

CITIC-UGD

Impulsando el
conocimiento
y la formación
de investigadores e
ingenieros en Misiones



COMPILADOR: DR. DIEGO GODOY
PRÓLOGO: ING. HÉCTOR RUIDÍAS

CITIC-UGD

CITIC-UGD

Impulsando el conocimiento
y la formación de investigadores
e ingenieros en Misiones

COMPILADOR: DR. DIEGO GODOY

PRÓLOGO: ING. HÉCTOR RUIDÍAS

CITIC-UGD : impulsando en conocimiento y la formación de investigadores e ingenieros en Misiones / Diego Godoy ... [et al.] ; compilación de Diego Godoy ; Prólogo de Héctor Ruidías. - 1a ed - Posadas : Universidad Gastón Dachary, 2024.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-25444-9-2

1. Inteligencia Artificial. 2. Medicina. 3. Protección del Medio Ambiente. I. Selva, Ricardo II. Godoy, Diego, comp. III. Ruidías, Héctor, prolog.

CDD 577

Editorial UGD
Universidad Gastón Dachary
Salta 1912, Posadas, Misiones, Argentina
editorial@ugd.edu.ar

Diseño: Brutal Creativos

Coordinación editorial: Lic. Marina Hlebovich

Edición, arte y maquetación integral: Loquepodemos Estudio Editorial

Asesoramiento editorial: Horacio Moreno

Publicación electrónica - distribución gratuita

Acceso web: <https://ugd.edu.ar/es/>



Licencia Creative Commons – Atribución/Reconocimiento - NoComercial - CompartirIgual 4.0 (by-nc-sa)

Se permite la generación de obras derivadas siempre que no se haga con fines comerciales. Tampoco se puede utilizar la obra original con fines comerciales. Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original.

Esta licencia no es una licencia libre. Algunos derechos reservados: <https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Índice

Prólogo	9
Introducción	11
Capítulo 1: IOT - Verificación automática del uso de elementos de seguridad utilizando visión computacional	33
Capítulo 2: IOT - Detección de botellas de PET en una Planta de reciclaje	65
Capítulo 3: Herramienta para la asistencia pre diagnóstico y seguimiento del tratamiento de la escoliosis idiopática del adolescente (EIA)	93
Sobre los autores	131

Prólogo

Ing. Héctor Ruidías



La Universidad Gastón Dachary ha alcanzado recientemente su 25° aniversario, lo que representa un importante hito en su trayectoria educativa. Aunque este logro la posiciona como una institución madura en términos de consolidación, también refleja una juventud notable en el contexto del sistema educativo argentino. Desde sus inicios, una de las primeras carreras que ofreció fue Ingeniería Informática, que hoy en día comparte la misma edad que la universidad y ha sido fundamental en su desarrollo.

A lo largo de estos años, la carrera de Ingeniería Informática ha graduado a 208 profesionales, quienes han contribuido significativamente al mundo tecnológico. Los/as egresados/as no solo han adquirido competencias técnicas valiosas, sino que también

han transformado la vida de personas mediante la implementación de soluciones innovadoras en informática.

El cuerpo docente actual de la carrera incluye a varios de estos egresados, quienes, junto a otros profesores, forman una comunidad académica rica y diversa. Este entorno permite que los ideales de la Ingeniería Informática se transmitan a las nuevas generaciones de estudiantes, quienes heredan una sólida formación técnica y un compromiso con la innovación. A lo largo de los años he podido observar esta evolución desde diferentes posiciones: como estudiante, como egresado y como docente investigador.

Sin embargo, debemos asumir el desafío que supone para los estudiantes de nuestra provincia cursar una carrera con un perfil productivo e industrial en un contexto geográfico y económico como el nuestro. La provincia de Misiones, aunque privilegiada por su belleza natural, se encuentra alejada de los principales centros tecnológicos y productivos del país, lo que genera en muchos estudiantes la sensación de estar lejos de donde *pasan las cosas*. Este sentimiento persiste incluso en un mundo globalizado y conectado, donde es posible acceder a grandes oportunidades *sin salir de casa*.

En este punto no puedo dejar de señalar la bisagra histórica que supuso la irrupción del COVID 19 y consecuentemente las medidas de confinamiento. Excede el alcance de este prólogo extplayarse al respecto, pero tal vez uno de los fenómenos que más

ha destacado fue sin duda la adopción del “trabajo remoto”. Esto a su vez generó una enorme presión sobre el mercado IT, provocando un aumento de la demanda de profesionales casi en forma exponencial.

Sin embargo, finalizadas las medidas de confinamiento, esta situación no pudo sostenerse y muchas empresas tecnológicas impulsaron despidos masivos y una “vuelta a la oficina” que hasta hoy perdura.

Toda esta dinámica tuvo efectos también en los procesos de reclutamiento de personal IT y los cambios culturales que se dieron en la industria contribuyeron a relativizar aún más la formación académica, a la vez que se denunciaba nuevamente la escasez de mano de obra calificada en el desarrollo de software. En respuesta proliferaron los *Bootcamps* y las “Academias online”, espacios de formación alternativos a los centros educativos tradicionales que prometían inserción laboral inmediata y salarios altos.

Las universidades y otros centros académicos formales detectaron una suerte de amenaza en dicho fenómeno, pero también una oportunidad de adaptarse e impulsar ofertas educativas *aggiornadas* a los nuevos formatos.

Transcurrido un tiempo desde aquellos eventos nos encontramos con nuevas perspectivas, donde también se erigió la Inteligencia Artificial como un factor determinante en las relaciones

entre el mercado y la academia, entre la demanda de empleo y la formación de perfiles profesionales.

Pese a ello, la Universidad Gastón Dachary ha logrado mantener viva la importancia de los encuentros presenciales, ritualizando el aprendizaje a través de reuniones semanales, discusión de textos y modelaje de situaciones reales. Estas prácticas fortalecen el vínculo entre estudiantes y docentes, recordando que, aunque el mundo avanza hacia la virtualidad, el contacto humano sigue siendo un elemento esencial en la formación de futuros ingenieros.

A menudo me pregunto, y entiendo que puede ser una pregunta recurrente para las familias de carreras de sistemas, ¿cómo podemos ser relevantes en un contexto en constante cambio a la vez que mantenemos los valores de la formación académica y universitaria? Entiendo que no es la respuesta sino el ejercicio del preguntarnos dónde reside el verdadero valor que nos mueve a buscar permanentemente el diálogo con los estudiantes, en el diseño de nuevos espacios de construcción de saberes y competencias.

El Centro de Investigación de Tecnologías de la Información y las Comunicaciones (CITIC) lleva 18 años propiciando tales encuentros. Con una estructura poco notoria dentro de UGD ha tenido presencia en diferentes foros de la comunidad académica nacional e internacional, y ha servido de vehículo para que muchos egresados desarrollaran el pensamiento técnico y científico con

el objetivo de resolver problemas reales de la sociedad, alineados con líneas de investigación innovadoras.

Quienes editan y publican este libro buscan ofrecer una mirada de la importancia estratégica del CITIC en la Universidad y, en particular, dentro de la carrera de Ingeniería Informática. Pero podríamos agregar que, más importante aún, esta publicación coincide con la celebración de un espacio de integración de una comunidad académica que mira hacia la sociedad y sus problemas, en un ejercicio permanente de diálogo con la misma, sin por ello ir en detrimento de la exploración científica como fin en sí mismo.

Podemos esperar que el CITIC continúe creciendo e impactando positivamente su entorno, formando graduados, siendo el hábitat para que los docentes de Ingeniería Informática puedan desarrollar sus investigaciones y que, en definitiva, sea punta de lanza de la innovación en la región y el mundo, anhelo que no resulta exagerado si tenemos en cuenta que se orienta al ejercicio de la investigación científica en primer lugar y en segundo término a la transferencia de resultados al medio.

La Colección Ciencia Abierta de la Editorial UGD reúne publicaciones interdisciplinarias orientadas a la difusión de la producción científica local, con un enfoque accesible y colaborativo.

Introducción

Dr. Diego Godoy



Este año el Centro de Investigación de Tecnología de la Información y las Comunicaciones (CITIC) de la Universidad Gastón Dachary (UGD) cumple 18 años de trayectoria, consolidándose como un espacio para la producción de conocimiento en el campo de la informática y la tecnología. Fundado el 27 noviembre de 2006, el CITIC representa un ámbito en el cual docentes, estudiantes y becarios desarrollan investigaciones y ponen en práctica diversos saberes. Para enumerar algunos de sus resultados podemos mencionar que 36 de los 208 egresados formularon sus proyectos de investigación y redactaron su Trabajo Final de carrera, 8 docentes han participado de 7 proyectos de investigación (convocatorias internas UGD) y ha sido lugar de trabajo para 3

becarios del Comité Ejecutivo de Desarrollo e Innovación de la provincia de Misiones CEDIT

Esta publicación tiene el objetivo de, por un lado, describir una breve historia del Centro y, por otro, mostrar una síntesis de tres trabajos colaborativos de investigación presentados en congresos nacionales e internacionales, lo que pretende introducir al lector a la labor que implica detectar un problema y desarrollar una solución en el marco del proyecto *Tecnologías para desarrollos sostenibles de ciudades inteligentes apoyados por machine learning*, adjudicado en concurso interno de proyectos de investigación en la XII convocatoria, Resolución rectoral UGD 15/A/2021. Parte de esos proyectos derivaron en trabajos finales de la carrera de Ingeniería Informática.

El primer trabajo se denomina “Verificación automática del uso de elementos de seguridad utilizando visión computacional” y aborda la problemática del uso de los elementos de protección personal para el mundo del trabajo (EPP). Debido a que existen normativas vigentes que dictan el uso obligatorio de diversos tipos de EPP, dependiendo de la actividad que se realice. El correcto uso de estos elementos puede ser la diferencia entre un incidente y un accidente con graves consecuencias para las personas involucradas. En distintas industrias, los supervisores dedican una gran parte de su tiempo a controlar que los trabajadores utilicen correctamente sus EPP en todo momento. Con el fin de contribuir

con el monitoreo de la seguridad de los trabajadores, este texto aborda el desarrollo de un prototipo que permite detectar el uso de EPP de forma automática y, consecuentemente, dar aviso a los supervisores. El prototipo está desarrollado en Python y la detección de los EPP se realiza a través de una red neuronal profunda (YOLOv5). Las inferencias preliminares muestran que puede ser adaptado en una gran variedad de escenarios brindando resultados prometedores.

El segundo de ellos se denomina “Detección de botellas de PET en una Planta de reciclaje”. Se presenta una solución basada en inteligencia artificial que permite automatizar muchos procesos que hoy en día se realizan manualmente mediante la utilización de redes neuronales, permitiendo la redistribución de esfuerzos y recursos, y también el aumento de la productividad. En particular este trabajo trata de indicar cómo se utiliza e implementa, al igual que en el trabajo anterior, una red neuronal denominada YOLOv5 para la detección de distintos tipos de residuos plásticos (botellas de PET) que llegan a una planta de reciclaje en la ciudad de Posadas, Misiones. Se han conseguido identificar una confianza de hasta un 65% de esos materiales reciclables.

El último trabajo tiene como título “Herramienta para la asistencia prediagnóstico y seguimiento del tratamiento de la escoliosis idiopática del adolescente (EIA)”. Aquí se presentan los avances en el estudio, diseño y desarrollo de una herramienta

para la asistencia en el prediagnóstico y el seguimiento de desviaciones anormales corporales posteriores del cuerpo humano. Un ejemplo de éstas es la Escoliosis Idiopática Adolescente (EIA), que afecta a muchos adolescentes en la actualidad. El objetivo general del trabajo es desarrollar una herramienta que permita a los especialistas de la región llevar a cabo el prediagnóstico y seguimiento de la desviación anormal en la columna. Potencialmente, luego de validar los resultados y respetando los protocolos de uso, la herramienta podría ser utilizada por las personas para tener un diagnóstico temprano y acudir al especialista para realizar análisis más exhaustivos. Consecuentemente, se desarrolló un prototipo utilizando el procesamiento de imágenes y redes neuronales convolucionales (RNC), que permite identificar una desviación de la columna vertebral del paciente y calcular de forma automática algunos parámetros que habitualmente utilizan los especialistas.

El libro dedica un capítulo por cada trabajo con el objetivo de dar detalle de cada proyecto, presentar la problemática, la solución desarrollada y los resultados finales.

ANTECEDENTES DEL CENTRO

Es necesario recordar que la UGD comenzó siendo un Instituto Universitario. En ese entonces sólo existían carreras del Departamento de Administración y Comercialización y de Informática. Los docentes que trabajan desde ese momento inicial

recuerdan que las primeras indagaciones referidas a problemas de investigación en sistemas informáticos tenían que ver con ser soporte a la carrera de Administración. Se comenzaron a trabajar ideas para abordar la investigación, pero todavía de una manera muy incipiente porque el esfuerzo estaba en robustecer el plantel docente, los planes de estudio y planificar las visitas de profesores externos provenientes de la Universidad Nacional de La Plata (UNLP) y de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNICEN).

La circulación de docentes visitantes mantuvo vivo el intercambio, la fuerte impronta de la UNLP, de la UNICEN y de la Universidad Tecnológica Nacional (UTN) fueron una buena base para la formación de todos los docentes locales. Entre esos investigadores podemos nombrar a la Dra. Alicia Díaz y al Dr. Gustavo Rossi de la UNLP, el Dr. Zunino de la UNICEN y el Dr. Marcelo Karanik de la UTN, Facultad Regional Resistencia. La influencia de los profesores visitantes que forjaron la carrera estaba muy presente y eso hizo que haya fuentes de inspiración cercanas.

