilde{I}_t los fotogramas pre-procesados, \mathcal{D}_t las detecciones generadas, V' el video de salida con anotaciones y R el registro tabular de resultados.

Para mantener la fluidez de la aplicación, el sistema utiliza hilos de ejecución que permiten procesar el video sin bloquear la interfaz gráfica. Esta decisión es relevante en aplicaciones de escritorio que ejecutan tareas intensivas de visión artificial, ya que el usuario necesita recibir retroalimentación permanente sobre el estado del proceso, la ruta de salida y el avance de la inferencia.

13.7. Flujo funcional del sistema

La operación de CRACKFINDER puede entenderse como una secuencia funcional que transforma un video capturado sobre pavimento en productos de salida útiles para análisis técnico. Esta secuencia integra la interacción del usuario con la aplicación, las etapas de procesamiento de imágenes y aprendizaje profundo, y la consolidación de resultados visuales y tabulares. La Figura 13.3 resume esta lógica de funcionamiento en nueve etapas consecutivas.



Figura 13.3: Flujo funcional de CRACKFINDER. El sistema parte de un video de entrada capturado sobre pavimento, realiza la configuración del análisis, aplica preprocesamiento e inferencia con YOLOv8, clasifica las irregularidades detectadas y genera salidas visuales y tabulares para análisis técnico.

El flujo inicia con la adquisición del *video de entrada*, correspondiente a una secuencia visual obtenida sobre la superficie del pavimento mediante una cámara montada en un vehículo. Esta primera etapa define la materia prima del análisis, pues la calidad del registro visual condiciona la capacidad del sistema para reconocer irregularidades de interés.

A continuación, el usuario realiza la *carga del video* en la interfaz de la aplicación. Este paso permite seleccionar el archivo que será procesado y establecer la conexión entre la evidencia capturada en campo y el entorno de análisis implementado en el software. Posteriormente, el sistema pasa a una etapa de *configuración*, en la cual se define la ruta de almacenamiento de resultados y, cuando es necesario, el intervalo temporal específico sobre el cual se ejecutará la inferencia. Esta posibilidad de limitar el análisis a un tramo particular del video resulta útil cuando solo una parte del recorrido contiene información relevante o cuando se desea reducir el tiempo de procesamiento.

Una vez definida la configuración, CRACKFINDER aplica una fase de *preprocesamiento*. En esta etapa se corrigen variaciones suaves de iluminación mediante una estrategia de homogenización que mejora la visibilidad de detalles locales sobre la superficie del pavimento. Su propósito es atenuar el efecto de sombras, cambios de brillo y condiciones no uniformes de captura, favoreciendo así la identificación de defectos de bajo contraste.

Después del preprocesamiento se ejecuta la etapa de *detección*, en la que el sistema aplica el modelo YOLOv8 sobre los fotogramas seleccionados. El modelo localiza regiones con probabilidad de corresponder a deterioros y asigna a cada una de ellas una predic-

ción con su respectiva confianza. Estas detecciones son luego organizadas en una fase de *clasificación*, donde las irregularidades se agrupan en las tres clases definidas para el desarrollo: grietas, huecos y piel de cocodrilo. Esta separación permite interpretar el deterioro detectado no solo como una localización espacial, sino también como una categoría funcionalmente útil para la evaluación de la vía.

El resultado inmediato del proceso se expresa en una *salida visual*, consistente en un video con las detecciones superpuestas sobre la secuencia original. Esta representación facilita la validación cualitativa del comportamiento del algoritmo y conserva evidencia visual del análisis realizado. De manera complementaria, el sistema produce un *registro tabular* que almacena el conteo de irregularidades y su ubicación temporal dentro del video. Esta salida permite transformar la información visual en datos estructurados, adecuados para seguimiento, revisión o integración en informes técnicos.

Finalmente, ambas salidas convergen en una *salida final* orientada al análisis técnico. En esta etapa, CRACKFINDER entrega un conjunto de productos que pueden apoyar procesos de inspección de pavimentos, documentación de hallazgos y priorización de mantenimiento. En conjunto, el flujo funcional representado en la Figura 13.3 muestra que el sistema no se limita a ejecutar un detector sobre imágenes aisladas, sino que organiza una cadena operativa completa que va desde la captura del video hasta la generación de resultados interpretables para la toma de decisiones.

Desde una perspectiva formal, esta secuencia puede resumirse como:

$$V \rightarrow C \rightarrow P \rightarrow D \rightarrow K \rightarrow (S_v, S_t), \quad (13.13)$$

donde V representa el video de entrada, C la configuración del análisis, P la etapa de pre-procesamiento, D la detección de regiones de interés, K la clasificación de irregularidades, S_v la salida visual y S_t la salida tabular. Esta formulación sintetiza el comportamiento general del sistema y pone de manifiesto la relación entre las etapas que componen la operación de CRACKFINDER.

