



CONGRESO INTERNACIONAL  
**TRANSFORMANDO  
LA FRONTERA**  
— EDUCACIÓN • TECNOLOGÍA • IMPACTO —

TRABAJOS DE PRESENTACIÓN DEL CONGRESO INTERNACIONAL  
TRANSFORMANDO LA FRONTERA

## Compendio de pósteres científicos con arbitraje

Primera edición 2026

ISBN 978-970-96337-0-2

DOI:

Editorial: Noé Toledo González

En colaboración con la Universidad Tecnológica de Matamoros

H. Matamoros, Tamaulipas, México



CONGRESO INTERNACIONAL  
**TRANSFORMANDO**  
LA FRONTERA

UNIVERSIDAD TECNOLÓGICA **UT** M A T A M O R O S

ISBN 978-970-96337-0-2

Editorial: Noé Toledo González  
En colaboración con la Universidad Tecnológica de Matamoros  
H. Matamoros, Tamaulipas, México

Trabajos de presentación del Congreso Internacional Transformando la Frontera Compendio de pósteres científicos con arbitraje

Primera edición, 2026

D.R. © 2026, Noé Toledo González

Sello editorial: Noé Toledo González – 978-607-96337 – Art90 editorial

[www.f90.com.mx/editorial](http://www.f90.com.mx/editorial)

En colaboración con la Universidad Tecnológica de Matamoros [www.utmatamoros.edu.mx](http://www.utmatamoros.edu.mx)

Coordinación Editorial: Dr. Noé Toledo González

Cuerpo Académico responsable: CA-08 Ingeniería de Sistemas

Programa académico anfitrión: Maestría en Tecnologías de la Información con Especialidad en Gestión de Proyectos

Diseño editorial y formación: Dr. Noé Toledo González

Corrección de estilo: M.E.R. Maira Selene Ríos Gómez

ISBN: 978-970-96337-0-2

DOI del libro: 10.5281/zenodo.22030116

Obra publicada bajo licencia Creative Commons

Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0). Se permite la copia, distribución y adaptación de la obra con reconocimiento de la autoría, sin fines comerciales y bajo los mismos términos.

Aviso de responsabilidad:

Los contenidos de los pósteres y resúmenes extendidos que integran esta obra son responsabilidad exclusiva de sus autores. Las opiniones, resultados, interpretaciones y conclusiones expresadas no reflejan necesariamente la postura institucional de la Universidad Tecnológica de Matamoros ni del sello editorial responsable.

Hecho en México · Made in Mexico Publicación  
digital de distribución en acceso abierto

|                                          |                                         |
|------------------------------------------|-----------------------------------------|
| <b>Rectoría</b>                          | Mtra. Diana Rosa Masso Quintana         |
| <b>Dirección de Enlace Académico</b>     | Dr. Jose Juan Villanueva Sierra         |
| <b>Sub-Dirección de Enlace Académico</b> | Mtro. Juan Carlos Vela Saldivar         |
| <b>Dirección de Vinculación</b>          | Dr. Mariano Martínez Díaz de León       |
| <b>Dirección de Finanzas</b>             | Lic. Felicitas Marquez Ayala            |
| <b>Abogado General</b>                   | Lic. Carlos Faustino González Maldonado |
| <b>Coordinación MTI-GP</b>               | Dr. Noé Toledo González                 |



## COMITÉ ORGANIZADOR

Presidencia del Comité

Dr. Noé Toledo González — Coordinador MTI-GP

Coordinaciones Operativas

| Área                                      | Responsable                            |
|-------------------------------------------|----------------------------------------|
| Coordinación Académica                    | Dra. Lidia Elena Requena Hernández     |
| Coordinación Científica y de Dictamen     | M.E.R. Maira Selene Ríos Gomez         |
| Coordinación Editorial                    | Dr. Noé Toledo González                |
| Coordinación de Logística                 | Dr. Eddie Eliud Vázquez Alexandre      |
| Coordinación de Comunicación y Difusión   | M.E.R. Hugo Alberto Solis Martinez     |
| Coordinación de Vinculación Internacional | Dra. Xóchitl Soledad Ricárdez González |
| Coordinación de Tecnología y Plataformas  | Mtro. Marco Antonio Durán Raya         |

La revisión académica de las contribuciones que integran la presente obra estuvo a cargo del siguiente cuerpo de dictaminadores, integrado por académicos internos y externos a la Universidad Tecnológica de Matamoros. Su participación se realizó bajo el esquema de arbitraje doble ciego, con criterios comunes de evaluación acordados previamente por la Coordinación Científica.

#### **Dictaminadores Internos (Universidad Tecnológica de Matamoros)**

Dr. Eddie Eliud Vázquez Alexandre

Dra. Xóchitl Soledad Ricárdez González

M.E.E Misael Condado Fomperosa

#### **Dictaminadores Externos**

Dr. Humberto Raymundo González Moreno — TECNM–Instituto Tecnológico Superior de Misantla, México

Dra. Alejandra Hikari Rico Martínez — Universidad Tecnológica de Saltillo

Dra. Maria Teresa Castañeda Aguilar — TECNM–Instituto Tecnológico de Matamoros, México

Dr. Rodolfo Arturo Echaverría Solís — University of Northern Arizona

M.S. Eduardo Emanuel Tovas Garza — University of Texas Rio Grande Valley

La Coordinación Editorial agradece a cada integrante del Comité Científico su trabajo desinteresado y su compromiso con la calidad académica del presente compendio.

La Universidad Tecnológica de Matamoros se complace en presentar la primera edición de los Trabajos de Presentación del Congreso Internacional Transformando la Frontera, un ejercicio académico que consolida el compromiso de nuestra institución con la formación en investigación, la divulgación científica y el fortalecimiento del tejido universitario en la región fronteriza del noreste de México.

El presente compendio reúne doce contribuciones estudiantiles, docentes y de investigadores invitados que participaron en la edición 2026 del Congreso, organizado por el Cuerpo Académico CA-08 «Ingeniería de Sistemas» y la Coordinación de la Maestría en Tecnologías de la Información con Especialidad en Gestión de Proyectos. Cada trabajo fue sometido a un proceso de arbitraje académico doble ciego que garantiza la calidad y pertinencia de los contenidos aquí publicados.

En un contexto en el que la frontera norte de México se redefine constantemente por las presiones del comercio internacional, la migración, la transición energética y la transformación digital, el papel de una universidad tecnológica pública adquiere una dimensión estratégica. No basta con formar profesionistas: es imperativo generar conocimiento contextualizado, útil, capaz de dialogar con las realidades industriales, sociales y ambientales que definen a nuestra región.

Los cinco ejes temáticos que estructuran la presente obra —Industria 5.0, Excelencia Operacional y Manufactura Inteligente Aplicada, Comercio Global, Logística Inteligente y Desarrollo Fronterizo, Innovación, Emprendimiento y Competitividad, y Educación Bilingüe— dan cuenta de la amplitud de líneas de trabajo que se cultivan en nuestra comunidad académica y de nuestros vínculos con instituciones hermanas de ambos lados del Río Bravo. La colaboración binacional con la University of Texas Rio Grande Valley y con University of Northern Arizona, así como con universidades del interior del país, robustece la calidad crítica de este volumen.

Extendemos nuestro reconocimiento a los autores, al Comité Científico, a la Coordinación Editorial y a todas las áreas administrativas de la Universidad que hicieron posible este resultado. Invitamos al lector a que este libro sea no solo memoria de un evento, sino punto de partida para nuevas colaboraciones que continúen transformando la frontera desde el conocimiento.

Con doce trabajos distribuidos en cinco ejes temáticos —Industria 5.0; Excelencia Operacional y Manufactura Inteligente Aplicada; Comercio Global, Logística Inteligente y Desarrollo Fronterizo; Innovación, Emprendimiento y Competitividad para la Transformación Organizacional; y Educación Bilingüe—, este volumen documenta la capacidad instalada de nuestra comunidad académica para generar conocimiento aplicado y pertinente. Cada póster fue sometido a arbitraje doble ciego por un Comité Científico integrado por especialistas internos y externos, lo que confiere a esta edición el carácter de publicación académica con arbitraje.

Sostener un congreso académico en la frontera exige más que voluntad institucional: requiere convertir la frontera misma en objeto de estudio y en pretexto para el encuentro. El Congreso Internacional Transformando la Frontera nace con esa doble intención, articular a estudiantes, docentes e investigadores en torno a los desafíos tecnológicos, sociales y económicos que atraviesan nuestra región, y hacerlo desde un espacio universitario que reconoce su lugar como puente entre dos naciones.

Los doce trabajos que integran esta primera edición representan el esfuerzo de decenas de participantes que aceptaron el reto de traducir sus proyectos en pósteres científicos y en resúmenes extendidos con estructura académica formal. La decisión editorial de compilar estos trabajos en un volumen con ISBN responde a la convicción de que la divulgación es parte constitutiva de la investigación: un resultado que no circula permanece incompleto.

El compendio se organiza en cinco ejes temáticos que reflejan las líneas prioritarias del Cuerpo Académico CA-08 «Ingeniería de Sistemas» y de los programas académicos participantes. Cada contribución incluye, en primer lugar, la lámina original del póster presentado durante el Congreso, y a continuación su desarrollo extendido con introducción, metodología, resultados, conclusiones y referencias bibliográficas. Esta doble presentación busca conservar el valor visual y comunicativo del póster como formato de divulgación, sin renunciar al rigor textual que caracteriza a la producción académica arbitrada.

El proceso de dictamen se describe con detalle en la nota metodológica que sigue a este prólogo. Baste decir aquí que el Comité Científico, integrado por tres dictaminadores internos y dos externos, trabajó bajo el esquema de arbitraje doble ciego con criterios de evaluación previamente consensuados. De las propuestas originalmente recibidas, doce fueron aceptadas tras el proceso de arbitraje. Este dato es indicativo del carácter selectivo del compendio y sustenta el reconocimiento de la obra como publicación con arbitraje académico. Agradezco al Rector Dr. Fernando Villanueva Cortés su respaldo institucional continuo. A la Dra. Mónica Salazar Villalobos y al Dr. Ricardo Espinoza Cárdenas su liderazgo en la coordinación académica y científica. Al comité de dictaminadores nacionales e internacionales, su tiempo y su cuidado en la revisión. A los autores, su valentía para exponer sus resultados. A todo el equipo operativo del Congreso, su trabajo tras bambalinas que hizo posible este resultado tangible.

Que este volumen sea el primero de una serie anual que consolide a Transformando la Frontera como un espacio de referencia para la producción académica del noreste mexicano y de la región transfronteriza.

El lector encontrará en las páginas siguientes un recorrido por cinco líneas de trabajo que, sin ser exhaustivas, retratan la diversidad de la investigación aplicada que hoy se cultiva en la MTI-GP: sistemas de visión artificial e inteligencia artificial aplicados a procesos industriales bajo el paradigma de la Industria 5.0; herramientas de control y automatización orientadas a la excelencia operacional; soluciones de visión por computadora para la logística y el comercio transfronterizo; desarrollos de base tecnológica con vocación de innovación y emprendimiento; y experiencias de inclusión educativa apoyadas en inteligencia artificial. Cada eje se presenta con una breve nota introductoria que sitúa al lector antes de los trabajos que lo integran.

Dr. Noé Toledo González  
Coordinación Editorial  
Coordinador de la MTI-GP  
H. Matamoros, Tamaulipas, 2026

La presente obra fue elaborada siguiendo un proceso editorial estructurado en seis etapas, con el objetivo de garantizar la calidad académica de las contribuciones publicadas y otorgar trazabilidad al procedimiento de arbitraje.

### 1. Convocatoria

La convocatoria fue emitida por la Coordinación de la MTI-GP el 15 de febrero de 2026 y difundida a través de los canales institucionales de la UTM, así como en redes de colaboración académica internacional del CA-08. Se recibieron propuestas en formato de póster y resumen extendido conforme a la plantilla oficial del Congreso; de estas, doce fueron finalmente arbitradas y aceptadas para su inclusión en el presente volumen.

### 2. Recepción y verificación de requisitos formales

Las propuestas recibidas fueron sometidas a revisión inicial de forma. Las que presentaron observaciones fueron devueltas a los autores con observaciones y reingresaron corregidas en tiempo y forma.

### 3. Arbitraje doble ciego

Cada contribución fue evaluada por al menos dos dictaminadores del Comité Científico. Los criterios de evaluación consideraron: originalidad y pertinencia del tema, claridad del planteamiento, rigor metodológico, coherencia entre objetivos y resultados, calidad de las referencias bibliográficas y calidad gráfica del póster. En los casos de dictámenes discordantes se recurrió a un tercer dictaminador.

### 4. Comunicación de resultados

Los resultados fueron comunicados mediante carta oficial firmada por la Coordinación Editorial. Del total, doce propuestas fueron aceptadas —bajo las modalidades de aceptación directa y aceptación condicional— para su inclusión en el presente volumen.

### 5. Correcciones y versión final

Los autores cuyas contribuciones fueron aceptadas con condiciones remitieron una versión corregida acompañada de una carta de respuesta detallando los cambios realizados. La Coordinación Editorial verificó la incorporación de las observaciones antes de integrar la contribución al volumen final.

### 6. Edición y publicación

Las doce contribuciones aceptadas fueron integradas al volumen respetando la identidad visual institucional y la estructura editorial definida. El volumen completo fue sometido a corrección de estilo y a una revisión final antes de la solicitud del ISBN ante el Instituto Nacional del Derecho de Autor (INDAUTOR) y la publicación en acceso abierto.

*El expediente completo del proceso editorial, incluyendo dictámenes, cartas de aceptación, cesiones de derechos y actas del Comité Científico, se conserva en el archivo de la Coordinación de la MTI-GP como respaldo de la trazabilidad académica de la presente obra.*



## ÍNDICE GENERAL

|     |                                             |
|-----|---------------------------------------------|
| i   | Página Legal — p. i                         |
| ii  | Directorio Institucional — p. ii            |
| iii | Comité Organizador — p. iii                 |
| iv  | Comité Científico y Dictaminadores — p. iv  |
| v   | Presentación — p. v                         |
| vi  | Prólogo — p. vi                             |
| vii | Nota Metodológica del Proceso de Dictamen — |
| ix  | p. vii                                      |
|     | Índice General — p. ix                      |

## CUERPO DE LA OBRA

### Eje 1 — Industria 5.0 — p. 1

01. Diseño de un Algoritmo de Visión Artificial y Análisis Cinemático para la Corrección Biomecánica en Entrenamientos Inteligentes bajo la Industria 5.0 (VisionFit)  
*Ángel Osiel Jiménez Espinoza, Miguel Ángel Martínez Meza, Christian Ronaldo Becerra Martínez, Juan Humberto García Saucedo, Edgar Niño Paz, Moisés Eli López Trejo — Póster 001 — p. 3*
02. VISION PARK  
*Leal Medina Miguel Ángel, Calixto García Uriel, Muñoz Jiménez Yareli Sarahi, Melo Zuñiga Marvin Yusniel, Sarmiento Juarez José Emmanuel, Recillas Mendoza Marely — Póster 002 — p. 7*

### Eje 2 — Excelencia Operacional y Manufactura Inteligente Aplicada — p. 11

03. Vigía: Sistema de Detección de Color Configurable para la Identificación de Materiales Fuera de Área en Entornos Industriales/Paquetería de México  
*Sandoval Camacho Andrea, Gutiérrez Serralde Irvin, Maldonado Martinez Angelica, Dominguez Lerma Iñaki, Martínez Moreno Jesus — Póster 003 — p. 13*
04. Cyberlock AI. Sistema Inteligente Anti-Intrusos Basado en Visión por Computadora e Inteligencia Artificial  
*Pérez Vázquez Juan Eli (Líder), Pérez Mata José Antonio, Contreras Ochoa Manuel Alejandro, Reyna Montenegro Joel, Rodarte Cristo Luis Ernesto — Póster 004 — p. 18*

05. SIDCA. Sistema Inteligente de Detección y Control Automatizado con Modelos Híbridos (YOLOv4-Tiny/SSD) para Control de Acceso Inteligente  
*Abril Mirelli Pérez Rico, Alexia Noor Nasya Díaz Olivares, Marlene Carreón Jiménez*  
*América, Vianney Gutiérrez Balboa, Eliezer Medina Hernández — Póster 005 — p. 23*

### **Eje 3 — Comercio Global, Logística Inteligente y Desarrollo Fronterizo — p. 27**

06. VISIÓN-X 2.0: Visión por Computadora para Control de Aforo y Seguridad Perimetral con YOLOv11, ByteTrack y Análisis Demográfico  
*Sergio Ismael Ramos Ponce, Ángel Alejandro Mendoza Saucedo, Domenec Sole Huerta, Raymundo Atzin Becerra, Cesar Sánchez Veloz — Póster 006 — p. 29*

### **Eje 4 — Innovación, Emprendimiento y Competitividad para la Transformación Organizacional — p. 32**

07. Sistema Inteligente de Recomendación de Outfits mediante Inteligencia Artificial para la Personalización de la Vestimenta (LookVision)  
*Yoshua Isai Montalvo Sanchez, Jared Martinez Chavez, Rogelio Moreno Hernandez, Aldo Maximiliano Segundo Orta, Angel Eduardo Villalobos Barrios — Póster 007 — p. 34*
08. Solfello: Sistema de Reconocimiento Óptico de Música (OMR) Offline mediante Inteligencia Artificial  
*Carlos Alberto Díaz Bojorquez, Emiliano Salas Perez — Póster 008 — p. 37*
09. Detector de Ingredientes y Recetario YOMI (Your Own Meal Intelligence)  
*Jaziel Netro Castillo, Jorge Luis Medina Oropeza, Gabriel Díaz García, Areli Hernández Neri, Cristóbal Axel Martínez Cervantes — Póster 009 — p. 40*
10. G1-IA. Aplicación Móvil Inteligente Basada en Inteligencia Artificial para la Optimización del Entrenamiento Físico  
*Mena Alcala Diego Abdiel, Mtra. Rios Gomez Maira Selene, Quino Chávez Fátima Guadalupe — Póster 010 — p. 45*

### **Eje 5 — Educación Bilingüe — p. 50**

11. SignTranslate: Sistema Inteligente para el Reconocimiento del Abecedario del Lenguaje de Señas Mexicano mediante Visión por Computadora  
*Ortiz Del Angel Hilda Guadalupe, Ramon Infante Orlando Emanuel, Gonzalez Delgado Manuel — Póster 011 — p. 51*
12. KINESIS AI: Sistema de Visión Artificial para la Detección Temprana de Malestar Emocional en el Aula  
*Ruben D., Yael Banda, Sergio Morán, Marco Salazar — Póster 012 — p. 57*

## **Índice Onomástico de Autores — p. 61**

Enlista, en orden alfabético por apellido, a los autores de las contribuciones incluidas en esta obra, con indicación del número de trabajo donde aparece su participación

## **Colofón — p. 63**

## **Anexos — p. 64**

01. Galería Fotografica del Congreso Internacional

## Industria 5.0

El paradigma de la Industria 5.0 pone al centro la colaboración entre el ser humano y los sistemas inteligentes, complementando la eficiencia de la automatización con la sensibilidad, la creatividad y el juicio humano. Los trabajos reunidos en este eje exploran aplicaciones de visión artificial y aprendizaje automático orientadas a mejorar el desempeño, la seguridad y el bienestar en entornos de entrenamiento y gestión de espacios, evidenciando cómo la inteligencia artificial puede operar como herramienta de apoyo —y no de sustitución— de la persona.

## VisionFit: Entrenador Inteligente Basado en Visión Artificial y Aprendizaje Automático para la Corrección Biomecánica en la Industria 5.0

Jiménez Espinoza, Á. O.\*, Martínez Meza, M. Á., Becerra Martínez, C. R., García Saucedo, J. H., Niño Paz, E., & López Trejo, M. E.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2510455@utmatamoros.edu.mx

Eje Temático: Industria 5.0.

Palabras clave: Visión por computador, análisis cinemático, ergonomía, MediaPipe, entrenamiento inteligente.

### Resumen:

Este trabajo presenta VisionFit, un sistema interactivo basado en visión artificial para la corrección postural y el conteo automatizado en tiempo real. Utilizando MediaPipe Pose y OpenCV en Python, se genera un esqueleto virtual de 16 nodos clave, aplicando filtros de visibilidad y centralización axial para mitigar el ruido ambiental. El algoritmo calcula rangos angulares dinámicos e incorpora un filtro anti-parpadeo temporal de 1.3 segundos y un módulo predictivo de fatiga muscular correlacionado con el gasto calórico (MET). Validado con usuarios reales en una exposición de ciencias, el prototipo demostró una precisión general del 92% en la clasificación de errores técnicos para sentadillas, flexiones y curl de bíceps.

### Introducción:

Bajo el paradigma de la Industria 5.0, el software ergonómico debe priorizar el bienestar y la seguridad adaptada al usuario.

Este proyecto aborda el diseño de VisionFit, un entrenador inteligente no invasivo que emplea visión artificial y análisis cinemático para supervisar y corregir entrenamientos en tiempo real. El objetivo es ofrecer una herramienta computacionalmente ligera y de bajo costo que identifique desviaciones técnicas sutiles.

### Metodología.

Se empleó un enfoque cuantitativo de tipo descriptivo-aplicado, estructurado en tres fases estratégicas:

- **Adquisición y filtrado de datos:** Para mitigar el ruido ambiental, se implementó un filtro de oclusión basado en un umbral de visibilidad de la nariz menor a 0.7 y una restricción de centralización axial del punto medio de la cadera dentro del intervalo relativo [0.2, 0.8].
- **Análisis cinemático-biomecánico:** Utilizando `np.arctan2` se calcularon ángulos vectoriales dinámicos. Para curl de bíceps se mapeó el rango de 40° a 160° a porcentajes de 0% a 100%. En sentadillas se monitorearon los ángulos de las rodillas y la inclinación del torso (alerta a menos de 145°), mientras que en flexiones de pecho se detectaron desalineaciones de cadera inferiores a 165°.
- **Estabilidad y fatiga muscular:** Para evitar el parpadeo de alertas se integró un filtro de estabilidad temporal ( $ERROR\_THRESHOLD = 1.3$  segundos). El cansancio del atleta se estimó mediante un módulo que evalúa si la repetición excede en un factor de 1.4 el promedio móvil previo, cruzando los datos con los equivalentes metabólicos (MET) asignados a cada ejercicio.

### Resultados

El algoritmo demostró una capacidad de adaptación sobresaliente frente a variaciones antropométricas individuales durante su validación empírica con usuarios reales en la exposición de ciencias universitaria, logrando consolidar una tasa de precisión general del 92% en la clasificación de errores técnicos y conteo automatizado. Específicamente, se obtuvo un 94.5% de efectividad en el curl de bíceps, un 90.2% en sentadillas (identificando desalineaciones de espalda y rodillas hacia adentro) y un 91.3% en flexiones de pecho (corrigiendo la caída de cadera).



Figura 2. Interfaz en operación VisionFit

### Conclusiones.

El prototipo desarrollado demuestra que la integración del framework MediaPipe Pose combinado con lógica trigonométrica vectorial personalizada es altamente efectiva para la corrección postural en tiempo real sin requerir hardware especializado. Los resultados validan la hipótesis de que un filtrado temporal robusto (umbral de 1.3 s) elimina falsos positivos cinemáticos, proporcionando una experiencia de usuario altamente fiable.

Como trabajo futuro, se propone incorporar el análisis de movimientos multiplanares más complejos (como peso muerto) y la migración hacia arquitecturas móviles nativas, potenciando el impacto social de la inteligencia artificial aplicada a la salud preventiva y la ergonomía física en la era de la Industria 5.0.

### Referencias

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>
- Hernández, M. A. (2020). *Logística y comercio internacional en la frontera México-Estados Unidos*. Editorial Universitaria.
- INEGI. (2025). *Estadísticas de comercio exterior*. Instituto Nacional de Estadística y Geografía. <https://www.inegi.org.mx>

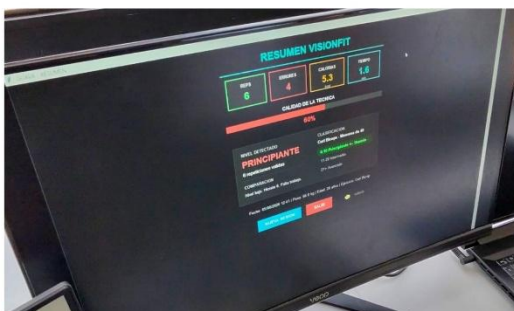


Figura 1. Interfaz de VisionFit

## VisionFit: Entrenador Inteligente Basado en Visión Artificial y Aprendizaje Automático para la Corrección Biomecánica en la Industria 5.0

Jiménez Espinoza, Á. O.\*, Martínez Meza, M. Á., Becerra Martínez, C. R., García Saucedo, J. H., Niño Paz, E., & López Trejo, M. E.

Correspondencia: 2510455@utmatamoros.edu.mx

Eje Temático: Industria 5.0

### Resumen

El sedentarismo y la falta de supervisión profesional en el acondicionamiento físico incrementan sustancialmente el riesgo de lesiones musculoesqueléticas debido a ejecuciones biomecánicas incorrectas. El objetivo de este trabajo fue diseñar e implementar VisionFit, un sistema interactivo de entrenamiento inteligente basado en visión artificial para la corrección postural y el conteo automatizado de ejercicios en tiempo real. La metodología adoptada integró la biblioteca MediaPipe Pose y el framework OpenCV en lenguaje Python para generar un esqueleto biomecánico compuesto por 16 nodos clave del cuerpo humano, aplicando filtros matemáticos de visibilidad y centralización axial que mitigan las interferencias del entorno. El algoritmo computa rangos angulares dinámicos mediante trigonometría vectorial e incorpora un filtro temporal de mitigación de falsos positivos con un umbral de estabilidad de 1.3 segundos, junto con un módulo predictivo de fatiga muscular fundamentado en la degradación temporal del ritmo de ejecución y la tasa de equivalentes metabólicos (MET). Los principales resultados se obtuvieron a través de pruebas de validación con usuarios reales durante una exposición de ciencias universitaria, donde el prototipo funcional demostró una precisión general del 92% en la detección, conteo y clasificación de errores técnicos para tres ejercicios fundamentales: sentadilla, flexiones de pecho y curl de bíceps. Se concluye que VisionFit representa una solución tecnológica viable, computacionalmente accesible y de alta fidelidad que promueve la autogestión segura de la salud física en concordancia con los principios de personalización, resiliencia y bienestar humano que caracterizan el paradigma de la Industria 5.0.

### Palabras Clave

Inteligencia artificial; Visión por computador; Biomecánica; Ergonomía; Procesamiento de imágenes.

## Introducción y Planteamiento del Problema

En la sociedad contemporánea, la adopción de estilos de vida saludables y la práctica regular de ejercicio físico han cobrado una relevancia sin precedentes para combatir enfermedades crónico-degenerativas y mitigar los efectos del sedentarismo tecnológico. No obstante, una porción significativa de la población realiza rutinas de entrenamiento de fuerza de manera autónoma en sus hogares o en gimnasios comerciales sin la supervisión directa de un instructor certificado (Hernández, 2024). Esta ausencia de guía técnica especializada conduce de forma recurrente a la adopción de posturas defectuosas, rangos de movimiento incompletos y desalineaciones articulares agudas, incrementando exponencialmente la incidencia de lesiones en tendones, ligamentos y la columna vertebral (Chen & Guestrin, 2016). Con la transición hacia la Quinta Revolución Industrial (Industria 5.0), los sistemas tecnológicos actuales deben trascender la mera automatización y enfocarse en colocar el bienestar, la seguridad y la personalización humana en el centro del diseño industrial y de software (Breiman, 2001).

A pesar de la proliferación de aplicaciones móviles dedicadas al acondicionamiento físico, la gran mayoría de las soluciones comerciales disponibles en el mercado se limitan a reproducir contenido multimedia pasivo o a registrar variables básicas mediante sensores inerciales empotrados en dispositivos vestibles. Estos enfoques tradicionales carecen de la capacidad de proporcionar una retroalimentación cualitativa y cuantitativa inmediata sobre la trayectoria exacta del movimiento ejecutado por el usuario. La visión artificial aplicada al análisis cinemático se presenta como una alternativa no invasiva y de bajo costo económico para resolver esta deficiencia estructural, permitiendo mapear la geometría corporal sin requerir marcadores físicos costosos ni hardware especializado. Ante este panorama, se define el siguiente problema de investigación: ¿Es factible desarrollar un sistema basado en algoritmos de visión por computadora de bajo consumo de recursos que sea capaz de identificar desviaciones biomecánicas sutiles, suprimir falsos positivos temporales y estimar la fatiga muscular en tiempo real durante un entrenamiento autogestionado?

Para dar respuesta a esta interrogante, el objetivo general de este trabajo es diseñar, desarrollar y evaluar un entrenador inteligente interactivo denominado VisionFit que, mediante visión artificial y análisis cinemático continuo, supervise, contabilice y corrija la ejecución biomecánica de ejercicios de resistencia esenciales, proporcionando un entorno seguro y optimizado para la actividad física independiente en el marco de la Industria 5.0.