En los años de vida del Instituto se realizaron esfuerzos por promover la investigación en informática que fueron debilitándose a medida que los profesores visitantes disminuyeron su participación en las cátedras, este proceso fue de alguna manera restringiendo las posibilidades de los alumnos de vincularse con las prácticas de investigación.

En 1999, un pequeño equipo comenzó a trabajar en conjunto con una profesora visitante, la Dra. Alicia Díaz, quien junto a tres estudiantes de ese entonces (Pablo Castro, Héctor Takayashiki y Héctor Ruidías) impulsaron un grupo de investigación en Sistemas Hipermediales. El equipo se sostuvo un tiempo, pero no logró continuidad. Esto se debía a que la carrera aún era muy joven y la energía estaba puesta en consolidar el cursado y el plantel docente. En los años 2000 no había egresados aún y se necesitaba cubrir los cargos docentes de la totalidad de las materias. Los primeros egresados terminaron su carrera en 2004 y todavía no estaban dadas las condiciones para desarrollar proyectos de investigación. Muy diferente a lo que sucedía en Buenos Aires donde ya existía un sistema, un entramado de profesionales, docentes e investigadores en materia de las TICs. En ese entonces era la única carrera de grado en informática de la provincia de Misiones y norte de Corrientes, con lo cual radicar a un profesor/investigador formado en ese momento era poco viable.

EVOLUCIÓN DEL CITIC

En este apartado queremos diagramar hitos y comentar brevemente la memoria del Centro (CITIC), poniendo el énfasis en el espacio como un lugar de intercambio en la formación de jóvenes ingenieros/as y su posterior formación en carreras de posgrado.

El Centro nació por la necesidad de generar un entorno de construcción de conocimiento colectivo, donde alumnos, becarios, docentes y estudiantes de posgrado pudieran compartir el trabajo con investigadores y llevar a cabo proyectos relacionados con las TICS en la UGD.

Las acciones que realizaron los docentes investigadores para mantener vivo el CITIC durante este tiempo fueron: participar en convocatorias internas de la Secretaría de Investigación de la UGD; presentar resultados de proyectos en congresos y revistas especializadas; y la formación de recursos humanos. Todas estas acciones permitieron mantener el Centro con actividad, acompañando los cambios de planes de estudios y las recomendaciones de las evaluaciones externas de la Comisión Nacional de Evaluación y Acreditación Universitaria (CONEAU).

**PRIMERA ETAPA 2004-2010: SENTANDO
BASES DE LA CREACIÓN FORMAL**

Inicialmente en esta etapa se puede identificar al Centro como un anhelo. El impulso de formar un espacio donde se pudiera generar trabajo compartido, armado de prototipos y presentaciones de proyectos fue dado, en ese entonces, por estudiantes avanzados de la carrera.

Como mencionamos anteriormente, en 2004 las dos universidades referentes en informática eran la UNLP y la UNICEN, am-

bas tenían Institutos de Investigación y eran una gran fuente de inspiración para replicar en la provincia de Misiones. Sin embargo, la ubicación geográfica, la juventud de la carrera y la falta de masa crítica de investigadores de ese momento no hacían posible desarrollar actividades de investigación equivalentes a las que se desarrollaban en las grandes universidades.

Un momento importante de esta etapa fue la primera participación de alumnos y docentes de la carrera de Ingeniería Informática en el EST - Concurso de Trabajos Estudiantiles en la XXXIII JAIIO, realizada en la ciudad de Córdoba en el año 2004. Los estudiantes Héctor Ruidías, Diego Godoy y Carlos Toledo, guiados por los profesores Patricia Vila Torre y Marcelo Karanik recibieron una mención especial por uno de los trabajos presentados.

Al regreso de ese viaje, el equipo trajo la convicción de impulsar estrategias de trabajo diferencial que pudieran organizar, fortalecer, unificar esfuerzos, direccionar e institucionalizar las actividades de investigación en el Área las TICs, y fue presentado el proyecto denominado “Creación del Centro de Investigación de Tecnologías de la Información y Comunicaciones”, impulsado por el Ing. Diego Alberto Godoy, docente del entonces Departamento de Informática, al entonces Rector el Ing. Luis Lichowski y al Consejo Asesor, quienes dieron su visto bueno para la creación por Resolución N° 516/06 el 27 de noviembre de 2006. En su resolución de creación se designó como director al autor del proyecto,

el entonces Ing. Diego Alberto Godoy, y como miembros del Comité Científico del CITIC a los siguientes profesionales: Mg. Antonio Foti (Universidad Nacional de la Matanza); Dr. Costa Montenegro (Universidad de Vigo); Ing. Roberto Suenaga (Coordinador del Departamento de Informática).

El proyecto proponía contribuir a la formación de Investigadores en Áreas de las Tecnologías de la Información y Comunicaciones (TICs), así como promover el establecimiento y consolidación de equipos de investigación para la generación de conocimiento publicable y transferible.

En 2008 el equipo se presentó a la primera convocatoria ordinaria a proyectos y, a partir de ese momento, empezó una metodología nueva y viva. Esto se dio a través de construir una propuesta, cumplir la labor del cronograma e ir redactando informes parciales y finales. La escritura, como ejercicio conjunto y permanente, hizo que la actividad permanezca y continúe. Por su parte, el plantel docente comenzó a cursar posgrados fuera de la provincia y esto también hizo que el equipo fuera consolidándose.

La participación en la red Cyted TRICOMFOR - Tecnologías de redes inalámbricas para comunicación y formación, coordinada por el Ing. Roberto Suenaga quien, abrió la posibilidad de realizar estadias en la Universidad de Vigo en España.

SEGUNDA ETAPA 2010-2016: CONSOLIDANDO EL ESPACIO

Si continuamos enumerando los hitos que colaboraron en fortalecer al Centro como un espacio de investigación, podemos mencionar la importancia de los viajes y presentaciones a congresos. El “mostrar y compartir los resultados” en jornadas nacionales, motivaba a los estudiantes y a profesores. El acompañamiento de la institución para asistir a los congresos que se realizaban a lo largo y ancho del país resultó clave en este periodo. El efecto de viajar, compartir con pares y recibir devoluciones de los trabajos fue empoderando al equipo de la UGD.

Una de las cuestiones que sostiene la modalidad de trabajo del CITIC es la necesidad que tienen los alumnos de realizar un Trabajo Final. El trabajo para guiar un TFC es una práctica de diálogo y búsqueda constante de estrategias de registro y preguntas de investigación a partir de un problema concreto. Por su parte, los investigadores del CITIC guían y acompañan la elaboración del proyecto de TFC en el que los estudiantes realizan el trabajo de campo, de simulación y/o de laboratorio, recolección de datos, desarrollos y otras actividades necesarias para la concreción de los proyectos.

Otra de las cuestiones que hizo que el espacio siga vigente durante tantos años fue la formación en simultáneo de docentes en carreras de posgrado, tal es el caso de Diego Godoy quien en

ese entonces dictaba la materia Modelos y Simulación en la carrera, y comenzó a investigar sobre cómo relacionar los temas dados en una Maestría en Ingeniería de software con la simulación. Eso desencadenó una línea que se plasmó en un proyecto de investigación que estudió cómo simular el desarrollo de un proyecto de software, particularmente con metodologías ágiles. Del mismo proyecto surgieron una tesis de maestría y varios trabajos finales de carrera. En esa misma línea, continuando con la simulación y la ingeniería de software y desde un proyecto de Ciudades inteligentes, con el advenimiento de nuevas tecnologías de redes de sensores inalámbricos y la participación del Dra. Rebeca Pilar Díaz Redondo (UVigo) quien dirigió la primera tesis doctoral del CITIC defendida en 2016. Dicho trabajo generó la primera publicación en congresos internacionales y publicaciones en IEEE. En esta etapa fue fundamental el acompañamiento de la entonces Rectora Prof. Alba Alicia Perez Chilavet. Otros docentes participantes fueron el Ing. Hernán Bareiro, el Ing. Fabián Favet, el Ing. Juan de Dios Benitez, el Ing. Edgardo Belloni y la Dra. Liliana Dieckov (†)

TERCERA ETAPA 2016-2024: ARMANDO REDES Y PUBLICACIONES INTERNACIONALES

Este periodo se caracteriza por atravesar fases de acreditación y reacreditación de la carrera por parte de la CONEAU. Estos procesos han sido claves para interactuar con expertos que devol-

vieron recomendaciones e informes. Estas evaluaciones y miradas externas impulsaron ideas, preguntas de investigación y armado de Trabajos Finales de Carrera (TFC), y mantuvieron al equipo docente pensando y repensando temáticas, y acompañando a los y las estudiantes para finalizar sus trayectos.

La participación de la red CYTED CITIES - Ciudades inteligentes totalmente integrales, eficientes y sostenibles, la cual pretendía aglutinar las experiencias de grupos de investigación multidisciplinares para que se puedan establecer unas bases de “ciudad de futuro”, y replicable a cualquier escala, en los diferentes países integrantes de la acción, permitió más roce internacional y la posibilidad de presentaciones y participación en congresos como el Ibero-American Congress of Smart Cities. Varios de estos trabajos fueron seleccionados como los mejores del congreso para ser publicados en editoriales internacionales como Springer. La incorporación del Dr. E. Marcelo Albornoz, quien es investigador adjunto de CONICET del Research Institute for Signals, Systems and Computational Intelligence Sinc(i) y Profesor Adjunto de Universidad Nacional del Litoral fue y es trascendental para estos logros, debido a sus conocimientos de visión artificial y *machine learning*. Asimismo, en los últimos años, y dentro del proyecto “Tecnologías para desarrollos sostenibles de ciudades inteligentes apoyados por *machine learning*”, adjudicado en las doceavas convocatorias de proyectos de investigación internas de UGD, surgió la

incorporación del Mg. Lucas Kucuk y el Licenciado en Informática Carlos Kornuta, quienes continúan con su formación doctoral y trabajando en el CITIC.

2024: AL FUTURO

La Universidad ha dado el gran paso de nuclear a todos los investigadores del Departamento de Ingeniería y Ciencias de la Producción en el CITIC, como el Dr. Jhonatan Shuster, la Mg. Karina Eckert, el Ing. Lucas Rodoni y la Ing. Micaela Olivera. siendo un total de ocho investigadores e investigadoras en formación. Estos profesores desempeñarán diferentes proyectos relacionados con *smart cities*, Internet de las cosas, *machine learning*, visión artificial, metaverso y tecnologías aplicadas a la educación. Los títulos de los proyectos son: “Tecnologías para Desarrollos Sostenibles de Ciudades Inteligentes apoyados por *Machine Learning*”, “Aplicación de técnicas semánticas y de aprendizaje de máquinas a la Ingeniería de Software: Aplicación de MLOps en *fnetworking*” y “Aplicaciones de Tecnologías Inmersivas y Robótica a la Industria y la Academia”.

Capítulo 1

IOT - Verificación automática del
uso de elementos de seguridad
utilizando visión computacional



**Diego Alberto Godoy, Ricardo Selva,
Nicolás Ibarra, Cesar Gallardo, Carlos Kornuta,
Enrique Marcelo Albornoz**

INTRODUCCIÓN

Las personas involucradas en una organización son el activo más valioso, ya que de ellos depende que se realice una determinada actividad productiva como una tarea operativa manual o como el manejo de determinadas máquinas. Diariamente, estos empleados están expuestos a los peligros del ambiente del trabajo, en el que un accidente puede traducirse en heridas, lesiones, discapacidades e incluso un fallecimiento (Asfahl, 2010). Además, en algunos ambientes laborales están presentes ciertos agentes de riesgo que, en el largo plazo, pueden generar diversas enfermedades profesionales (Chávez, Vicuña, Camacho, 2020). La misión

de los EPP (Elementos de Protección Personal) es la de evitar o mitigar las consecuencias de los accidentes y agentes de riesgo, aumentando la seguridad del trabajador. De este modo, la falta de uso o la utilización incorrecta de los mismos supone un gran riesgo para los empleados y empleadores, estos últimos pueden incluso ser sancionados por el ente de control, aunque no se registraran accidentes. El uso de EPP es obligatorio para las operaciones de cualquier industria. En Argentina, existen normativas vigentes en la Ley N° 19587 de la Superintendencia de Riesgos del Trabajo (SRT) (Argentina, XXXX) , que exigen a los empleadores brindar a los empleados los EPP correspondientes, de acuerdo con las tareas que realicen, así como también controlar que utilicen esos elementos de manera correcta durante todo el desarrollo de la jornada laboral. Algunos de los EPP que pueden ser exigidos son el casco, el protector facial, el protector ocular, los guantes, el chaleco refractario y los zapatos. Cada EPP puede ser específico para determinados tipos de tareas. Por ejemplo, los guantes de látex están diseñados para ser utilizados en laboratorios, donde es necesario aislar las manos de elementos patógenos u otras sustancias. Existen también guantes anticorte para manipular elementos cortopunzantes, guantes térmicos para manipular elementos muy calientes o muy fríos, y guantes de trabajo que mejoran la adherencia de los elementos a la mano y protegen contra los abrasivos.