13.8. Generación de resultados

El sistema produce dos salidas principales: un video con detecciones superpuestas y un archivo de hoja de cálculo con registros temporales de las imperfecciones. El video conserva la secuencia original, pero incorpora cajas delimitadoras y etiquetas sobre las regiones donde el modelo identifica irregularidades. Esta salida permite una validación visual directa de la respuesta del algoritmo.



Figura 13.4: Ejemplo de salida visual generada por CRACKFINDER. Las detecciones se superponen al fotograma original mediante cajas delimitadoras y etiquetas de clase.

La Figura 13.4 muestra un ejemplo de inferencia en el que el sistema identifica regiones compatibles con grietas en la superficie del pavimento. Esta salida es útil para revisar el comportamiento del modelo, verificar la localización de las detecciones y conservar evidencia visual del análisis realizado.

La segunda salida corresponde a un archivo en formato Excel. Este registro contiene el minuto y segundo del video original en el que se detecta al menos una imperfección, junto con el número de ocurrencias por clase. El sistema evita registrar fotogramas sin detecciones, con el propósito de reducir información innecesaria y facilitar la revisión posterior.

	A	B	C	D	E	F
1	Minuto	Segundo	Huecos	Grietas	l de cocodrilo	
2	3	56	0	2	0	
3	3	56	0	1	0	
4	3	57	0	1	0	
5	4	57	0	1	1	
6	4	58	0	1	0	
7	4	58	0	2	0	
8						
9						
10						
11						

Figura 13.5: Registro tabular de resultados generado por CRACKFINDER. La hoja de cálculo almacena el tiempo de detección y el conteo de imperfecciones por clase.

La Figura 13.5 ilustra la estructura del archivo de resultados. Para cada instante registrado, se almacenan las columnas de minuto, segundo y conteos por clase. Este esquema facilita la construcción de reportes técnicos, la comparación entre recorridos y la integración futura con sistemas de gestión de mantenimiento vial.

Desde el punto de vista formal, el registro de resultados puede definirse como:

$$R = \left\{ (m_t, s_t, n_t^{hueco}, n_t^{grieta}, n_t^{piel}) \mid n_t^{hueco} + n_t^{grieta} + n_t^{piel} > 0 \right\}, \quad (13.14)$$

donde m_t y s_t representan el minuto y segundo asociados al fotograma procesado, mientras que los términos n_t^{hueco} , n_t^{grieta} y n_t^{piel} corresponden a los conteos de cada clase. Esta estructura conserva únicamente los instantes con evidencia de deterioro, lo que permite concentrar el análisis en segmentos relevantes del video.

13.9. Trazabilidad, documentación y propiedad intelectual

El desarrollo de CRACKFINDER incluyó documentación técnica y de usuario, manuales de instalación y uso, descripción funcional, arquitectura del software, dependencias, re-

querimientos y esquema de generación de resultados [4], [5]. Esta documentación fortalece la reproducibilidad del desarrollo y facilita su revisión por usuarios técnicos, entidades responsables de mantenimiento vial o instancias institucionales de transferencia tecnológica.

El software fue desarrollado como resultado de investigación aplicada en el marco del proyecto VISIONARY. El acta institucional de inicio registra una duración de 24 meses, entre el 29 de julio de 2024 y el 29 de julio de 2026, y una financiación conjunta entre la Universidad de Antioquia e InterTelco S.A.S. [2]. Asimismo, el convenio específico de investigación estableció como parte de los resultados esperados el registro de software asociado a la identificación de irregularidades en pavimento [1].

La solicitud de registro ante la Dirección Nacional de Derecho de Autor fue radicada para la obra *CRACKFINDER – detección de irregularidades en pavimento*. En este capítulo se menciona dicha radicación únicamente como evidencia de trámite y no como certificado definitivo de registro concedido [3]. Esta precisión es importante para mantener coherencia documental y evitar afirmaciones que excedan el estado actual del proceso.

13.10. Discusión

CRACKFINDER representa una aproximación aplicada al problema de inspección automatizada de pavimentos mediante visión artificial. Su principal aporte consiste en integrar, dentro de una misma aplicación, la adquisición de video, el preprocesamiento orientado a homogenizar iluminación, la detección con aprendizaje profundo, la visualización de resultados y la generación de reportes temporales. Esta integración permite transformar un registro visual continuo en información estructurada sobre la presencia de irregularidades.