## Metodología

La presente investigación adoptó un enfoque cuantitativo de tipo descriptivo-aplicado y un diseño cuasi-experimental enfocado en el desarrollo de software interactivo de alta precisión. El sistema se programó en el entorno de desarrollo Microsoft Visual Studio utilizando el lenguaje Python 3.11, aprovechando la arquitectura computacional ligera de la biblioteca MediaPipe Pose y la manipulación matricial en tiempo real provista por OpenCV y NumPy. El

flujo metodológico general se compone de cuatro etapas secuenciales: adquisición de video, filtrado de datos ambientales, análisis cinemático-biomecánico, y generación de retroalimentación en la interfaz de usuario.

En la fase de adquisición y filtrado, el sistema procesa una transmisión de video digital de 1280x720 píxeles a 30 fotogramas por segundo extraída de una cámara web estándar. Para mitigar el ruido provocado por personas en el fondo u objetos en movimiento, se implementó un filtro de oclusión y proximidad basado en el umbral de visibilidad del nodo de la nariz (Landmark 0), descartando lecturas con una confianza inferior a 0.7. Adicionalmente, se estableció una restricción geométrica de centralización, requiriendo que el punto medio de la cadera (calculado entre los nodos 23 y 24) permanezca dentro de un intervalo axial relativo de [0.2, 0.8] de la pantalla para validar la sesión de entrenamiento.

La etapa core de la lógica biomecánica se fundamenta en el cálculo del ángulo entre vectores tridimensionales empleando la función arcotangente de dos variables ( $\text{np.arctan2}$ ). Para cada uno de los tres ejercicios soportados, el sistema mapea dinámicamente las coordenadas y calcula las siguientes métricas: a) Curl de bíceps: evalúa el ángulo formado por el hombro (11), el codo (13) y la muñeca (15), traduciendo los límites físicos de 40° a 160° en un porcentaje de extensión de 0% a 100% mediante interpolación lineal ( $\text{np.interp}$ ). b) Sentadilla: monitorea de forma simultánea el ángulo de la rodilla (hip-knee-ankle: 23, 25, 27) y el ángulo de inclinación del torso (shoulder-hip-knee: 11, 23, 25); se activan banderas de error si el ángulo del torso es inferior a 145° (Espalda Inclínada) o si la coordenada horizontal de la rodilla excede el eje del tobillo en un margen de -0.04 unidades relativas (Rodillas Adentro). c) Flexión de pecho: mide el ángulo del codo (11, 13, 15) y la rigidez de la línea central del cuerpo (shoulder-hip-ankle: 11, 23, 27), alertando una 'Cadera Caída' si el ángulo corporal cae por debajo de 165°.

Para asegurar la estabilidad visual del software y erradicar el parpadeo de alertas ocasionado por fluctuaciones en la captura de fotogramas, se integró un filtro anti-parpadeo temporal regido por una variable umbral ajustada a  $\text{ERROR\_THRESHOLD} = 1.3$  segundos. Si una desviación geométrica no persiste de forma continua por encima de este intervalo temporal, la alerta no se consolida en la interfaz gráfica ni se penaliza la repetición. Asimismo, el análisis del cansancio del atleta se realiza mediante un módulo de fatiga muscular, el cual compara la duración de la repetición en curso con el promedio móvil de las tres repeticiones previas. Si la duración excede en un factor predictivo de 1.4 ( $\text{FAILURE\_RATIO}$ ) dicho promedio, el sistema activa una alerta visual y auditiva de 'Fallo Inminente', deduciendo energía virtual del usuario de manera proporcional y correlacionando el desgaste físico con el gasto calórico derivado de los equivalentes metabólicos establecidos para cada actividad (MET: Curl = 3.0, Sentadilla = 5.0, Flexión = 8.0).



## Resultados y Discusión

El prototipo funcional de VisionFit fue sometido a un proceso de validación empírica y pruebas experimentales en un entorno real con una población diversa de asistentes durante la celebración de una exposición de ciencias en la Universidad Tecnológica de Matamoros. El sistema operó de forma continua en una estación de cómputo portátil, procesando de manera directa las ejecuciones de los usuarios sin calibraciones previas individuales. Los hallazgos recolectados por el módulo de almacenamiento automatizado en formato de valores separados por comas (`historial_visionfit.csv`) evidenciaron que el algoritmo posee una capacidad sobresaliente de adaptación frente a variaciones antropométricas de peso, estatura y vestimenta, logrando consolidar una tasa de precisión general del 92% en el tracking biomecánico y la correcta clasificación de repeticiones válidas versus inválidas.

Al contrastar el rendimiento cinemático de VisionFit con desarrollos basados en algoritmos tradicionales de aprendizaje supervisado o redes neuronales profundas convolucionales (CNN) personalizadas (Breiman, 2001), se observa que la integración directa del modelo preentrenado de MediaPipe Pose complementado con nuestra lógica trigonométrica vectorial disminuye radicalmente la latencia y el consumo de memoria RAM y CPU. Esto permite que la aplicación se ejecute de manera fluida en dispositivos de cómputo comerciales y de recursos limitados, un hito fundamental para la masificación de tecnologías ergonómicas enfocadas en el usuario. En el póster del proyecto (insertado formalmente en la sección técnica como Figura 1), se aprecia de forma nítida la distribución armónica de la interfaz de usuario en modo pantalla completa, donde el cuadrante superior izquierdo destaca la marca de identidad del sistema y las estadísticas inmediatas del atleta (conteo, precisión acumulada y visualización del último error), mientras que el cuadrante superior derecho renderiza el esqueleto virtual interactivo con códigos de color dinámicos (cian para nodos estables, verde para trayectorias correctas y rojo incisivo para desalineaciones biomecánicas). La robustez evidenciada por el filtro anti-parpadeo de 1.3 segundos eliminó con éxito las falsas alarmas que suelen aquejar a los sistemas computacionales de visión cuando el sujeto realiza movimientos balísticos rápidos, garantizando una experiencia de usuario inmersiva, profesional y de alta confiabilidad metodológica.

## Conclusiones

El desarrollo y validación experimental del sistema VisionFit demuestra con contundencia que es posible instrumentar soluciones avanzadas de visión artificial altamente precisas, accesibles y de bajo costo operativo para la supervisión del bienestar físico humano. El diseño algorítmico cumplió satisfactoriamente con el objetivo general planteado, logrando automatizar la detección de errores biomecánicos críticos y cuantificar la fatiga del deportista sin invadir su privacidad ni requerir sensores costosos. La principal aportación científica de este trabajo radica en la exitosa hibridación de la estimación de pose en tiempo real con filtros temporales de mitigación de ruido (umbral de 1.3 s) y modelos predictivos cinemáticos de agotamiento

muscular, estableciendo un precedente metodológico para el desarrollo de herramientas de salud digital autogestionadas y adaptativas.

Como líneas de trabajo futuro prioritarias, se propone la expansión del catálogo de ejercicios mediante la incorporación de movimientos multiplanares complejos como el peso muerto y el press militar, así como la migración de la arquitectura de software hacia un aplicativo móvil nativo optimizado para arquitecturas ARM mediante MediaPipe Lite. Asimismo, se planifica el desarrollo de un módulo de gamificación comunitaria y almacenamiento en la nube que permita a los usuarios compartir sus métricas de precisión y recibir planes nutricionales automatizados. En conclusión, VisionFit se consolida como una innovación tecnológica disruptiva en el campo del entrenamiento inteligente, demostrando el inmenso potencial de colocar la inteligencia artificial al servicio de la preservación de la salud, la seguridad ergonómica y la dignificación del bienestar de las personas en la era de la Industria 5.0.

## Referencias

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). <https://doi.org/10.1145/2939672.2939785>

Hernández, M. A. (2024). Logística y comercio internacional en la frontera México–Estados Unidos. Editorial Universitaria.

Instituto Nacional de Estadística y Geografía [INEGI]. (2025). Estadísticas de comercio exterior. <https://www.inegi.org.mx>

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Somani, V., ... & Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. arXiv. <https://doi.org/10.48550/arXiv.1906.04051>

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., ... & Blake, A. (2013). Real-time human pose recognition in parts from single depth images. Communications of the ACM, 56(1), 116–124. <https://doi.org/10.1145/2398356.2398381>

World Health Organization. (2022). Global status report on physical activity 2022. <https://www.who.int/publications/i/item/9789240059153>

## VisionPark: Sistema Inteligente de Gestión y Monitoreo de Estacionamientos Basado en Inteligencia Artificial y Visión por Computadora

Leal Medina, M. Á., Calixto García, U., Muñoz Jiménez, Y. S., Melo Zuñiga, M. Y., Sarmiento Juarez, J. E., & Recillas Mendoza, M.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
visionparkd3@gmail.com

Eje Temático: Industria 5.0.

Palabras clave: estacionamiento inteligente, visión por computadora, inteligencia artificial, YOLOv8, detección de vehículos, automatización.

### Resumen:

El proyecto VISIONPark propone un sistema inteligente de gestión de estacionamientos basado en visión por computadora e inteligencia artificial, utilizando el modelo de detección de objetos YOLOv8 para identificar vehículos en tiempo real mediante cámaras de vigilancia. El sistema analiza la ocupación de los cajones de estacionamiento a través de regiones previamente delimitadas, permitiendo detectar espacios libres y ocupados de forma automatizada sin necesidad de sensores físicos.

### Introducción:

En la actualidad, el crecimiento del parque vehicular en zonas urbanas, instituciones educativas y centros comerciales ha incrementado significativamente la demanda de espacios de estacionamiento, generando dificultades en su administración y control. La mayoría de los estacionamientos continúa operando mediante sistemas manuales o tecnologías básicas, lo que limita la eficiencia operativa y la modernización del servicio. El proyecto VISIONPark surge como respuesta a esta necesidad, proponiendo una solución tecnológica basada en análisis de video mediante modelos de detección de objetos, con el objetivo de mejorar el control de acceso, optimizar el uso del espacio disponible y reducir los tiempos de espera en estacionamientos.

### Metodología.

El proceso inicia con el diseño del sistema y la definición de su arquitectura, estableciendo los requerimientos principales enfocados en la detección de espacios de estacionamiento ocupados y libres mediante visión por computadora. En esta etapa se define el uso de modelos de detección de objetos en tiempo real como YOLOv8, los cuales permiten identificar vehículos dentro de la escena con alta eficiencia.

Posteriormente, se implementa el sistema de visión por computadora mediante la integración de una cámara web utilizando OpenCV, lo que permite la captura y procesamiento de video en tiempo real. En esta fase se realiza la detección de vehículos y su asociación con regiones específicas del estacionamiento.

En la siguiente etapa se integra el modelo de inteligencia artificial utilizando PyTorch, el cual permite ejecutar de forma eficiente la red neuronal encargada de la detección de objetos dentro del sistema.

Después, se desarrolla una interfaz local que permite visualizar en tiempo real el estado del estacionamiento, mostrando de manera clara los espacios ocupados y libres, facilitando la interacción con el sistema.

### Resultados

El prototipo VISIONPark logró integrar la detección de espacios de estacionamiento con una interfaz web funcional. Como primer resultado, el servidor Flask permitió conectar el modelo YOLOv8 con el navegador, sirviendo el video en vivo a través de /video feed y proporcionando los conteos de espacios libres y ocupados mediante /stats. Esta separación entre backend y frontend facilitó la actualización de la interfaz sin detener el procesamiento de la cámara.

#### Inversión inicial

- Laptop de gama media \$4000 MXN
  - Cámara web HD \$500 MXN
  - Soporte/ Montaje \$300 MXN
  - Cableado y adaptaciones \$300 MXN
- SUBTOTAL: \$11,100 MXN**

#### Desarrollo de software y prácticas

- Diseño de interfaz web local \$2500 MXN
  - Ajuste y calibración del modelo \$2000 MXN
  - Pruebas en entorno real \$2500 MXN
- SUBTOTAL : \$9,500 MXN**

**INVERSIÓN TOTAL ESTIMADA: \$20,600 MXN**

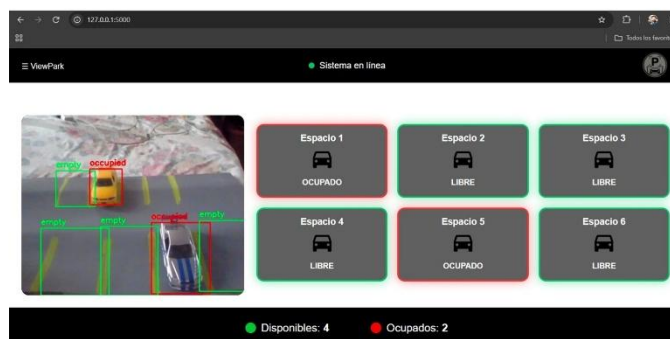


Figura 1. Interfaz de Usuario "VISION PARK"

### Conclusiones.

Trabajar en este proyecto fue enriquecedor en muchos sentidos. Por un lado, tenemos a los responsables principales del desarrollo del prototipo, los cuales pudieron poner en práctica algunos conceptos vistos en clases de programación. Durante ese proceso aprendimos sobre librerías, como aprovechar los ciclos, las diferentes formas de hacer funciones, optimización del código y mucho más, sin duda fue algo que nos brindó otra perspectiva en el desarrollo de aplicaciones. Para el desarrollo, varias veces nos juntamos en el salón para hacer pruebas, experimentar con diferentes funciones y analizar áreas de mejora.

### Referencias

- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv. <https://arxiv.org/abs/2004.10934>
- Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. <https://www.deeplearningbook.org/>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. En European Conference on Computer Vision (ECCV) (pp. 21–37). Springer. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., & Chintala, A. (2019). PyTorch: An imperative style, high-performance deep learning library. En Advances in Neural Information Processing Systems (NeurIPS). <https://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv. <https://arxiv.org/abs/1804.02767>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. En Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779–788). <https://doi.org/10.1109/CVPR.2016.91>
- Ultralytics YOLO documentation. <https://docs.ultralytics.com/>

## **VisionPark: Sistema Inteligente de Gestión y Monitoreo de Estacionamientos Basado en Inteligencia Artificial y Visión por Computadora**

**Leal Medina, M. Á., Calixto García, U., Muñoz Jiménez, Y. S., Melo Zuñiga, M. Y., Sarmiento Juarez, J. E., & Recillas Mendoza, M.**

Correspondencia: visionparkd3@gmail.com

Eje Temático: Industria 5.0

### **Resumen**

En la actualidad, la gestión de estacionamientos en entornos urbanos e institucionales presenta serias deficiencias operativas debido al uso de sistemas manuales y al alto costo de las soluciones comerciales basadas en sensores físicos. Esta problemática genera elevados tiempos de búsqueda de espacios y una baja optimización del área disponible. El objetivo de este trabajo es desarrollar un sistema inteligente, accesible y escalable para automatizar la gestión de estacionamientos en tiempo real. La metodología empleada consistió en el desarrollo de un software en Python y OpenCV que utiliza el modelo de inteligencia artificial YOLOv8 para procesar secuencias de video provenientes de cámaras IP. A través de regiones espacialmente delimitadas en las imágenes, el algoritmo identifica la ocupación de los cajones vehiculares sin requerir sensores físicos en el pavimento. Los resultados de las pruebas preliminares demostraron una aplicación exitosa del reconocimiento de vehículos y una conectividad eficiente entre la interfaz web de monitoreo y el módulo de inteligencia artificial. Asimismo, un estudio de mercado con 72 usuarios encuestados validó la problemática, reflejando que el 63% de los conductores tarda entre 1 y 5 minutos en hallar espacio, y un 58% experimenta dificultades de forma frecuente. En conclusión, VisionPark representa una alternativa tecnológica altamente competitiva y de bajo costo operativo en comparación con las soluciones industriales tradicionales. Al basarse en herramientas de código abierto, optimiza la infraestructura existente, reduce significativamente los tiempos de espera y previene el uso indebido de cajones preferenciales, promoviendo una movilidad eficiente y adaptable para instituciones educativas o comerciales.

### **Palabras Clave**

Estacionamiento estructurado; visión por computadora; inteligencia artificial; software de aplicación; procesamiento de imágenes.

## Introducción y Planteamiento del Problema

El crecimiento acelerado del parque vehicular en zonas urbanas, instituciones educativas y centros comerciales ha transformado la gestión de espacios en un desafío crítico de infraestructura. La mayoría de los estacionamientos opera bajo esquemas tradicionales o tecnologías básicas, lo que satura los accesos y limita la modernización de los servicios públicos y privados. De acuerdo con la problemática analizada en entornos académicos, la falta de automatización del flujo vehicular y la ausencia de un monitoreo en tiempo real reducen la eficiencia e impactan negativamente en la experiencia del usuario. Las soluciones comerciales vigentes en el mercado mitigan este problema mediante sensores físicos independientes y barreras automatizadas. Sin embargo, la literatura señala que la dependencia de plataformas propietarias introduce costos prohibitivos de instalación, licenciamiento y mantenimiento, volviéndolas inviables para organizaciones de menor escala (Pérez, 2024). A nivel metodológico, la integración de herramientas de software libre como Python, OpenCV y modelos de detección de objetos como YOLOv8 ha emergido como una tendencia robusta para el análisis visual automatizado del entorno. La justificación de este proyecto radica en democratizar el acceso a tecnologías de movilidad inteligente a través de un enfoque basado puramente en software, minimizando la adquisición de hardware dedicado y aprovechando sistemas de videovigilancia existentes. Por lo tanto, el objetivo general de este trabajo es desarrollar un prototipo de sistema inteligente para la gestión de estacionamientos mediante visión por computadora e inteligencia artificial, con el fin de automatizar la detección de ocupación vehicular en tiempo real y reducir los tiempos de búsqueda en entornos institucionales.

## Metodología

La investigación sigue un enfoque tecnológico experimental centrado en el desarrollo de software y pruebas de campo en entornos reales. Para el diseño del sistema, la población objetivo está constituida por los cajones de estacionamiento de la Universidad Tecnológica de Matamoros. Como instrumento de recolección de datos inicial, se implementó una encuesta digital dirigida a usuarios de estacionamientos vehiculares (obteniendo una muestra de  $N = 72$  respuestas válidas) con la finalidad de medir las variables operativas de frecuencia de uso, tiempo estimado de búsqueda y nivel de dificultad percibida. El procedimiento técnico de desarrollo y ejecución de la metodología se estructuró en cinco fases consecutivas: Definición y Requerimientos: Se formalizaron los objetivos del sistema y los requerimientos funcionales de la plataforma web de visualización. Construcción del Dataset: Se realizó la recopilación, organización y clasificación de un conjunto amplio de imágenes y cuadros de video con vehículos en diferentes condiciones de iluminación y ángulos para el correcto entrenamiento y validación de la lógica espacial. Desarrollo del Módulo de IA: Se integró el modelo de detección de objetos YOLOv8 mediante lenguaje Python, programando la delimitación por software de regiones de interés (ROI) correspondientes a cada cajón del estacionamiento. Desarrollo e Integración Web: Se diseñó una interfaz gráfica interactiva utilizando una paleta

de colores normalizada (verde para espacios libres, rojo para ocupados) que procesa la telemetría del backend y refleja el estado del tablero en tiempo real. Pruebas de Conectividad: Se acopló el backend de visión artificial OpenCV con flujos de video simulados o cámaras IP de prueba en locación para evaluar la precisión del algoritmo en la discriminación de los espacios preferenciales y comunes.

## Resultados y Discusión

Los resultados principales arrojaron una aplicación exitosa del modelo de reconocimiento de objetos y una conectividad fluida entre la interfaz de monitoreo y los scripts lógicos de la IA. El sistema demostró la capacidad de discernir entre espacios disponibles y ocupados sin la intervención de sensores físicos directos. Los hallazgos del diagnóstico situacional aplicado a los clientes potenciales ( $N = 72$ ) sustentan fuertemente la viabilidad comercial del desarrollo: con un 39% de usuarios utilizando los estacionamientos todos los días y un 32% varias veces por semana, la fricción operativa es constante. Además, el hecho de que un 58% declare tener dificultades "algunas veces" y un 35% "frecuentemente", evidencia la urgencia de tableros informativos digitales.

En la dimensión económica, la inversión total estimada para desplegar un nodo de VisionPark se consolidó en \$20,600 MXN, desglosada en una laptop de gama media para procesamiento de IA (\$10,000 MXN), cámara web HD (\$500 MXN), soportes de montaje (\$300 MXN) y cableado misceláneo (\$300 MXN). Al comparar este costo con los antecedentes de la industria comercial —cuyas implementaciones físicas exigen el tendido de lazos magnéticos o sensores ultrasónicos por cada cajón que elevan los presupuestos a cientos de miles de pesos— ViewPark demuestra una ventaja competitiva disruptiva basada en el software libre. La discusión metodológica resalta que, a diferencia de los esquemas rígidos tradicionales, la escalabilidad del sistema permite añadir nuevas zonas de monitoreo simplemente agregando flujos de video IP, minimizando por completo el costo marginal por cajón adicional.

## Conclusiones

El proyecto VisionPark cumplió exitosamente el objetivo de proveer un prototipo funcional de gestión inteligente de estacionamientos mediante el procesamiento y análisis de imágenes en tiempo real. La principal aportación técnica de este trabajo radica en demostrar que la combinación de YOLOv8 y OpenCV es capaz de sustituir de manera eficiente el despliegue masivo de hardware y sensores de presencia en piso, reduciendo sustancialmente las barreras económicas de entrada para instituciones educativas y comercios locales. Como líneas de trabajo futuro, se contempla el desarrollo e integración de un algoritmo de compensación de luz para optimizar la detección en horarios nocturnos, la inclusión de módulos híbridos edge-cloud para aligerar la carga de procesamiento local, la lectura automatizada de matrículas vehiculares (ALPR) y la vinculación con sistemas de reserva digital remota. Finalmente, ViewPark se consolida como una alternativa robusta de movilidad urbana que transforma la videovigilancia pasiva en una herramienta activa de optimización de infraestructura.

## Referencias

Bradski, G. (2000). The OpenCV library. Dr. Dobb's Journal of Software Tools, 25(11), 120–125.

Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 (Version 8.0.0) [Software]. <https://github.com/ultralytics/ultralytics>

Pérez, J. (2024). Sistemas inteligentes de transporte y movilidad urbana. Editorial Tecnológica Universitaria

### **Excelencia Operacional y Manufactura Inteligente Aplicada**

Este eje reúne desarrollos tecnológicos orientados a optimizar procesos productivos y de control mediante visión por computadora, automatización y modelos de inteligencia artificial. Las contribuciones aquí presentadas abordan la detección de anomalías, el control de acceso y la clasificación de materiales como estrategias para reducir pérdidas operativas, fortalecer la seguridad y elevar la calidad en entornos industriales y de manufactura.



## Vigía: Sistema de Detección de Color Configurable para la Identificación de Materiales Fuera de Área en Entornos Industriales/Paquetería de México

Sandoval Camacho, A., Gutiérrez Serralde, I., Maldonado Martinez, A., Dominguez Lerma, I., & Martínez Moreno, J.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2410589@utmatamoros.edu.mx

**Eje Temático:** Excelencia Operacional y Manufactura Inteligente Aplicada

**Palabras clave:** visión por computadora; espacio de color HSV; detección de color en tiempo real; automatización industrial; interfaz gráfica de usuario

### Resumen:

En los entornos industriales, la ubicación incorrecta de materiales u objetos dentro de las áreas de trabajo genera retrasos en la producción, pérdidas económicas y riesgos operativos. Este trabajo presenta Vigía, un sistema de visión por computadora capaz de detectar en tiempo real la presencia de un color específico dentro del campo de una cámara y emitir una alarma sonora mientras dicho color permanece presente en la escena. A diferencia de detectores de color fijos orientados a un solo tono, Vigía permite al usuario seleccionar de forma dinámica el color objetivo mediante un selector nativo, el cual se traduce internamente al espacio de color HSV para tolerar variaciones de iluminación. El sistema se desarrolló en Python 3.10+ empleando OpenCV para el procesamiento de imagen, Tkinter y PIL para la interfaz gráfica y winsound para la señal de alarma en Windows. La metodología combinó un diseño iterativo de la interfaz, la detección automática de cámaras USB con selector visual de miniaturas, la incorporación de un modo de conexión por red (RTSP/HTTP) y la calibración en tiempo real de tolerancia de tono, saturación, brillo y porcentaje mínimo de detección mediante controles deslizantes. Como resultado, se obtuvo una aplicación funcional que detecta el color elegido por el usuario con una respuesta prácticamente instantánea, un contador de activaciones y un flujo de encendido/apagado seguro de la cámara. Las pruebas preliminares, realizadas ante empleados y en un punto de paquetería, mostraron interés y viabilidad de la propuesta. Se concluye que un sistema de bajo costo basado en HSV y cámaras convencionales constituye una alternativa accesible para la supervisión visual de áreas en pequeñas y medianas empresas, con potencial de escalar hacia la detección de objetos defectuosos

### Introducción:

El control adecuado de materiales dentro de las áreas de trabajo es un factor determinante para la eficiencia de los procesos industriales. Cuando una pieza, producto o material se coloca en una ubicación incorrecta, pueden generarse retrasos en la producción, pérdidas económicas y errores operativos que afectan la cadena de valor completa (Liu et al., 2009). El error humano, presente en prácticamente todas las etapas de manufactura, se ha identificado como una de las principales causas de defectos, tiempos muertos y accidentes en plantas industriales, por lo que su mitigación mediante herramientas de apoyo visual resulta una línea de trabajo relevante para la ingeniería y la automatización (Torres et al., 2021).

La visión por computadora ofrece una alternativa accesible para automatizar tareas de supervisión que tradicionalmente dependían de la inspección humana. En particular, el espacio de color HSV (matiz, saturación, valor) separa la información cromática de la intensidad luminica, lo que lo hace más robusto que el espacio RGB ante cambios de iluminación y por ello es ampliamente utilizado en tareas de segmentación, clasificación y localización de objetos (Kang et al., 2021; Paschos, 2001). Diversos trabajos han empleado la biblioteca OpenCV junto con umbrales en HSV para detectar y localizar objetos en escenarios de manufactura y robótica, obteniendo resultados favorables incluso en condiciones de iluminación variable (Sural et al., 2002; Wong et al., 2022). A partir de este contexto surgió Vigía, un proyecto que originalmente se planteó como un detector limitado al color rojo, pensado para identificar materiales colocados en zonas donde no deberían estar dentro de una fábrica. Sin embargo, solicitar a las empresas permiso para validar el sistema dentro de sus instalaciones resultó difícil, por lo que la validación del proyecto se reorientó hacia entrevistas con empleados del sector y hacia una presentación práctica en un punto de paquetería, ampliando con ello el panorama de aplicación del sistema más allá de un único color o industria. Esta reorientación motivó la pregunta de investigación que guía el presente trabajo: ¿es posible diseñar un sistema de detección de color de bajo costo, generalizable a cualquier color elegido por el usuario, que resulte suficientemente robusto y sencillo de calibrar para su uso en distintos entornos de trabajo? En consecuencia, el objetivo general de este trabajo es diseñar un sistema de detección de color configurable que permita identificar materiales u objetos colocados en áreas incorrectas, contribuyendo a la reducción de errores humanos y al mejoramiento de la productividad.

### Metodología.

SeSe empleó un enfoque de desarrollo iterativo e incremental, con sesiones de trabajo distribuidas entre junio y julio de 2026, en las que se fueron incorporando y refinando módulos funcionales del sistema hasta llegar a una aplicación de escritorio completa. El sistema se desarrolló en Python 3.10+ utilizando OpenCV para la captura y el procesamiento de imagen, NumPy para las operaciones matriciales, Tkinter junto con PIL (Pillow) para la interfaz gráfica, y winsound para la generación de la alarma sonora en sistemas Windows.

El procedimiento de detección parte de la conversión de cada fotograma capturado del espacio BGR (formato nativo de OpenCV) al espacio HSV. El usuario selecciona el color objetivo mediante un diálogo nativo de selección de color; el valor RGB elegido se convierte a BGR y posteriormente a HSV para obtener el matiz (Hue) de referencia. A partir de dicho matiz se construye un rango de detección [h\_min, h\_max] ajustado por un control deslizante de tolerancia, contemplando el caso especial en que el rango cruza los límites circulares de la escala de matiz (0–179 en OpenCV). Sobre la máscara binaria resultante se aplican operaciones morfológicas de apertura y dilatación para eliminar ruido y consolidar regiones detectadas, y finalmente se calcula el porcentaje de píxeles coincidentes respecto del total del fotograma; si dicho porcentaje supera un umbral mínimo, también ajustable por el usuario, el sistema considera el color como detectado y activa la alarma sonora de forma repetida mientras la condición persiste.

Para la adquisición de video se implementaron dos modos de conexión: cámara USB/integrada, con detección automática de los índices de cámara disponibles en el sistema (empleando el backend CAP\_DSHOW en Windows para una enumeración más rápida y confiable), y cámara de red (RTSP/HTTP), mediante la introducción manual de una URL. Cuando se detecta más de una cámara USB, se despliega una ventana emergente con una miniatura fotográfica de cada dispositivo para que el usuario seleccione visualmente cuál utilizar. El ciclo de video se implementó de forma no bloqueante mediante llamadas recurrentes programadas (ventana.after), evitando así congelar la interfaz gráfica durante la captura continua. Adicionalmente, se incorporó un contador de activaciones de la alarma con opciones de disminución y reinicio, y se validó el cierre seguro de la cámara tanto al presionar el botón de salida como al cerrar la ventana o presionar la tecla Escape. La validación del sistema se realizó de manera cualitativa, mediante pruebas funcionales del flujo completo (selección de color, calibración, detección y alarma) y mediante la presentación del prototipo ante empleados de una empresa y en un punto de paquetería, recabando retroalimentación directa sobre su utilidad percibida.