En los últimos años se han realizado avances en la detección automática de EPP. Por ejemplo, en Massiris et al (2021), la problemática planteada en el trabajo es la mala utilización de los EPP por parte de los trabajadores y visitantes en los sitios donde se llevan a cabo proyectos de construcción. Este planteo, además, está respaldado por estudios e informes europeos y norteamericanos, que indican que la mayoría de las lesiones laborales producidas por impacto en la cabeza se deben a la falta de uso del casco de seguridad. El objetivo planteado es diseñar y probar una metodología para la detección de la correcta utilización de EPP en vídeos sin cámara fija, a través de técnicas de Visión por Computador (CV) y aprendizaje automático, para prevenir accidentes laborales. En lo que corresponde a los aspectos metodológicos diremos que plantean el uso de un algoritmo y la red OpenPose para detectar los puntos antropométricos de un individuo y, en función de ellos, determinar a qué parte del cuerpo corresponde. El objetivo principal de este enfoque es determinar en primer lugar los puntos antropométricos, para luego verificar si en virtud de los puntos obtenidos, estos cuentan con el EPP correspondiente a esa región del cuerpo. Por ejemplo, obteniendo los puntos que vinculan a la cabeza del individuo, se procede a detectar si la persona lleva puesto el casco. También, al detectar los puntos antropométricos del tronco del individuo se verifica si está utilizando un chaleco refractario. Una desventaja de la investigación es que al de-

tectar EPP en función de los puntos antropométricos no se tienen en cuenta las clasificaciones parciales producidas por las posiciones de las personas mientras se encuentran trabajando ya que es probable que debido a la dinámica del trabajo, en oportunidades, las tomas de puntos antropométricos no capturen el cuerpo entero del individuo. También cabe destacar que otro aspecto no muy positivo es la cantidad de recursos que se deben emplear para llevar a cabo dicho experimento, tanto en términos de cómputos para el procesamiento de los puntos antropométricos, como de la implementación de la red y los recursos necesarios para capturar las muestras de las personas trabajando.

Otro trabajo relacionado (Sandru et al, 2020) menciona que los trabajadores situados en ambientes peligrosos y tóxicos muchas veces prefieren no utilizar los EPP correspondientes al realizar sus tareas. Esto puede ocurrir porque son incómodos o porque pueden afectar la exactitud de las acciones propias de la tarea ejercida. Se ha demostrado que una forma de obtener resultados precisos en la identificación de personas y elementos de protección personal es mediante la utilización de un detector y un clasificador. De esta forma, el prototipo se encarga de reconocer a la persona en una primera instancia y, sobre ese resultado, se realiza la identificación de los EPP. Siguiendo este último modelo, el trabajo propone un nuevo enfoque para mejorar la detección de elementos de protección personal utilizando tres componentes:

un detector de personas, un estimador de poses y uno o varios clasificadores de EPP (uno por cada tipo de EPP a detectar).

El detector de personas se basa en el detector de objetos de YOLOv3, que está preentrenado con el *dataset* MS COCO. Este detector ofrece buen rendimiento y velocidad para realizar clasificaciones. Por cada entrada devuelve los recuadros de los objetos reconocidos, con sus correspondientes etiquetas. En este caso, los objetos de importancia pertenecían a la categoría personas. El estimador de poses es utilizado para obtener la pose de cada persona identificada y está basado en Cao, et al (2016). Para cada entrada, el estimador procesa la imagen para recuperar la pose del cuerpo, brindando una representación por medio de un esqueleto, que está compuesto por nodos y enlaces que representan segmentos del cuerpo. Gracias a esta información es posible crear mapas de las partes de la persona que son de interés. El clasificador es una CNN que posee una arquitectura basada en la red neuronal MobileNetV2, la cual es equivalente a la red AlexNet en términos de precisión, pero mejor en términos de velocidad. Esta red está entrenada con el *dataset* ImageNet. Para el proceso de entrenamiento, se inicia con los modelos de detección de personas, estimador de poses preentrenados y uno o más clasificadores de EPP. La cantidad de clasificadores dependerá de la cantidad de elementos de protección personal a identificar. Se utilizaron 3 da-

tasets, cada uno de ellos sobre un determinado EPP, y en este caso el trabajo se centró en identificar cascos, máscaras y botas.

Cada imagen ingresa al detector de personas, que tiene como salida una imagen con recuadros que encierran a cada persona detectada. Estos cuadros fijan los límites para el procesamiento en los siguientes pasos. Luego se emplea el estimador de poses que tiene como salida un “esqueleto” del cuerpo que indica su posición y la ubicación de las distintas partes del cuerpo; de esta salida se recortan automáticamente las regiones del cuerpo que podrían contener EPP y se envían al clasificador correspondiente (la cabeza al clasificador de cascos, las manos al clasificador de guantes, etc). Cada imagen de entrada es procesada por una serie de bloques convolucionales para luego realizar la clasificación, se utiliza el detector de personas en imágenes de prueba para obtener los límites a clasificar y, en este caso, no se utiliza el estimador de poses, sino que se envían recortes de las subáreas al clasificador correspondiente (el área superior al clasificador de cascos, el área inferior al clasificador de botas, etc). Cada clasificador brinda una predicción binaria para cada sub área recibida, indicando si la parte del cuerpo analizada está utilizando el elemento de protección personal correspondiente.

Parte de la Gestión de Residuos Sólidos Urbanos (GIRSU) (Misiones. Posadas (municipio), 2024). se realiza en una planta de reciclaje localizada en el Centro Verde Municipal, que contribuye

con el desarrollo sostenible, se encarga de la reducción de residuos enviados a los centros de depósito finales, promoviendo la reutilización y reciclaje de los mismos (Figura 1). Esta planta se presenta como un ambiente industrial con muchas maquinarias, que plantean diferentes riesgos para la salud. Los operarios deben usar en todo momento sus elementos de protección personal, como cascos, guantes, zapatos y chalecos, entre otros.



Fig. 1. Trabajadores operando en la planta de reciclaje

El objetivo de este trabajo es abordar la identificación automática de EPP en una planta de residuos reciclables a través de visión computacional, aportando un modelo tendiente a mitigar los efectos negativos de los accidentes laborales en una industria.

LOS PRIMEROS PASOS

En esta sección, en primer término se detallan el análisis del modelo de datos, la elaboración de un prototipo de solución para la problemática abordada y la evaluación de ese prototipo utilizando la metodología CRISP-DM, por su flexibilidad y adaptabilidad a las características del proyecto. En segundo lugar, se realiza una descripción detallada de las imágenes que son utilizadas por un modelo de red neuronal para identificar si los trabajadores llevan puestos correctamente los EPP.

COMPRESIÓN DEL MODELO DE DATOS

Describimos a continuación las actividades teniendo en cuenta el estado del arte sobre imágenes de trabajadores utilizando los EPP en ambientes operativos. En primer lugar, se detallan la recopilación y preparación del *dataset* de imágenes para luego proceder al etiquetado a fin de realizar el modelado del prototipo

Recopilación y preparación del dataset de imágenes

Comenzamos con las imágenes del *dataset* Hard Hat Workers (China. Northeastern University, 2022) del que solamente se pudieron utilizar 1280 imágenes debido a limitaciones del plan de precios de Roboflow al momento de realizar la importación. Por otro lado, se tomaron 136 fotos de los trabajadores realizando sus

funciones en la planta de reciclado del centro verde (Figura 2) para evaluar el comportamiento del modelo con imágenes del escenario objetivo. En estas imágenes se pueden observar trabajadores cumpliendo sus tareas en la cinta transportadora y también se solicitó la ayuda de algunos de ellos para tomar imágenes de frente, con su equipamiento completo, y también se les pidió que se quiten algunos elementos para poder realizar distintas pruebas.



Fig. 2. Ejemplos de imágenes tomadas del centro verde municipal

Finalmente, se capturaron un conjunto de imágenes que permitieran la validación del prototipo. El *dataset* elaborado comprende una gran variedad de ejemplos y elementos que sirven para el entrenamiento, validación y testeo del prototipo, tal como refleja la Figura 3. En algunos casos se observan imágenes con *augmentation* (con cabezas que están invertidas a fin de poder mejorar el rendimiento de la red neuronal).

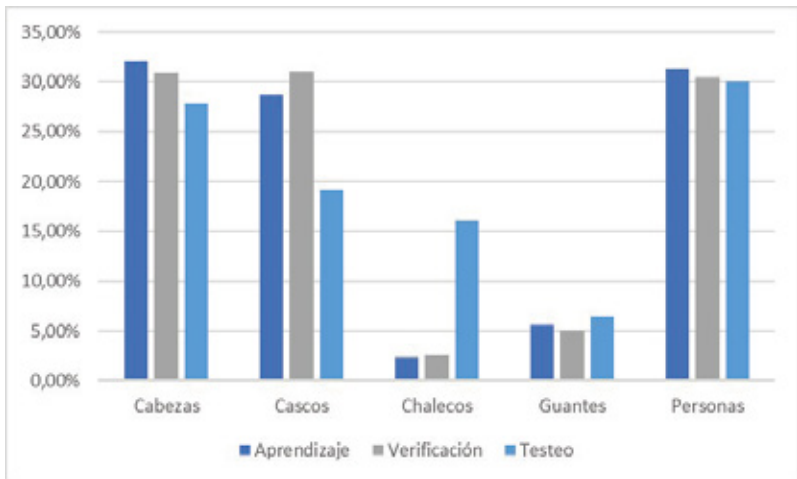


Fig. 3. Porcentaje de imágenes utilizadas para los procesos de entrenamiento, validación y testeo.

Etiquetado de los datos

Se determinó que se etiquetaría la clase casco delimitando el casco lo más precisamente posible, encerrando en las etiquetas

la menor área posible de la cabeza de los trabajadores. Otra clase etiquetada fue la de las cabezas que no tienen casco, y las que tienen casco puesto. El propósito de esta etiqueta es que el prototipo pueda determinar si la persona se encuentra en una posición donde sea visible el casco, ya que si se puede visualizar su cabeza, debería ser posible visualizar su casco.

Con el objetivo de determinar los sujetos sobre los que se deben realizar las verificaciones, se etiqueta el cuerpo entero de las personas que aparecen en las imágenes. Se intenta encerrar en la etiqueta el cuerpo entero evitando en lo posible incluir otras personas que puedan estar superpuestas.

Al igual que con los EPP previamente mencionados, se seleccionaron imágenes de trabajadores utilizando guantes de protección de varios tipos. Además, se determinó que etiquetar las manos descubiertas de las personas que no llevan guantes sería útil para determinar la falta de ese elemento protectorio, en una estrategia similar a la de los cascos y las cabezas.

Una vez definidas las clases y los lineamientos para el etiquetado de cada una, se procedió a realizarlo por imagen, observando con atención cada caso para definir los límites de cada etiqueta. Por cada persona se pudieron generar hasta 6 etiquetas, es decir, que como máximo se etiquetó un casco, una cabeza, un chaleco, dos guantes o manos y finalmente la persona. La Figura 4 presenta ejemplos de imágenes con su etiquetado final.



Fig. 4. Imágenes de trabajadores con sus correspondientes etiquetas

Se optó por repartir las imágenes proporcionalmente entre los *datasets* de aprendizaje, verificación y testeo. En las tablas 1, 2 y en la Figura 5 se puede observar el resultado final del *dataset*, que además cuenta con 65 imágenes sin etiquetas, de las cuales 46 forman parte del conjunto de entrenamiento y 19 del conjunto de verificación. Estas imágenes no contienen personas ni EPP para etiquetar y su propósito es el de intentar controlar y evitar los falsos positivos que pueda detectar la red durante el entrenamiento y corregir estas falencias.

Etiqueta	Aprendizaje	Verificación	Testeo
Cabezas	8054	866	90
Casos	7201	868	62
Chalcos	597	67	42
Guantes	1404	129	21
Personas	7853	793	97
Totales	25699	2630	323
Porcentaje	30%	9%	1%

Tabla 1. Descripción del *dataset*

	Aprendizaje	Verificación	Testeo
Cabezas	32,03%	30,34%	27,56%
Casos	28,08%	31,02%	19,20%
Chalcos	2,33%	2,57%	10,10%
Guantes	5,50%	4,95%	6,50%
Personas	31,27%	30,52%	30,01%
% Totales	100%	100%	100%

Tabla 2. Porcentaje de representación de cada clase en el *dataset*

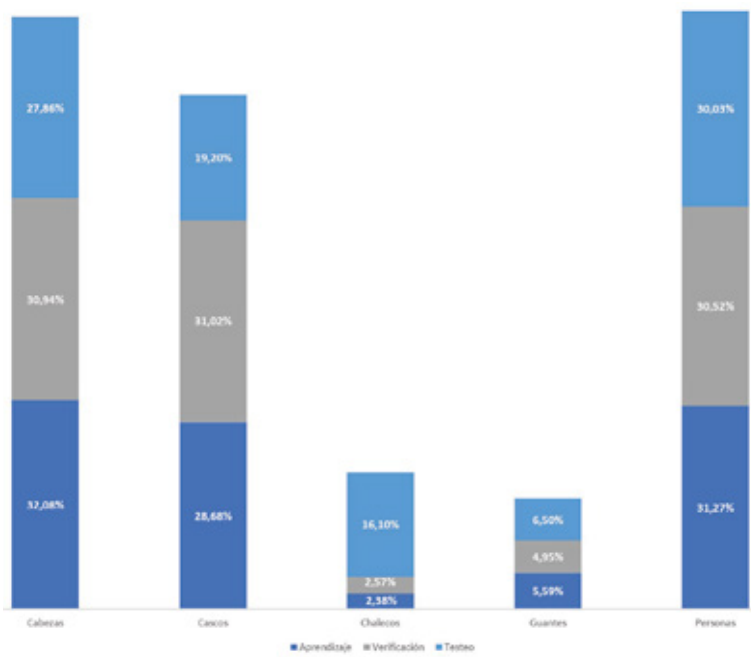


Fig. 5. Gráfico de representación de cada clase en los *datasets*

EL MODELADO DEL PROTOTIPO

Esta fase consiste en analizar y diseñar un prototipo de solución que sea capaz de satisfacer la necesidad de identificar los EPP. Uno de los trabajos de referencia obtiene un muy buen porcentaje de confianza (0.99) en la identificación de EPP en la industria de la construcción. Aunque fue necesaria la inversión de un alto costo computacional, para lograrlo se utilizan procesamiento

de videos de alta calidad, se obtienen las imágenes del mismo, se identifican los puntos antropométricos y luego se identifica si poseen los EPP. Con el objetivo de reducir los costos computacionales elevados que acarrea dicha implementación, se opta por resolver el problema de una manera más sencilla, empleando una red neuronal convolucional de detección de objetos en una imagen. Para esta implementación se decide utilizar la red YOLOv5 debido a su gran precisión, bajo costo computacional y completa documentación (ultralytics/yolov5, 2022). La salida de la red es procesada por un pequeño algoritmo que decide si una persona está utilizando sus EPP.