El sistema no reemplaza la valoración técnica especializada de infraestructura vial, pero puede actuar como una herramienta de apoyo para reducir tiempos de inspección, organizar evidencia visual y priorizar tramos que requieren revisión detallada. Su utilidad depende de la calidad de las imágenes, la similitud entre las condiciones de captura y el conjunto de entrenamiento, la estabilidad de la cámara, la iluminación y la diversidad de pavimentos incluidos en la base de datos.

Una limitación propia de este tipo de sistemas es la sensibilidad ante condiciones no representadas durante el entrenamiento. Sombras fuertes, lluvia, superficies húmedas, reflejos, oclusiones, marcas viales desgastadas y cambios abruptos de textura pueden afectar la respuesta del modelo. Por ello, el software debe entenderse como una plataforma en evolución, susceptible de ser robustecida mediante nuevas campañas de captura, ampliación del conjunto de datos, validación en distintos escenarios y ajuste de umbrales de confianza.

Desde la perspectiva de transferencia tecnológica, la articulación con InterTelco S.A.S. permitió orientar el desarrollo hacia un entorno de uso potencial en videoanalítica aplicada a infraestructura. La participación de estudiantes en el desarrollo, junto con la documentación del sistema y el trámite institucional de propiedad intelectual, fortalece la relación entre formación, investigación aplicada y generación de productos tecnológicos.

13.11. Conclusiones

Se desarrolló un sistema de software para la detección automatizada de irregularidades en pavimento a partir de videos capturados desde un vehículo. El sistema integra preprocesamiento por homogenización de iluminación, inferencia mediante YOLOv8, visualización de detecciones y generación automática de reportes en video y hoja de cálculo.

La arquitectura modular implementada en Python permite separar la interfaz gráfica de la lógica de procesamiento, lo que facilita el mantenimiento, la trazabilidad y la posibilidad de ampliaciones futuras. La generación de resultados temporales por clase convierte la detección visual en información estructurada, útil para reportes técnicos y posibles sistemas de gestión vial.

El desarrollo se realizó en el marco del convenio de investigación entre la Universidad de Antioquia e InterTelco S.A.S., dentro del proyecto VISIONARY, y cuenta con solicitud de registro de soporte lógico ante la Dirección Nacional de Derecho de Autor. En conjunto, CRACKFINDER evidencia una ruta de trabajo en la que la visión artificial se articula con necesidades reales de supervisión vial, formación de estudiantes y transferencia tecnológica desde la ingeniería.

Referencias

- [1] Universidad de Antioquia e InterTelco S.A.S., *Convenio específico de investigación No. 21330002-0002-2024: Uso de técnicas de visión artificial para sistemas de seguridad y supervisión de vías*, Documento institucional del convenio de investigación entre la Universidad de Antioquia e InterTelco S.A.S., 29 de jul. de 2024.
- [2] Universidad de Antioquia, *Acta de registro y de propiedad intelectual del proyecto VISIONARY: Advanced Artificial Vision Systems for Enhanced Road Safety and Infrastructure Supervision*, Acta No. 2024-74783, 13 de nov. de 2024.

- [3] Dirección Nacional de Derecho de Autor, *Constancia de radicación de solicitud de registro de soporte lógico: CRACKFINDER – detección de irregularidades en pavimento*, Radicado No. 1-2026-39681, 2 de mar. de 2026.
- [4] Universidad de Antioquia, *CRACKFINDER: detección de irregularidades en pavimento. Manual técnico*, Documento técnico del software CRACKFINDER, 2025.
- [5] Universidad de Antioquia, *CRACKFINDER: detección de irregularidades en pavimento. Manual de usuario*, Documento técnico de usuario del software CRACKFINDER, 2025.
- [6] Ultralytics, *Ultralytics YOLOv8 documentation*, Documentación técnica del modelo YOLOv8, 2024.

Investigaciones 2024-2025

Grupos de investigación GeoLimna y GEPAR



Este libro reúne una selección de desarrollos, investigaciones y soluciones tecnológicas construidas entre 2024 y 2025 en la convergencia entre visión artificial, instrumentación, analítica de datos, automatización y monitoreo ambiental.

Los capítulos aquí reunidos muestran cómo problemas reales del entorno pueden traducirse en sistemas medibles, algoritmos verificables, dispositivos funcionales y herramientas útiles para la investigación, la docencia y la transferencia de conocimiento.

La obra refleja una forma de trabajo interdisciplinaria en la que la ingeniería dialoga con la limnología, la ecología, la agricultura de precisión, la infraestructura y el análisis ambiental. En ese cruce, los grupos GeoLimna y GEPAR han consolidado una línea de producción académica y tecnológica orientada a la comprensión de fenómenos y a la construcción de soluciones concretas.

Universidad de Antioquia

Facultad de Ingeniería

Grupos de investigación GeoLimna y GEPAR

Medellín, Colombia

2026