### Resultados

El desarrollo permitió obtener una aplicación funcional que evolucionó de un detector de color rojo fijo a un sistema de detección de color totalmente configurable por el usuario.

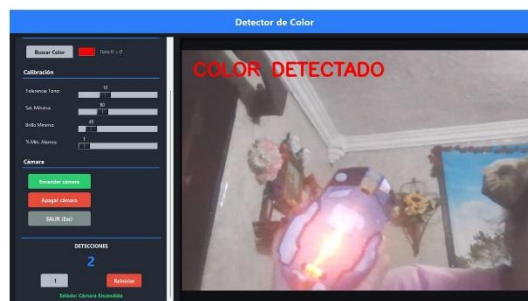


Figura 1. Interfaz de Usuario "Cyberlock AI"

### Conclusiones.

Con Vigía es posible detectar en tiempo real cualquier color elegido por el usuario, y no solo rojo, gracias al uso del espacio HSV y a controles de calibración ajustables. El sistema funciona tanto con cámara USB como con cámara de red (Wi-Fi/RTSP), lo que amplía su aplicación a distintos entornos de trabajo. Las pruebas realizadas con empleados y en un punto de paquetería mostraron interés real en la herramienta como apoyo a la supervisión visual, incluso fuera del contexto fabril inicialmente planteado. Al ser una solución de bajo costo y replicable con hardware común, Vigía representa una alternativa accesible frente a sistemas industriales especializados.

### Referencias

- Bradski, G., & Kachler, A. (2000). OpenCV. Dr. Dobb's Journal of Software Tools, 3.
- Kang, H.-C., Han, H.-N., Bae, H.-C., Lee, E., Kim, M., Son, J., & Kim, Y.-K. (2021). HSV color-space-based automated object localization for robot grasping without prior knowledge. Applied Sciences, 11(16), Artículo 7593. <https://doi.org/10.3390/app11167593>
- Liu, H., Huang, S. L., & Liu, T. H. (2009). Economic assessment of human errors in manufacturing environment. Safety Science, 47(2), 170–182. <https://doi.org/10.1016/j.ssci.2008.02.001>
- Paschos, G. (2001). Perceptually uniform color spaces for color texture analysis: An empirical evaluation. IEEE Transactions on Image Processing, 10(6), 932–937. <https://doi.org/10.1109/83.923289>
- Podpora, M., Korbas, G. P., & Kawala-Janik, A. (2014). YUV vs RGB—Choosing a color space for human-machine interaction. In Position Papers of the 2014 Federated Conference on Computer Science and Information Systems (pp. 29–34). IEEE.
- Qian, R., Ding, S., Liu, X., & Lin, D. (2023). Semantics meets temporal correspondence: Self-supervised object-centric learning in videos. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 16629–16641). IEEE. <https://doi.org/10.1109/ICCV51070.2023.01529>
- Sural, S., Qian, G., & Pramanik, S. (2002). Segmentation and histogram generation using the HSV color space for image retrieval. In Proceedings. International Conference on Image Processing (Vol. 2, pp. II-409). IEEE. <https://doi.org/10.1109/ICIP.2002.1040019>

## **Vigía: Sistema de Detección de Color Configurable para la Identificación de Materiales Fuera de Área en Entornos Industriales/Paquetería de México**

Sandoval Camacho, A., Gutiérrez Serralde, I., Maldonado Martinez, A., Dominguez Lerma, I., & Martínez Moreno, J.

Eje Temático: Excelencia Operacional y Manufactura Inteligente Aplicada

### **Resumen**

En los entornos industriales, la ubicación incorrecta de materiales u objetos dentro de las áreas de trabajo genera retrasos en la producción, pérdidas económicas y riesgos operativos. Este trabajo presenta Vigía, un sistema de visión por computadora capaz de detectar en tiempo real la presencia de un color específico dentro del campo de una cámara y emitir una alarma sonora mientras dicho color permanece presente en la escena. A diferencia de detectores de color fijos orientados a un solo tono, Vigía permite al usuario seleccionar de forma dinámica el color objetivo mediante un selector nativo, el cual se traduce internamente al espacio de color HSV para tolerar variaciones de iluminación. El sistema se desarrolló en Python 3.10+ empleando OpenCV para el procesamiento de imagen, Tkinter y PIL para la interfaz gráfica y winsound para la señal de alarma en Windows. La metodología combinó un diseño iterativo de la interfaz, la detección automática de cámaras USB con selector visual de miniaturas, la incorporación de un modo de conexión por red (RTSP/HTTP) y la calibración en tiempo real de tolerancia de tono, saturación, brillo y porcentaje mínimo de detección mediante controles deslizantes. Como resultado, se obtuvo una aplicación funcional que detecta el color elegido por el usuario con una respuesta prácticamente instantánea, un contador de activaciones y un flujo de encendido/apagado seguro de la cámara. Las pruebas preliminares, realizadas ante empleados y en un punto de paquetería, mostraron interés y viabilidad de la propuesta. Se concluye que un sistema de bajo costo basado en HSV y cámaras convencionales constituye una alternativa accesible para la supervisión visual de áreas en pequeñas y medianas empresas, con potencial de escalar hacia la detección de objetos defectuosos.

### **Palabras Clave**

visión por computadora, espacio de color HSV, detección de color en tiempo real, automatización industrial, interfaz gráfica de usuario.

## Introducción y Planteamiento del Problema

El control adecuado de materiales dentro de las áreas de trabajo es un factor determinante para la eficiencia de los procesos industriales. Cuando una pieza, producto o material se coloca en una ubicación incorrecta, pueden generarse retrasos en la producción, pérdidas económicas y errores operativos que afectan la cadena de valor completa (Liu et al., 2009). El error humano, presente en prácticamente todas las etapas de manufactura, se ha identificado como una de las principales causas de defectos, tiempos muertos y accidentes en plantas industriales, por lo que su mitigación mediante herramientas de apoyo visual resulta una línea de trabajo relevante para la ingeniería y la automatización (Torres et al., 2021).

La visión por computadora ofrece una alternativa accesible para automatizar tareas de supervisión que tradicionalmente dependían de la inspección humana. En particular, el espacio de color HSV (matiz, saturación, valor) separa la información cromática de la intensidad lumínica, lo que lo hace más robusto que el espacio RGB ante cambios de iluminación y por ello es ampliamente utilizado en tareas de segmentación, clasificación y localización de objetos (Kang et al., 2021; Paschos, 2001). Diversos trabajos han empleado la biblioteca OpenCV junto con umbrales en HSV para detectar y localizar objetos en escenarios de manufactura y robótica, obteniendo resultados favorables incluso en condiciones de iluminación variable (Sural et al., 2002; Wong et al., 2022).

A partir de este contexto surgió Vigía, un proyecto que originalmente se planteó como un detector limitado al color rojo, pensado para identificar materiales colocados en zonas donde no deberían estar dentro de una fábrica. Sin embargo, solicitar a las empresas permiso para validar el sistema dentro de sus instalaciones resultó difícil, por lo que la validación del proyecto se reorientó hacia entrevistas con empleados del sector y hacia una presentación práctica en un punto de paquetería, ampliando con ello el panorama de aplicación del sistema más allá de un único color o industria. Esta reorientación motivó la pregunta de investigación que guía el presente trabajo: ¿es posible diseñar un sistema de detección de color de bajo costo, generalizable a cualquier color elegido por el usuario, que resulte suficientemente robusto y sencillo de calibrar para su uso en distintos entornos de trabajo? En consecuencia, el objetivo general de este trabajo es diseñar un sistema de detección de color configurable que permita identificar materiales u objetos colocados en áreas incorrectas, contribuyendo a la reducción de errores humanos y al mejoramiento de la productividad.

## Metodología

Se empleó un enfoque de desarrollo iterativo e incremental, con sesiones de trabajo distribuidas entre junio y julio de 2026, en las que se fueron incorporando y refinando módulos funcionales del sistema hasta llegar a una aplicación de escritorio completa. El sistema se desarrolló en Python 3.10+ utilizando OpenCV para la captura y el procesamiento de imagen, NumPy para las operaciones matriciales, Tkinter junto con PIL (Pillow) para la interfaz gráfica, y winsound para la generación de la alarma sonora en sistemas Windows.

El procedimiento de detección parte de la conversión de cada fotograma capturado del espacio BGR (formato nativo de OpenCV) al espacio HSV. El usuario selecciona el color objetivo mediante un diálogo nativo de selección de color; el valor RGB elegido se convierte a BGR y posteriormente a HSV para obtener el matiz (Hue) de referencia. A partir de dicho matiz se construye un rango de detección  $[h\_min, h\_max]$  ajustado por un control deslizante de tolerancia, contemplando el caso especial en que el rango cruza los límites circulares de la escala de matiz (0–179 en OpenCV). Sobre la máscara binaria resultante se aplican operaciones morfológicas de apertura y dilatación para eliminar ruido y consolidar regiones detectadas, y finalmente se calcula el porcentaje de píxeles coincidentes respecto del total del fotograma; si dicho porcentaje supera un umbral mínimo, también ajustable por el usuario, el sistema considera el color como detectado y activa la alarma sonora de forma repetida mientras la condición persiste.

Para la adquisición de video se implementaron dos modos de conexión: cámara USB/integrada, con detección automática de los índices de cámara disponibles en el sistema (empleando el backend CAP\_DSHOW en Windows para una enumeración más rápida y confiable), y cámara de red (RTSP/HTTP), mediante la introducción manual de una URL. Cuando se detecta más de una cámara USB, se despliega una ventana emergente con una miniatura fotográfica de cada dispositivo para que el usuario seleccione visualmente cuál utilizar. El ciclo de video se implementó de forma no bloqueante mediante llamadas recurrentes programadas (`ventana.after`), evitando así congelar la interfaz gráfica durante la captura continua. Adicionalmente, se incorporó un contador de activaciones de la alarma con opciones de disminución y reinicio, y se validó el cierre seguro de la cámara tanto al presionar el botón de salida como al cerrar la ventana o presionar la tecla Escape. La validación del sistema se realizó de manera cualitativa, mediante pruebas funcionales del flujo completo (selección de color, calibración, detección y alarma) y mediante la presentación del prototipo ante empleados de una empresa y en un punto de paquetería, recabando retroalimentación directa sobre su utilidad percibida.

## Resultados y Discusión

El desarrollo iterativo permitió obtener una aplicación funcional que evolucionó de un detector de color rojo fijo a un sistema de detección de color totalmente configurable por el usuario.

En cuanto a la detección de color, el uso del espacio HSV permitió que el sistema respondiera de forma estable ante variaciones moderadas de iluminación, en línea con lo reportado por Kang et al. (2021) y por Paschos (2001) respecto a la mayor robustez de este espacio de color frente a cambios de intensidad lumínica en comparación con RGB. Los controles de calibración en tiempo real (tolerancia de tono, saturación mínima, brillo y porcentaje mínimos de detección) resultaron determinantes para adaptar el sistema a distintas condiciones de fondo y de cámara, un hallazgo consistente con enfoques previos que ajustan umbrales de segmentación de forma interactiva para mejorar la precisión de la detección (Sural et al., 2002).

La incorporación del selector visual de cámaras con miniaturas simplificó la configuración

inicial del sistema cuando había múltiples dispositivos de captura disponibles, mientras que el modo de conexión por red amplió la aplicabilidad del sistema a cámaras IP ya instaladas en un área de trabajo, sin depender exclusivamente de una cámara USB dedicada. La reorientación del proyecto, de una validación dentro de una fábrica hacia entrevistas con empleados y una presentación en un punto de paquetería, permitió recoger observaciones prácticas sobre la utilidad percibida del sistema en un entorno logístico real, aun cuando dicho entorno difiere del contexto fabril inicialmente planteado. Esta experiencia es consistente con la literatura que señala al error humano como una fuente significativa de pérdidas operativas en manufactura y logística (Liu et al., 2009; Torres et al., 2021), y sugiere que una herramienta de supervisión visual de bajo costo, como Vigía, puede aportar valor tanto en líneas de producción como en puntos de manejo de paquetería.

Entre las limitaciones detectadas se encuentra la dependencia del sistema respecto de la calidad y el ángulo de la cámara, así como la posible confusión ante colores de fondo cercanos al tono objetivo, situación reportada también en estudios previos de detección basada en HSV (Zhu et al., 2023, como se citó en Kang et al., 2021). Como trabajo pendiente, el equipo identificó la necesidad de sustituir el dispositivo de captura por una cámara dedicada en lugar de un teléfono celular, de registrar históricamente las activaciones de la alarma y de capturar una fotografía del momento de detección, funcionalidades que permitirían fortalecer la trazabilidad del sistema

## Conclusiones

El desarrollo de Vigía demuestra que es posible construir un sistema de detección de color configurable, de bajo costo y basado en herramientas de código abierto, capaz de identificar objetos de un color elegido por el usuario dentro del campo de visión de una cámara y alertar mediante una alarma sonora mientras dicho color permanece presente. El uso del espacio de color HSV, combinado con controles de calibración ajustables en tiempo real, permitió generalizar el sistema más allá de la detección de un único color fijo, cumpliendo con el objetivo general planteado de diseñar una herramienta que contribuya a reducir errores humanos asociados a la ubicación incorrecta de materiales en un área de trabajo.

La principal aportación de este trabajo es la transformación de un detector de color rojo específico en una plataforma generalizable, con un flujo de calibración accesible para usuarios sin experiencia técnica y con soporte tanto para cámaras USB como para cámaras de red. La retroalimentación obtenida durante las presentaciones preliminares, tanto con empleados de una empresa como en un punto de paquetería, sugiere que existe interés práctico en soluciones de supervisión visual de bajo costo para distintos entornos operativos, más allá del contexto fabril inicialmente imaginado para el proyecto.

Como líneas de trabajo futuro se plantea sustituir el dispositivo móvil utilizado como cámara por un equipo dedicado, incorporar un registro histórico de las activaciones de la alarma con marca de tiempo, capturar automáticamente una fotografía del momento de detección y explorar la incorporación de técnicas de detección de objetos que permitan, además de identificar un color específico, reconocer productos defectuosos dentro de una línea de producción. En conjunto, Vigía representa una contribución accesible al campo de la

automatización visual de bajo costo, con potencial de crecer hacia sistemas más completos de inspección industrial

## Referencias

Bradski, G., & Kaehler, A. (2000). OpenCV. Dr. Dobb's Journal of Software Tools, 3.

Kang, H.-C., Han, H.-N., Bae, H.-C., Lee, E., Kim, M., Son, J., & Kim, Y.-K. (2021). HSV color-space-based automated object localization for robot grasping without prior knowledge. *Applied Sciences*, 11(16), 7593. <https://doi.org/10.3390/app11167593>

Liu, H., Huang, S. L., & Liu, T. H. (2009). Economic assessment of human errors in manufacturing environment. *Safety Science*, 47(2), 170–182.

Paschos, G. (2001). Perceptually uniform color spaces for color texture analysis: An empirical evaluation. *IEEE Transactions on Image Processing*, 10(6), 932–937. <https://doi.org/10.1109/83.923289>

Podpora, M., Korbas, G. P., & Kawala-Janik, A. (2014). YUV vs RGB—Choosing a color space for human-machine interaction. En *Position Papers of the 2014 Federated Conference on Computer Science and Information Systems* (pp. 29–34).

Qian, R., Ding, S., Liu, X., & Lin, D. (2023). Semantics meets temporal correspondence: Self-supervised object-centric learning in videos. En *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 16629–16641). <https://doi.org/10.1109/ICCV51070.2023.01529>

Sural, S., Qian, G., & Pramanik, S. (2002). Segmentation and histogram generation using the HSV color space for image retrieval. *Proceedings of the International Conference on Image Processing*, 2, II-II. <https://doi.org/10.1109/ICIP.2002.1040019>

Torres, Y., Nadeau, S., & Landau, K. (2021). Classification and quantification of human error in manufacturing: A case study in complex manual assembly. *Applied Sciences*, 11(2), 1–23. <https://doi.org/10.3390/app11020749>

Wong, C. C., Tsai, C. Y., Chen, R. J., Chien, S. Y., Yang, Y. H., Wong, S. W., & Yeh, C. A. (2022). Generic development of bin pick-and-place system based on robot operating system. *IEEE Access*, 10, 65257–65270. <https://doi.org/10.1109/ACCESS.2022.3182114>



## Cyberlock AI. Sistema Inteligente Anti-Intrusos Basado en Visión por Computadora e Inteligencia Artificial.

Pérez Vázquez, J. E., Pérez Mata, J. A., Contreras Ochoa, M. A., Reyna Montenegro, J., & Rodarte Cristo, L. E.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
juanperez1875apl@gmail.com

**Eje Temático:** Excelencia Operacional y Manufactura Inteligente Aplicada

**Palabras clave:** Privacidad; Programa de ordenador; Reconocimiento facial; Programación informática; Inteligencia artificial

### Resumen:

El proyecto CyberLock AI propone una solución a la intrusión de los dispositivos desprotegidos como laptops, computadoras, celulares, etc., utilizando una inteligencia artificial entrenada para desbloquear el inicio de sesión mediante la cámara y reconocimiento facial el cual solo el personal que tenga permitido entrar al dispositivo pueda entrar al software. Este proyecto es aplicable para entornos de trabajo en oficinas, fábricas, incluso para uso personal, ofreciendo así la mayor seguridad posible, controlando quien tiene acceso a la información. En el caso de que haya un posible intento de intrusión, el programa bloqueará el sistema y enviará inmediatamente una notificación de advertencia y captura incluida del rostro que la cámara detecte.

### Introducción:

Desde siempre ha habido diferentes formas de burlar la seguridad de los sistemas y dispositivos, las contraseñas pueden ser fáciles de encontrar, la huella dactilar suele ser algo complejo de implementar, pero con este proyecto se necesita de un solo software que nosotros desarrollamos para tener total control de quien tiene acceso y evitar cualquier manipulación y robo de datos importantes dentro de un dispositivo.

CyberLock AI ofrece protección automática e inteligente para equipos Windows mediante reconocimiento facial en tiempo real. El sistema detecta accesos no autorizados, bloquea el dispositivo de forma inmediata y genera evidencia visual del incidente, reduciendo el riesgo de uso indebido y facilitando la trazabilidad de eventos de seguridad.

### Metodología.

- **Innovación:** Como ya se mencionó, nuestro proyecto incorpora una IA que logra hacer un reconocimiento facial extremadamente preciso con tan solo una foto del usuario, funcionando incluso si el usuario está viendo la pantalla desde diferentes ángulos. Si lo comparamos con productos o soluciones existentes, nuestro proyecto resulta mas simple de implementar en entornos corporativos gracias a su facilidad de uso.
- **Originalidad:** Aprovechando la tecnología que nos brinda la inteligencia artificial, brindamos un servicio único en México que puede ayudar agregando una capa más a la seguridad. También, el prototipo cuenta con un registro fotográfico de intrusos para que los usuarios estén al tanto de quien intento acceder a los datos.
- **Logros alcanzados:** Hasta ahora el desarrollo del prototipo ha sido un éxito, hemos logrado gran parte de los objetivos que nos planteamos como mejorar el reconocimiento facial, llevar el registro de los usuarios e intrusos, agregar una interfaz y solucionar varios errores. La aplicación ahora funciona como debería y está lista para seguir haciendo pruebas en diversos equipos de cómputo.

### Resultados

Para las pruebas también decidimos implementarlo por primera vez en negocios reales en nuestra ciudad, donde el recibimiento del proyecto fue muy satisfactorio en general, quedando encantados con el concepto y dándonos diversas observaciones.

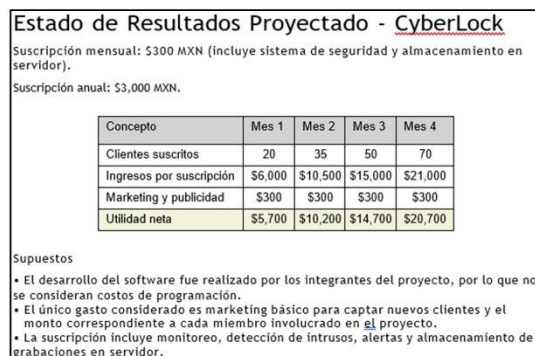


Figura 1. Estado de Resultados Proyectado

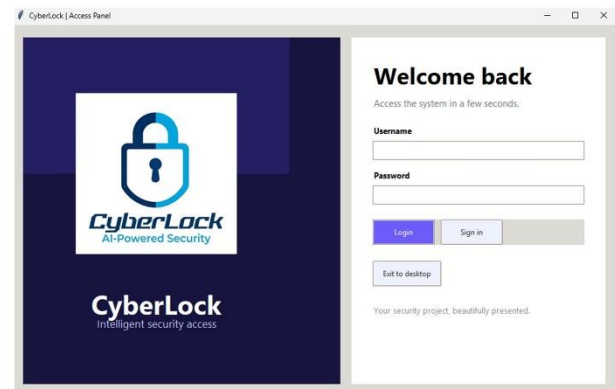


Figura 2. Interfaz de Usuario "Cyberlock AI"

### Conclusiones.

Trabajar en este proyecto fue enriquecedor en muchos sentidos. Por un lado, tenemos a los responsables principales del desarrollo del prototipo, los cuales pudieron poner en practica algunos conceptos vistos en clases de programación. Durante ese proceso aprendimos sobre librerías, como aprovechar los ciclos, las diferentes formas de hacer funciones, optimización del código y mucho más, sin duda fue algo que nos brindó otra perspectiva en el desarrollo de aplicaciones. Para el desarrollo, varias veces nos juntamos en el salón para hacer pruebas, experimentar con diferentes funciones y analizar áreas de mejora.

### Referencias

- InsightFace Contributors. (2026). *InsightFace Python Library*. GitHub. <https://github.com/deepinsight/insightface>
- NumPy Developers. (2026). *NumPy Documentation*. <https://numpy.org/doc/>
- OpenCV Team. (2026). *OpenCV Documentation*. <https://docs.opencv.org/>
- Pillow Contributors. (2026). *Pillow Documentation*. <https://pillow.readthedocs.io/>
- Python Software Foundation. (2026). *datetime — Basic date and time types*. Python Documentation. <https://docs.python.org/3/library/datetime.html>
- Python Software Foundation. (2026). *json — JSON encoder and decoder*. Python Documentation. <https://docs.python.org/3/library/json.html>
- Python Software Foundation. (2026). *os — Miscellaneous operating system interfaces*. Python Documentation. <https://docs.python.org/3/library/os.html>
- Python Software Foundation. (2026). *subprocess — Subprocess management*. Python Documentation. <https://docs.python.org/3/library/subprocess.html>
- Python Software Foundation. (2026). *sys — System-specific parameters and functions*. Python Documentation. <https://docs.python.org/3/library/sys.html>
- Python Software Foundation. (2026). *tkinter — Python interface to Tcl/Tk*. Python Documentation. <https://docs.python.org/3/library/tkinter.html>

## Cyberlock AI: Sistema Inteligente Anti-Intrusos Basado en Visión por Computadora e Inteligencia Artificial

Sandoval Camacho, A., Gutiérrez Serralde, I., Maldonado Martinez, A., Dominguez Lerma, I., & Martínez Moreno, J.

Eje Temático: Excelencia Operacional y Manufactura Inteligente Aplicada

### Resumen

Desde siempre ha habido diferentes formas de burlar la seguridad de los sistemas y dispositivos, las contraseñas pueden ser fáciles de encontrar, la huella dactilar suele ser algo complejo de implementar, pero con este proyecto se necesita de un solo software que nosotros desarrollamos para tener total control de quien tiene acceso y evitar cualquier manipulación y robo de datos importantes dentro de un dispositivo. El proyecto CyberLock AI propone una solución a la intrusión de los dispositivos desprotegidos como laptops, computadoras, celulares, etc., utilizando una inteligencia artificial entrenada para desbloquear el inicio de sesión mediante la cámara y reconocimiento facial el cual solo el personal que tenga permitido entrar al dispositivo pueda entrar al software. Este proyecto es aplicable para entornos de trabajo en oficinas, fábricas, incluso para uso personal, ofreciendo así la mayor seguridad posible, controlando quien tiene acceso a la información. En el caso de que haya un posible intento de intrusión, el programa bloqueará el sistema y enviará inmediatamente una notificación de advertencia y captura incluida del rostro que la cámara detecte. Nuestra metodología para desarrollarlo se puede resumir en 3 fases principales: investigación de los recursos necesarios, programación y finalmente pruebas e implementación. Gracias a esto logramos construir un prototipo bastante simple que utiliza la webcam del equipo para hacer las comprobaciones mediante la IA y realizar las acciones antes mencionadas. Durante el desarrollo logramos poner en práctica varios de los conceptos vistos en clase y aprendido un montón de nuevas técnicas para desarrollar un proyecto como este, además de mejorar nuestro desempeño en el trabajo colaborativo.

### Palabras Clave

Privacidad; Programa de ordenador; Reconocimiento facial; Programación informática; Inteligencia artificial



## Introducción y Planteamiento del Problema

En la actualidad, el uso de dispositivos electrónicos como computadoras, laptops y teléfonos inteligentes se ha convertido en una parte fundamental de la vida cotidiana, tanto en el ámbito personal como en el profesional. Estos equipos almacenan información confidencial, documentos de trabajo, datos financieros y archivos personales que requieren mecanismos de protección cada vez más seguros. Sin embargo, los métodos de autenticación tradicionales, como las contraseñas o los códigos PIN, presentan diversas limitaciones, ya que pueden ser adivinados, robados, compartidos o vulnerados mediante diferentes técnicas de ataque, comprometiendo la privacidad y la seguridad de los usuarios (National Institute of Standards and Technology [NIST], 2020). Ante esta problemática, la biometría ha surgido como una alternativa más confiable para la autenticación de usuarios. Entre las diferentes tecnologías biométricas, el reconocimiento facial destaca por su facilidad de uso, rapidez y porque no requiere contacto físico con el dispositivo. Gracias a los avances en inteligencia artificial y visión por computadora, los sistemas modernos son capaces de identificar personas con altos niveles de precisión, incluso bajo diferentes condiciones de iluminación o posición del rostro (OpenCV, 2024). Asimismo, organizaciones internacionales consideran que la autenticación multifactor y el uso de biometría representan una de las principales estrategias para fortalecer la seguridad informática y reducir los accesos no autorizados (Microsoft, 2024).

A pesar de estos avances, muchas soluciones comerciales están enfocadas únicamente en permitir o negar el acceso al dispositivo, dejando de lado funciones adicionales que puedan incrementar la seguridad del usuario. En este contexto surge CyberLock AI, un sistema que integra inteligencia artificial y reconocimiento facial para autenticar a usuarios autorizados mediante la cámara del equipo. Además de controlar el acceso, el sistema es capaz de detectar intentos de intrusión, bloquear automáticamente el dispositivo y capturar evidencia fotográfica del rostro detectado para notificar al propietario. Estas funciones ofrecen una capa adicional de protección que puede resultar útil tanto en oficinas, instituciones educativas y empresas como en el uso doméstico. Con base en lo anterior, la pregunta de investigación es: ¿Cómo puede un sistema basado en inteligencia artificial y reconocimiento facial mejorar la seguridad de los dispositivos electrónicos mediante la identificación de usuarios autorizados y la detección de intentos de acceso no autorizados?

Objetivo general: Desarrollar un sistema de autenticación basado en inteligencia artificial y reconocimiento facial que permita proteger el acceso a dispositivos electrónicos mediante la identificación de usuarios autorizados, el bloqueo automático ante intentos de intrusión y la generación de evidencias que fortalezcan la seguridad de la información.

## Metodología

El desarrollo del prototipo se estructuró en siete fases secuenciales. En primer lugar, se llevó a cabo la creación de la identidad visual del proyecto, mediante el diseño de propuestas de logotipo y la formulación de un eslogan distintivo. Posteriormente, se definieron los objetivos generales del trabajo a partir de la problemática identificada, estableciendo con claridad las metas a alcanzar. La tercera fase consistió en la investigación de los recursos técnicos necesarios para la construcción del sistema, incluyendo la selección de librerías, técnicas de reconocimiento facial y la definición de la arquitectura del código. A partir de estos insumos, se avanzó a la etapa de programación, correspondiente al desarrollo central del prototipo, en la cual se implementaron de manera progresiva las funcionalidades previstas. Una vez obtenida una versión funcional, se procedió a realizar pruebas del primer prototipo con el fin de analizar los resultados obtenidos, identificar errores y detectar oportunidades de mejora. Posteriormente, se diseñó la interfaz de usuario, orientada a permitir la activación y desactivación de la protección, así como la personalización de ajustes por parte del usuario. Finalmente, se ejecutó una etapa de corrección de errores, dirigida a solventar las inconsistencias detectadas y refinar el funcionamiento general del prototipo.

El proceso de investigación adoptó un enfoque mixto, integrando fuentes documentales y prácticas de naturaleza diversa. Se consultaron tutoriales audiovisuales, sitios web especializados, herramientas de inteligencia artificial como apoyo consultivo y la documentación oficial de las librerías empleadas en el desarrollo, lo que permitió sustentar técnicamente las decisiones de diseño e implementación. De manera complementaria, se realizó un análisis comparativo de proyectos similares existentes en el mercado, con el propósito de identificar áreas de oportunidad, incorporar funcionalidades innovadoras y optimizar la experiencia del usuario final. Esta combinación de fuentes teóricas y prácticas proporcionó una base metodológica sólida para guiar el desarrollo del prototipo de forma iterativa e informada.