El diagrama de arquitectura del prototipo (Figura 6) ilustra los procesos de entrenamiento y funcionamiento, sus entradas y las respectivas salidas para cada uno. El objetivo es brindar al lector una estructura general y abstracta de los parámetros necesarios para que el prototipo de solución se ejecute de manera efectiva. La etapa de entrenamiento de la red requiere de un conjunto de imágenes etiquetadas con sus respectivas clases a fin de producir un conjunto de pesos que permitan ser utilizados a posteriori. Los pesos obtenidos en el aprendizaje y las imágenes capturadas por un dispositivo son utilizados para poner en funcionamiento la red. Dicha entrada se procesa dentro del componente y en función de los pesos introducidos a la red y la foto-

grafía utilizada como imagen de entrada determinará si la persona en la muestra posee los EPP reglamentarios.

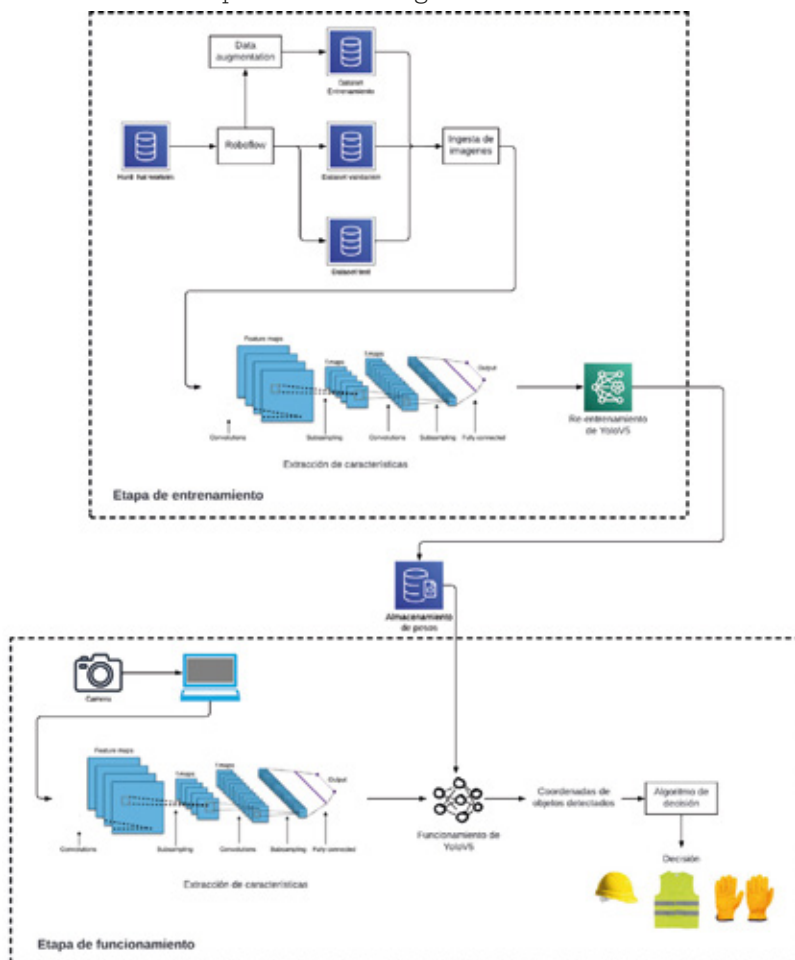


Fig. 6. arquitectura del prototipo propuesto

Las detecciones sirven como entrada al algoritmo encargado de decidir si las personas están utilizando sus EPP, en la Figura 7 se observa un diagrama de actividades UML que resume el funcionamiento del algoritmo, que toma cada persona detectada y, en base a sus coordenadas, encuentra las detecciones de EPP en ella. Estas detecciones se almacenan en una lista para consultas posteriores.

El criterio de decisión está dado si al menos una de las coordenadas del segmento está contenida en el área que representa a la persona. Esto se diseñó así para no descartar guantes o manos que pueden estar un poco alejadas del cuerpo y, a la vez, filtrar otras detecciones que pueden pertenecer a otra persona que pueda estar cerca.



Fig. 7. Diagrama de actividades de decisión

EL PROTOTIPO EN ACCIÓN

En este apartado se hace la descripción de las pruebas realizadas, además de la interpretación de sus resultados para evaluar el desempeño del prototipo diseñado, en distintos ambientes que

abarquen en gran medida las escenas en donde se podría utilizar la red neuronal convolucional entrenada. Por otro lado, también se evalúa este desempeño con el objetivo de encontrar las debilidades del modelo y sugerir trabajos futuros a realizar.

Con el propósito de obtener resultados más comprensibles, se procedió a modificar ligeramente el formato de salida del algoritmo de decisión. Además de las alertas escritas, al final de la ejecución del algoritmo se toma la imagen original y se marca con un cuadro a cada persona analizada. El recuadro es verde con la leyenda “Ok” cuando no se detectó la falta de algún EPP y rojo cuando el algoritmo estima que la persona no está usando alguno de los elementos pertinentes, en este último caso, la leyenda contiene los elementos en falta, abreviados en dos letras (CA para cascos, CH para chalecos y GU para guantes).

PRUEBAS EN MODO PORTAL

En la Figura 8 se puede observar la salida del prototipo con el modo portal activado. A la izquierda se evidencia una detección correcta debido a que la persona tiene sus guantes visibles, pero no está utilizando casco ni tampoco chaleco, como indica correctamente la leyenda “CA CH”. En la imagen de la derecha se puede observar otra detección correcta, el prototipo no marca ninguna falta de EPP ya que la persona tiene todos sus elementos de protección a la vista. independientemente de su posición.



**Fig. 8. Resultados correctos en modo portal.
Salida del algoritmo de decisión**

La Figura 9 presenta a la izquierda las detecciones de la red y a la derecha el resultado final del prototipo. En esta prueba se obtuvo un resultado incorrecto como se puede ver en las imágenes la persona está utilizando correctamente todo su equipo, sin embargo, en la imagen de la derecha se ilustra que el prototipo detecta que faltan sus guantes. Este problema se puede deber al contraste de brillo en la imagen, las áreas que están al sol provocaron que el resto de la imagen se vea más oscura y con un poco de ruido, lo que pudo provocar la falla en la detección de ambos guantes. Otro factor que pudo haber resultado en el error es la posición de la persona con respecto a la cámara, la persona se encuentra algo torcida, pero no debería ser un motivo de fallas.



Fig. 9. Resultados erróneos y correctos en modo portal

PRUEBAS EN LA CINTA TRANSPORTADORA

La Figura 10 presenta una prueba sencilla, donde se ve a un único operario de cerca, con un contraste aceptable y sus manos claramente visibles. A la izquierda se encuentra la salida gráfica de la red, donde es destacable que ninguno de los objetos de la

cinta confundió a la red y todas las detecciones fueron correctas con buenos valores de confianza. A la derecha se encuentra la salida gráfica del prototipo, donde se puede observar que el mismo detecta que falta el chaleco y los guantes del trabajador. En esta prueba se destaca la detección del casco, ya que su color es distinto a los cascos que normalmente se usan en el centro verde, esto indica que el prototipo es capaz de adaptarse a ligeros cambios en los EPP sin necesidad de reentrenar la red.



Fig. 10. Resultado correcto sobre operario de la cinta

Se procede a probar el prototipo en condiciones más difíciles, las figuras 11 y 12 presentan un escenario de detección de un operario de espaldas y personas que se ven parcialmente. En

el caso del operario de espaldas, si bien no se pudo determinar el uso de guantes, se detectó correctamente el casco y el chaleco. Se destaca además la correcta detección del operario de la derecha a pesar de poder verse solo la mitad de su cuerpo, demostrando que el prototipo es capaz de funcionar en escenarios donde algún objeto obstruye parcialmente la visión sobre los trabajadores. En las personas de la izquierda se empiezan a notar algunas fallas sobre el prototipo, es aceptable que no se detecte a la persona que está de espaldas ya que solo se ve menos de la mitad de su cuerpo. Es esperable que la red detecte al menos la mano, sin embargo, a juzgar por la leyenda, no fue detectada ya que solo se marca la falta de casco y chaleco. Si bien no se pueden ver las manos de la persona marcada, por el funcionamiento del algoritmo de decisión se debía detectar la falta de guantes en esa persona. Las detecciones que se observan en la imagen de la izquierda son totalmente correctas y con buenos valores de confianza, con la salvedad de la mano de la persona de la izquierda que no fue detectada, así como tampoco el guante de la persona de la derecha, aunque esto se debe a que en la foto salió movida. Si bien en la salida gráfica no se muestran correctamente los valores de confianza de la persona de la derecha, en la salida de texto se observan valores de 0.24 para el casco y 0.47 para el chaleco los cuales son muy buenos considerando que solo es visible una parte de ellos.



Fig. 11. Resultados sobre operarios de la cinta



Fig. 12. Resultados mixtos con personas superpuestas. Salida del algoritmo de decisión

PRUEBAS EN OTROS AMBIENTES

Las siguientes pruebas (Figura 13) se realizan con imágenes tomadas de internet de distintos sitios y las fotografías representan a trabajadores de obras públicas de otros entornos, fuera del centro verde municipal, como trabajadores de obras públicas de la ciudad de Posadas, trabajadores de Vialidad Provincial y de Vialidad Nacional. El objetivo de estas pruebas es probar la generalización del prototipo y la posibilidad de poder reutilizarlo adaptándolo a otros escenarios.

PRUEBAS DEL DETECTOR EN UNA APLICACIÓN MÓVIL

Las siguientes imágenes son una demostración de cómo funciona la aplicación utilizando la cámara del celular. Es posible verificar el buen desempeño de la red en una aplicación, aunque la misma también tenga ciertas cuestiones por corregir, se evidencia que es factible la utilización de un modelo basado en visión computacional para la detección de EPP. La Figura 14 expone un conjunto de imágenes de prueba que pasaron por el algoritmo de decisión de la aplicación móvil. En las primeras tres imágenes se describen los resultados positivos de la red ya que en las primeras dos imágenes el prototipo reconoce a los operarios y sus EPP, mientras que en la tercera y cuarta imágenes, también es un caso positivo ya que el operario no se encuentra utilizando el chaleco refractario y la red logra detectar su ausencia. Por último, utilizando las mismas imágenes se realiza la detección en tiempo real¹ utilizando la cámara del teléfono móvil, captando las imágenes propuestas para evidenciar el funcionamiento de la aplicación.

1 <https://www.youtube.com/watch?v=xhQAPWqrqtY>



Fig. 14. Pruebas de detección con la aplicación móvil

CONCLUSIÓN

Se desarrolló un prototipo de un sistema de visión computacional para la identificación de EPP en escenarios industriales, en particular se probó en el centro verde municipal de la ciudad de Posadas. A partir de los resultados expuestos, es posible decir que resulta factible su implementación y se destacan las características que deben ser tenidas en cuenta para futuras investigaciones y desarrollos.

Es destacable que con los parámetros del entrenamiento y el *dataset* utilizado se lograron grandes resultados en una primera iteración de un proyecto de estas características, que apuntan a optimizar el proceso de control sobre la seguridad laboral de los operarios de una industria. La red se comporta muy bien en una aplicación móvil, otra evidencia de que el proyecto tiene un alto grado de factibilidad y que es posible en un futuro la implementación de un modelo semejante en un escenario como el centro verde municipal. Es por ello que se hace hincapié en continuar evaluando y mejorando el modelo propuesto a fin de realizar entre todos un aporte significativo al conocimiento local.

Considerando lo mencionado, los resultados cuantitativos de confianza, que fueron obtenidos por la red en la línea de clasificación del centro verde municipal fueron buenos, aunque aún se presentan fallas y ciertos aspectos a mejorar. Esto da espacio

a seguir reentrenando el modelo con los pesos obtenidos, a modo de mejora continua.

BIBLIOGRAFÍA CITADA

Argentina. (2019). *Equipos y elementos de protección personal*. Recuperado de https://www.argentina.gob.ar/sites/default/files/04_guia_equipos_y_elementos_de_proteccion_personal_ok.pdf.

Asfahl, C. R., (2010). *Seguridad industrial y administración de la salud*. Falta Lugar: Pearson Educación.

Cao, Z. ; Simon, T.; Wei, S.-E. ; Sheikh, Y. (2016). Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. Recuperado de <https://arxiv.org/abs/1611.08050>.

Chávez, L. E. S.; Vicuña, J. W. U.; Camacho, L. B. C. (2020). Evaluación de la Exposición a agentes de riesgo físico en centros de salud. *Polo del Conocimiento: Revista científico – profesional*, 5(10), 424-439.

China. Northeastern University (2022) Hard Hat Workers Object Detection Dataset. Recuperado de <https://public.roboflow.com/object-detection/hard-hat-workers>.

Massiris, M.; Fernández, J. A.; Bajo, J.; Delrieux, C. (2021). Sistema automatizado para monitorear el uso de equipos de protección personal en la industria de la construcción. *Rev. iberoam. autom. inform. ind.*, 18(1), 68-74.