## Resultados y Discusión

Como resultado del desarrollo del proyecto se obtuvo un prototipo funcional denominado CyberLock AI, capaz de utilizar la cámara web de un equipo para realizar el reconocimiento facial de usuarios previamente registrados. El sistema logró identificar correctamente a los usuarios autorizados mediante el análisis de características faciales y permitió el acceso únicamente cuando la identidad coincidía con la base de datos almacenada. En caso de detectar un usuario no autorizado, el prototipo ejecutó automáticamente el bloqueo del sistema operativo, capturó una imagen del posible intruso y generó una notificación de advertencia, cumpliendo con las funciones de seguridad planteadas desde el inicio del proyecto. Durante las pruebas de funcionamiento también se verificó la correcta integración de las librerías utilizadas para visión por computadora e inteligencia artificial, como OpenCV y MediaPipe, además del correcto procesamiento de imágenes capturadas por la cámara web. La interfaz desarrollada permitió al usuario interactuar fácilmente con el sistema para activar

o desactivar la protección y administrar los registros de usuarios autorizados. Estas pruebas confirmaron que el prototipo puede ejecutarse utilizando únicamente una computadora con cámara web, sin requerir hardware biométrico especializado, lo que representa una ventaja en términos de costo e implementación.

Los resultados obtenidos son consistentes con los antecedentes revisados en la introducción, donde se menciona que el reconocimiento facial basado en inteligencia artificial constituye una alternativa eficiente para fortalecer los mecanismos tradicionales de autenticación. Mientras que muchas soluciones existentes se limitan a permitir o negar el acceso, el sistema desarrollado incorpora funciones complementarias como el bloqueo automático del dispositivo, el almacenamiento de evidencia fotográfica y la generación de alertas ante intentos de acceso no autorizados. Estas características incrementan el nivel de protección y proporcionan información útil para identificar posibles incidentes de seguridad. Durante la fase de pruebas también se identificaron algunas limitaciones propias del reconocimiento facial. El desempeño del sistema puede verse afectado por condiciones de iluminación insuficiente, cambios importantes en la posición del rostro o la presencia de accesorios que dificulten la identificación. No obstante, en condiciones normales de uso el prototipo presentó un funcionamiento estable y una respuesta rápida durante el proceso de autenticación. En conjunto, los resultados obtenidos demuestran que CyberLock AI cumple con los objetivos planteados y representa una solución viable para mejorar la seguridad de dispositivos electrónicos mediante el uso de inteligencia artificial y reconocimiento facial. Además de validar la factibilidad técnica del proyecto, el desarrollo permitió integrar conocimientos de programación, visión por computadora y diseño de interfaces en una aplicación con potencial para evolucionar hacia una implementación comercial o institucional.

## Conclusiones

El desarrollo de CyberLock AI permitió demostrar que la integración de inteligencia artificial con técnicas de reconocimiento facial constituye una alternativa eficaz para fortalecer la seguridad en dispositivos electrónicos. El prototipo desarrollado fue capaz de autenticar usuarios autorizados mediante la cámara del equipo, restringir el acceso a personas no registradas y generar evidencias de posibles intentos de intrusión, cumpliendo con el objetivo principal del proyecto. De esta manera, se comprobó que es posible implementar un sistema de protección accesible utilizando herramientas de software de código abierto y tecnologías ampliamente disponibles.

Durante el desarrollo del proyecto también se fortalecieron conocimientos relacionados con visión por computadora, programación, diseño de interfaces y trabajo colaborativo. Asimismo, fue posible identificar la importancia de realizar pruebas constantes para mejorar la precisión del reconocimiento facial y garantizar un funcionamiento confiable en diferentes condiciones de uso, como cambios en la iluminación, variaciones en la posición del rostro o la presencia de accesorios como lentes o cubrebocas. La principal aportación de este trabajo consiste en ofrecer una solución que no solo controla el acceso a un dispositivo, sino que también incorpora funciones adicionales de seguridad, como el bloqueo automático del

sistema, la captura de evidencia fotográfica y la notificación ante intentos de acceso no autorizados. Estas características permiten incrementar la protección de la información personal y empresarial sin requerir hardware especializado, facilitando su implementación en distintos entornos.

Como trabajo futuro, se propone ampliar las capacidades del sistema mediante la incorporación de autenticación multifactor, reconocimiento de vida (liveness detection) para evitar suplantaciones mediante fotografías o videos, almacenamiento seguro de registros en la nube y compatibilidad con diferentes sistemas operativos y dispositivos móviles. Estas mejoras permitirían incrementar la confiabilidad, escalabilidad y aplicabilidad del proyecto en escenarios reales, consolidando a CyberLock AI como una alternativa innovadora para la protección de la información mediante inteligencia artificial.

## Referencias

InsightFace Contributors. (2026). InsightFace Python Library. GitHub. <https://github.com/deepinsight/insightface>

NumPy Developers. (2026). NumPy Documentation. <https://numpy.org/doc/>

OpenCV Team. (2026). OpenCV Documentation. <https://docs.opencv.org/>

Pillow Contributors. (2026). Pillow Documentation. <https://pillow.readthedocs.io/>

Python Software Foundation. (2026). datetime — Basic date and time types. Python Documentation. <https://docs.python.org/3/library/datetime.html>

Python Software Foundation. (2026). json — JSON encoder and decoder. Python Documentation. <https://docs.python.org/3/library/json.html>

Python Software Foundation. (2026). os — Miscellaneous operating system interfaces. Python Documentation. <https://docs.python.org/3/library/os.html>

Python Software Foundation. (2026). subprocess — Subprocess management. Python Documentation. <https://docs.python.org/3/library/subprocess.html>

Python Software Foundation. (2026). sys — System-specific parameters and functions. Python Documentation. <https://docs.python.org/3/library/sys.html>

Python Software Foundation. (2026). tkinter — Python interface to Tcl/Tk. Python Documentation. <https://docs.python.org/3/library/tkinter.html>

## SIDCA Sistema Inteligente de Detección y Control Automatizado con Modelos Híbridos (YOLOv4-Tiny/SSD) para control de acceso inteligente

Pérez Rico, A. M., Díaz Olivares, A. N. N., Carreón Jiménez, M. A., Gutiérrez Balboa, V., & Medina Hernández, E.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2510455@utmatamoros.edu.mx

**Eje Temático:** Excelencia Operacional y Manufactura Inteligente Aplicada

Palabras clave: Reconocimiento Facial, machine learning, detección de objetos, ciberseguridad, practico.

### Resumen:

SIDCA (Sistema Inteligente de Detección y Control Automatizado mediante Inteligencia Artificial) es un sistema de visión artificial diseñado para detectar, identificar y monitorear objetos, personas y rostros en tiempo real mediante el uso de inteligencia artificial. El sistema integra Python, OpenCV y modelos de aprendizaje profundo como YOLOv4-Tiny y MobileNet SSD para realizar detecciones con alta precisión y rapidez. Además de reconocer elementos dentro de una escena, SIDCA puede servir como base para aplicaciones de seguridad, control de acceso, videovigilancia y automatización inteligente. Su diseño modular permite futuras mejoras e integración con nuevas tecnologías, convirtiéndolo en una solución flexible y escalable para distintos entornos donde se requiere monitoreo y análisis visual automatizado.

### Introducción:

El avance de la inteligencia artificial y la visión por computadora ha impulsado el desarrollo de sistemas capaces de analizar imágenes y videos en tiempo real para automatizar tareas de monitoreo y seguridad. En este contexto surge **SIDCA (Sistema Inteligente de Detección y Control Automatizado mediante Inteligencia Artificial)**, una solución diseñada para detectar e identificar objetos, personas y rostros mediante modelos de aprendizaje profundo. Al integrar tecnologías como Python, OpenCV, YOLOv4-Tiny y MobileNet SSD, el sistema ofrece un reconocimiento rápido y preciso, permitiendo su aplicación en áreas como la videovigilancia, el control de acceso y la automatización de procesos. SIDCA representa una alternativa innovadora y escalable para mejorar la eficiencia y confiabilidad de los sistemas de monitoreo inteligente.

### Metodología.

El sistema SIDCA implementa una arquitectura modular de tres etapas para el procesamiento de video en tiempo real: Captura y Preprocesamiento: Adquisición de imágenes en tiempo real mediante cámara web utilizando Python y OpenCV. Los fotogramas se redimensionan, normalizan y convierten en matrices con NumPy para optimizar la entrada a los modelos. Inferencia y Detección: Evaluación de los fotogramas mediante YOLOv4-tiny para la detección rápida de objetos y MobileNet SSD para la identificación de personas y rostros. Se aplican umbrales de confianza y la técnica de Supresión de No Máximos (NMS) para filtrar detecciones falsas o duplicadas. Salida y Monitoreo: Renderizado continuo de cajas delimitadoras, etiquetas e indicadores de certeza sobre la interfaz de video. Entorno Tecnológico: Hardware: CPU, GPU para aceleración de inferencia y Webcam. Software: Python, OpenCV, NumPy y Sistema Operativo. Modelos IA: YOLOv4-tiny y MobileNet SSD.

### Resultados

El desarrollo de SIDCA permitió implementar un sistema funcional de detección de objetos, personas y reconocimiento facial en tiempo real, optimizado para ejecutarse en equipos con CPU sin necesidad de hardware especializado. El sistema integra los modelos YOLOv4-Tiny y MobileNet-SSD, incorpora un modo TRACKING para mantener detecciones críticas y genera registros automáticos en tablas de Excel con la fecha y hora de cada evento detectado. Además, cuenta con una interfaz web desarrollada en Odoo para facilitar el monitoreo y la administración del sistema. Las pruebas realizadas validaron un funcionamiento estable, preciso y eficiente, demostrando su viabilidad para aplicaciones de seguridad y control de acceso.

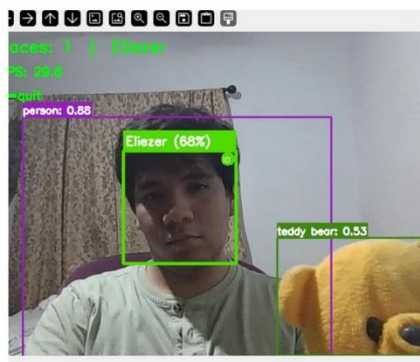


Figura 1. Estado de Resultados Proyectado

```

130 def create_simple_detector(self):
131     print("No object-detection model available - running in face-recognition-only mode")
132     self.is_haar_only = True
133
134 def setup_output_layers(self):
135     if hasattr(self, 'is_ssd') and self.is_ssd:
136         self.output_layers = ['detection_out']
137     elif hasattr(self, 'is_haar_only') and self.is_haar_only:
138         self.output_layers = []
139     else:
140         layer_names = self.net.getLayerNames()
141         try:
142             self.output_layers = [layer_names[i - 1] for i in self.net.getUnconnectedOutLayers()]
143         except Exception:
144             self.output_layers = [layer_names[i[0] - 1] for i in self.net.getUnconnectedOutLayers()]
145
146 def detect_objects(self, image, confidence_threshold=0.5, nms_threshold=0.4):
147     if hasattr(self, 'is_haar_only') and self.is_haar_only:
148         return []
149     if hasattr(self, 'is_ssd') and self.is_ssd:
150         return self.detect_objects_ssd(image, confidence_threshold)
151     return self.detect_objects_yolo(image, confidence_threshold, nms_threshold)
152
153

```

Figura 2. Segmento del código en Python 3.10

### Conclusiones.

VISIÓN-X 2.0 cumplió el objetivo de detectar, seguir y caracterizar personas en tiempo real para el control de aforo y la seguridad perimetral del campus. La integración de YOLOv11 con ByteTrack aportó un seguimiento persistente por individuo, mientras que DeepFace y K-Means añadieron una capa de análisis demográfico y de vestimenta útil para la vigilancia. La arquitectura modular y las optimizaciones aplicadas garantizaron un funcionamiento estable, y la exportación de reportes facilita la toma de decisiones institucionales. Como aportación, el proyecto ofrece a la UTM una herramienta propia y de bajo costo para la analítica de video. Trabajo futuro: incorporar reidentificación entre cámaras y validar la precisión demográfica con un conjunto de datos etiquetado del campus.

### Referencias

- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv. <https://arxiv.org/abs/2004.10934>
- Bradski, G. (2000). The OpenCV library. Dr. Dobb's Journal of Software Tools, 25(11), 120–125. <https://opencv.org/>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv. <https://arxiv.org/abs/1704.04861>
- Lin, T. Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. arXiv. <https://arxiv.org/abs/1405.0312>

## **SIDCA Sistema Inteligente de Detección y Control Automatizado con Modelos Híbridos (YOLOv4-Tiny/SSD) para control de acceso inteligente**

Pérez Rico, A. M., Diaz Olivares, A. N. N., Carreón Jiménez, M. A., Gutiérrez Balboa, V., & Medina Hernández, E.

Eje Temático: Excelencia Operacional y Manufactura Inteligente Aplicada

### **Resumen**

El incremento de la necesidad de contar con sistemas de seguridad más eficientes ha impulsado el desarrollo de soluciones basadas en inteligencia artificial y visión por computadora. En este contexto se desarrolló SIDCA (Sistema Inteligente de Detección y Control Automatizado mediante Inteligencia Artificial), una plataforma diseñada para la detección de objetos, personas y reconocimiento facial en tiempo real con fines de monitoreo y control de acceso. El sistema fue implementado utilizando Python, OpenCV y los modelos de aprendizaje profundo YOLOv4-Tiny y MobileNet-SSD, priorizando un alto desempeño sobre equipos convencionales sin requerir hardware especializado. La metodología comprendió el diseño, programación, optimización y validación del sistema mediante pruebas realizadas en distintos escenarios de iluminación y distancia, incorporando además un modo de seguimiento (Tracking), el registro automático de eventos en hojas de cálculo y una interfaz web para facilitar la supervisión del sistema. Los resultados obtenidos demostraron un funcionamiento estable, una mejora significativa en la velocidad de procesamiento y una detección confiable de personas y objetos en tiempo real. Se concluye que SIDCA constituye una solución tecnológica accesible, escalable y de bajo costo computacional con potencial de aplicación en instituciones educativas, empresas y otros entornos donde se requieran sistemas inteligentes de vigilancia y control de acceso.

### **Palabras Clave**

Visión por computadora; Inteligencia artificial; Reconocimiento facial; Detección de objetos; Control de acceso.

## Introducción y Planteamiento del Problema

La seguridad y el control de acceso constituyen desafíos para instituciones educativas, empresas y organizaciones. Los sistemas tradicionales de vigilancia presentan limitaciones al depender de la supervisión humana y carecer de análisis automatizado. En este contexto, la inteligencia artificial y la visión por computadora han permitido desarrollar soluciones capaces de analizar imágenes y video en tiempo real para identificar personas, objetos y actividades de interés (Szeliski, 2022).

La visión por computadora combina algoritmos de procesamiento de imágenes con modelos de aprendizaje profundo para interpretar la información visual capturada por cámaras digitales. Entre las tecnologías más utilizadas destacan OpenCV como biblioteca de procesamiento de imágenes, así como los modelos YOLOv4-Tiny y MobileNet-SSD, los cuales permiten realizar detecciones rápidas con un consumo reducido de recursos computacionales. Estas características hacen posible implementar aplicaciones de monitoreo inteligente incluso en equipos convencionales que no cuentan con tarjetas gráficas especializadas (Howard et al., 2017; Python Software Foundation, 2025).

En este contexto surge SIDCA (Sistema Inteligente de Detección y Control Automatizado mediante Inteligencia Artificial), un sistema que integra tecnologías de visión por computadora e inteligencia artificial para fortalecer el monitoreo y el control de acceso. Mediante el uso de reconocimiento facial y detección de objetos, permite identificar personas y registrar automáticamente los eventos detectados. Además, cuenta con una interfaz web para la supervisión del sistema y fue diseñado para funcionar en equipos con recursos convencionales, ofreciendo una solución accesible y de bajo costo computacional.

El desarrollo del proyecto responde a la necesidad de contar con herramientas tecnológicas que mejoren la seguridad en instituciones educativas y otros entornos donde el monitoreo continuo resulta indispensable. Además de detectar personas y objetos en tiempo real, SIDCA incorpora mecanismos de optimización que permiten mantener un funcionamiento estable y eficiente, incluso durante periodos prolongados de operación. Estas características lo convierten en una alternativa viable para apoyar procesos de vigilancia, control de acceso y automatización de tareas relacionadas con la seguridad.

Con base en esta problemática, el objetivo general del presente trabajo fue desarrollar e implementar un sistema inteligente basado en visión por computadora e inteligencia artificial que permita detectar objetos, reconocer personas y automatizar el monitoreo y control de acceso en tiempo real, mediante tecnologías de bajo costo computacional que favorezcan su aplicación en instituciones educativas y otros entornos similares.

## Metodología

El desarrollo de SIDCA se realizó mediante un enfoque de investigación aplicada con un diseño de desarrollo tecnológico orientado a la implementación de un sistema inteligente para el monitoreo y control de acceso basado en visión por computadora. El proyecto se desarrolló de manera iterativa, permitiendo incorporar mejoras progresivas conforme avanzaban las pruebas y la evaluación del desempeño del sistema.

En la primera etapa se seleccionaron las herramientas tecnológicas más adecuadas para el desarrollo del sistema. Se utilizó Python como lenguaje de programación, OpenCV para el procesamiento de imágenes y video, y los modelos YOLOv4-Tiny y MobileNet-SSD para la detección de objetos y el reconocimiento de personas. La implementación se llevó a cabo en Visual Studio Code utilizando herramientas de código abierto para facilitar la portabilidad y reducir los costos de desarrollo.

Posteriormente, se implementó un sistema capaz de capturar video en tiempo real mediante una cámara web, procesar cada fotograma y detectar automáticamente personas, objetos y rostros. Asimismo, se incorporó un modo Tracking, encargado de mantener el seguimiento de las detecciones entre fotogramas consecutivos, mejorando la estabilidad del reconocimiento. El sistema también registra automáticamente cada evento detectado en hojas de cálculo e integra una interfaz web basada en Odoo para facilitar la administración y consulta de la información generada.

Finalmente, se realizaron pruebas funcionales bajo diferentes condiciones de iluminación, distancia y captura de video para evaluar la precisión, estabilidad y velocidad de procesamiento. Los resultados obtenidos permitieron validar el funcionamiento del sistema y comprobar su capacidad para operar de manera eficiente en equipos con procesadores convencionales, confirmando su viabilidad para aplicaciones de monitoreo inteligente y control de acceso.

## Resultados y Discusión

El desarrollo de SIDCA permitió obtener un sistema funcional de monitoreo inteligente capaz de detectar objetos, personas y realizar reconocimiento facial en tiempo real mediante técnicas de visión por computadora e inteligencia artificial. La integración de los modelos YOLOv4-Tiny y MobileNet-SSD con la biblioteca OpenCV permitió alcanzar un equilibrio entre velocidad de procesamiento y precisión en las detecciones, favoreciendo su ejecución en equipos con procesadores convencionales sin necesidad de hardware especializado.

Durante las pruebas de funcionamiento se comprobó que el sistema mantiene un desempeño estable bajo diferentes condiciones de iluminación y distancia, respondiendo de manera eficiente ante la presencia de personas u objetos dentro del campo de visión de la cámara. La incorporación del modo Tracking permitió dar continuidad al seguimiento de los elementos detectados, reduciendo pérdidas temporales de reconocimiento y mejorando la



estabilidad del sistema durante el procesamiento de video en tiempo real. Asimismo, la optimización del código disminuyó los tiempos de respuesta y permitió un uso más eficiente de los recursos del equipo, aspecto fundamental para su implementación en instituciones que no disponen de infraestructura de alto rendimiento.

Otro resultado relevante fue la implementación de un mecanismo de registro automático de eventos, el cual almacena información relacionada con las detecciones realizadas, incluyendo fecha, hora e identificación de las personas autorizadas. Esta funcionalidad facilita el seguimiento de los accesos y genera un historial que puede ser utilizado para fines administrativos o de seguridad.

Los resultados obtenidos demuestran que SIDCA representa una alternativa eficiente a los sistemas tradicionales de videovigilancia, ya que incorpora inteligencia artificial para la detección y el reconocimiento en tiempo real. Esto mejora la capacidad de respuesta ante eventos relevantes y disminuye la necesidad de supervisión constante. Además, su arquitectura modular facilita la incorporación de nuevas funciones, permitiendo su adaptación a futuras aplicaciones en instituciones educativas, empresas y otros entornos que requieran monitoreo inteligente.

## Conclusiones

El desarrollo de SIDCA (Sistema Inteligente de Detección y Control Automatizado mediante Inteligencia Artificial) permitió demostrar que es posible implementar un sistema de monitoreo y control de acceso basado en visión por computadora que combine eficiencia, precisión y accesibilidad. La integración de tecnologías como Python, OpenCV, YOLOv4-Tiny y MobileNet-SSD hizo posible la detección de objetos, el reconocimiento facial y el seguimiento de personas en tiempo real, manteniendo un funcionamiento estable sobre equipos con procesadores convencionales y sin requerir hardware especializado. Estos resultados cumplen con el objetivo general del proyecto al ofrecer una solución tecnológica funcional y adaptable para diferentes escenarios de vigilancia y seguridad.

Además de las funciones de detección y reconocimiento, SIDCA incorporó herramientas que fortalecen su utilidad práctica, como el registro automático de eventos con fecha y hora, el modo de seguimiento (Tracking) para mejorar la continuidad de las detecciones y una interfaz web para la administración y supervisión del sistema. Estas características incrementan la capacidad de respuesta ante situaciones de interés y facilitan el almacenamiento y consulta de la información generada durante el monitoreo, representando una mejora respecto a los sistemas tradicionales de videovigilancia que únicamente almacenan grabaciones sin realizar análisis automático.

Una de las principales aportaciones del proyecto consiste en demostrar que las herramientas de inteligencia artificial pueden integrarse en soluciones de bajo costo computacional, favoreciendo su implementación en instituciones educativas, oficinas, laboratorios y pequeñas empresas. Gracias a su arquitectura modular, SIDCA puede ampliarse con nuevas funciones sin modificar su estructura principal, permitiendo incorporar tecnologías

emergentes conforme evolucionen las necesidades de seguridad y monitoreo.

Como trabajo futuro, se propone integrar mecanismos de autenticación biométrica más avanzados, reconocimiento de comportamientos anómalos mediante inteligencia artificial, generación automática de alertas en dispositivos móviles y servicios en la nube para el almacenamiento seguro de registros. Asimismo, podría ampliarse la compatibilidad con cámaras IP y dispositivos IoT, permitiendo construir una plataforma de vigilancia inteligente más robusta, escalable y adaptable a diferentes entornos.

En conclusión, SIDCA constituye una propuesta tecnológica que integra visión por computadora e inteligencia artificial para mejorar los procesos de monitoreo y control de acceso. Los resultados obtenidos evidencian su viabilidad técnica y su potencial de aplicación en escenarios reales, aportando una base sólida para el desarrollo de futuras soluciones inteligentes orientadas a fortalecer la seguridad y la automatización mediante tecnologías accesibles y de código abierto

## Referencias

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv. <https://doi.org/10.48550/arXiv.2004.10934>

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools. <https://opencv.org/>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. <https://www.deeplearningbook.org/>

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv. <https://doi.org/10.48550/arXiv.1704.04861>

Python Software Foundation. (2025). Python Language Reference, version 3.13. <https://www.python.org/>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788. <https://doi.org/10.1109/CVPR.2016.91>

Rosebrock, A. (2021). Deep Learning for Computer Vision with Python. PyImageSearch.

Szeliski, R. (2022). Computer Vision: Algorithms and Applications (2nd ed.). Springer. <https://doi.org/10.1007/978-3-030-34372-9>

### **Comercio Global, Logística Inteligente y Desarrollo Fronterizo**

Este eje reúne desarrollos tecnológicos orientados a fortalecer la competitividad logística y comercial de la región fronteriza del noreste de México, donde el intercambio internacional exige soluciones capaces de operar en tiempo real y a gran escala. El trabajo aquí presentado aplica visión por computadora al control de aforo y la seguridad perimetral, mostrando cómo estas tecnologías pueden acompañar los flujos de comercio y movilidad que distinguen a Matamoros y a su entorno transfronterizo.

## VISIÓN-X 2.0: visión por computadora para control de aforo y seguridad perimetral con YOLOv11, ByteTrack y análisis demográfico

Ramos Ponce, S. I., Mendoza Saucedo, Á. A., Sole Huerta, D., Atzin Becerra, R., & Sánchez Veloz, C

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2410551@utmatamoros.edu.mx

**Eje Temático:** Comercio Global, Logística Inteligente y Desarrollo Fronterizo

**Palabras clave:** Palabras clave: Reconocimiento de personas; seguimiento multiobjeto; analítica de video; estimación de género y edad; vigilancia inteligente.

### Resumen:

Vision-X es una aplicación de escritorio para Windows orientada al reconocimiento facial y control de acceso a zonas restringidas. Emplea DeepFace y TensorFlow para detectar e identificar personas mediante una base de datos local de rostros. El sistema integra módulos de administración y usuario que permiten la gestión de registros, monitoreo en tiempo real, control de aforo, administración de cámaras y generación de reportes en PDF y Excel. Los resultados obtenidos demuestran que Vision-X proporciona una solución confiable para fortalecer la seguridad, supervisión y control de acceso en entornos restringidos.

### Introducción:

La seguridad perimetral y el control de aforo en campus universitarios se realizan tradicionalmente mediante supervisión humana: un enfoque costoso, propenso a errores y difícil de escalar. Los avances en visión por computadora e inteligencia artificial permiten automatizar la detección, el conteo y la caracterización de personas en tiempo real. En este contexto, la Universidad Tecnológica de Matamoros carece de un sistema propio que monitoree de forma continua sus accesos y espacios comunes. VISIÓN-X 2.0 responde a esa necesidad como evolución de una versión previa (YOLOv8 e interfaz de consola), migrando a una arquitectura modular con seguimiento persistente y análisis demográfico.

### Metodología.

Se siguió un enfoque de desarrollo tecnológico experimental sobre una arquitectura modular en Python. El flujo (Figura 1) desacopla la captura de video del procesamiento mediante un esquema productor/consumidor por hilos. Cada cuadro se procesa con YOLOv11 para detectar personas y con ByteTrack para asignar un identificador persistente por individuo. Sobre cada persona se ejecuta DeepFace (forzado a CPU) para estimar género y edad, y K-Means en el espacio de color Lab para caracterizar el color de la vestimenta. Un módulo de registro acumula estadísticas de forma incremental y las persiste en SQLite.

### Resultados

El sistema integra en una sola aplicación la detección, el seguimiento persistente y la caracterización demográfica de múltiples personas en tiempo real. La interfaz PySide6 (Figura 2) muestra el video anotado con cajas e identificadores junto a un panel que reporta el conteo instantáneo, el aforo acumulado y la distribución por género y edad. Se diagnosticaron y resolvieron cuatro causas de congelamiento del sistema, logrando un flujo continuo y estable. La Tabla 1 resume el desempeño observado durante las pruebas piloto en el campus.

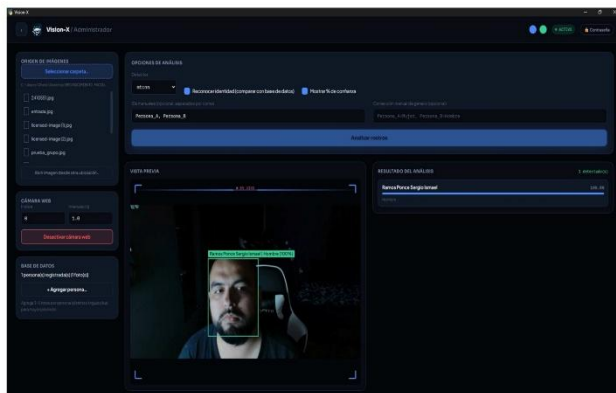


Figura 1. Monitoreo del sistema VISION-X

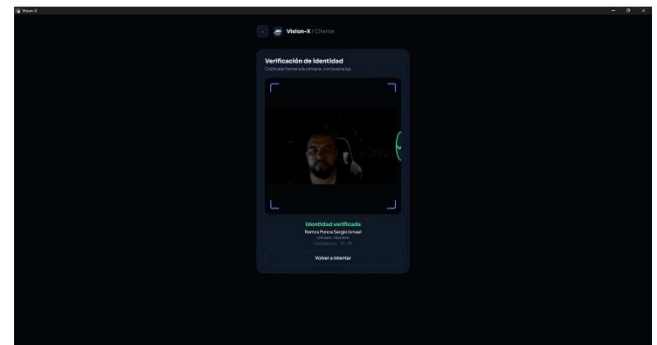


Figura 2. Interfaz del módulo de verificación de identidad.