Argentina. Misiones. Posadas (Municipio). (2024). Centro Verde Municipal. Recuperado de <https://posadas.gov.ar/sustentable/centro-verde-municipal/>.

Sandru, A. ; Duta, G.-E. ; Georgescu, M. I. ; Ionescu, R. T. (2020). Su-
PER-SAM: Using the Supervision Signal from a Pose Estimator to
Train a Spatial Attention Module for Personal Protective Equipment
Recognition. Recuperado de <https://arxiv.org/abs/2009.12339>.

ultralytics/yolov5 (2022). Recuperado de <https://github.com/ultralytics/yolov5>.

Capítulo 2

IOT - Detección de botellas de PET en una Planta de reciclaje



**Diego Alberto Godoy, Enrique Marcelo Albornoz,
Ricardo Selva, Nicolás Ibarra, Cesar Gallardo**

INTRODUCCIÓN

Hace ya muchos años que el mundo es desafiado por problemas que comprometen al medio ambiente, como el deterioro de la capa de ozono (Romo-Gómez et al, 2019), la deforestación masiva de árboles, la caza furtiva de animales o la contaminación ambiental. De este último problema existen varios tipos, como la contaminación del suelo, del agua, del aire, la radiactiva, la acústica o la contaminación química, siendo los primeros tipos mencionados los que más suelen verse. Esto no solo contamina el medio ambiente sino que también causa enfermedades peligrosas y graves, y en el peor de los casos, la muerte (Arcos Medina et al, 2018).

Un aspecto positivo del reciclaje es la reutilización de las materias primas, ya que reciclarlas contribuye a frenar la explotación de los recursos naturales. Por ejemplo, al reciclar papel y cartón se contribuye a disminuir la deforestación y con esto se ayuda a preservar el medio ambiente (Citlalic Gonzales Martinez, 2001). Por estas razones, el reciclaje de basura resulta ser una tarea de gran importancia no solo para nuestro medio ambiente, sino también para nuestros vecinos y para la comunidad. Sin embargo, hay algunos aspectos que se deben tener en cuenta antes de pensar en una solución. Según estudios realizados por la organización Natural Resources Defense Council (NRDC) en América Latina (Moyer, 2018), si la basura se separara adecuadamente antes de llegar al basurero municipal, se podría reciclar en promedio casi el 92% de la misma, pero como los ciudadanos no suelen diferenciar sus residuos, solamente es posible reciclar el 30%. En el 2020, en la ciudad de Posadas se produjeron en promedio 8.000 toneladas de residuos por mes¹ que equivale a aproximadamente 266 toneladas por día, pero teniendo en cuenta el porcentaje mencionado es probable que sólo se procesen unas 80 toneladas al día (de las cuales se recicla una parte bastante baja). Este número crece constantemente debido al incremento de la población y de las grandes empresas. El proceso de reciclado se lleva a cabo a tra-

1 Ver <https://posadas.gov.ar/sustentable/servicio-de-recoleccion-general-de-residuos-solidos-urbanos/>

vés de una planta de recolección de residuos ubicada en el centro verde municipal en la Ciudad de Posadas, que forma parte de la Gestión Integral de Residuos Sólidos Urbanos (GIRSU).

Como problemática principal, a pesar de que existe una ordenanza correspondiente a la correcta separación de residuos en la ciudad de Posadas (la Ordenanza VI, N° 29, dictada por el Honorable Concejo Deliberante), las personas no están acostumbradas a realizar una correcta diferenciación de la basura y usualmente mezclan residuos orgánicos con residuos inorgánicos. Esta es la principal causa por la cual no se puede reciclar la mayor parte de la basura diaria, ya que los residuos orgánicos se descomponen y vuelven inservibles a aquellos residuos que se podrían reciclar. Por esta razón, antes del proceso de clasificación de un tipo de residuo en particular (como plástico, cartón, etc.), en GIRSU se realiza un proceso de preclasificación manual, donde se determina qué residuos se pueden reciclar y cuáles no.

Automatizar este proceso de clasificación es una tarea compleja ya que se debe emular al cerebro humano para lograr reconocer imágenes. Es una actividad inmersa en un proceso rápido y continuo para los operarios de GIRSU, lo que implica que una computadora debe ser entrenada para reconocer elementos que se presentan en distintas posiciones, tamaños, colores y superpuestos con otros elementos. Una posible solución podría ser aislar cada elemento sorteando las irregularidades y clasificarlos

para simplificar el proceso. Todo esto debe realizarse en cuestión de segundos ya que la cinta transportadora no se detiene y se deben procesar grandes volúmenes de residuos.

Existen técnicas de procesamiento digital de imágenes con Redes Neuronales Convolucionales que pueden dar soporte a la identificación de residuos de manera automática. Sin embargo, para aplicarlas es necesario establecer un conjunto de restricciones y condiciones sobre determinadas operaciones de clasificación de residuos, considerando cómo llegan los residuos a GIRSU hoy en día.

TRABAJOS RELACIONADOS

En Castrillón Medina (2019a) se aborda la clasificación de residuos haciendo uso de Machine Learning utilizando Google Cloud Platform, Microsoft Azure y un modelo diseñado por los autores. El principal problema que describe es la inexistencia de algún tratamiento o proceso de gestión para la inevitable generación masiva de basura por parte de las personas en las ciudades. El objetivo fue desarrollar una API basada en técnicas de visión por computadora, soportada en servicios Cloud, que permita realizar el reconocimiento de residuos y determinar la pertenencia de un elemento al grupo de residuos reciclables.

Se toman dos enfoques para el proceso de clasificación de residuos. El primero se basa en Auto Machine Learning, el cual

junto con servicios Cloud permite diseñar automáticamente un modelo de Machine Learning para un *dataset* determinado. En concreto, se utilizaron Cloud Vision API y AutoML (que son productos de Google Cloud Platform) y Custom Vision (que es un producto de Microsoft Azure). El segundo se basa en el diseño manual de un modelo de visión por computadora haciendo uso de redes neuronales convolucionales y Transfer Learning, que permite tomar un modelo previamente entrenado y adaptarlo a un problema en particular. Se realizaron varios experimentos combinando arquitecturas como VGG16, VGG19 y MobileNet, al mismo tiempo que se ajustaban hiper parámetros como el número de épocas y el tamaño de *batch*.

El set de datos utilizado, con 2527 imágenes de residuos, está disponible online de manera gratuita en Thung (2024). Dicho *dataset* está constituido por 501 imágenes de vidrio, 594 de papel, 403 de cartón, 482 de plástico, 401 de metal y 137 de residuos no reciclables.

Para las pruebas, las imágenes fueron tomadas sobre un fondo blanco con luz tanto artificial como natural. El tamaño de cada imagen es de 512 x 384 píxeles. Los dispositivos utilizados para tomar las imágenes fueron: Apple iPhone 7 Plus, Apple iPhone 5S y un Apple iPhone SE. Los resultados obtenidos fueron un 76% de acierto para el modelo diseñado por los autores, mientras que hubo un 93% y 98% de aciertos con herramientas Cloud.

Tanto el código fuente de los experimentos realizados, así como los resultados completos pueden ser encontrados en Castrillón Medina (2019b).

A diferencia del trabajo mencionado se utilizó Deep Learning para la identificación de residuos reciclados en el Centro Verde Municipal. Para el entrenamiento de la red neuronal se utilizarán 300 imágenes de cada tipo de residuo, mientras que para las pruebas, 100 imágenes de cada uno. En ambos procesos, de entrenamiento y pruebas, las fotografías se tomarán con una cámara GoPro.

Por otro lado, en Thokrairak et al (2020) se propone la clasificación automática de plásticos, botellas de vidrio y latas utilizando la red neuronal convolucional SSD-MobileNet. La motivación de este trabajo fue la gran generación de residuos en las playas, donde la mayoría de los residuos son reciclables pero la gente no tiene el interés o conocimiento para arrojar los residuos en el contenedor correspondiente. El set de datos a utilizar contiene un conjunto de tres tipos de residuos: botellas de plástico, botellas de vidrios y latas de metal, donde para cada tipo se evalúan 317 imágenes con una gran variedad de marcas. Para botellas de plástico se utilizaron 31 marcas distintas, para las botellas de vidrio 19 marcas y para las latas 18. En primer lugar se realizó el etiquetado del *dataset* para luego realizar el entrenamiento. Para ello se utilizó una computadora con las siguientes características:

CPU Intel Core i7-6700, 8 GB RAM y el GPU GEFORCE GTX 1070 TI 8 GB GDDR5.

Para las imágenes de prueba se utilizó una cámara Raspberry Pi. Se realizaron cinco corridas de prueba para cada tipo de residuo, con 20 imágenes en cada una. Como resultado, para las latas de metal se tuvo un acierto del 86%, para las botellas de plástico se tuvo un 95% de aciertos y para las botellas de vidrio se obtuvo un porcentaje de acierto del 82%.

Este trabajo, al igual que en el presentado en este artículo, implementa una red neuronal convolucional (CNN): la SSD-MobileNet.² Una diferencia a tener en cuenta son los residuos a clasificar, que son botellas de plástico, botellas de vidrio y latas de metal, mientras que en el trabajo a desarrollar se identificarán distintos tipos de plásticos, como lo son el *pet* verde, *pet* cristal, *pet* celeste y plástico soplado. Además, en el trabajo mencionado se utilizaron en promedio 300 imágenes por tipo de residuo a clasificar, mientras que en el presente trabajo se utilizarán 400 imágenes por tipo de residuo entre las etapas de entrenamiento y pruebas. El proceso a seguir a partir de las imágenes será similar, ya que consiste en el etiquetado de las mismas para su posterior entrenamiento.

2 Ver Howard et al (2018)

METODOLOGÍA

En esta sección se utiliza la metodología Crisp-DM, que es ampliamente utilizada en proyectos de minería de datos en múltiples industrias (Schröer et al 2021).

ANÁLISIS DEL MODELO DE DATOS

En esta primera fase se pretende entender y determinar qué tipos de imágenes serán las mejores como entrada al prototipo para que éste pueda funcionar correctamente. Dichas imágenes deben ser realistas y reflejar apropiadamente el escenario y las condiciones bajo las cuales trabajará el prototipo. Para realizar esto se utilizaron imágenes existentes provenientes de distintos *datasets* de residuos, así como también imágenes de autoría propia tomadas en el escenario de aplicación.

DATOS EXISTENTES

Los dos primeros conjuntos que se utilizaron para realizar el relevamiento de imágenes fueron de Bottles and Cans Images³ y Garbage Classification⁴ donde existen una gran variedad de *pet* cristal, celeste, verde y soplado, dispuestos en fondos blancos que

3 Ver Starter: Bottles and Cans Images en <https://www.kaggle.com/code/kernelel/starter-bottles-and-cans-images-aa8808c7-4>.

4 Ver Garbage Classification en <https://www.kaggle.com/datasets/asdasdasas-das/garbage-classification>.

sirven de punto de partida para proporcionar a la red conocimiento sobre los datos que se deben identificar.



Fig. 1. Ejemplos de imágenes contenidas en *dataset* Garbage Classification

Datos adquiridos

Los datos adquiridos son aquellos recopilados directamente del centro verde municipal, así como también imágenes que fueron tomadas desde los hogares de distintas personas donde se expusieron a los *pet* a situaciones que simulen el escenario de aplicación.

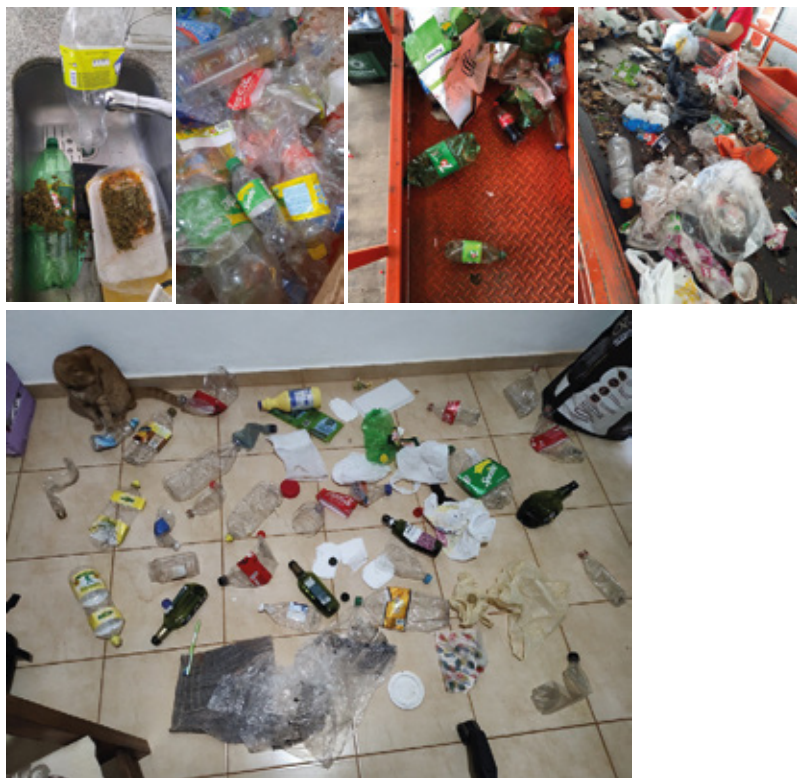


Fig. 2. Ejemplos de imágenes tomadas con cámaras fotográficas (imágenes propias)

Preparación de los datos

Una vez que las imágenes han sido definidas, siguen dos procesos fundamentales para que el prototipo pueda reconocerlas correctamente y obtener una buena precisión, el etiquetado y formación del *dataset*.

Etiquetado de imágenes

Este proceso consiste en clasificar cada imagen del conjunto, asignándole una etiqueta que la represente. De esta forma, el prototipo aprenderá los patrones deseados durante la etapa de entrenamiento.

Hay que tener en cuenta que la asignación de etiquetas se realiza de esta manera ya que las redes neuronales convolucionales utilizan aprendizaje supervisado. En este caso se utilizó Robo-flow⁵ como herramienta de etiquetado.

Definición del dataset

En esta fase se conforman, con las imágenes ya etiquetadas, distintos subconjuntos de imágenes que tendrán distintos propósitos: entrenamiento, validación y testeo del prototipo.

Los ejemplos de entrenamiento y validación fueron seleccionados de manera aleatoria mientras que para el testeo se eligieron exclusivamente ejemplos del escenario de aplicación, es decir, la planta de reciclaje.

⁵ En <https://roboflow.com/>.

	Cantidad de imágenes	%	Primeras	Últimas por Clase	% de cada clase en cada dataset
Imagenes Wario	530	69	Celeste	350	41
			Crystal	302	40
			Venice	52	7
			SaplaCo	32	3
Val.Ju. i 51	297	28	Celeste	127	32
			Crystal	132	39
			Venice	26	8
			SaplaCo	12	3
Texaco	29	3	Celeste	11	12
			Crystal	27	61
			Venice	3	7
			SaplaCo	2	10
Totales	317	100			1

Tabla 1. Disposición de imágenes del dataset según su tipo y clase

MODELADO

En esta etapa se analizaron los distintos resultados brindados por cada CNN candidata de implementación en el prototipo. Se realizó un análisis en función del tiempo de ejecución y la precisión obtenida a fin de seleccionar el modelo pre entrenado que mejor se adapte a las particularidades del dominio.

Número de épocas: 25

Redes Neuronales	Tiempo de ejecución (min)	Época de mejor precisión	UAR (val)	UAR (test)
ALEXNET	3	6	0.87	0.71
GOOGLENET	3.3	13	0.56	0.3
INCEPTION	3	3	0.33	0.34
RESNET18	3	15	0.87	0.79
VGG16	9.6	6	0.87	0.79
YOLO V5		18	0.84	0.85

Tabla 2. Resultados de entrenamiento en 25 épocas

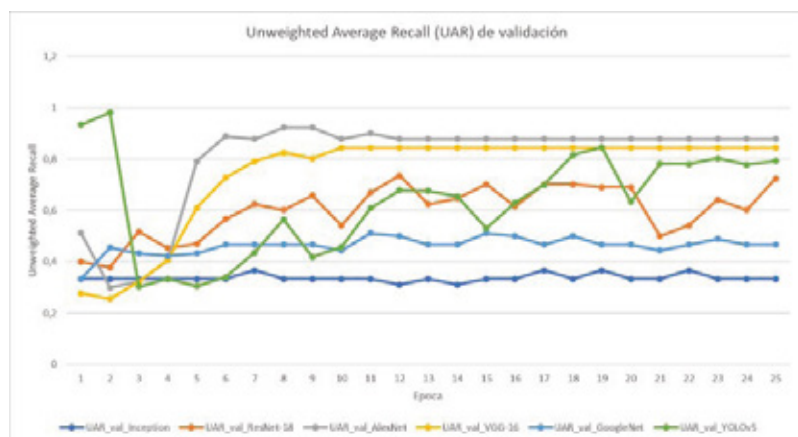


Fig. 3. Progreso de la precisión de validación durante el entrenamiento en 25 épocas

Los datos obtenidos surgen de las métricas *recall* y UAR (unweighted average recall), donde el *recall* se define como la cantidad de aciertos en la detección de una clase sobre la cantidad total de imágenes de la clase en cuestión. Mientras que UAR es el promedio de todas las *recall* por clases. Las mismas se puede ver en la siguientes ecuaciones 1 a 4.

$$\text{recall}_{\text{pet celeste}} = \frac{\text{cantidad de aciertos clase celeste}}{\text{cantidad de imágenes de clase celeste}} \rightarrow (1)$$

$$\text{recall}_{\text{pet cristal}} = \frac{\text{cantidad de aciertos clase cristal}}{\text{cantidad de imágenes de clase cristal}} \rightarrow (2)$$

$$\text{recall}_{\text{pet verde}} = \frac{\text{cantidad de aciertos clase verde}}{\text{cantidad de imágenes de clase verde}} \rightarrow (3)$$

$$\text{UAR} = \frac{\text{recall}_{\text{pet celeste}} + \text{recall}_{\text{pet cristal}} + \text{recall}_{\text{pet verde}}}{3} \rightarrow (4)$$

ARQUITECTURA DE LA SOLUCIÓN

El prototipo propuesto está implementado en Python y consta de dos etapas. En la primera se toman las imágenes, que pueden provenir de distintas fuentes, y en la segunda etapa se implementa la red neuronal convolucional elegida y entrenada para el dominio del problema. Dicha red se encargará de clasificar el objeto presente en la imagen. La Figura 6 muestra gráficamente a grandes rasgos los procesos realizados por el prototipo.

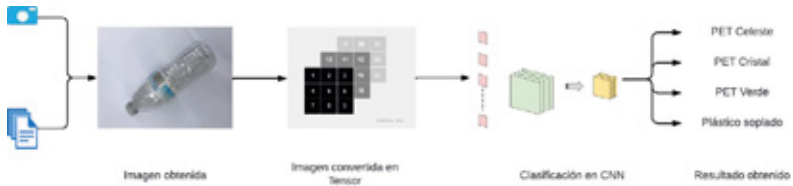


Fig. 4. arquitectura de la solución propuesta

Datos de entrada

El prototipo puede tomar imágenes de diversas fuentes, como cámaras de video o de vigilancia, donde se toman capturas con ciertos intervalos de tiempo establecidos; o bien se pueden tomar imágenes existentes en el sistema en formato jpg, jpeg o png. Estas imágenes se transforman en Tensors de Pytorch⁶ para su posterior tratamiento en la siguiente etapa.

Inferencia

En esta etapa comienza el funcionamiento del prototipo. El mismo se realiza por medio del lenguaje de programación Python utilizando la librería Pytorch y la red neuronal convolucional YOLOv5.

Una vez que el modelo pasó por la etapa de entrenamiento, funciona analizando imágenes de testeo que fueron tomadas del escenario de aplicación.

6 The Pytorch Foundation (<https://pytorch.org/foundation>).

ESCENARIOS DE PRUEBAS

Clasificación de botellas a una distancia cercana, sin nada a su alrededor

En las pruebas iniciales de laboratorio fueron utilizados los distintos *pet* cristal, verde y celeste, dispuestos en superficies cercanas, limpias y con fondo blanco. Esto se realizó con el fin de buscar una primera aproximación de resultados y evaluar la factibilidad del proyecto. Las siguientes imágenes fueron clasificadas después del entrenamiento y demuestran un acercamiento inicial muy apropiado evaluando las métricas obtenidas.



Fig. 5. Imágenes de prueba en laboratorio

Clasificación de botellas a una distancia lejana, sin nada a su alrededor

En una segunda instancia de pruebas se han evaluado distintos escenarios presentados en las figuras 6 y 7, en las que se evidencia que el prototipo es capaz de identificar los elementos objetivo para el proyecto.

En la figura de la izquierda, la confianza es relativamente baja aunque se observa de manera empírica que el modelo es capaz de identificar las botellas *pet* cristal que se encontraban en el suelo y en la superficie de la mesa. Por otro lado, en la figura de la derecha, la red logra reconocer con una confianza del 51%, dando indicios de que es posible reconocer estos residuos reciclables en superficies alejadas y con un escenario computacionalmente complejo de reconocer.



Fig. 6. Botellas lejanas sin objetos a su alrededor

Una posible mejora para la figura 7 es aplicar técnicas de procesamiento de imágenes para mejorar la *performance* de aprendizaje de la red aplicando, por ejemplo, una rotación de imágenes en distintos tipos de ángulos, creando mosaicos de la misma imagen para mostrar el mismo escenario hasta 4 imágenes de la original en una nueva muestra que contiene las mismas.



Fig. 7. Botellas lejanas sin objetos a su alrededor

Clasificación de botellas a una distancia lejana, con basura a su alrededor

Otro escenario complejo es el siguiente (figura 8 y figura 9), donde las botellas no solo están dispersas por el suelo verde, pudiendo prestar a confusión con las botellas *pet* cristal verde y cristal. Pero teniendo en cuenta esta disposición de métricas visibles

en la figura ocurrió todo lo contrario, existe una confianza promedio del 23% dando otro claro ejemplo de que es posible la identificación de residuos en ambientes con mucho ruido.



Fig. 8. Escenario complejo para PET verde y cristal

Es deseable destacar los siguientes resultados, dado que se aproximan al escenario planteado, en el que los residuos se encuentran dispersos en una superficie con otros elementos a su

alrededor, por ejemplo, otros residuos que no están al alcance de este proyecto. A pesar de todo, se observa que la confianza obtenida en algunos casos es relativamente baja en comparación con otros reconocimientos del mismo prototipo. Por lo tanto, por más que las métricas no sean lo suficientemente altas para pensar en una implementación, sí se puede pensar en trabajos futuros, ya que es evidente que es factible invertir conocimiento en el prototipo propuesto.



Fig. 9. Botellas lejanas con basura a su alrededor

Clasificación de botellas a una distancia cercana, con basura a su alrededor

Las siguientes imágenes fueron tomadas de la cinta transportadora de la planta de residuos reciclables y reflejan que en ambientes relativamente limpios de otra basura es posible poder

reconocer los elementos objetivos con un porcentaje de confianza de 0.65.



Fig. 10. Botellas en cinta transportadora de basura

Las figuras 10 y 11 demuestran que es posible la identificación de los *pets* en el centro verde municipal de la ciudad de Posadas, abriendo las puertas a nuevas mejoras y sirviendo de puntapié inicial para la automatización de los procesos industriales a través de inteligencia artificial en el municipio.



Fig. 11. Botella cercana con basura a su alrededor

CONCLUSIÓN Y TRABAJOS FUTUROS

Una red pre entrenada puede ahorrar mucho tiempo y recursos si se quiere construir un prototipo de identificación de imágenes. Esto ocurre principalmente porque ya cuenta con valores establecidos en los pesos de las conexiones de sus neuronas y no es necesario entrenar la red desde cero sino que, mediante el entrenamiento con imágenes que reflejen el escenario de aplicación, es suficiente para obtener buenos valores de confianza.

Como lo indican las pruebas, YOLOv5 resultó ser una red neuronal convolucional que ofrece muy buenos resultados en cuanto a la múltiple identificación de objetos. Sin embargo, como todo modelo de predicción, puede verse afectado por factores

como la luminosidad del entorno, la captura parcial o completa del objeto, la superposición entre cuerpos o el enfoque y la distancia a los mismos. Estas circunstancias pueden confundir al modelo, generando resultados erróneos como falsos positivos, clases equivocadas o bajos porcentajes de acierto. Al mismo tiempo, el prototipo presenta la posibilidad de mejora de los resultados mediante la entrega, de manera exhaustiva, de entradas que reflejen todas las condiciones y circunstancias con las cuales se podría encontrar a la hora de realizar una evaluación.

Por todo lo mencionado anteriormente se espera como trabajo futuro el estricto entrenamiento del modelo a través de imágenes que brinden una fiel representación del escenario, así como también un aumento en la cantidad de épocas ejecutadas, con el objetivo de aumentar su precisión al reconocer residuos reciclables. También evidencia la posibilidad de implementación del prototipo en un ambiente real, como la planta de reciclaje en el Centro Verde Municipal, lo cual servirá para analizar la información generada por la masiva cantidad de residuos procesados y establecer estrategias de reciclaje. Además, el prototipo abre horizontes para el desarrollo de un sistema de clasificación de residuos completamente automatizado, que redundaría en la disminución del tiempo de exposición de los trabajadores a sustancias dañinas y el aumento de la productividad de la entidad beneficiada.

BIBLIOGRAFÍA CITADA

- Arcos Medina, G.; Armijo-Arcos, F.; Oñate, A., Pastor, D.; Jerves-Cobos, R. (2018) Simulación para Estimación de Muertes por Cáncer de Pulmón por Contaminación Ambiental de PM2.5. *Revista Ciencia Unemi*, 11(27), 97-110.
- Argentina. Misiones. Posadas (Concejo Deliberante). (2021). Ordenanza N° VI, 29.
- Argentina. Misiones. Posadas (municipio) (2024). Servicio de Recolección general de Residuos Sólidos Urbanos. Recuperado de <https://posadas.gov.ar/sustentable/servicio-de-recoleccion-general-de-residuos-solidos-urbanos/>.
- Castrillón Medina, M. (2019a). REDO: Sistema de reconocimiento de desechos reciclables. Recuperado de https://bibliotecadigital.udea.edu.co/bitstream/10495/16766/1/CastrillonManuela_2019_SistemaReconocimientoResiduos.pdf.
- Castrillón Medina, M. (2019b). Sistema de reconocimiento de desechos reciclables. Recuperado de <https://github.com/ManuCastriillonM/redo>. (Actualmente no disponible).
- Citlalic Gonzales Martinez, A. (2001) Costos y beneficios ambientales del reciclaje en México. Una aproximación monetaria. *Gaceta Ecológica*, (58), 17-26.
- Howard, A. G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M., Hartwig, A. (2018). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision. Recuperado de <https://arxiv.org/abs/1704.04861>.