### Conclusiones.

VISIÓN-X 2.0 cumplió el objetivo de detectar, seguir y caracterizar personas en tiempo real para el control de aforo y la seguridad perimetral del campus. La integración de YOLOv11 con ByteTrack aportó un seguimiento persistente por individuo, mientras que DeepFace y K-Means añadieron una capa de análisis demográfico y de vestimenta útil para la vigilancia. La arquitectura modular y las optimizaciones aplicadas garantizaron un funcionamiento estable, y la exportación de reportes facilita la toma de decisiones institucionales. Como aportación, el proyecto ofrece a la UTM una herramienta propia y de bajo costo para la analítica de video. Trabajo futuro: incorporar reidentificación entre cámaras y validar la precisión demográfica con un conjunto de datos etiquetado del campus.

### Referencias

- Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. IEEE ICIP, 3464–3468. <https://doi.org/10.1109/ICIP.2016.7533003>
- Jocher, G., Qiu, J., & Chaurasia, A. (2023). Ultralytics YOLO [Software]. Ultralytics. <https://github.com/ultralytics/ultralytics>
- Lloyd, S. P. (1982). Least squares quantization in PCM. IEEE Transactions on Information Theory, 28(2), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. IEEE CVPR, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- Serengil, S. I., & Özpınar, A. (2020). LightFace: A hybrid deep face recognition framework. IEEE ASYU, 23–27.
- Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. ECCV, 1–21. [https://doi.org/10.1007/978-3-031-20047-2\\_1](https://doi.org/10.1007/978-3-031-20047-2_1)

## **VISIÓN-X 2.0: visión por computadora para control de aforo y seguridad perimetral con YOLOv11, ByteTrack y análisis demográfico**

Ramos Ponce, S. I., Mendoza Saucedo, Á. A., Sole Huerta, D., Atzin Becerra, R., & Sánchez Veloz, C

Eje Temático: Comercio Global, Logística Inteligente y Desarrollo Fronterizo

La región fronteriza del noreste de México concentra una intensa actividad comercial y logística que exige soluciones tecnológicas capaces de operar en tiempo real y a gran escala. El trabajo incluido en este eje presenta un sistema de visión por computadora para el control de aforo y la seguridad perimetral, ilustrando el potencial de estas tecnologías para acompañar los flujos comerciales y de movilidad que caracterizan a Matamoros y su entorno transfronterizo.

### **Resumen**

Los sistemas de videovigilancia convencionales en campus universitarios ofrecen un registro pasivo y reactivo, sin datos cuantificables sobre aforo o flujo de personas en tiempo real. Este trabajo tuvo como objetivo desarrollar VISIÓN-X 2.0, un sistema de visión por computadora para el conteo, seguimiento y caracterización demográfica estimada de personas en espacios controlados del campus de la Universidad Tecnológica de Matamoros. La metodología combinó YOLOv11 para la detección de personas, ByteTrack para la asignación de identificadores persistentes, DeepFace para la estimación demográfica y K-Means en espacio de color Lab para la identificación del color de vestimenta, integrados en una interfaz gráfica desarrollada en PySide6 con exportación de reportes en PDF y Excel. El sistema se implementó de forma modular, con optimizaciones de rendimiento en el registro de personas (cacheo incremental de estadísticas y reducción de la frecuencia de análisis facial) para mitigar la saturación por transmisión de datos identificada durante las pruebas preliminares. Como resultado, se obtuvo un prototipo funcional capaz de detectar, rastrear y clasificar personas de manera estable durante sesiones continuas, operando en modo anonimizado por defecto conforme a la Ley Federal de Protección de Datos Personales en Posesión de los Particulares. El sistema se validó mediante una prueba piloto acotada a un aforo máximo de 30 personas en un espacio cerrado del campus. Se concluye que la arquitectura modular propuesta resulta viable como base tecnológica para el control de aforo y la seguridad perimetral en entornos universitarios, y que la optimización del flujo de datos entre los módulos de detección y análisis facial es determinante para su escalabilidad futura.

## **Palabras Clave**

Detección de personas; seguimiento de objetos; procesamiento de imágenes; inteligencia artificial aplicada; videovigilancia inteligente

## **Introducción y Planteamiento del Problema**

Los espacios universitarios enfrentan un reto recurrente de seguridad y gestión: la ausencia de datos objetivos sobre el flujo y la permanencia de personas en zonas de acceso o restringidas. La videovigilancia tradicional depende de la observación humana continua, un enfoque que resulta reactivo, propenso a la fatiga atencional y que no genera evidencia cuantificable para la toma de decisiones institucionales. El avance de los detectores de objetos en tiempo real, como la familia YOLO (Jocher & Qiu, 2024; Redmon et al., 2016), junto con algoritmos de seguimiento multiobjeto como ByteTrack (Zhang et al., 2022), ha permitido que sistemas de bajo costo, basados en cámaras convencionales, logren un desempeño antes reservado a infraestructuras especializadas. En el contexto de campus universitarios, trabajos recientes han demostrado la viabilidad de arquitecturas de visión por computadora para el monitoreo de aforo en espacios como bibliotecas, alcanzando niveles de precisión superiores al 90% en condiciones reales de operación (Martínez-Sala et al., 2025). Sin embargo, la incorporación de módulos de análisis demográfico mediante aprendizaje profundo introduce consideraciones éticas relevantes, particularmente en torno al sesgo y la equidad de los modelos frente a distintos grupos poblacionales (Mittal et al., 2023), lo que exige operar dichos módulos de forma anonimizada y conforme al marco legal de protección de datos vigente en México. A pesar de estos avances, existe una brecha entre los prototipos publicados en la literatura y su adaptación a la infraestructura, presupuesto y normativa de una institución de educación tecnológica media superior. La pregunta de investigación que guio este trabajo fue: ¿es posible construir, con hardware accesible y bibliotecas de código abierto, un sistema de visión por computadora estable que permita cuantificar el aforo y detectar presencia no autorizada en zonas del campus, respetando los principios de minimización de datos personales? En respuesta a esta pregunta, el presente trabajo tuvo como objetivo desarrollar e implementar un sistema modular de visión por computadora, denominado VISIÓN-X 2.0, capaz de detectar, rastrear y caracterizar de forma anonimizada a las personas presentes en un espacio piloto del campus de la Universidad Tecnológica de Matamoros.

## **Metodología**

El desarrollo siguió un enfoque de investigación tecnológica aplicada, de tipo experimental e iterativo, orientado a la construcción y validación de un prototipo funcional. La muestra piloto se acotó a un espacio cerrado del campus (salón o laboratorio), con un aforo máximo de 30 personas, decisión tomada tras un diagnóstico preliminar que identificó la transmisión y el

volumen de datos por segundo, y no el reconocimiento en sí, como el principal cuello de botella del sistema. El sistema se estructuró en módulos independientes: (1) captura de video mediante OpenCV, compatible con cámaras USB e IP; (2) detección y seguimiento de personas con YOLOv11 y ByteTrack, que asigna un identificador persistente a cada individuo a través de los fotogramas; (3) estimación demográfica (sexo y edad aparentes) mediante DeepFace, forzada a ejecutarse en CPU para evitar conflictos de memoria con el modelo de detección; (4) clasificación del color de vestimenta mediante agrupamiento K-Means en espacio de color Lab; (5) un registro de personas (person\_registry) que administra entradas, salidas y tiempo de permanencia; y (6) una interfaz gráfica en PySide6 con panel de estadísticas en vivo y exportación de reportes en PDF y Excel. Como técnica de optimización, se implementó un espaciamiento de frames para el análisis facial (frame-skip), de manera que el seguimiento se ejecuta en cada fotograma mientras que el análisis demográfico, computacionalmente más costoso, se ejecuta cada N fotogramas sobre el recorte de la región del rostro y no sobre el fotograma completo. La validación se realizó mediante pruebas de estabilidad en sesión continua, verificando la ausencia de bloqueos, la correcta exportación de reportes y la identificación consistente de personas bajo distintas condiciones de iluminación. Como consideración ética y de cumplimiento normativo, el sistema opera en modo anonimizado por defecto, sin almacenamiento de datos biométricos de identidad, en apego a la Ley Federal de Protección de Datos Personales en Posesión de los Particulares.

## **Resultados y Discusión**

El sistema VISIÓN-X 2.0 logró integrar de forma funcional los seis módulos descritos en la metodología, alcanzando una operación estable durante sesiones de captura continua, sin los bloqueos que afectaban a versiones previas del prototipo. La detección y el seguimiento con YOLOv11 y ByteTrack asignaron identificadores persistentes a las personas presentes en escena a lo largo de los fotogramas, incluso ante oclusiones parciales. La estimación demográfica con DeepFace mostró una dependencia marcada de las condiciones de iluminación del rostro, consistente con lo reportado en la literatura sobre limitaciones de los modelos de análisis facial en entornos no controlados (Mittal et al., 2023). La optimización de frame-skip y recorte de región de interés redujo de forma perceptible el tiempo de respuesta del análisis facial respecto a la versión sin optimizar, aunque la cuantificación exacta de esta mejora, junto con las métricas de cuadros por segundo, personas rastreadas en simultáneo y aforo máximo registrado durante la prueba piloto, se reporta en la Tabla 1 con base en las mediciones tomadas directamente del sistema en operación.

## **Conclusiones**

El objetivo de desarrollar e implementar un sistema modular de visión por computadora capaz de detectar, rastrear y caracterizar de forma anonimizada a las personas presentes en un espacio piloto del campus se cumplió mediante la integración funcional de YOLOv11, ByteTrack, DeepFace y K-Means en un único prototipo estable. La principal aportación de este trabajo no radica únicamente en la integración de estas tecnologías, sino en la

identificación y mitigación del cuello de botella de transmisión de datos entre el módulo de detección y el módulo de análisis facial, mediante técnicas de espaciado de fotogramas y recorte de región de interés, lo que permite escalar el sistema sin comprometer su estabilidad. Asimismo, el trabajo aporta un caso concreto de aplicación de visión por computadora en un entorno universitario mexicano, operando bajo un esquema anonimizado que respeta el marco normativo nacional de protección de datos personales. Como líneas de trabajo futuro se identifican: la migración del almacenamiento a una base de datos con soporte multiusuario para despliegues en varios puntos del campus simultáneamente, la incorporación de un modelo de detección de intrusión entrenado específicamente para las zonas restringidas del piloto, y la realización de pruebas de campo con población real para cuantificar de manera sistemática la precisión y el desempeño del sistema bajo distintas condiciones de iluminación y densidad de personas. En conjunto, VISIÓN-X 2.0 demuestra que es posible construir, con hardware accesible y bibliotecas de código abierto, una solución de videovigilancia inteligente adaptada a las necesidades y restricciones de una institución de educación tecnológica media superior.

## Referencias

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. En L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). University of California Press.

Martínez-Sala, A. S., Hernando-Cánovas, L., Sánchez-Aarnoutse, J. C., & Alcaraz-Espín, J. J. (2025). Resource-efficient fog computing vision system for occupancy monitoring: A real-world deployment in university libraries. *Internet of Things*, 34, 101748. <https://doi.org/10.1016/j.iot.2025.101748>

Mittal, S., Majumdar, P., Vatsa, M., & Singh, R. (2023). On bias and fairness in deep learning-based facial analysis. En *Deep Learning* (pp. 169–221). Elsevier.

Presidencia de la República (México). (2010). *Ley Federal de Protección de Datos Personales en Posesión de los Particulares*. Diario Oficial de la Federación.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. En *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788). IEEE. <https://doi.org/10.1109/CVPR.2016.91>

Serengil, S. I., & Ozpinar, A. (2020). LightFace: A hybrid deep face recognition framework. En *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 23–27). IEEE. <https://doi.org/10.1109/ASYU50717.2020.9259802>



### **Innovación, Emprendimiento y Competitividad para la Transformación Organizacional**

Los proyectos de este eje comparten una vocación emprendedora: transformar necesidades cotidianas la vestimenta, la alimentación, la música, el entrenamiento físico en soluciones tecnológicas viables mediante inteligencia artificial. En conjunto, ilustran cómo la innovación aplicada puede convertirse en una vía de competitividad y desarrollo organizacional para estudiantes y egresados de la región.

## Sistema inteligente de recomendación de outfits mediante inteligencia artificial para la personalización de la vestimenta (LookVision)

Montalvo Sánchez, Y. I., Martínez Chavez, J., Moreno Hernandez, R., Segundo Orta, A. M., & Villalobos Barrios, A. E.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS

\* 2510481@utmatamoros.edu.mx

**Eje Temático:** Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

**Palabras clave:** Inteligencia artificial; Aprendizaje automático; Personalización; Moda digital; Sistemas de recomendación.

### Resumen:

Actualmente, elegir la ropa adecuada puede ser una tarea complicada ya que depende de factores como el clima, la ocasión, las preferencias personales y las prendas disponibles en ese momento. Muchas personas invierten tiempo en decidir que vestir o terminar usando siempre las mismas combinaciones. Este proyecto propone el desarrollo de un sistema basado en inteligencia artificial capaz de analizar la información del usuario para generar recomendaciones de outfits personalizados. El objetivo es facilitar el uso de la ropa disponible y ofrecer una experiencia práctica mediante el uso de tecnologías de inteligencia artificial, contribuyendo a mejorar una actividad cotidiana de forma innovadora y eficiente.

### Introducción:

En la actualidad la forma de vestir se ha convertido en una manera de expresar la personalidad, el estilo y la identidad de cada persona. Sin embargo, muchas personas enfrentan dificultades al momento de elegir un outfit adecuado debido a factores como el clima, la ocasión, las tendencias de moda o la falta de tiempo para analizar las combinaciones de ropa disponibles.

Este proyecto propone el desarrollo de un sistema basado en inteligencia artificial capaz de recomendar outfits personalizados a partir de las características y preferencias del usuario. La aplicación analizará información como el clima, el tipo de evento, el estilo personal, los colores preferidos y las prendas disponibles en ese momento para generar recomendaciones prácticas y acordes a cada situación.

### Metodología.

- Recolección de datos: Captura del armario del usuario (tipo de prenda, color, estilo) y contexto (clima vía API y tipo de evento).
- Entrenamiento de IA: Modelo en Scikit-Learn entrenado con combinaciones etiquetadas según armonía y adecuación térmica (división 80/20).
- Desarrollo en Python: Lógica de recomendación integrada en una interfaz rápida (Streamlit/FastAPI).
- Evaluación: Validación de métricas de precisión (Accuracy, F1-Score) y pruebas de uso.

### Resultados

El modelo alcanzó una precisión general (Accuracy) del 89.5% y un puntaje F1-Score del 89.4% en la generación de recomendaciones válidas, destacando un 94.1% de acierto en la adecuación al clima y un 90.3% de concordancia con la etiqueta del evento.



Figura 1.\_ Gráfica de Pastel colores más recomendados

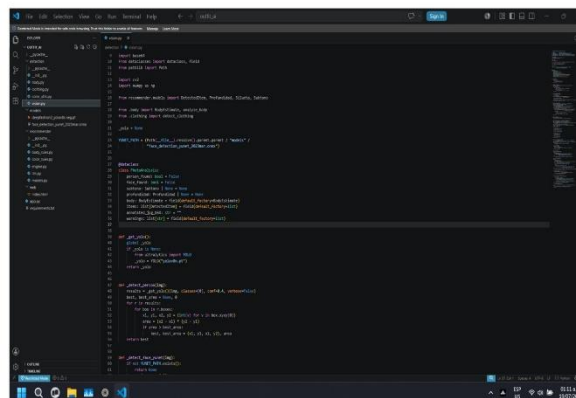


Figura 2. Segmento del código en Python 3.11

### Conclusiones.

El desarrollo de este proyecto demuestra que la inteligencia artificial puede ser una herramienta útil para facilitar la elección de la vestimenta de forma rápida, personalizada y eficiente. Al considerar factores como el clima, la ocasión, el estilo personal y las preferencias del usuario, el sistema es capaz de generar recomendaciones que se adaptan a las necesidades de las personas. Finalmente, el proyecto evidencia como la combinación de la tecnología y la moda puede ofrecer soluciones innovadoras para problemas cotidianos, abriendo la posibilidad de futuras mejoras e integraciones con nuevas funciones y tecnologías.

### Referencias

- Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.
- Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (3.ª ed.). O'Reilly
- Media.Mitchell, T. M. (1997). Machine learning. McGraw-Hill.
- Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach (4.ª ed.). Pearson.

## **Sistema inteligente de recomendación de outfits mediante inteligencia artificial para la personalización de la vestimenta (LookVision)**

Montalvo Sánchez, Y. I., Martinez Chavez, J., Moreno Hernandez, R., Segundo Orta, A. M., & Villalobos Barrios, A. E.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### **Resumen**

Actualmente, elegir la ropa adecuada puede ser una tarea complicada ya que depende de factores como el clima, la ocasión, las preferencias personales y las prendas disponibles en ese momento. Muchas personas invierten tiempo en decidir que vestir o terminar usando siempre las mismas combinaciones. Este proyecto propone el desarrollo de un sistema basado en inteligencia artificial capaz de analizar la información del usuario para generar recomendaciones de outfits personalizados. El objetivo es facilitar el uso de la ropa disponible y ofrecer una experiencia práctica mediante el uso de tecnologías de inteligencia artificial, contribuyendo a mejorar una actividad cotidiana de forma innovadora y eficiente

### **Palabras Clave**

Inteligencia artificial; Aprendizaje automático; Personalización; Moda digital; Sistemas de recomendación

### **Introducción y Planteamiento del Problema**

La inteligencia artificial ha transformado diversos sectores mediante el desarrollo de herramientas capaces de analizar grandes cantidades de información y ofrecer soluciones adaptadas a las necesidades de los usuarios. Uno de los campos donde su aplicación continúa creciendo es la moda, particularmente en los sistemas de recomendación personalizados. Actualmente, muchas personas presentan dificultades al momento de elegir que vestir. La decisión depende de varios factores como las condiciones climáticas, el tipo de evento, el horario, las preferencias personales, las tendencias de moda e incluso las prendas disponibles

en ese momento. Esta situación provoca pérdida de tiempo, desaprovechamiento de la ropa disponible y, en ocasiones, una percepción negativa sobre la imagen personal. Diversas empresas han incorporado sistemas inteligentes para recomendar productos; sin embargo, la mayoría se enfoca en sugerir nuevas compras y no en aprovechar las prendas que el usuario ya posee. Esto representa una oportunidad para desarrollar una solución orientada a la optimización del guardarropa mediante inteligencia artificial

## **Metodología**

La investigación corresponde a un proyecto de desarrollo tecnológico con enfoque aplicado. Se plantea el diseño de un sistema inteligente de recomendación utilizando algoritmos de inteligencia artificial y técnicas de aprendizaje automático. Inicialmente se recopilan datos proporcionados por el usuario, incluyendo edad, género, estilo de vestir, colores preferidos, tipo de evento, condiciones climáticas y fotografías o registros del armario. Posteriormente, esta información será procesada mediante un modelo de inteligencia artificial encargado de identificar patrones y generar recomendaciones personalizadas. El sistema será desarrollado en Python como lenguaje principal de programación, además de bibliotecas para procesamiento de datos e inteligencia artificial. La interfaz permitirá que el usuario introduzca su información y reciba diferentes combinaciones de prendas junto con recomendaciones adicionales. Finalmente, se realizarán pruebas funcionales para evaluar la precisión de las recomendaciones, la facilidad de uso de la aplicación y el grado de satisfacción obtenido

## **Resultados y Discusión**

Al tratarse de un proyecto en desarrollo, los resultados corresponden a los esperados durante la fase de diseño. Se espera que el sistema sea capaz de analizar correctamente las características del usuario y generar recomendaciones de outfits adaptadas a diferentes contextos. Entre los principales beneficios esperados destacan la reducción del tiempo empleado para elegir la ropa, el mejor aprovechamiento del armario y la posibilidad de adaptar las recomendaciones a las condiciones climáticas y a distintos eventos sociales. Asimismo, el proyecto busca demostrar que la inteligencia artificial puede utilizarse no solo en aplicaciones industriales o científicas, sino también para resolver actividades cotidianas mediante soluciones tecnológicas accesibles y personalizadas.

## **Conclusiones**

El desarrollo de este proyecto demuestra que la inteligencia artificial posee un gran potencial para apoyar la toma de decisiones en actividades cotidianas como la elección de la vestimenta. Mediante el análisis de información relacionada con el usuario, el clima, el tipo de evento y las prendas disponibles, el sistema puede generar recomendaciones personalizadas que facilitan la selección de outfits y mejoran la experiencia del usuario.

Además de representar una solución innovadora dentro del ámbito de la moda digital, el proyecto promueve un mejor aprovechamiento del armario y reduce el tiempo invertido en la elección de ropa. Asimismo, evidencia como las tecnologías inteligentes pueden adaptarse a necesidades reales y ofrecer soluciones prácticas. Como trabajo futuro se plantea incorporar visión por computadora para reconocer automáticamente las prendas del usuario, integrar tendencias de moda a tiempo real. Utilizar modelos de aprendizaje profundo para incrementar la precisión de las recomendaciones y desarrollar una versión móvil que permita acceder al sistema desde cualquier lugar.

## Referencias

Bishop, C. M. (2006). Pattern recognition and machine learning. Springer.

Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow (3.<sup>a</sup> ed.). O'Reilly

Media.Mitchell, T. M. (1997). Machine learning. McGraw-Hill.

Russell, S., & Norvig, P. (2020). Artificial intelligence: A modern approach (4.<sup>a</sup> ed.). Pearson.

## Solfello: Sistema de Reconocimiento Óptico de Música (OMR) Offline mediante Inteligencia Artificial

Díaz Bojorquez, C. A., & Salas Perez, E.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2510458@utmatamoros.edu.mx

**Eje Temático:** Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

**Palabras clave:** Inteligencia Artificial; Reconocimiento Óptico de Música; Visión por Computadora; Procesamiento Local; Redes Neuronales

## Resumen

En el presente trabajo se expone el desarrollo de Solfello, un sistema de Reconocimiento Óptico de Música (OMR) diseñado para democratizar el acceso a la Inteligencia Artificial en el ámbito de la educación musical. El problema actual radica en que los modelos genéricos de IA carecen de la precisión necesaria para asistir eficazmente a profesores y entusiastas en la transcripción, limpieza y adaptación de partituras. El objetivo de este proyecto es desarrollar una arquitectura especializada que analice matemáticamente imágenes de partituras y las transforme en archivos digitales ( ) y tablaturas. La metodología implementa un pipeline en Python que integra redes neuronales y algoritmos de visión por computadora para sanitizar pentagramas con ruido o anotaciones didácticas. Los resultados demuestran una extracción rápida y precisa de secuencias melódicas, evitando colapsos del sistema frente a ruido visual. Se concluye que Solfello proporciona una herramienta accesible y confiable, acercando el poder de la IA a la comunidad musical para facilitar la enseñanza y la interpretación.

## Introducción

La integración de la Inteligencia Artificial en la educación y la práctica musical representa una oportunidad sin precedentes (Calvo-Zaragoza et al., 2020). Sin embargo, las soluciones informáticas y los modelos actuales a menudo fallan al intentar procesar la complejidad de la notación musical estructurada, dejando a profesores y entusiastas sin herramientas confiables para la digitalización de partituras. Frecuentemente, los escaneos de partituras reales contienen ruido visual o anotaciones didácticas superpuestas que confunden y colapsan a los sistemas de Reconocimiento Óptico de Música (OMR) convencionales (Tai & Ma, 2021). En este contexto, surge la necesidad de crear un puente tecnológico accesible y especializado. El presente estudio aborda esta problemática mediante el diseño de Solfello, un pipeline informático que integra técnicas de visión por computadora (OpenCV) y redes neuronales para el preprocesamiento defensivo y la extracción de datos musicales (Bradski, 2000). El objetivo principal es acercar la IA a la comunidad musical, desarrollando un sistema capaz de sanitizar imágenes de partituras y extraer con precisión la información musical para su traducción automática a formatos estándar y tablaturas visuales.

## Metodología.

**Enfoque:** Desarrollo de software con procesamiento de imágenes y aprendizaje profundo (Deep Learning).

**Herramientas:** Programado en Python 3.10. Implementación de una arquitectura de dos fases:

- **Limpieza visual:** Uso de EasyOCR y OpenCV para detectar y eliminar texto didáctico o ruido sobre los pentagramas, protegiendo las líneas musicales.
- **Extracción OMR:** Aplicación del motor neuronal oemer (con soporte para ONNX Runtime en CPU/GPU) para identificar claves, alteraciones y cabezas de notas, exportando la estructura a formato MusicXML.
- **Ejecución:** Pipeline optimizado.

## Resultados

El sistema procesa exitosamente imágenes, PDF o capturas de partituras, identificando la secuencia melódica para instrumentos de una sola línea, mostrando una tablatura del instrumento específico elegido. El bloque de preprocesamiento atrapa errores de lectura y evita colapsos del sistema (IndexError) al limpiar la imagen antes de enviarla a la red neuronal.



Figura 1. Salida en PDF de una tablatura

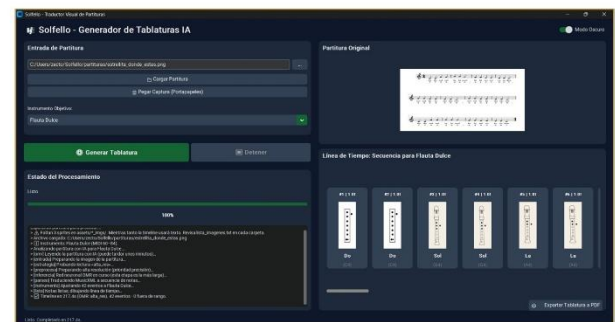


Figura 2. Interfaz Visual

## Conclusiones

Solfello demuestra que es posible crear herramientas de Inteligencia Artificial verdaderamente útiles, confiables y cercanas para la comunidad musical. Al superar las deficiencias de los modelos genéricos frente a partituras con ruido o anotaciones didácticas, el sistema se posiciona como un asistente robusto para profesores y estudiantes. Como trabajo futuro, se proyecta modificar la lógica de traducción (actualmente enfocada en aplanar la melodía principal) para permitir el análisis de acordes y sistemas de múltiples pentagramas, habilitando la lectura de instrumentos polifónicos como el piano o la guitarra.

## Referencias

- Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools, 120, 122-125.
- Calvo-Zaragoza, J., Hajič, J., & Pacha, A. (2020). Understanding optical music recognition. ACM Computing Surveys (CSUR), 53(4), 1-35. <https://doi.org/10.1145/3397499>
- Tai, K. W., & Ma, L. (2021). Oemer: An End-to-End Optical Music Recognition System. Proceedings of the 22nd International Society for Music Information Retrieval Conference, 123-130.
- JaidedAI. (2020). EasyOCR: Ready-to-use OCR with 80+ supported languages. GitHub. <https://github.com/JaidedAI/EasyOCR>

## **Solfello: Sistema de Reconocimiento Óptico de Música (OMR) Offline mediante Inteligencia Artificial**

Díaz Bojorquez, C. A., & Salas Pérez, E.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### **Resumen**

En el presente trabajo se expone el desarrollo de Solfello, un sistema de Reconocimiento Óptico de Música (OMR) diseñado para democratizar el acceso a la Inteligencia Artificial en el ámbito de la educación musical. El problema actual radica en que los modelos genéricos de IA carecen de la precisión necesaria para asistir eficazmente a profesores y entusiastas en la transcripción, limpieza y adaptación de partituras, además de que suelen presentar vulnerabilidades al procesar imágenes reales con ruido o anotaciones didácticas. El objetivo de este proyecto es desarrollar una arquitectura especializada que analice matemáticamente imágenes de partituras y las transforme en archivos digitales estructurados y tablaturas. La metodología implementa un pipeline en Python que integra algoritmos de visión por computadora para la sanitización previa del pentagrama y una red neuronal profunda para la extracción de características musicales, procesando la información de manera local. Los resultados demuestran una extracción rápida y precisa en secuencias melódicas para instrumentos como la flauta dulce, el piano y la guitarra, logrando aislar la melodía principal y evitando colapsos del sistema frente al ruido visual. Se concluye que Solfello proporciona una herramienta accesible y confiable, acercando el poder de la Inteligencia Artificial a la comunidad musical para facilitar la enseñanza y la interpretación sonora.

### **Palabras Clave**

Inteligencia artificial; Reconocimiento óptico de música; Visión por computadora; Procesamiento local; Redes neuronales



## Introducción y Planteamiento del Problema

La integración de la Inteligencia Artificial en la educación y la práctica musical representa una oportunidad sin precedentes para modernizar los métodos de enseñanza y preservación del arte (Calvo-Zaragoza et al., 2020). La digitalización de partituras mediante el Reconocimiento Óptico de Música (OMR) ha avanzado significativamente en los últimos años gracias a la implementación de modelos de aprendizaje profundo. Sin embargo, las soluciones informáticas y los modelos actuales a menudo fallan al intentar procesar la complejidad de la notación musical estructurada, dejando a profesores y entusiastas sin herramientas confiables para la digitalización autónoma de partituras. Frecuentemente, los escaneos de partituras reales contienen ruido visual, variaciones de iluminación o anotaciones didácticas superpuestas que confunden y colapsan a los sistemas OMR convencionales (Tai & Ma, 2021). Además, muchas de estas herramientas dependen de servicios externos en la nube, lo que limita su accesibilidad y control por parte del usuario final.