- Moyer, E.: Día del Reciclaje: ¿Qué tanto se recicla en América Latina? (2018). Recuperado de <https://www.nrdc.org/es/bio/erika-moyer/dia-reciclaje-tanto-recicla-america-latina>.
- Romo-Gómez, C.; Camacho-López, C.; Marmolejo-Santillán, Y.; Ota-zo-Sánchez, E. M. (2019). Actividades antrópicas: deterioro de la capa de ozono estratosférico. *Pädi. Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, 7(13), 1-5.
- Schröer, C., Kruse, F., Marx Gómez, J.(2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526-534.
- Thokrairak, S.; Thibuy, K.; Jitngernmadan, P. (2020). Valuable Waste Classification Modeling based on SSD-MobileNet, 2020 - 5th International Conference on Information Technology (InCIT), Chonburi, Thailand, 228-232, doi: 10.1109/InCIT50588.2020.9310928.
- Thung, G. (2024). Dataset of images of trash; Torch-based CNN for garbage image classification. Recuperado de <https://github.com/garythung/trashnet>.

Capítulo 3

Herramienta para la asistencia
pre diagnóstico y seguimiento
del tratamiento de la escoliosis
idiopática del adolescente (EIA)



**Diego Alberto Godoy, Angeles Belén Piotroski,
Damián Horacio Stetson, Carlos Kornuta,
Lucas Kucuk, Enrique Marcelo Albornoz**

INTRODUCCIÓN

La escoliosis es una curvatura lateral de la columna vertebral que ocurre con mayor frecuencia durante el período de crecimiento anterior a la pubertad. Aproximadamente el 3% de los adolescentes tienen escoliosis, y si bien la mayoría de los casos de escoliosis son leves, algunas deformidades de la columna vertebral continúan empeorando a medida que los niños crecen. La escoliosis grave puede ser incapacitante. Además, una curvatura especialmente grave en la columna vertebral reduce el espacio dentro del pecho, lo cual dificulta el funcionamiento correcto de los pulmones. La escoliosis idiopática adolescente (AIS) es una de-

formidad tridimensional de la columna vertebral. Si bien el AIS puede progresar durante el crecimiento y causar una deformidad de la superficie, generalmente no es sintomático (Yi et al, 2020).

La escoliosis es una enfermedad idiopática que padecen los niños y adolescentes. Normalmente, los médicos usan imágenes de rayos X para medir el ángulo de Cobb y luego evaluar la gravedad de la enfermedad. Se debe tener en cuenta que, durante la medición de la radiografía, se estima que se puede tener un error de medición variable de +/- 5 grados. Hoy en día con el uso de los sistemas de medición digital el error de medición disminuye, pero se debe tener presente ya que el error intra e inter observador varía entre 3 y 5 grados (Cobb, 1948).

En la actualidad existen pruebas de que el uso del corsé es eficaz para evitar la progresión de las formas más graves de escoliosis y que disminuye la necesidad de cirugía. Sin embargo, las pruebas de cribado que se disponen y se proponen en la literatura tienen valores predictivos positivos (VPP) muy bajos y su aplicación universal supondría un alto número de derivaciones y pruebas innecesarias (Esparza Olcina, 2015).

La disponibilidad de herramientas y sistemas que provean indicadores para la detección oportuna de EIA, podría evitar algunas de las consecuencias antes mencionadas. A tal efecto, el desarrollo de un algoritmo que utilice redes neuronales profundas para la detección y reconocimiento de escoliosis utilizando

imágenes de la espalda descubierta de pacientes, sería una herramienta de soporte a los profesionales de la salud en la detección de escoliosis y la clasificación de gravedad.

DEFINICIÓN DE ESCOLIOSIS

Es una curva con desplazamiento lateral de la columna vertebral, que le da a la columna un aspecto de “S” o de “C” en vez de una “I” recta.



Fig. 1. Columna con aspecto de “S” (El Aakel El Ouelkade, 2014)

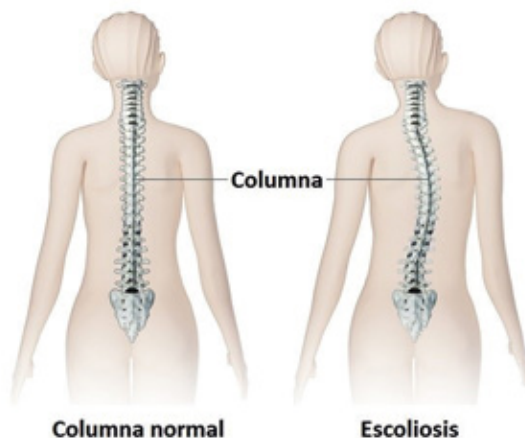


Fig. 2. Columna con aspecto de “C” (Ramos-Álvarez et al, 2023)

La escoliosis puede provocar que los huesos de la columna giren (roten), y así un hombro, omóplato (escápula) o cadera queden más arriba de un lado que del otro.

En etapas tempranas, la enfermedad no causa dolor, disfunción neurológica ni problemas respiratorios. Pero la preocupación por el aspecto estético de la espalda es lo que a menudo se detecta como inquietud primaria del paciente y de los padres.

Existe una clasificación para los diferentes tipos de escoliosis, que puede presentarse a cualquier edad. La *escoliosis infantil* que se manifiesta en niños menores de 3 años y puede ser el resultado de vértebras con forma anormal de nacimiento (congénita), diversos síndromes, trastornos neurológicos o causas desconocidas (idiopática).

La *escoliosis juvenil* ocurre en niños que tienen entre 3 y 10 años de edad, pero no es común. Finalmente, la *escoliosis idiopática adolescente* que ocurre después de los 10 años y es la manifestación más recurrente de la enfermedad.



Fig. 3. Identificación del ángulo de Cobb en una radiografía

Existe otro tipo de clasificación que se refiere al ángulo de desvío de la columna, como se muestra en la Tabla 1.

Tipo de escoliosis	Ángulo de Cobb	Tratamiento Orientativo
Leve	10°-25°	Ejercicios de Fisioterapia
Moderado	25°-45°	Fisioterapia + Tratamiento Ortopédico (Corsé)
Grave	>45°	Posibilidad de tratamiento quirúrgico

Tabla 1. Clasificación de la escoliosis en función del Ángulo de Cobb

INTELIGENCIA ARTIFICIAL

El término inteligencia artificial es el más general y engloba a los campos de *Machine Learning* y *Deep Learning*, junto con otras técnicas como los algoritmos de búsqueda, el razonamiento simbólico, el razonamiento lógico y la estadística. Nació en los años 1950, cuando un grupo de pioneros de la computación comenzaron a preguntarse si se podía hacer que las computadoras pensaran. Una definición concisa de la inteligencia artificial sería: el esfuerzo para automatizar las tareas intelectuales que normalmente realizan los seres humanos. Específicamente, el *Machine Learning* –también llamado aprendizaje automático– se refiere a un amplio conjunto de técnicas informáticas que nos permiten dar a las computadoras la capacidad de aprender sin ser explícitamente programadas. Por su parte el *Deep Learning* –o aprendizaje profundo– es un subcampo dentro del *Machine Learning*, que utiliza distintas estructuras de redes neuronales para lograr el aprendizaje de sucesivas capas de representaciones cada vez más significativas de los datos. El profundo en *Deep Learning* hace referencia a la cantidad de capas de representaciones que se utilizan en el modelo; en general se suelen utilizar decenas o incluso cientos de capas de representación, que aprenden automáticamente a medida que el modelo es entrenado con los datos (López Briega, 2017).

REDES NEURONALES

Una red neuronal es un modelo simplificado que emula el modo en que el cerebro humano procesa la información, las unidades de procesamiento se organizan en capas. Normalmente hay tres partes en una red neuronal: una capa de entrada, con unidades que representan los campos de entrada; una o varias capas ocultas; y una capa de salida, con una unidad o unidades que representan el campo o los campos de destino. Las unidades se conectan con fuerzas de conexión variables (o ponderaciones). Los datos de entrada se presentan en la primera capa y los valores se propagan desde cada neurona hasta cada neurona de la capa siguiente, al final se envía un resultado desde la capa de salida (IBM, 2021).

Las redes neuronales convolucionales –CNN por sus siglas en inglés– son un tipo de redes neuronales artificiales donde las “neuronas” corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria (V1) de un cerebro biológico. Este tipo de red es una variación de un perceptrón multicapa pero, debido a que su aplicación es realizada en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones. Las CNN están formadas por múltiples capas de filtros convolucionales de una o más dimensiones.

Después de cada capa, por lo general se añade una función para realizar un mapeo causal no-lineal (Matich, 2001).

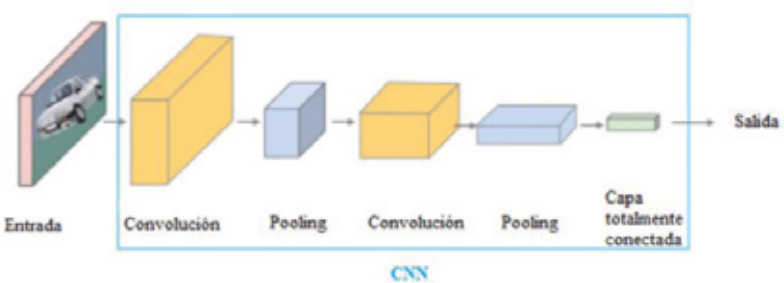


Fig. 4. Funcionamiento de la CNN (Bagnato, 2020)

Existen una serie de trabajos relacionados con la propuesta de este artículo. El denominado *Vertebra-Focused Landmark Detection for Scoliosis Assessment* es una propuesta en la que los ángulos de Cobb se miden de acuerdo con los puntos de referencia de las vértebras, proponiendo un método novedoso de detección de puntos de referencia centrado en las vértebras. El modelo localiza primero los centros de las vértebras, sobre la base de los cuales luego traza los puntos de referencia de las cuatro esquinas de la vértebra a través del desplazamiento de la esquina aprendida. De esta forma, el método es capaz de mantener el orden de los puntos de referencia. Los resultados de la comparación demuestran los méritos del método tanto en la medición del ángulo de Cobb

como en la detección de puntos de referencia en imágenes de rayos X ambiguas y de bajo contraste (Yi et al, 2020).

Otro ejemplo es *Development and validation of deep learning algorithms for scoliosis screening using back images*, que desarrolla y valida algoritmos de aprendizaje profundo para la detección automática de escoliosis utilizando imágenes de la espalda descubierta (Yang, 2019). Este proyecto tiene una mayor relación porque es de utilidad para obtener el estado del arte de las redes neuronales.

Otra propuesta que analizamos es *Reconocimiento de imágenes para detección temprana de alteraciones posturales*, una investigación que tiene como objetivo detectar de forma temprana desviaciones de columna aplicando un software libre a través del que se realizó un algoritmo para la detección del ángulo de desviación de la columna vertebral, ya sea en pacientes que presenten una desviación pronunciada como en pacientes que tengan un bajo grado de desviación, permitiendo que el médico pueda determinar el tipo de tratamiento que es necesario para evitar complicaciones futuras (Romero Rodríguez, 2019).

OBJETIVOS

El objetivo general de este trabajo es proponer una herramienta de software para la asistencia pre diagnóstico y seguimiento del tratamiento de la escoliosis idiopática del adolescente (EIA), a partir del procesamiento de imágenes mediante redes neurona-

les profundas. Como metas más específicas se pueden enumerar el definir el estado del arte de las redes neuronales profundas y su uso en la detección de posturas corporales, determinar las características de la escena para la recolección de muestras fotográficas, determinar los requerimientos de software para el análisis y detección de la escoliosis idiopática del adolescente a través del procesamiento de imágenes con CNN, desarrollar una herramienta de software que cumplan con los requerimientos especificados y validar la herramienta a partir de los resultados obtenidos en las pruebas, con la valoración de un profesional de la salud.

LOS MÉTODOS Y LOS MATERIALES

Para el desarrollo del trabajo se recurrió a fuentes de información como ser en trabajos académicos y sitios web especializados en el tema de estudio. El trabajo se dividirá en etapas: en primer lugar, se analizaron en profundidad las redes neuronales profundas y se explicarán los conceptos relacionados al diagnóstico y seguimiento de tratamientos de la escoliosis. En la segunda etapa se realizó la recolección de muestras con la ayuda de kinesiólogos y médicos especialistas, de las cuales se necesitan alrededor de 100 imágenes de personas de espalda. En la tercera etapa se realizó el diseño del modelo general del sistema siguiendo el modelado orientado a objetos, basándose en los estándares provistos por el Lenguaje Unificado de Modelado (UML). La cuarta

etapa se realizó la validación, la misma se llevará a cabo con un porcentaje de las muestras recolectadas y etiquetadas, separadas para este fin. Y en la quinta etapa, se probó el desempeño de la herramienta, comparando los resultados, que constan de porcentajes de probabilidad de acierto, con las etiquetas de las muestras, para luego ser analizados.

Las herramientas utilizadas en el proyecto fueron:

- Python como lenguaje de desarrollo.
- Visual studio code como IDE.

Se evaluarán diferentes alternativas basadas en redes neuronales profundas para la segmentación y detección de esqueletos y posturas de las personas, como son:

- TensorFlow-Pose.
- Open Pose based neural network.
- MediaPipe Pose.
- Librería Pandas, Seaborn, PyTorch, Tkinter.

DISEÑO Y DESARROLLO DEL SISTEMA

En esta sección se presentan todas las cuestiones referidas a la implementación de la herramienta desarrollada con el propósito de explicar cómo es la arquitectura de la herramienta y cuáles son sus partes fundamentales. Con el diagrama de secuencia se expone el funcionamiento del sistema, donde a través de los mensajes, se explica cómo se realiza todo el circuito de la

herramienta propuesta. En un diagrama de flujo se presenta la secuencia de los pasos para obtener el prediagnóstico. Al final de esta sección se explican las partes más relevantes del código desarrollado, desde cómo se realiza el cálculo de cada ángulo hasta obtener un prediagnóstico final. Además, se explica cómo mediante la utilización de MediaPipe Pose se reconocen los puntos de referencia fundamentales para comenzar con el proceso.