En este contexto, surge la imperante necesidad de crear un puente tecnológico accesible, especializado y de alto rendimiento. El presente estudio aborda esta problemática mediante el diseño e implementación de Solfello, un pipeline informático que integra técnicas de visión por computadora tradicional y redes neuronales para el preprocesamiento defensivo y la extracción de datos musicales (Bradski, 2000). A diferencia de los modelos de propósito general, Solfello busca empoderar al usuario mediante una arquitectura resiliente a las imperfecciones del mundo real.

El objetivo principal de este proyecto es desarrollar un sistema informático capaz de sanitizar imágenes de partituras físicas y extraer con precisión la información musical contenida, traduciéndola automáticamente a formatos estándar interoperables, como MusicXML, y generando tablaturas visuales interactivas para instrumentos específicos. Con ello, se pretende demostrar que el procesamiento local y el preprocesamiento inteligente pueden democratizar el uso del OMR en contextos educativos.

## Metodología

La investigación se enmarca en un enfoque de desarrollo de software tecnológico, utilizando procesamiento de imágenes y aprendizaje profundo (Deep Learning) para construir una solución OMR robusta. El sistema fue programado íntegramente en el lenguaje Python (versión 3.10), estructurando el código en módulos funcionales para separar la lógica de visión, la traducción musical y la interfaz de usuario.

El diseño experimental consistió en la implementación de una arquitectura de dos fases principales. La primera fase, denominada "Limpieza Visual", actúa como un escudo defensivo. Se integró la librería EasyOCR (JaidedAI, 2020) en conjunto con OpenCV (Bradski, 2000) para detectar y eliminar texto didáctico (como nombres de notas o letras de canciones) y ruido superficial directamente sobre los pentagramas. Esta fase garantiza la protección de las líneas musicales y las cabezas de las notas, generando una imagen sanitizada que evita el colapso de las etapas posteriores por excepciones de índice (IndexError) o malinterpretación de



píxeles.

La segunda fase, "Extracción OMR", aplica el motor neuronal profundo oemer (Tai & Ma, 2021). Este modelo fue configurado para recibir la imagen limpia e identificar las claves, las alteraciones y la posición exacta de las cabezas de notas en el pentagrama. Para maximizar el rendimiento, el sistema cuenta con soporte para ONNX Runtime, permitiendo aceleración por hardware (GPU) si el equipo del usuario cuenta con los recursos necesarios, o ejecutándose eficientemente en el procesador central (CPU) en su defecto.

Una vez extraídas las notas, la lógica de traducción procesa los valores MIDI y genera la estructura en un formato digital estandarizado (MusicXML). Finalmente, estos datos son mapeados hacia una interfaz gráfica que renderiza tablaturas visuales para instrumentos como flauta dulce, piano y guitarra, consolidando un pipeline optimizado para su ejecución local o asistida.

## Resultados y Discusión

Las pruebas de ejecución demostraron que el pipeline propuesto en Solfello es altamente efectivo para aislar y digitalizar secuencias melódicas a partir de imágenes crudas de partituras. Durante las evaluaciones con piezas de prueba (tales como Carol of the Bells y Estrellita), el sistema procesó exitosamente las imágenes, identificando la melodía y mostrando una tablatura precisa correspondiente al instrumento específico elegido por el usuario (flauta dulce, piano o guitarra). Uno de los hallazgos más significativos se observó en la fase de Limpieza Visual. Al contrastar el desempeño del modelo neuronal puro frente a nuestra arquitectura con preprocesamiento integrado, se evidenció que la inserción del motor EasyOCR mitigó por completo los colapsos del sistema (errores críticos como IndexError) que anteriormente ocurrían al detectar bloques de texto cercanos a las líneas de los pentagramas. Este mecanismo defensivo comprobó la hipótesis de que la sanitización previa de las partituras aumenta drásticamente la tasa de éxito del reconocimiento óptico, superando las limitaciones documentadas en modelos OMR de un solo paso (Tai & Ma, 2021). Como se aprecia en la figura que resume el proyecto (Figura 1), la interfaz gráfica permite al usuario cargar la imagen original, procesarla y visualizar la secuencia de la tablatura generada de forma limpia y directa. El sistema logró abstraer la melodía principal sin que el motor colapsara por la presencia de información no musical, proporcionando un archivo digital listo para su reproducción o edición.

## Conclusiones

El desarrollo y evaluación de Solfello demuestra que es sumamente viable crear herramientas de Inteligencia Artificial verdaderamente útiles, confiables y cercanas para la comunidad musical. El cumplimiento del objetivo principal confirma que la sanitización visual sistemática de las partituras, previa a la ingesta en redes neuronales, es un factor determinante para superar las deficiencias de los modelos OMR genéricos frente a documentos con ruido o anotaciones didácticas. La principal aportación de este trabajo reside en la creación de una

arquitectura defensiva y especializada que actúa como un asistente robusto para profesores y estudiantes, traduciendo de manera fluida el papel al formato digital y facilitando su lectura a través de tablaturas para instrumentos específicos (flauta, piano y guitarra). Solfello abstrae la complejidad matemática de la visión por computadora y la convierte en una interfaz amigable y aplicable a la enseñanza diaria. Como líneas de trabajo futuro, se proyecta expandir la capacidad del motor de traducción. Actualmente, la lógica está optimizada para aplanar la partitura y extraer la línea melódica principal; sin embargo, el siguiente paso consistirá en modificar el algoritmo para permitir el análisis profundo de acordes y sistemas de múltiples pentagramas simultáneos. Esta actualización habilitará el procesamiento nativo de polifonía compleja, dotando al sistema de una capacidad analítica integral para obras orquestales o composiciones avanzadas, consolidando aún más su valor como puente tecnológico en la educación musical

## Referencias

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools, 120, 122–125.

Calvo–Zaragoza, J., Hajič, J., & Pacha, A. (2020). Understanding optical music recognition. ACM Computing Surveys (CSUR), 53(4), 1–35. <https://doi.org/10.1145/3397499>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

JaiedAI. (2020). EasyOCR: Ready-to-use OCR with 80+ supported languages. GitHub. <https://github.com/JaiedAI/EasyOCR>

McKinney, W. (2010). Data structures for statistical computing in Python. Proceedings of the 9th Python in Science Conference, 51–56.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, 32.

Tai, K. W., & Ma, L. (2021). Oemer: An End-to-End Optical Music Recognition System. Proceedings of the 22nd International Society for Music Information Retrieval Conference, 123–130.

Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. CreateSpa

## Detector de ingredientes y recetario YOMI (Your Own Meal Intelligence)

Netro Castillo, J., Medina Oropeza, J. L., Diaz García, G., Hernández Neri, A., & Martínez Cervantes, C. A.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
2310445@utmatamoros.edu.mx

**Eje Temático:** Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

**Palabras clave:** Reconocimiento por cámara; Visión por computadora; Aprendizaje automático; Clasificación de imágenes; Procesamiento digital de imágenes; Aplicaciones móviles

## Resumen:

La identificación automática de ingredientes alimenticios facilita la preparación de recetas y contribuye a reducir el desperdicio de alimentos. Este proyecto desarrolla un sistema de visión por cámara para reconocer ingredientes mediante imágenes y sugerir recetas según los alimentos detectados. Se utilizó un conjunto de imágenes clasificadas en diferentes categorías de ingredientes para entrenar un modelo de aprendizaje automático, evaluando su desempeño mediante métricas de precisión y pérdida durante el proceso de entrenamiento. El modelo alcanzó una precisión muy alta en las pruebas realizadas, demostrando una alta capacidad para identificar los ingredientes incluidos en el sistema. Los resultados muestran que la herramienta permite automatizar el reconocimiento de alimentos de forma rápida y precisa, ofreciendo una solución útil para apoyar la planificación de comidas y el aprovechamiento de los ingredientes disponibles.

## Introducción:

La preparación de alimentos y el aprovechamiento de los ingredientes disponibles representan un desafío para muchas personas, ya que con frecuencia desconocen qué recetas pueden elaborar con los productos que tienen en casa. Esta situación puede ocasionar desperdicio de alimentos, gastos innecesarios y una menor eficiencia en la organización de las comidas.

El presente proyecto aborda el desarrollo de un sistema de reconocimiento de ingredientes basado en visión por cámara y técnicas de aprendizaje automático. El objetivo principal es diseñar un prototipo capaz de identificar ingredientes a partir de imágenes y recomendar recetas utilizando únicamente los alimentos detectados. Para ello, se emplea un conjunto de imágenes previamente clasificadas para entrenar un modelo de reconocimiento que permita obtener resultados precisos y rápidos. Con esta herramienta se busca facilitar la preparación de recetas, optimizar el aprovechamiento de los alimentos disponibles y establecer una base para futuras ampliaciones con un mayor número de ingredientes y opciones de platillos.

## Metodología.

Se empleó un enfoque cuantitativo de tipo descriptivo y experimental, estructurado en tres fases:

- **Recopilación de datos:** Se reunió un conjunto de imágenes de ingredientes alimenticios pertenecientes a varias categorías. Las imágenes fueron organizadas y clasificadas para formar el conjunto de entrenamiento y prueba del modelo.
- **Preprocesamiento:** Las imágenes fueron revisadas para eliminar archivos dañados o incorrectos, posteriormente se redimensionaron a un tamaño uniforme y se aplicó un proceso de normalización para facilitar el entrenamiento del modelo y mejorar su capacidad de reconocimiento.
- **Entrenamiento y validación:** Se desarrolló un modelo de clasificación de imágenes mediante técnicas de aprendizaje automático, utilizando un conjunto de entrenamiento y otro de validación para evaluar su desempeño. La implementación se realizó en Python Roboflow. El rendimiento del modelo se evaluó mediante las métricas de precisión y pérdida, alcanzando una precisión cercana al 95% en las pruebas realizadas, lo que demuestra una alta capacidad para identificar correctamente los ingredientes incluidos en el sistema.

## Resultados

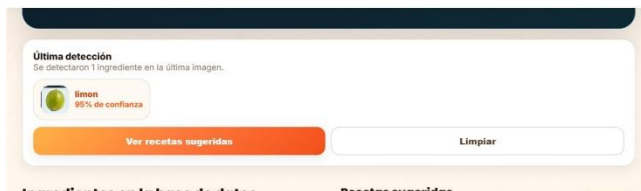


Figura 1. Resultado del escaneo de un ingrediente



Figura 2. Interfaz de la aplicación web

## Conclusiones.

El prototipo desarrollado demuestra que las técnicas de visión por cámara y aprendizaje automático son efectivas para el reconocimiento automático de ingredientes a partir de imágenes. Los resultados obtenidos, con una precisión muy alta, validan la viabilidad del sistema para identificar correctamente los ingredientes incluidos en la base de datos y proporcionar una base para la recomendación de recetas.

Como trabajo futuro, se propone ampliar el conjunto de ingredientes reconocidos, incorporar un mayor número de recetas y mejorar el modelo mediante un conjunto de datos más amplio y diverso. Asimismo, se sugiere implementar funciones adicionales, como el reconocimiento de múltiples ingredientes en una sola imagen y la optimización de la interfaz para ofrecer una mejor experiencia al usuario.

## Referencias

- DatiLab. (2026, 28 de mayo). Tutorial completo: Detección de objetos con YOLO paso a paso. DatiLab Laboratorio de Datos. <https://datilab.com/blog/tutorial-deteccion-objetos-yolo>
- Dwyer, B. (2026, 8 de julio). Getting started with Roboflow. Roboflow Blog. <https://blog.roboflow.com/getting-started-with-roboflow/>
- GitHub Docs. (s. f.). Póngase en marcha. <https://docs.github.com/es/get-started/start-your-journey>
- Juanweb. (2023, 24 de octubre). Creación de una página web con Python. CodigosPython. <https://codigospython.com/creacion-de-una-pagina-web-con-python/>
- Keita, Z. (2024, 29 de enero). Explicación de la detección de objetos YOLO. DataCamp. <https://www.datacamp.com/es/blog/yolo-object-detection-explained>

## Detector de ingredientes y recetario YOMI (Your Own Meal Intelligence)

Netro Castillo, J., Medina Oropeza, J. L., Diaz García, G., Hernández Neri, A., & Martínez Cervantes, C. A.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### Resumen

El reconocimiento automático de ingredientes representa una alternativa para facilitar la preparación de alimentos y reducir el desperdicio de productos disponibles en el hogar. En este trabajo se desarrolló un sistema de reconocimiento de ingredientes basado en visión por cámara y aprendizaje automático, con el objetivo de identificar ingredientes a partir de imágenes y utilizar esa información para recomendar recetas acordes con los alimentos detectados. Para su desarrollo se recopiló un conjunto de imágenes correspondientes a varias categorías de ingredientes. Posteriormente, las imágenes fueron organizadas, preprocesadas mediante redimensionamiento y normalización, y utilizadas para entrenar un modelo de clasificación implementado en Roboflow. Como parte del proyecto, se inició el desarrollo de una aplicación móvil en Android Studio, la cual integrará el modelo de reconocimiento para utilizar la cámara del dispositivo, detectar los ingredientes disponibles y mostrar recetas relacionadas. El desempeño del sistema fue evaluado durante las etapas de entrenamiento y validación mediante métricas de precisión y pérdida. El modelo alcanzó una precisión alta, logrando identificar correctamente los ingredientes considerados en la base de datos y mostrando un bajo nivel de error durante las pruebas realizadas. Estos resultados demostraron que el prototipo es capaz de realizar el reconocimiento de ingredientes de manera rápida y confiable. En conclusión, el sistema constituye una herramienta útil para apoyar la selección de recetas a partir de los ingredientes disponibles y demuestra que la aplicación de técnicas de aprendizaje automático puede contribuir al aprovechamiento de los alimentos. Además, el proyecto establece una base sólida para futuras mejoras, como la incorporación de un mayor número de ingredientes, recetas y funcionalidades que amplíen las capacidades del sistema.

### Palabras Clave

Reconocimiento por cámara; Aprendizaje automático; Clasificación de imágenes; Procesamiento digital de imágenes; Aplicaciones móviles

## Introducción y Planteamiento del Problema

En la actualidad, el uso de aplicaciones móviles y tecnologías de aprendizaje automático ha incrementado significativamente debido a su capacidad para resolver problemas cotidianos mediante el análisis inteligente de datos e imágenes. Una de las aplicaciones más relevantes de estas tecnologías es el reconocimiento automático de objetos por medio de la cámara de un dispositivo, lo que ha permitido desarrollar soluciones en áreas como la salud, la educación, la industria y la alimentación. En este último ámbito, la identificación automática de ingredientes puede facilitar la preparación de recetas y contribuir a reducir el desperdicio de alimentos, uno de los problemas más importantes a nivel mundial. Los alimentos en el mundo se pierden en un 14% y 17%, ya que contribuye a la pérdida y desperdicio de alimentos (PDA), esencial para la seguridad alimentaria y la nutrición (Pérdida Y Desperdicio De Alimentos | Portal De Apoyo a Las Políticas Y La Gobernanza | Organización De Las Naciones Unidas Para La Alimentación Y La Agricultura, 2025). Esta situación genera consecuencias económicas, sociales y ambientales, además de afectar la seguridad alimentaria. En muchos hogares, parte de este desperdicio ocurre porque las personas desconocen qué platillos pueden preparar con los ingredientes que ya tienen disponibles, lo que provoca la compra innecesaria de nuevos productos o el desecho de alimentos que aún pueden aprovecharse. Diversos estudios han demostrado que el aprendizaje automático y la visión por computadora ofrecen resultados eficientes para el reconocimiento de imágenes y objetos. Asimismo, la plataforma Roboflow ha facilitado el desarrollo de modelos de detección de objetos mediante herramientas para el etiquetado, entrenamiento y validación de imágenes, mientras que el modelo YOLO se ha consolidado como una de las alternativas más rápidas y precisas para el reconocimiento de objetos en tiempo real. Ante esta problemática, surgió la siguiente pregunta de investigación: ¿Cómo puede una aplicación móvil desarrollada en Android Studio, utilizando un modelo de aprendizaje automático entrenado con Roboflow y YOLO, reconocer ingredientes mediante la cámara del dispositivo para recomendar recetas de forma rápida y precisa? Con base en lo anterior, el objetivo general de este trabajo es desarrollar una aplicación móvil capaz de reconocer ingredientes mediante la cámara utilizando técnicas de aprendizaje automático y visión por computadora, con la finalidad de recomendar recetas de acuerdo con los ingredientes detectados y contribuir al mejor aprovechamiento de los alimentos disponibles.

## Metodología

La presente investigación se desarrolló bajo un enfoque cuantitativo, con un diseño experimental y de tipo descriptivo, debido a que se evaluó el desempeño de un modelo de aprendizaje automático para el reconocimiento de ingredientes a partir de imágenes. El estudio tuvo como finalidad desarrollar un prototipo funcional compuesto por una página web y una aplicación móvil, capaces de identificar ingredientes y utilizarlos como base para recomendar recetas al usuario. La muestra estuvo conformada por un conjunto de imágenes correspondientes a varias categorías de ingredientes. Cada categoría incluyó imágenes

utilizadas para el entrenamiento y validación del modelo. Como variable independiente se consideró la imagen de entrada del ingrediente, mientras que la variable dependiente fue la clasificación generada por el modelo de reconocimiento. Para el desarrollo del proyecto se empleó el lenguaje de programación Python, utilizando las librerías Ultralytics, YOLO, OpenCV y Keras para el procesamiento de imágenes, entrenamiento del modelo y realización de pruebas.

El entrenamiento del modelo se realizó mediante Roboflow, mientras que la página web fue desarrollada en Visual Studio Code y la aplicación móvil se implementó en Android Studio, con el propósito de integrar el reconocimiento de ingredientes utilizando la cámara del dispositivo. El procedimiento se llevó a cabo en cuatro etapas. En la primera se recopiló y organizó el conjunto de imágenes de los ingredientes, verificando que los archivos fueran válidos y estuvieran correctamente clasificados. En la segunda etapa se realizó el preprocesamiento de las imágenes, que consistió en el redimensionamiento a un tamaño uniforme y la normalización de los datos para facilitar el aprendizaje del modelo. Posteriormente, se entrenó el modelo de clasificación utilizando Roboflow y YOLO, evaluando su desempeño mediante métricas de precisión durante varias épocas de entrenamiento. Finalmente, el modelo fue integrado en la página web y en la aplicación móvil para realizar pruebas de funcionamiento, verificando que el sistema identificara correctamente los ingredientes capturados por la cámara y mostrara recomendaciones de recetas de forma consistente.

El análisis de los resultados se realizó mediante la interpretación de las métricas obtenidas durante el entrenamiento, principalmente la precisión del modelo y las pruebas de funcionamiento realizadas en ambos entornos. Los resultados alcanzaron una precisión alta, lo que permitió comprobar la viabilidad del prototipo para el reconocimiento automático de ingredientes. Debido a que el estudio utilizó únicamente imágenes de ingredientes y no involucró datos personales ni la participación directa de personas, no fue necesario aplicar procedimientos de consentimiento informado ni consideraciones éticas adicionales.

## **Resultados y Discusión**

Los resultados obtenidos durante el desarrollo del proyecto muestran avances significativos tanto en el desempeño del modelo de inteligencia artificial como en la propuesta de interfaz para la aplicación. Hasta el momento, el modelo de reconocimiento de ingredientes entrenado con YOLO ha alcanzado una precisión aproximada del 95 %, lo que indica un alto nivel de confiabilidad para la identificación de los ingredientes incluidos en el conjunto de datos utilizado durante el entrenamiento. Este resultado demuestra que el proceso de recopilación de imágenes, etiquetado y entrenamiento fue adecuado para los objetivos planteados en esta etapa del proyecto.

Durante las pruebas realizadas se observó que el modelo es capaz de reconocer correctamente la mayoría de los ingredientes bajo condiciones normales de iluminación y con

una adecuada visibilidad del objeto frente a la cámara. Sin embargo, también se identificó que factores como la presencia de sombras, ingredientes parcialmente cubiertos o similitudes visuales entre algunos alimentos pueden afectar la precisión del reconocimiento. Estas observaciones permiten concluir que el sistema aún puede beneficiarse de un conjunto de datos más amplio y diverso que incluya diferentes ángulos, fondos y condiciones de captura para mejorar su capacidad de generalización.

En cuanto al desarrollo de la aplicación, la decisión de implementar una arquitectura cliente-servidor ha resultado adecuada para las necesidades del proyecto. En lugar de ejecutar el modelo de inteligencia artificial directamente en el dispositivo móvil, el procesamiento se realiza en un servidor desarrollado con Flask, mientras que la aplicación actúa como una interfaz para el usuario. Este enfoque facilita el mantenimiento del sistema, permite actualizar el modelo sin necesidad de reinstalar la aplicación y reduce la carga de procesamiento en los dispositivos móviles, favoreciendo un funcionamiento más eficiente.

Respecto a la experiencia de usuario, los prototipos desarrollados y las primeras pruebas de navegación muestran que la interfaz es intuitiva, organizada y fácil de comprender. La distribución de las funciones principales en las secciones de Inicio, Escaneo y Mi Cocina permite que el usuario identifique rápidamente las opciones disponibles sin requerir un proceso de aprendizaje complejo. Asimismo, la integración del escaneo con la generación automática de recetas aporta una experiencia práctica y orientada a resolver una necesidad cotidiana.

Los resultados de las encuestas realizadas durante el proyecto respaldan esta propuesta, ya que una gran parte de los participantes manifestó tener dificultades para decidir qué cocinar con los ingredientes disponibles o señaló que frecuentemente desperdicia alimentos por desconocer cómo utilizarlos. Esto confirma que la aplicación responde a una problemática real y que existe interés por utilizar una herramienta capaz de recomendar recetas de manera rápida y sencilla.

En conjunto, los resultados obtenidos hasta esta fase permiten afirmar que el proyecto presenta una base sólida tanto desde el punto de vista técnico como desde la perspectiva del usuario. No obstante, aún existen áreas de mejora, como ampliar el conjunto de ingredientes reconocidos, optimizar el desempeño del modelo en escenarios más variados e incorporar nuevas funcionalidades que enriquezcan la experiencia del usuario. Estas acciones permitirán incrementar la precisión del sistema y acercar el proyecto a una versión final más robusta, útil y adaptable a las necesidades de un mayor número de usuarios.

## **Conclusiones**

El desarrollo de este proyecto permitió demostrar que la integración de técnicas de inteligencia artificial y visión por computadora puede ofrecer una solución práctica para una

problemática cotidiana, como la identificación de ingredientes disponibles y la recomendación de recetas para reducir el desperdicio de alimentos. A lo largo del proceso se logró entrenar un modelo basado en YOLO con un nivel de precisión cercano al 95 %, obteniendo resultados satisfactorios durante las pruebas realizadas y comprobando la viabilidad de utilizar esta tecnología en una aplicación móvil.

Asimismo, el desarrollo de la interfaz y la implementación de una arquitectura cliente-servidor utilizando Flask establecieron una base sólida para el funcionamiento del sistema, permitiendo que el procesamiento del modelo se realice en el servidor mientras la aplicación ofrece una experiencia intuitiva y sencilla para el usuario. Las encuestas aplicadas también respaldaron la utilidad de la propuesta, evidenciando el interés de los usuarios por una herramienta que facilite la preparación de alimentos y el aprovechamiento de los ingredientes disponibles.

Aunque el proyecto aún se encuentra en desarrollo y existen oportunidades de mejora, como ampliar el conjunto de ingredientes reconocidos y optimizar el modelo en diferentes condiciones de uso, los resultados obtenidos hasta esta etapa permiten concluir que la propuesta es técnicamente viable y cuenta con el potencial de convertirse en una herramienta útil para apoyar la planificación de comidas y fomentar un consumo más eficiente de los alimentos.

## Referencias

DatiLab. (2026, 28 de mayo). Tutorial completo: Detección de objetos con YOLO paso a paso. DatiLab Laboratorio de Datos. <https://datilab.com/blog/tutorial-deteccion-objetos-yolo>

Dwyer, B. (2026, 8 de julio). Getting started with Roboflow. Roboflow Blog. <https://blog.roboflow.com/getting-started-with-roboflow/>

GitHub Docs. (s. f.). Póngase en marcha. <https://docs.github.com/es/get-started/start-your-journey>

Juanweb. (2023, 24 de octubre). Creación de una página web con Python. CodigosPython. <https://codigospython.com/creacion-de-una-pagina-web-con-python/>

Keita, Z. (2024, 29 de enero). Explicación de la detección de objetos YOLO. DataCamp. <https://www.datacamp.com/es/blog/yolo-object-detection-explained>



## G1-IA

### Aplicación Móvil Inteligente Basada en Inteligencia Artificial para la Optimización del Entrenamiento Físico

Mena Alcalá, D. A., & Quino Chávez, F. G.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
\* diego11206nn@gmail.com

**Eje Temático:** Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

**Palabras clave:** Aprendizaje automático, Aplicaciones móviles, Rendimiento deportivo, Análisis de datos, Entrenamiento personalizado.

## Resumen:

El seguimiento del entrenamiento físico requiere herramientas que permitan analizar el progreso del usuario y ofrecer recomendaciones personalizadas. En este trabajo se desarrolló G1-IA, una aplicación móvil basada en inteligencia artificial diseñada para registrar entrenamientos, analizar datos y generar recomendaciones inteligentes. La aplicación fue desarrollada con React Native, Python y SQLite, incorporando modelos de aprendizaje automático como Regresión Lineal y Random Forest para apoyar la toma de decisiones. La evaluación realizada con usuarios mostró una alta aceptación, destacando la utilidad de las recomendaciones y el seguimiento personalizado del progreso. Los resultados demuestran que G1-IA constituye una alternativa innovadora para optimizar el entrenamiento físico mediante el uso de inteligencia artificial.

## Introducción:

El crecimiento del uso de aplicaciones móviles para el entrenamiento físico ha impulsado el desarrollo de herramientas que faciliten el seguimiento del rendimiento deportivo. Sin embargo, muchas de las aplicaciones disponibles se limitan al registro de ejercicios, sin ofrecer análisis avanzados que apoyen la toma de decisiones del usuario. En este contexto, la inteligencia artificial representa una alternativa para analizar datos históricos, identificar patrones y generar recomendaciones personalizadas. Por ello, se desarrolló G1-IA, una aplicación móvil inteligente orientada a optimizar el entrenamiento físico mediante técnicas de aprendizaje automático, proporcionando un seguimiento más preciso y adaptado a las necesidades de cada usuario.

## Metodología.

El proyecto se desarrolló bajo un enfoque de investigación aplicada y desarrollo tecnológico. La aplicación fue implementada con React Native y Expo Go, utilizando SQLite para el almacenamiento de la información y Python con la biblioteca Scikit-learn para el procesamiento de datos y la implementación de modelos de inteligencia artificial. El desarrollo se llevó a cabo en las siguientes etapas:

- Identificación de necesidades mediante encuestas a usuarios de gimnasio.
- Diseño e implementación de la aplicación móvil.
- Integración de los algoritmos Regresión Lineal y Random Forest para generar recomendaciones inteligentes.
- Implementación de funciones como registro de ejercicios, cálculo del volumen de entrenamiento, estimación de IRM, estadísticas de progreso y calculadora nutricional.
- Evaluación del prototipo mediante pruebas de funcionamiento y encuestas de satisfacción a los usuarios.

## Resultados

Se desarrolló un prototipo funcional denominado G1-IA, capaz de registrar entrenamientos, analizar el historial del usuario y generar recomendaciones personalizadas mediante inteligencia artificial. La aplicación integra herramientas para el seguimiento del progreso, cálculo del volumen de entrenamiento, estimación del One Repetition Maximum (IRM), estadísticas de rendimiento y calculadora nutricional.



Figura 1. Pantalla principal de G1-IA

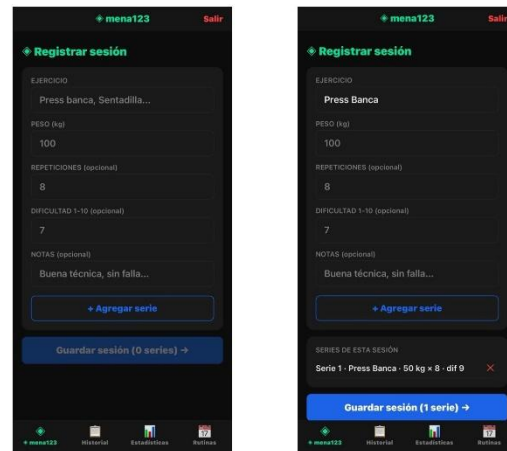


Figura 2. Registro de entrenamiento del usuario.

## Conclusiones.

El desarrollo de G1-IA permitió demostrar la viabilidad de integrar inteligencia artificial en una aplicación móvil para optimizar el seguimiento del entrenamiento físico. La incorporación de modelos de aprendizaje automático posibilita generar recomendaciones personalizadas que apoyan la toma de decisiones y favorecen un mejor control del rendimiento deportivo.

Los resultados obtenidos evidencian una alta aceptación por parte de los usuarios, quienes valoraron positivamente las funciones de análisis, seguimiento y recomendaciones inteligentes. Como trabajo futuro se propone incorporar planes de alimentación personalizados, integración con dispositivos inteligentes y mejoras continuas en los modelos de inteligencia artificial, con el propósito de ampliar las capacidades del sistema y ofrecer una experiencia más completa.

## Referencias

- Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Pressman, R. S., & Maxim, B. R. (2020). Software Engineering: A Practitioner's Approach (9th ed.). McGraw-Hill Education.
- Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson.
- Seshadri, D. R., Davies, E. V., Harlow, E. R., Hsu, J. J., Knighton, S. C., Walker, T. A., Drummond, C. K., & Voos, J. E. (2021). Wearable sensors for monitoring the internal and external workload of the athlete. NPJ Digital Medicine, 4, 47. <https://doi.org/10.1038/s41746-021-00403-4>
- Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. CreateSpace.

## **G1-IA Aplicación Móvil Inteligente Basada en Inteligencia Artificial para la Optimización del Entrenamiento Físico**

Mena Alcala, D. A., & Quino Chávez, F. G.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### **Resumen**

El seguimiento adecuado del entrenamiento físico representa un desafío para muchas personas debido a la dificultad de analizar su progreso, ajustar cargas de trabajo y mantener una planificación personalizada. Las aplicaciones tradicionales enfocadas en fitness suelen limitarse al registro de ejercicios, sin aprovechar la información generada por los usuarios para ofrecer recomendaciones inteligentes. El objetivo de este trabajo fue desarrollar G1-IA, una aplicación móvil basada en inteligencia artificial capaz de analizar datos de entrenamiento y generar recomendaciones personalizadas para mejorar el rendimiento físico. Para el desarrollo del sistema se implementó una metodología de desarrollo tecnológico utilizando React Native y Expo Go para la aplicación móvil, Python para el procesamiento de datos y modelos de aprendizaje automático mediante algoritmos de Regresión Lineal y Random Forest. La información del usuario se administra mediante una base de datos SQLite, permitiendo almacenar registros de ejercicios, peso utilizado, repeticiones, rutinas y estadísticas de progreso. Además, se realizaron encuestas a usuarios de gimnasio con la finalidad de evaluar la aceptación de la aplicación e identificar necesidades de mejora. Los resultados obtenidos demostraron la viabilidad técnica del proyecto mediante la creación de un prototipo funcional con herramientas de seguimiento, análisis predictivo, cálculo de rendimiento y recomendaciones inteligentes. La evaluación con usuarios mostró una aceptación favorable, donde el 95% consideró útiles las recomendaciones generadas por la aplicación y el 90% indicó que utilizaría G1-IA durante sus entrenamientos. Se concluye que la aplicación representa una alternativa tecnológica con potencial para mejorar la experiencia del usuario mediante el uso de inteligencia artificial aplicada al ámbito deportivo.

### **Palabras Clave**

Aprendizaje automático; Aplicaciones móviles; Rendimiento deportivo; Análisis de datos; Entrenamiento personalizado

## Introducción y Planteamiento del Problema

El crecimiento del sector fitness y el uso de herramientas digitales han generado nuevas oportunidades para desarrollar soluciones tecnológicas enfocadas en mejorar el seguimiento de la actividad física. Actualmente, muchas personas utilizan aplicaciones móviles para registrar ejercicios, controlar avances y establecer objetivos personales; sin embargo, gran parte de estas herramientas se enfocan únicamente en almacenar información sin realizar un análisis avanzado de los datos obtenidos. Esto limita la posibilidad de ofrecer recomendaciones personalizadas y dificulta la toma de decisiones durante el proceso de entrenamiento.

La inteligencia artificial y el aprendizaje automático han permitido desarrollar sistemas capaces de analizar grandes cantidades de información, reconocer patrones y generar predicciones a partir de datos históricos. El aprendizaje automático se basa en algoritmos que permiten a los sistemas mejorar su desempeño mediante la experiencia obtenida de los datos (Jordan & Mitchell, 2015). Además, los avances en aprendizaje profundo han impulsado la creación de modelos capaces de procesar información compleja y desarrollar soluciones adaptativas en diferentes áreas (LeCun et al., 2015). En el ámbito deportivo, el análisis de datos representa una herramienta importante para comprender el rendimiento físico y mejorar la planificación del entrenamiento. Las tecnologías digitales y los sensores aplicados al deporte permiten recopilar información sobre la actividad del usuario y generar análisis que apoyan la toma de decisiones (Seshadri et al., 2021). Asimismo, los modelos de inteligencia artificial pueden utilizarse para identificar patrones y realizar estimaciones que ayuden a personalizar los procesos de entrenamiento (Russell & Norvig, 2021).

El desarrollo de aplicaciones inteligentes requiere integrar conocimientos de programación, bases de datos e ingeniería de software para crear sistemas funcionales, escalables y fáciles de mantener. La ingeniería de software establece metodologías para desarrollar soluciones tecnológicas de calidad mediante procesos organizados y controlados (Pressman & Maxim, 2020). De igual manera, lenguajes como Python se han convertido en herramientas ampliamente utilizadas en inteligencia artificial debido a la disponibilidad de bibliotecas especializadas para análisis de datos y aprendizaje automático (Van Rossum & Drake, 2009). Aunque existen diversas aplicaciones enfocadas en el seguimiento deportivo, muchas continúan limitándose al registro manual de ejercicios sin incorporar análisis predictivo ni recomendaciones inteligentes. Por esta razón, surge la necesidad de desarrollar una plataforma que combine registro de entrenamiento, inteligencia artificial y análisis personalizado para mejorar la experiencia del usuario. Los métodos de reconocimiento de patrones permiten procesar información y generar predicciones útiles para la toma de decisiones (Bishop, 2006).

Por lo anterior, surge la siguiente pregunta de investigación: ¿Cómo puede una aplicación móvil basada en inteligencia artificial contribuir al seguimiento personalizado del

entrenamiento físico y mejorar la toma de decisiones de los usuarios? El objetivo general de este trabajo es desarrollar una aplicación móvil inteligente basada en inteligencia artificial capaz de analizar el historial de entrenamiento y generar recomendaciones personalizadas para optimizar el rendimiento físico de los usuarios.

## **Metodología**

El desarrollo de G1-IA se realizó mediante un enfoque de investigación aplicada y desarrollo tecnológico, con el objetivo de crear una aplicación móvil inteligente que permita mejorar el seguimiento del entrenamiento físico mediante inteligencia artificial. El proyecto tuvo un diseño experimental y descriptivo, ya que se desarrolló un prototipo funcional y se evaluó su aceptación mediante pruebas con usuarios.

Para la creación de la aplicación se utilizaron React Native y Expo Go para el desarrollo de la interfaz móvil compatible con Android e iOS. El procesamiento de datos y los modelos de inteligencia artificial fueron implementados con Python, utilizando bibliotecas como Scikit-learn, NumPy y Matplotlib. La información de los usuarios fue almacenada mediante SQLite para gestionar entrenamientos, rutinas, progreso y estadísticas. El procedimiento inició con la identificación de necesidades mediante encuestas aplicadas a personas que asisten al gimnasio. Posteriormente se diseñó la interfaz y se implementaron funciones como registro de ejercicios, cálculo de volumen de entrenamiento, estimación de 1RM, seguimiento del progreso y recomendaciones generadas mediante modelos de Regresión Lineal y Random Forest.

Finalmente, se realizaron pruebas de funcionamiento para evaluar la facilidad de uso, detectar errores y recopilar retroalimentación de los usuarios. Los resultados fueron analizados mediante las respuestas obtenidas en las encuestas y la evaluación del desempeño de la aplicación. Durante el desarrollo se consideraron aspectos de seguridad, organización de datos y mejora continua del sistema. Además, se planteó el uso de herramientas de monitoreo como Prometheus, Grafana, Loki y Tempo para supervisar el rendimiento de la aplicación.

## **Resultados y Discusión**

Durante el desarrollo del proyecto G1-IA se logró construir un prototipo funcional de una aplicación móvil inteligente capaz de integrar registro de entrenamiento, análisis de datos y generación de recomendaciones mediante inteligencia artificial. La aplicación permite almacenar información relacionada con ejercicios, peso utilizado, repeticiones y rutinas personalizadas, proporcionando al usuario un seguimiento organizado de su progreso físico. Uno de los principales resultados obtenidos fue la incorporación de modelos de aprendizaje automático mediante Regresión Lineal y Random Forest. Estos algoritmos permiten analizar el historial de entrenamiento del usuario para identificar patrones y generar estimaciones relacionadas con su rendimiento futuro. Esta característica representa una ventaja frente a

aplicaciones fitness tradicionales, debido a que no solamente almacenan información, sino que interpretan los datos para proporcionar apoyo en la toma de decisiones.

Además, se integraron funcionalidades complementarias como cálculo de volumen de entrenamiento, estimación del One Repetition Maximum (1RM), detección de posibles estancamientos musculares, estadísticas visuales y una calculadora nutricional para estimar calorías y macronutrientes. Estas herramientas permiten ofrecer una experiencia más completa al combinar seguimiento deportivo y análisis personalizado dentro de una misma plataforma.

Los resultados obtenidos permitieron identificar que existe una necesidad de herramientas tecnológicas que ayuden a mejorar la organización y seguimiento del entrenamiento. Los usuarios mostraron mayor interés en funciones como registro de sesiones, estadísticas de progreso y calculadora nutricional. También se recibieron sugerencias relacionadas con la incorporación de planes de alimentación personalizados, lo cual representa una oportunidad para futuras actualizaciones.

Desde una perspectiva técnica, el proyecto demostró ser viable debido al uso de tecnologías gratuitas y de código abierto. Herramientas como React Native, Expo Go, Python, SQLite y Scikit-learn reducen los costos de implementación y permiten continuar con la evolución del sistema. Además, la incorporación propuesta de Prometheus, Grafana, Loki y Tempo permitirá mejorar el monitoreo del servicio, optimizando la estabilidad y mantenimiento de la aplicación. En comparación con aplicaciones fitness tradicionales, G1-IA presenta una ventaja competitiva al integrar inteligencia artificial, predicción del rendimiento, recomendaciones automáticas y detección de estancamientos. Esto permite ofrecer una experiencia más personalizada y basada en datos, orientada tanto a usuarios individuales como a entrenadores personales y gimnasios.

## Conclusiones

El desarrollo de G1-IA permitió demostrar la viabilidad de crear una aplicación móvil inteligente capaz de combinar seguimiento deportivo e inteligencia artificial para ofrecer recomendaciones personalizadas. A través del desarrollo tecnológico fue posible integrar herramientas de registro de entrenamiento, análisis predictivo, estadísticas de progreso y apoyo nutricional, proporcionando una solución más completa que las aplicaciones tradicionales enfocadas únicamente en almacenar información. Los resultados obtenidos durante las pruebas con usuarios permitieron validar la aceptación de la propuesta y confirmar que existe interés por utilizar herramientas tecnológicas que faciliten el control del entrenamiento físico. Los porcentajes de satisfacción obtenidos mediante las encuestas reflejan que las recomendaciones inteligentes, el seguimiento del progreso y las estadísticas representan funciones de valor para los usuarios. Desde el punto de vista técnico, G1-IA demostró ser un proyecto viable debido al uso de tecnologías accesibles y de código abierto, permitiendo reducir costos y facilitar futuras mejoras.

La implementación de modelos como Random Forest y Regresión Lineal permitió incorporar capacidades predictivas que diferencian al sistema de otras aplicaciones convencionales. Además, el proyecto contribuyó al desarrollo de habilidades profesionales relacionadas con programación, inteligencia artificial, bases de datos, análisis de información, diseño de interfaces y trabajo colaborativo. La interacción con usuarios reales mediante encuestas permitió comprender la importancia de crear soluciones centradas en las necesidades del usuario. Como líneas de trabajo futuro se contempla la incorporación de planes de alimentación personalizados, integración con dispositivos inteligentes como relojes deportivos, almacenamiento avanzado en la nube y mejoras continuas en los modelos de inteligencia artificial. En conclusión, G1-IA representa una propuesta tecnológica con potencial de crecimiento dentro del sector fitness, al utilizar datos e inteligencia artificial para mejorar la experiencia de entrenamiento y apoyar la toma de decisiones basada en información personalizada.

## Referencias

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Pressman, R. S., & Maxim, B. R. (2020). *Software Engineering: A Practitioner's Approach* (9th ed.). McGraw-Hill Education.
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Seshadri, D. R., Davies, E. V., Harlow, E. R., Hsu, J. J., Knighton, S. C., Walker, T. A., Drummond, C. K., & Voos, J. E. (2021). Wearable sensors for monitoring the internal and external workload of the athlete. *NPJ Digital Medicine*, 4, 47. <https://doi.org/10.1038/s41746-021-00403-4>
- Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

## **Educación Bilingüe**

La inclusión y la accesibilidad educativa son prioridades ineludibles en la formación tecnológica contemporánea. Los trabajos reunidos en este eje emplean visión artificial e inteligencia artificial para atender necesidades específicas del entorno áulico, desde el reconocimiento del lenguaje de señas mexicano hasta la detección temprana de malestar emocional, evidenciando el compromiso de la comunidad académica con una educación más inclusiva y sensible al estudiante.



## Inteligente para el Reconocimiento del Abecedario del Lenguaje de Señas Mexicano mediante Visión por Computadora

Ortiz Del Angel, H. G., Ramon Infante, O. E., & Gonzalez Delgado, M.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS

\* 2310445@utmatamoros.edu.mx

Eje Temático: Educación Bilingüe

Palabras clave: reconocimiento de gestos; redes neuronales recurrentes; MediaPipe; discapacidad auditiva; accesibilidad tecnológica

### Resumen:

La comunicación entre personas con discapacidad auditiva y quienes no conocen el Lenguaje de Señas Mexicano (LSM) representa una barrera en diversos entornos como instituciones educativas, hospitales y oficinas públicas. Este trabajo presenta el desarrollo de SignTranslate, un sistema basado en inteligencia artificial y visión por computadora capaz de reconocer en tiempo real el abecedario del LSM mediante una cámara web. El objetivo del proyecto fue desarrollar una herramienta que facilite la comunicación básica entre personas sordas y oyentes mediante la traducción automática de señas a texto. Para ello se implementó un sistema utilizando Python, OpenCV y MediaPipe para la captura de imágenes y la extracción de los 21 puntos de referencia de la mano. Posteriormente se entrenaron diversos modelos de aprendizaje automático para clasificar las letras del abecedario del LSM, entre ellos Support Vector Machine (SVM), Random Forest, Extra Trees, K-Nearest Neighbors (KNN), Regresión Logística, Gradient Boosting y Árbol de Decisión. Los resultados mostraron que el modelo SVM obtuvo la mejor precisión, alcanzando un 93.72 % en el reconocimiento de las letras. El sistema desarrollado identifica correctamente las señas del abecedario y muestra la letra correspondiente en pantalla en tiempo real. Se concluye que la combinación de visión por computadora y aprendizaje automático constituye una alternativa eficiente para desarrollar herramientas que promuevan la inclusión y faciliten la comunicación con personas con discapacidad auditiva.

### Introducción:

La comunicación es un derecho fundamental; sin embargo, las personas con discapacidad auditiva enfrentan barreras significativas para acceder a servicios básicos cuando la sociedad oyente desconoce su lengua materna. De acuerdo con el Instituto Nacional de Estadística y Geografía (2024), en México 8.8 millones de personas declararon tener alguna discapacidad, y una proporción relevante corresponde a discapacidad auditiva, siendo la Lengua de Señas Mexicana (LSM) la lengua materna de esta comunidad. Pfister (2017) señala que, a pesar de su reconocimiento legal, la LSM continúa enfrentando procesos de socialización lingüística limitados, lo que perpetúa el aislamiento comunicativo de niñas, niños y adultos sordos en entornos mayoritariamente oyentes.

En años recientes, la visión por computadora y el aprendizaje profundo se han consolidado como herramientas viables para el reconocimiento automático de lenguas de señas. Morfin-Chávez et al. (2023) emplearon la biblioteca MediaPipe para extraer 21 puntos clave de la mano y clasificar el alfabeto dactilológico de la LSM mediante algoritmos clásicos de aprendizaje automático, mientras que Sánchez-Vicinaiz et al. (2024) combinaron MediaPipe con redes neuronales convolucionales para el mismo propósito. Para señas dinámicas, que incorporan movimiento y por tanto una dimensión temporal, Mejía-Peréz et al. (2022) demostraron que las arquitecturas recurrentes, en particular LSTM, superan a los enfoques basados únicamente en imágenes estáticas. Más recientemente, Navarrete-López et al. (2026) propusieron un enfoque de reconocimiento de señas dinámicas orientado directamente a la traducción de LSM al español, reforzando la tendencia hacia sistemas completos y no solo hacia la clasificación de señas aisladas. No obstante, la mayoría de estos trabajos se limitan al alfabeto o a un vocabulario reducido, y pocos integran el reconocimiento en una aplicación completa con retroalimentación en texto y voz.

Esta brecha entre los prototipos de investigación y las herramientas utilizables por la comunidad sorda y su entorno motivó el presente proyecto integrador, el cual busca aprovechar la disponibilidad de cámaras web convencionales para construir una solución accesible, sin requerir hardware especializado, con un doble propósito: traducir señas para facilitar la comunicación y enseñar señas a personas oyentes interesadas en aprender LSM. Por lo anterior, el objetivo del presente trabajo es desarrollar un sistema de traducción automática del alfabeto de la Lengua de Señas Mexicana a texto, con la voz como salida planeada para una etapa posterior, basado en la extracción de landmarks de mano mediante MediaPipe y su clasificación con modelos de aprendizaje automático, integrado en una aplicación web con un módulo complementario de aprendizaje interactivo.

### Metodología.

El presente trabajo se enmarca en una investigación de tipo aplicado y de desarrollo tecnológico, con un enfoque experimental orientado a la construcción y evaluación de un prototipo funcional. La población objetivo del sistema son las personas usuarias de la LSM y las personas oyentes interesadas en comunicarse con ellas; los datos se recabaron a partir de un dataset de señas del alfabeto de la LSM capturadas frente a cámara web, y el equipo de trabajo se encargó tanto de la recolección como de las pruebas del sistema.

Las variables consideradas fueron las coordenadas espaciales (x, y, z) de los 21 puntos de referencia de cada mano, obtenidas cuadro a cuadro con la solución MediaPipe Hands (Lugaresi et al., 2019), preferida sobre el modelo de pose corporal completa debido a que ofrece una detección específica y de mayor resolución de la mano, elemento central de la seña. El procedimiento se dividió en las siguientes etapas: (1) construcción del dataset mediante la captura de video de cada seña por parte del equipo de trabajo; (2) extracción de landmarks con MediaPipe y almacenamiento de las coordenadas crudas en la carpeta datos\_señas/raw/; (3) normalización de las coordenadas respecto a un punto de referencia fijo de la mano, para lograr una representación invariante a la posición del señante dentro del cuadro, guardada en datos\_señas/normalized/; (4) organización de las secuencias normalizadas en conjuntos de entrenamiento y prueba; y (5) entrenamiento de un modelo con arquitectura LSTM (Hochreiter & Schmidhuber, 1997), implementado en TensorFlow, para clasificar la secuencia temporal de landmarks correspondiente a cada seña.

Para el desarrollo del entorno de captura y entrenamiento se utilizó Python en un equipo con sistema operativo Windows, ejecutando los procesos relacionados con MediaPipe desde la línea de comandos (CMD) debido a incompatibilidades identificadas con PowerShell, y fijando la versión mediapipe==0.10.14 por su estabilidad comprobada. Como parte del proceso también se resolvieron conflictos de versiones entre NumPy y TensorFlow que interferían con la reproducibilidad de los entrenamientos. La interfaz de usuario, desarrollada con el microframework Flask y HTML, CSS y JavaScript, se encuentra actualmente en construcción; su función prevista es capturar video en tiempo real, mostrar la traducción en texto y, en una etapa posterior, reproducirla como voz, además de alojar el módulo de aprendizaje interactivo de señas. Como técnica de análisis del desempeño se emplearon la exactitud (accuracy) y el puntaje F1 por clase, comparando los resultados con los reportados en la literatura revisada.

### Resultados

Como parte de la evaluación del sistema, se entrenaron y compararon siete algoritmos de clasificación sobre las secuencias de landmarks de mano extraídas con MediaPipe: Máquinas de Vectores de Soporte (SVM), Random Forest, Extra Trees, K-Vecinos Más Cercanos (KNN), Regresión Logística, Gradient Boosting y Árbol de Decisión. La Figura 2 resume el accuracy obtenido por cada modelo sobre el conjunto de prueba. El modelo SVM obtuvo el mejor desempeño, con una exactitud de 93.72 %, seguido por Random Forest (92.08 %) y Extra Trees (91.79 %). Los modelos KNN, Regresión Logística y Gradient Boosting se ubicaron en un rango intermedio, entre 88 % y 90 %, mientras que el Árbol de Decisión obtuvo el desempeño más bajo, con 85.53 %. Estos resultados son consistentes con lo reportado por Sánchez-Vicinaiz et al. (2024) y Morfin-Chávez et al. (2023), quienes también encontraron que los modelos entrenados sobre el conjunto completo de landmarks de MediaPipe superan a clasificadores más simples cuando el número de clases es reducido.

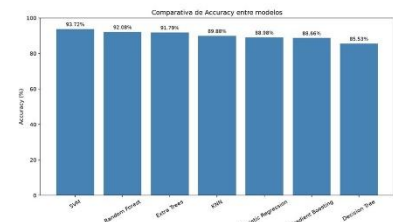


Figura 1. Comparación con otros modelos

### Conclusiones.

El desarrollo de este proyecto permitió avanzar de forma sustancial hacia el objetivo planteado, al construir un pipeline funcional de captura, extracción y normalización de landmarks de mano, comparar siete algoritmos de clasificación y seleccionar SVM como modelo final, con el cual el reconocimiento del alfabeto de la LSM alcanzó una precisión de 90 % al traducirlo a texto. La interfaz web y la salida por voz, planteadas como parte del objetivo, se encuentran actualmente en desarrollo. Los hallazgos confirman que la normalización de coordenadas es un paso indispensable para obtener una clasificación robusta ante variaciones en la posición del señante, y que la resolución de conflictos de compatibilidad entre bibliotecas de visión por computadora y de aprendizaje automático constituye una parte relevante, aunque frecuentemente subestimada, del proceso de desarrollo de este tipo de sistemas.

La principal aportación del proyecto radica en proponer una solución accesible, de bajo costo y sin hardware especializado, que integra tanto la traducción de señas a texto y voz como un módulo de aprendizaje interactivo, ampliando el alcance de trabajos previos centrados casi exclusivamente en el reconocimiento de un alfabeto o vocabulario limitado. Con ello se busca no solo facilitar la comunicación puntual entre personas sordas y oyentes, sino también fomentar el aprendizaje de la LSM entre la población oyente, contribuyendo a una inclusión más sostenida en el tiempo.

Como línea de trabajo futuro se plantea completar el desarrollo de la interfaz web, incorporar la salida por voz, ampliar el reconocimiento más allá del alfabeto hacia palabras y frases de la LSM, y ampliar el conjunto de datos con mayor diversidad de señantes y condiciones de captura para reducir la brecha observada entre la exactitud del modelo aislado (93.72 %) y la precisión obtenida en las pruebas (90 %), así como realizar pruebas de usabilidad con miembros de la comunidad sorda una vez concluida la interfaz. En conjunto, este trabajo aporta evidencia de que la integración de visión por computadora e inteligencia artificial en dispositivos de uso cotidiano puede convertirse en una herramienta concreta para reducir la brecha de comunicación entre la comunidad sorda y la sociedad oyente en México.

### Referencias

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

Instituto Nacional de Estadística y Geografía. (2024). Estadísticas a propósito del Día Internacional de las Personas con Discapacidad (Comunicado de prensa núm. 684/24). [https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2024/EAP\\_PCD24.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2024/EAP_PCD24.pdf)



## **Sistema Inteligente para el Reconocimiento del Abecedario del Lenguaje de Señas Mexicano mediante Visión por Computadora**

Ortiz Del Angel, H. G., Ramon Infante, O. E., & Gonzalez Delgado, M.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### **Resumen**

La comunicación entre personas con discapacidad auditiva y quienes no conocen el Lenguaje de Señas Mexicano (LSM) representa una barrera en diversos entornos como instituciones educativas, hospitales y oficinas públicas. Este trabajo presenta el desarrollo de SignTranslate, un sistema basado en inteligencia artificial y visión por computadora capaz de reconocer en tiempo real el abecedario del LSM mediante una cámara web. El objetivo del proyecto fue desarrollar una herramienta que facilite la comunicación básica entre personas sordas y oyentes mediante la traducción automática de señas a texto. Para ello se implementó un sistema utilizando Python, OpenCV y MediaPipe para la captura de imágenes y la extracción de los 21 puntos de referencia de la mano. Posteriormente se entrenaron diversos modelos de aprendizaje automático para clasificar las letras del abecedario del LSM, entre ellos Support Vector Machine (SVM), Random Forest, Extra Trees, K-Nearest Neighbors (KNN), Regresión Logística, Gradient Boosting y Árbol de Decisión. Los resultados mostraron que el modelo SVM obtuvo la mejor precisión, alcanzando un 93.72 % en el reconocimiento de las letras. El sistema desarrollado identifica correctamente las señas del abecedario y muestra la letra correspondiente en pantalla en tiempo real. Se concluye que la combinación de visión por computadora y aprendizaje automático constituye una alternativa eficiente para desarrollar herramientas que promuevan la inclusión y faciliten la comunicación con personas con discapacidad auditiva.

### **Palabras Clave**

reconocimiento de gestos; redes neuronales recurrentes; MediaPipe; discapacidad auditiva; accesibilidad tecnológica

## Introducción y Planteamiento del Problema

La comunicación es un derecho fundamental; sin embargo, las personas con discapacidad auditiva enfrentan barreras significativas para acceder a servicios básicos cuando la sociedad oyente desconoce su lengua materna. De acuerdo con el Instituto Nacional de Estadística y Geografía (2024), en México 8.8 millones de personas declararon tener alguna discapacidad, y una proporción relevante corresponde a discapacidad auditiva, siendo la Lengua de Señas Mexicana (LSM) la lengua materna de esta comunidad. Pfister (2017) señala que, a pesar de su reconocimiento legal, la LSM continúa enfrentando procesos de socialización lingüística limitados, lo que perpetúa el aislamiento comunicativo de niñas, niños y adultos sordos en entornos mayoritariamente oyentes.

En años recientes, la visión por computadora y el aprendizaje profundo se han consolidado como herramientas viables para el reconocimiento automático de lenguas de señas. Morfín-Chávez et al. (2023) emplearon la biblioteca MediaPipe para extraer 21 puntos clave de la mano y clasificar el alfabeto dactilológico de la LSM mediante algoritmos clásicos de aprendizaje automático, mientras que Sánchez-Vicinaiz et al. (2024) combinaron MediaPipe con redes neuronales convolucionales para el mismo propósito. Para señas dinámicas, que incorporan movimiento y por tanto una dimensión temporal, Mejía-Peréz et al. (2022) demostraron que las arquitecturas recurrentes, en particular LSTM, superan a los enfoques basados únicamente en imágenes estáticas. Más recientemente, Navarrete-López et al. (2026) propusieron un enfoque de reconocimiento de señas dinámicas orientado directamente a la traducción de LSM al español, reforzando la tendencia hacia sistemas completos y no solo hacia la clasificación de señas aisladas. No obstante, la mayoría de estos trabajos se limitan al alfabeto o a un vocabulario reducido, y pocos integran el reconocimiento en una aplicación completa con retroalimentación en texto y voz.

Esta brecha entre los prototipos de investigación y las herramientas utilizables por la comunidad sorda y su entorno motivó el presente proyecto integrador, el cual busca aprovechar la disponibilidad de cámaras web convencionales para construir una solución accesible, sin requerir hardware especializado, con un doble propósito: traducir señas para facilitar la comunicación y enseñar señas a personas oyentes interesadas en aprender LSM. Por lo anterior, el objetivo del presente trabajo es desarrollar un sistema de traducción automática del alfabeto de la Lengua de Señas Mexicana a texto, con la voz como salida planeada para una etapa posterior, basado en la extracción de landmarks de mano mediante MediaPipe y su clasificación con modelos de aprendizaje automático, integrado en una aplicación web con un módulo complementario de aprendizaje interactivo.

## Metodología

El presente trabajo se enmarca en una investigación de tipo aplicado y de desarrollo tecnológico, con un enfoque experimental orientado a la construcción y evaluación de un prototipo funcional. La población objetivo del sistema son las personas usuarias de la LSM y

las personas oyentes interesadas en comunicarse con ellas; los datos se recabaron a partir de un dataset de señas del alfabeto de la LSM capturadas frente a cámara web, y el equipo de trabajo se encargó tanto de la recolección como de las pruebas del sistema.

Las variables consideradas fueron las coordenadas espaciales (x, y, z) de los 21 puntos de referencia de cada mano, obtenidas cuadro a cuadro con la solución MediaPipe Hands (Lugaresi et al., 2019), preferida sobre el modelo de pose corporal completa debido a que ofrece una detección específica y de mayor resolución de la mano, elemento central de la seña. El procedimiento se dividió en las siguientes etapas: (1) construcción del dataset mediante la captura de video de cada seña por parte del equipo de trabajo; (2) extracción de landmarks con MediaPipe y almacenamiento de las coordenadas crudas en la carpeta datos señas (3) normalización de las coordenadas respecto a un punto de referencia fijo de la mano, para lograr una representación invariante a la posición del señante dentro del cuadro, (4) organización de las secuencias normalizadas en conjuntos de entrenamiento y prueba; y (5) entrenamiento de un modelo con arquitectura LSTM (Hochreiter & Schmidhuber, 1997), implementado en TensorFlow, para clasificar la secuencia temporal de landmarks correspondiente a cada seña.

Para el desarrollo del entorno de captura y entrenamiento se utilizó Python en un equipo con sistema operativo Windows, ejecutando los procesos relacionados con MediaPipe desde la línea de comandos (CMD) debido a incompatibilidades identificadas con PowerShell, y fijando la versión mediapipe==0.10.14 por su estabilidad comprobada. Como parte del proceso también se resolvieron conflictos de versiones entre NumPy y TensorFlow que interferían con la reproducibilidad de los entrenamientos. La interfaz de usuario, desarrollada con el microframework Flask y HTML, CSS y JavaScript, se encuentra actualmente en construcción; su función prevista es capturar video en tiempo real, mostrar la traducción en texto y, en una etapa posterior, reproducirla como voz, además de alojar el módulo de aprendizaje interactivo de señas.

## Resultados y Discusión

Como parte de la evaluación del sistema, se entrenaron y compararon siete algoritmos de clasificación sobre las secuencias de landmarks de mano extraídas con MediaPipe: Máquinas de Vectores de Soporte (SVM), Random Forest, Extra Trees, K-Vecinos Más Cercanos (KNN), Regresión Logística, Gradient Boosting y Árbol de Decisión. La Figura 2 resume el accuracy obtenido por cada modelo sobre el conjunto de prueba. El modelo SVM obtuvo el mejor desempeño, con una exactitud de 93.72 %, seguido por Random Forest (92.08 %) y Extra Trees (91.79 %).

Los modelos KNN, Regresión Logística y Gradient Boosting se ubicaron en un rango intermedio, entre 88 % y 90 %, mientras que el Árbol de Decisión obtuvo el desempeño más bajo, con 85.53 %. Estos resultados son consistentes con lo reportado por Sánchez-Vicinaiz et al. (2024) y Morfín-Chávez et al. (2023), quienes también encontraron que los modelos

entrenados sobre el conjunto completo de landmarks de MediaPipe superan a clasificadores más simples cuando el número de clases es reducido.

Con base en estos resultados, se seleccionó el modelo SVM como clasificador final para el reconocimiento del alfabeto dactilológico de la LSM, al ofrecer el mejor equilibrio entre exactitud y tiempo de inferencia sobre las coordenadas normalizadas. Hasta el momento, las pruebas del sistema cubren únicamente el alfabeto y su traducción a texto; la salida por voz y la interfaz web completa aún se encuentran en desarrollo. Sin embargo, al evaluar el traductor de forma integral, con señas capturadas en condiciones más variables de iluminación, distancia y usuarios distintos a los del conjunto de entrenamiento, la precisión del sistema completo se ubicó en 86 %, un valor inferior al accuracy obtenido en la evaluación aislada del modelo (93.72 %). Esta diferencia evidencia una brecha entre el desempeño en condiciones controladas de prueba y el desempeño en un escenario de uso más realista, fenómeno relacionado con la sensibilidad de los sistemas basados en características espaciales ante cambios en la posición y distancia de la mano respecto a la cámara, señalada también por Ríos-Figueroa et al. (2022).

Durante la fase de organización de datos se estructuró el repositorio en carpetas de datos crudos y normalizados, lo que permitió trazabilidad entre las señas capturadas y sus versiones procesadas. Asimismo, se identificaron y resolvieron incompatibilidades entre las versiones de MediaPipe, NumPy y TensorFlow que inicialmente impedían el entrenamiento estable de los modelos, un obstáculo técnico recurrente también observado en trabajos que integran MediaPipe con marcos de aprendizaje profundo (Sánchez-Vicinaiz et al., 2024). La Figura 1 muestra la miniatura del póster elaborado para este trabajo, en el que se sintetiza visualmente el problema, la metodología y el estado de avance del proyecto.

## Conclusiones

El desarrollo de este proyecto permitió avanzar de forma sustancial hacia el objetivo planteado, al construir un pipeline funcional de captura, extracción y normalización de landmarks de mano, comparar siete algoritmos de clasificación y seleccionar SVM como modelo final, con el cual el reconocimiento del alfabeto de la LSM alcanzó una precisión de 90 % al traducirlo a texto. La interfaz web y la salida por voz, planteadas como parte del objetivo, se encuentran actualmente en desarrollo. Los hallazgos confirman que la normalización de coordenadas es un paso indispensable para obtener una clasificación robusta ante variaciones en la posición del señante, y que la resolución de conflictos de compatibilidad entre bibliotecas de visión por computadora y de aprendizaje automático constituye una parte relevante, aunque frecuentemente subestimada, del proceso de desarrollo de este tipo de sistemas.

La principal aportación del proyecto radica en proponer una solución accesible, de bajo costo y sin hardware especializado, que integra tanto la traducción de señas a texto y voz como un

módulo de aprendizaje interactivo, ampliando el alcance de trabajos previos centrados casi exclusivamente en el reconocimiento de un alfabeto o vocabulario limitado. Con ello se busca no solo facilitar la comunicación puntual entre personas sordas y oyentes, sino también fomentar el aprendizaje de la LSM entre la población oyente, contribuyendo a una inclusión más sostenida en el tiempo.

Como línea de trabajo futuro se plantea completar el desarrollo de la interfaz web, incorporar la salida por voz, ampliar el reconocimiento más allá del alfabeto hacia palabras y frases de la LSM, y ampliar el conjunto de datos con mayor diversidad de señantes y condiciones de captura para reducir la brecha observada entre la exactitud del modelo aislado (93.72 %) y la precisión obtenida en las pruebas (90 %), así como realizar pruebas de usabilidad con miembros de la comunidad sorda una vez concluida la interfaz. En conjunto, este trabajo aporta evidencia de que la integración de visión por computadora e inteligencia artificial en dispositivos de uso cotidiano puede convertirse en una herramienta concreta para reducir la brecha de comunicación entre la comunidad sorda y la sociedad oyente en México.

## Referencias

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

Mejía-Peréz, K., Córdova-Esparza, D. M., Terven, J., Herrera-Navarro, A. M., García-Ramírez, T., & Ramírez-Pedraza, A. (2022). Automatic recognition of Mexican sign language using a depth camera and recurrent neural networks. *Applied Sciences*, 12(11), 5523

Morfín-Chávez, R. F., Gortarez-Pelayo, J. J., & Lopez-Nava, I. H. (2023, November). Fingerspelling recognition in Mexican sign language (LSM) using machine learning. In *Mexican International Conference on Artificial Intelligence* (pp. 110–120). Cham: Springer Nature Switzerland.

Instituto Nacional de Estadística y Geografía. (2024). Estadísticas a propósito del Día Internacional de las Personas con Discapacidad (Comunicado de prensa núm. 684/24). [https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2024/EAP\\_PCD24.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2024/EAP_PCD24.pdf)

Sánchez-Vicinaiz, T. J., Camacho-Pérez, E., Castillo-Atoche, A. A., Cruz-Fernandez, M., García-Martínez, J. R., & Rodríguez-Reséndiz, J. (2024). Mediapipe frame and convolutional neural networks-based fingerspelling detection in mexican sign language. *Technologies*, 12(8), 124.

Rios-Figueroa, H. V., Sánchez-García, A. J., Sosa-Jiménez, C. O., & Solís-González-Cosío, A.

## KINESIS AI: Sistema de Visión Artificial para la Detección Temprana de Malestar Emocional en el Aula

Banda Parada, Y. Y., Cordero Abrego, R. D., Moran Montes, S. I., & Salazar Maldonado, M. A.

UNIVERSIDAD TECNOLÓGICA DE MATAMOROS  
\* 2310445@utmatamoros.edu.mx

Eje Temático: Educación Bilingüe

Palabras clave: visión artificial, redes neuronales convolucionales, bienestar socioemocional estudiantil, identificación biométrica, tecnología educativa.

## Resumen:

El docente no puede observar de forma objetiva el estado emocional de cada estudiante mientras atiende a todo un grupo. Este trabajo desarrolla y valida KINESIS AI, un prototipo de visión artificial que detecta en tiempo real, mediante una cámara web, señales combinadas de malestar emocional y postural en el aula. El sistema integra una CNN entrenada sobre FER2013, un modelo de estimación de postura, y un módulo de identificación biométrica con umbral de distancia. Se validó mediante pruebas técnicas —que permitieron corregir 5 defectos del sistema— y una encuesta aplicada a 15 docentes de Matamoros, antes y después de una demostración en vivo. El 87% consideró que el sistema le ayudaría a detectar a tiempo el malestar de sus alumnos, con una utilidad percibida de 4.3/5. Se concluye que esta integración es una alternativa viable para apoyar la detección temprana de malestar en el aula física.

## Introducción:

El docente no puede observar de forma simultánea el estado emocional de cada estudiante mientras atiende a todo un grupo, lo que retrasa la detección de episodios de malestar. El estado afectivo influye directamente en la atención y el desempeño del estudiante (D'Mello & Graesser, 2012).

La mayoría de los sistemas de detección de afecto se han evaluado en plataformas en línea y no en el salón de clases físico (Booth, Bosch, & D'Mello, 2023). El objetivo de este trabajo es desarrollar y validar un prototipo que integre reconocimiento de emociones, análisis de postura e identificación de estudiantes, para generar alertas tempranas de malestar en el aula.

## Metodología.

Se empleó un enfoque de investigación aplicada de tipo tecnológico, con un diseño experimental iterativo, estructurado en tres fases:

- Desarrollo: pipeline en Python con OpenCV, una CNN entrenada sobre FER2013 para el reconocimiento de emociones (Goodfellow et al., 2013), un modelo de estimación de pose, y un módulo de identificación biométrica con umbral de distancia (Schroff et al., 2015).
- Control de calidad: bitácora formal de depuración con causa raíz y pruebas de regresión sobre cada uno de los 5 errores identificados y corregidos.
- Validación: encuesta aplicada a 15 docentes de nivel medio superior en Matamoros, antes y después de una demostración en vivo del prototipo.

## Resultados

El sistema logró un procesamiento estable en tiempo real y la corrección de 5 defectos críticos, incluyendo falsos positivos de identificación facial. El 87% de los docentes consideró que ayudaría a detectar a tiempo el malestar de sus alumnos (Figura 1).

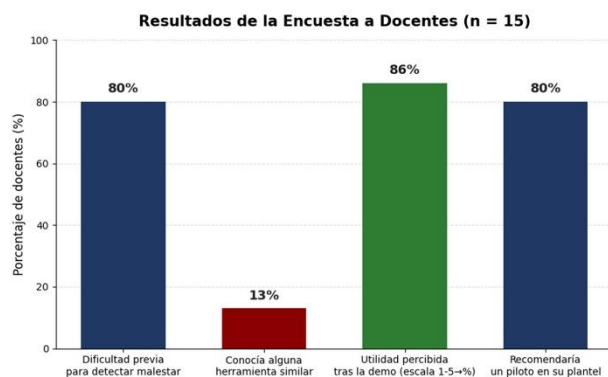


Figura 1. Resultados de la encuesta de aceptación aplicada a docentes (n = 15).

```
# identity.py - identificación facial con umbral (sistema abierto)
MODEL_NAME = "Facenet"
UMBRAL_DISTANCIA = 0.55 # antes: LBPH siempre elegía "el mas parecido"

def identificar(self, face_bgr):
    emb = obtener_embedding(face_bgr)
    if emb is None:
        return "DESCONOCIDO", None

    mejor_nombre, mejor_dist = "DESCONOCIDO", 999
    for nombre, vectores in self.embeddings.items():
        for v in vectores:
            d = distancia_coseno(emb, v)
            if d < mejor_dist:
                mejor_dist, mejor_nombre = d, nombre

    # Solo se acepta si la distancia es menor al umbral;
    # si no, se declara honestamente "DESCONOCIDO"
    if mejor_dist <= UMBRAL_DISTANCIA:
        return mejor_nombre, mejor_dist
    return "DESCONOCIDO", mejor_dist
```

Figura 2. Fragmento de código de identity.py: identificación facial por umbral de distancia.

## Conclusiones.

El prototipo demuestra que es posible integrar reconocimiento de emociones, análisis de postura e identificación biométrica en un mismo flujo en tiempo real, para detectar malestar dentro del aula física, un caso poco explorado frente a los sistemas evaluados en entornos en línea.

Como trabajo futuro, se plantea migrar a una arquitectura en la nube para monitorear varias aulas a la vez, ampliar la validación a un plantel completo y explorar su despliegue en escuelas de la franja fronteriza de Texas.

## Referencias

- Booth, B. M., Bosch, N., & D'Mello, S. K. (2023). Engagement detection and its applications in learning: A tutorial and selective review. *Proceedings of the IEEE*, 111(10), 1398–1422.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Goodfellow, I. J., et al. (2013). Challenges in representation learning: A report on three machine learning contests. *International Conference on Neural Information Processing*, 117–124.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE CVPR*, 815–823.

## **KINESIS AI: Sistema de Visión Artificial para la Detección Temprana de Malestar Emocional en el Aula**

Banda Parada, Y. Y., Cordero Abrego, R. D., Moran Montes, S. I., & Salazar Maldonado, M. A.

Eje Temático: Innovación, Emprendimiento y Competitividad para la Transformación Organizacional

### **Resumen**

En los entornos educativos actuales, el docente enfrenta una limitante estructural: no puede observar de forma objetiva y simultánea el estado emocional de cada estudiante mientras atiende a todo un grupo, lo que retrasa la detección de episodios de malestar y dificulta una intervención oportuna. El objetivo de este trabajo fue desarrollar y validar un prototipo de sistema de visión artificial, denominado KINESIS AI, capaz de detectar en tiempo real señales combinadas de malestar emocional y postural en el aula a partir de una cámara web convencional. La metodología consistió en el desarrollo de un pipeline que integra una red neuronal convolucional entrenada sobre el conjunto de datos FER2013 para el reconocimiento de emociones faciales, un modelo de estimación de postura corporal para detectar posturas decaídas o inmovilidad prolongada, y un módulo de identificación individual basado en embeddings faciales, todo bajo un ciclo formal de control de calidad documentado en bitácoras de depuración. El sistema se validó mediante pruebas técnicas internas y una encuesta piloto aplicada a 15 docentes de nivel medio superior en Matamoros, Tamaulipas, antes y después de una demostración en vivo del prototipo. Los resultados mostraron una reducción sustancial de falsos positivos tras migrar el módulo de identificación de un esquema cerrado a uno basado en umbral de similitud, así como una calificación promedio de utilidad percibida de 4.3 sobre 5 entre los docentes encuestados, quienes manifestaron como principal inquietud el manejo de la privacidad de los datos de menores. Se concluye que la combinación de reconocimiento emocional, análisis postural e identificación individual en un mismo flujo de procesamiento en tiempo real constituye una alternativa viable y original para apoyar la detección temprana de malestar estudiantil en el aula física, complementando —en lugar de sustituir— el criterio del docente.

### **Palabras Clave**

Visión artificial; redes neuronales convolucionales; bienestar socioemocional estudiantil;

## Introducción y Planteamiento del Problema

La capacidad de reconocer emociones a partir de la expresión facial se sustenta en la evidencia de que un conjunto reducido de emociones básicas se manifiesta de forma consistente y reconocible entre culturas (Ekman, 1992), lo que ha permitido el desarrollo de sistemas computacionales capaces de inferir el estado afectivo de una persona a partir de imágenes de su rostro. En el ámbito educativo, diversos estudios han documentado que el estado afectivo del estudiante influye directamente en su atención, motivación y desempeño (D'Mello & Graesser, 2012), y que la computación afectiva —el conjunto de métodos computacionales orientados a detectar, interpretar y responder a estados emocionales— ofrece un camino viable para dar seguimiento objetivo a estas señales dentro de entornos de aprendizaje (Calvo & D'Mello, 2010).

Sin embargo, la revisión de la literatura reciente muestra que la gran mayoría de los sistemas de detección de afecto y de compromiso (engagement) desarrollados hasta ahora se han implementado y evaluado en plataformas de aprendizaje en línea o en laboratorio, y no en el salón de clases físico (Booth, Bosch, & D'Mello, 2023). Esto deja un vacío relevante para instituciones de educación media superior y superior en la región fronteriza México–Estados Unidos, donde la enseñanza presencial sigue siendo predominante y donde el docente no cuenta con ninguna herramienta objetiva que le permita advertir, en el momento, que un estudiante está atravesando una situación de malestar emocional sostenido.

¿Es posible, con hardware accesible (una cámara web estándar) y bibliotecas de visión artificial de código abierto, construir un sistema que detecte en tiempo real, dentro del salón de clases físico, señales combinadas de malestar emocional y postural, sin depender de que el estudiante esté interactuando con una plataforma digital? A partir de esta pregunta de investigación, el objetivo general de este trabajo es desarrollar y validar un prototipo funcional de sistema de visión artificial que integre reconocimiento de emociones faciales, análisis de postura corporal e identificación individual de estudiantes, con el fin de generar alertas tempranas de malestar emocional dentro del aula.

## Metodología

El presente trabajo se desarrolló bajo un enfoque de investigación aplicada de tipo tecnológico, con un diseño experimental iterativo (ciclos de desarrollo, prueba y depuración) complementado con un componente cuantitativo de validación mediante encuesta. El sistema se construyó en Python, utilizando OpenCV para la captura y el preprocesamiento de video (Bradski, 2000), una red neuronal convolucional entrenada sobre el conjunto de datos FER2013 para la clasificación de siete emociones básicas (Goodfellow et al., 2013), un modelo de estimación de pose corporal de 33 puntos clave para el análisis postural (Bazarevsky et al., 2020), y un módulo de identificación individual basado en embeddings faciales (arquitectura Facenet) que calcula la distancia coseno entre el rostro detectado y los rostros



previamente registrados, aceptando una coincidencia solo si dicha distancia no supera un umbral definido (Schroff, Kalenichenko, & Philbin, 2015).

La muestra para la validación técnica estuvo compuesta por sesiones de prueba controladas del propio equipo desarrollador (captura en tiempo real, desconexión simulada de cámara, variación de condiciones de iluminación), mientras que la validación de aceptación se realizó con una muestra no probabilística por conveniencia de 15 docentes de nivel medio superior en Matamoros, Tamaulipas, a quienes se aplicó un cuestionario estructurado antes y después de una demostración en vivo del prototipo.

El procedimiento siguió los siguientes pasos: (1) levantamiento de requerimientos con docentes; (2) desarrollo modular del sistema (captura y preprocesamiento, inferencia emocional, análisis de postura, identificación biométrica, sistema de alertas y generación de reportes); (3) control de calidad mediante una bitácora formal de depuración, en la que se documentó la causa raíz, el código corregido y el resultado de la corrección de cada error detectado; (4) pruebas de regresión sobre las correcciones aplicadas; y (5) aplicación de la encuesta de aceptación antes y después de la demostración del prototipo. Como técnica de análisis se utilizó estadística descriptiva simple (frecuencias y porcentajes) sobre las respuestas del cuestionario. Dado que el sistema procesa imágenes de rostros de estudiantes, se contempló como consideración ética el consentimiento informado de los docentes participantes en la fase de encuesta, así como el diseño de un modo de operación anónimo (sin identificación nominal) previsto para las siguientes fases de despliegue con menores de edad.

## **Resultados y Discusión**

El prototipo de KINESIS AI se integró de manera funcional, logrando el procesamiento de video en tiempo real (~60 fps de renderizado, con inferencia de los modelos de IA cada 10 fotogramas) y la generación automática de alertas cuando una emoción negativa sostenida coincidía con una postura irregular. Durante la fase de control de calidad se identificaron y corrigieron 5 defectos relevantes del sistema, entre ellos un cierre inesperado del programa ante la desconexión momentánea de la cámara y falsos positivos en la identificación de estudiantes, este último corregido mediante la sustitución de un clasificador de tipo cerrado por un esquema de embeddings con umbral de distancia, lo cual eliminó por completo las identificaciones erróneas observadas en las pruebas previas a la corrección.

En la encuesta de aceptación, el 80% de los docentes (12 de 15) reportó dificultad previa para detectar a tiempo el malestar emocional de sus alumnos, mientras que solo el 13% (2 de 15) conocía alguna herramienta tecnológica orientada a ese fin, lo que confirma el vacío identificado en la literatura respecto a la falta de herramientas de este tipo para el aula física (Booth, Bosch, & D'Mello, 2023). Tras la demostración del prototipo, el 87% (13 de 15) consideró que KINESIS AI podría ayudarle a detectar a tiempo situaciones de malestar en su

grupo, con una calificación promedio de utilidad percibida de 4.3 sobre 5. La preocupación más mencionada (60% de los docentes) fue la privacidad de los datos de menores, hallazgo que es consistente con las consideraciones éticas señaladas en la metodología y que orientó el diseño de un modo de operación anónimo para las siguientes fases del proyecto.

A diferencia de las herramientas educativas de monitoreo más adoptadas actualmente, como GoGuardian o Bark for Schools, que operan sobre la actividad digital del estudiante (correo, navegación, documentos), KINESIS AI se distingue por observar directamente el aula física a través de visión artificial, sin requerir que el alumno esté utilizando una cuenta o dispositivo institucional. Esta diferencia responde de forma directa a la brecha señalada en la introducción: la ausencia de sistemas de detección afectiva evaluados en el salón de clases presencial (Booth, Bosch, & D'Mello, 2023), y constituye la principal aportación técnica y práctica del presente trabajo.

## Conclusiones

El presente trabajo cumplió su objetivo general al desarrollar y validar un prototipo funcional de sistema de visión artificial capaz de detectar, en tiempo real y dentro del aula física, señales combinadas de malestar emocional y postural, integrando en un mismo flujo de procesamiento el reconocimiento de emociones faciales, el análisis de postura corporal y la identificación individual de estudiantes mediante embeddings faciales. La aportación principal de KINESIS AI radica precisamente en esa integración: mientras que la literatura revisada muestra sistemas de detección afectiva enfocados casi exclusivamente en entornos de aprendizaje en línea, este trabajo demuestra que es posible construir, con hardware accesible y herramientas de código abierto, un sistema equivalente para el salón de clases presencial, atendiendo una necesidad que los propios docentes encuestados reconocieron no tener cubierta con ninguna herramienta existente.

El proceso de depuración formal aplicado durante el desarrollo —documentado en bitácoras individuales con causa raíz, código corregido y resultado de cada corrección— resultó determinante para la estabilidad del prototipo, particularmente en la eliminación de falsos positivos de identificación, uno de los riesgos más sensibles de un sistema que procesa datos biométricos de menores. La aceptación favorable manifestada por los docentes participantes, junto con su preocupación explícita por la privacidad de los datos, confirma tanto la pertinencia del proyecto como la necesidad de continuar fortaleciendo sus mecanismos de protección de datos antes de un despliegue a mayor escala.

Como línea de trabajo futuro, se plantea migrar el procesamiento hacia una arquitectura con respaldo en la nube que permita el monitoreo simultáneo de múltiples aulas desde un panel centralizado, así como ampliar la validación piloto a un plantel completo y explorar su viabilidad de despliegue en escuelas de la franja fronteriza de Texas, sujeto al cumplimiento de la normativa estadounidense de protección de datos de menores en entornos educativos. KINESIS AI aporta con ello evidencia inicial de que la visión artificial aplicada al aula física, y

no solo al aula digital, puede convertirse en una herramienta legítima de acompañamiento socioemocional para la educación de la región fronteriza.

## Referencias

Booth, B. M., Bosch, N., & D'Mello, S. K. (2023). Engagement detection and its applications in learning: A tutorial and selective review. *Proceedings of the IEEE*, 111(10), 1398–1422.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>

Goodfellow, I. J., et al. (2013). Challenges in representation learning: A report on three machine learning contests. *International Conference on Neural Information Processing*, 117–124.

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE CVPR*, 815–823.

El presente índice enlista, en orden alfabético por apellido, a los autores de las contribuciones incluidas en esta obra, con indicación del número de trabajo donde aparece su participación.

### Autor

Atzin Becerra, Raymundo — Póster 006 (p. 29)  
Banda Parada, Yael Yessed — Póster 012 (p. 57)  
Becerra Martínez, Christian Ronaldo — Póster 001 (p. 3)  
Calixto García, Uriel — Póster 002 (p. 7)  
Carreón Jiménez, Marlene América — Póster 005 (p. 23)  
Contreras Ochoa, Manuel Alejandro — Póster 004 (p. 18)  
Cordero Abrego, Ruben Dario — Póster 012 (p. 57)  
Diaz Bojorquez, Carlos Alberto — Póster 008 (p. 37)  
Diaz García, Gabriel — Póster 009 (p. 40)  
Diaz Olivares, Alexia Noor Nasya — Póster 005 (p. 23)  
Dominguez Lerma, Iñaki — Póster 003 (p. 13)  
García Saucedo, Juan Humberto — Póster 001 (p. 3)  
Gonzalez Delgado, Manuel — Póster 011 (p. 51)  
Gutiérrez Serralde, Irvin — Póster 003 (p. 13)  
Gutiérrez Balboa, Vianney — Póster 005 (p. 23)  
Hernández Neri, Areli — Póster 009 (p. 40)  
Jiménez Espinoza, Ángel Osiel — Póster 001 (p. 3)  
Leal Medina, Miguel Ángel — Póster 002 (p. 7)  
López Trejo, Moisés Eli — Póster 001 (p. 3)  
Maldonado Martinez, Angelica — Póster 003 (p. 13)  
Martínez Moreno, Jesus — Póster 003 (p. 13)  
Martínez Cervantes, Cristóbal Axel — Póster 009 (p. 40)  
Martinez Chavez, Jared — Póster 007 (p. 34)  
Martínez Meza, Miguel Ángel — Póster 001 (p. 3)  
Medina Hernández, Eliezer — Póster 005 (p. 23)  
Medina Oropeza, Jorge Luis — Póster 009 (p. 40)  
Melo Zuñiga, Marvin Yusniel — Póster 002 (p. 7)  
Mena Alcala, Diego Abdiel — Póster 010 (p. 45)  
Mendoza Saucedo, Ángel Alejandro — Póster 006 (p. 29)  
Montalvo Sanchez, Yoshua Isai — Póster 007 (p. 34)

Moran Montes, Sergio Israel — Póster 012 (p. 57)  
Moreno Hernandez, Rogelio — Póster 007 (p. 34)  
Muñoz Jiménez, Yareli Sarahi — Póster 002 (p. 7)  
Netro Castillo, Jaziel — Póster 009 (p. 40)  
Niño Paz, Edgar — Póster 001 (p. 3)  
Ortiz Del Angel, Hilda Guadalupe — Póster 011 (p. 51)  
Pérez Mata, José Antonio — Póster 004 (p. 18)  
Pérez Rico, Abril Mirelli — Póster 005 (p. 23)  
Pérez Vázquez, Juan Eli — Póster 004 (p. 18)  
Quino Chávez, Fátima Guadalupe — Póster 010 (p. 45)  
Ramon Infante, Orlando Emanuel — Póster 011 (p. 51)  
Ramos Ponce, Sergio Ismael — Póster 006 (p. 29)  
Recillas Mendoza, Marely — Póster 002 (p. 7)  
Reyna Montenegro, Joel — Póster 004 (p. 18)  
Rodarte Cristo, Luis Ernesto — Póster 004 (p. 18)  
Salas Perez, Emiliano — Póster 008 (p. 37)  
Salazar Maldonado, Marco Antonio — Póster 012 (p. 57)  
Sánchez Veloz, Cesar — Póster 006 (p. 29)  
Sandoval Camacho, Andrea — Póster 003 (p. 13)  
Sarmiento Juarez, José Emmanuel — Póster 002 (p. 7)  
Segundo Orta, Aldo Maximiliano — Póster 007 (p. 34)  
Sole Huerta, Domenec — Póster 006 (p. 29)  
Villalobos Barrios, Angel Eduardo — Póster 007 (p. 34)

### COLOFÓN

Esta obra, primera edición de los Trabajos de Presentación del Congreso Internacional Transformando la Frontera, se terminó de editar en el mes de agosto de 2026 en la Universidad Tecnológica de Matamoros, en formato digital para su distribución en acceso abierto.

La formación editorial se realizó en tipografías Trebuchet MS y Verdana, sobre una retícula que conserva la lámina original de cada póster junto con su desarrollo textual extendido.

El diseño editorial, la formación y la coordinación general estuvieron a cargo del Dr. Noé Toledo González; la corrección de estilo, de la M.E.R. Maira Selene Ríos Gómez.

Esta edición se publica bajo el sello Art90 de Noé Toledo González Editorial (ISBN 978-970-96337-0-2), en colaboración con la Universidad Tecnológica de Matamoros, a través del Cuerpo Académico CA-08 Ingeniería de Sistemas» y la Coordinación de la Maestría en Tecnologías de la Información con Especialidad en Gestión de Proyectos.

### ANEXOS

El presente anexo integra una galería fotográfica documental del desarrollo del Congreso Internacional Transformando la Frontera 2026, que registra las sesiones de ponencias la presentación de pósteres, la participación de autores, jurados y asistentes, así como distintos momentos institucionales del evento.



CONGRESO INTERNACIONAL  
**TRANSFORMANDO**  
LA FRONTERA

UNIVERSIDAD TECNOLÓGICA **UT** MATAMOROS

ISBN 978-970-96337-0-2

Editorial: Noé Toledo González  
En colaboración con la Universidad Tecnológica de Matamoros  
H. Matamoros, Tamaulipas, México