La arquitectura principal del sistema se divide en dos partes diferenciadas por plataforma y funcionalidad. Una parte móvil multiplataforma que es la interfaz gráfica para que el kinesiólogo o especialista pueda utilizar el sistema denominado *Scoliosis App*. Esto brinda una aplicación móvil que permite tomar la fotografía de la espalda del paciente, y enviarla al *BackEnd* para ser procesada.

La parte *BackEnd* recibe la imagen, la procesa y analiza. Se trata de una API contenida en un contenedor de Docker. Se decidió implementarlo de esta forma ya que hacer uso de los contenedores aumenta la portabilidad de la aplicación. Estas dos partes se comunican a través del lenguaje JSON, que se trata de un formato estándar de texto sencillo para el intercambio de datos, que en este caso son la imagen enviada para ser procesada y el PDF retornado con el prediagnóstico calculado.

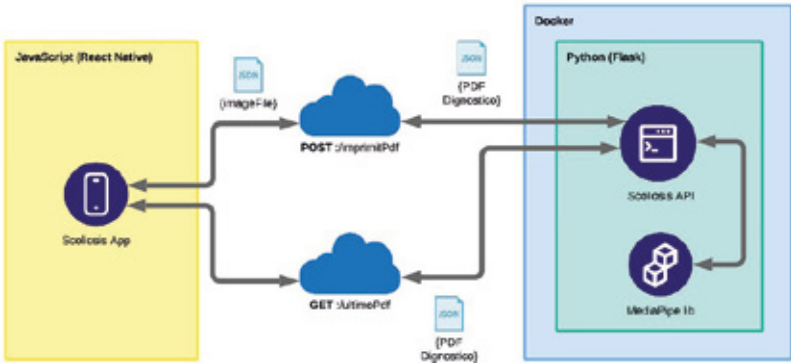


Fig. 5. Arquitectura del sistema

DIAGRAMA DE SECUENCIA

El médico o kinesiólogo comienza la secuencia como se puede notar en la figura 5, enviando una fotografía de espaldas de su paciente, la aplicación toma la fotografía y la envía al *api* Rest. Allí, el sistema realiza el análisis de la imagen. La imagen es procesada para obtener los puntos necesarios de referencias, las cuales son posiciones de orejas, hombros, codos y cintura. Una vez obtenidos los ángulos y centros de cada posición de referencia, se realiza el análisis de la postura para determinar si está o no equilibrada y se identifica si existe el triángulo de la talla.

Luego de analizar los ángulos obtenidos, se genera el diagnóstico. A la imagen recibida se le agregan los ángulos y rectas calculadas para que se vean de manera gráfica y se pueda com-

prender mejor. Finalmente se integran las imágenes y el prediagnóstico obtenido en un PDF para ser retornado a la aplicación y que el usuario, ya sea kinesiólogo o médico pueda visualizarlo e incluso imprimirlo.

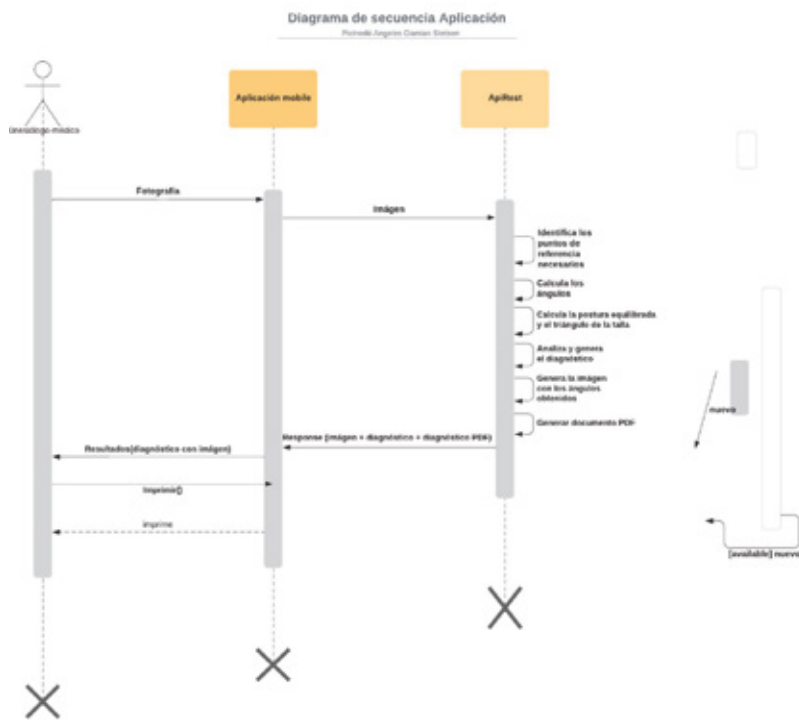


Fig. 6. Diagrama de secuencia del sistema

DIAGRAMA DE FLUJO PARA DETERMINAR EL PREDIAGNÓSTICO

Una vez calculados los ángulos se realiza un análisis para dictaminar el diagnóstico del paciente, evaluando todos los ángulos de los señaladores anatómicos definidos (hombros, codos y cintura) por separado. Como se muestra en el diagrama de flujos de la figura 6, el algoritmo comienza evaluando si el ángulo calculado es mayor a 0.5, ya que si es menor no posee una variación anormal en dicho indicador anatómico.

El valor 0.5 es un umbral predefinido que evalúa la mínima desviación en la columna que poseen comúnmente todas las personas y se encuentra dentro de los parámetros normales. Esto se definió junto a un kinesiólogo, quién comentó que se trata de la preexistencia de una escoliosis fisiológica que es inherente a la propia persona, y puede ser producida por el simple hecho de ser zurdo o diestro. Esto no necesariamente es una anomalía y es por ello que, si el valor del ángulo supera dicho mínimo umbral (0.5 grados) recién se considera que la desviación es importante y se encuentra fuera de lo normal.

Volviendo sobre el diagrama, luego se evalúa si es mayor a 1.5, ya que, si no lo es, el ángulo estimado se encuentra en el rango (0.5, 1.5) y el diagnóstico es “leve”. Sin embargo, si es mayor a 1.5 se re evalúa para determinar si es “moderado” o “grave”, para

esto se utiliza un umbral de 4 grados. En resumen, el algoritmo definido con el profesional presenta 4 umbrales para determinar desviaciones normales, leves, moderadas y graves.

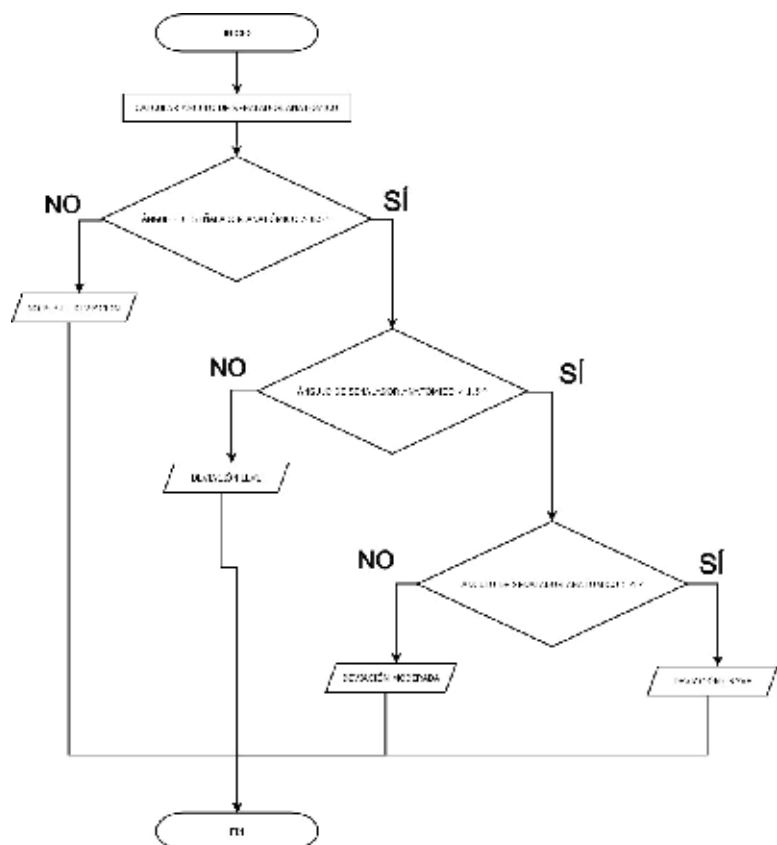


Fig. 7. Diagrama de flujo que permite estimar la condición de la columna

La evaluación presentada previamente se realiza para cada uno de los ángulos obtenidos de los señaladores anatómicos, es decir, el ángulo de los hombros, codos y cintura. Una vez obtenido el resultado de cada uno de los ángulos, se procede a determinar cuál será el prediagnóstico final del paciente, para ello se tiene en cuenta el señalador anatómico de mayor gravedad determinada por el algoritmo.

Finalmente, el diagnóstico se basa en este indicador anatómico de mayor gravedad, ya que es la forma en que el kinesiólogo lo hace. Respecto a esto, el especialista comentó que no se puede realizar un cálculo promediable de dicho resultado, ya que la gravedad del padecimiento si bien podría ser “negociable” en términos matemáticos, en términos físicos la gravedad está presente y complejiza enormemente la salud del paciente por más que existan otros indicadores de menor gravedad. Esta es una primera aproximación, pero está claro que la ecuación puede ser adaptada, modificada y evolucionada en trabajos posteriores.

MÉTODOS DE DETECCIÓN Y FUNCIONAMIENTO DEL SISTEMA

Para realizar la detección de los puntos del cuerpo el sistema necesita recibir una fotografía de una persona de espaldas, a priori puede ser con la espalda descubierta o con una remera, se recomienda que se visualicen los codos para una correcta de-

tección de los puntos. La imagen enviada debe tener extensión JPG, PNG o JPEG y en ella se debe visualizar completamente a la persona desde la cabeza hasta la cintura, y en amplitud de codo a codo, esto es importante para lograr reconocer los puntos necesarios y realizar el pre diagnóstico. Para esta etapa de desarrollo se pretende que la imagen posea un fondo claro y homogéneo donde se encuentre una única persona para simplificar las evaluaciones del sistema.

La imagen se convierte en un array de bytes RGB (necesario para que la red pueda manejar los píxeles de la imagen de entrada y generar los mapas de calor, detectar el fondo del cuerpo reconocible del paciente y generar los puntos claves. Para esto, se utiliza la librería Numpy de Python y luego la librería MediaPipe Pose la cual retorna como resultado una lista de puntos de referencia.

Si se usa una imagen de cuerpo completo, se obtienen 33 puntos de referencia. En este trabajo solo tomaremos aquellos puntos necesarios para realizar los cálculos que definimos previamente, son 8 puntos que determinan la posición de ambos codos, los hombros, las orejas y la cintura.

A partir de las dimensiones de la imagen, estos puntos son mapeados para que coincidan con la imagen recibida y así puedan ser mostrados como en la figura 8.

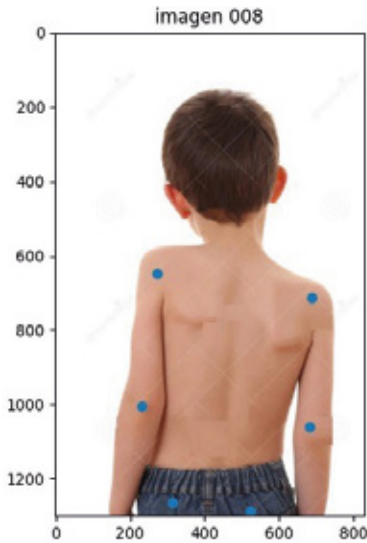


Fig. 8. Señaladores anatómicos identificados
(fuente: <https://es.dreamstime.com/>)

Una vez que se obtienen los 8 puntos normalizados se realizan los cálculos, a partir del teorema de Pitágoras. Primero, se calcula la distancia entre cada uno de los puntos:

Oreja derecha y oreja izquierda.

Hombro derecho y hombro izquierdo.

Codo derecho y codo izquierdo.

Cintura derecha y cintura izquierda.

Esta distancia, entre el punto A y el punto B, se corresponde a la hipotenusa de un triángulo rectángulo y a partir de ésta se computan los otros datos como se muestra en la figura 8.

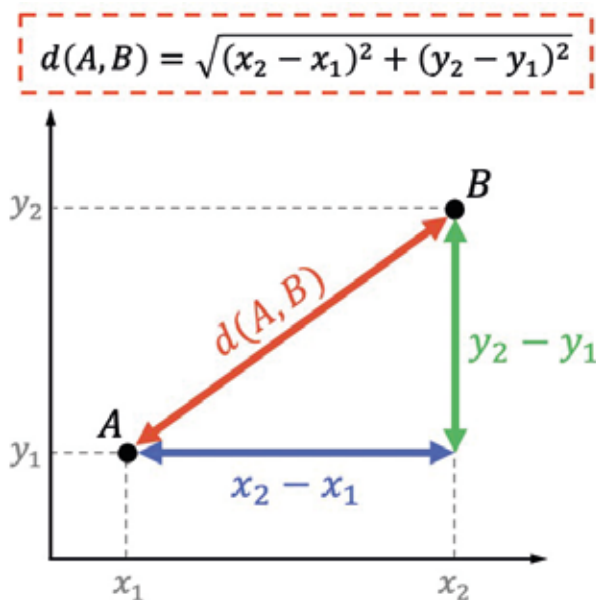


Fig. 9. Cálculo de distancia entre dos puntos

Entonces, se calcula la altura entre los puntos (h) y se obtienen los ángulos necesarios para el diagnóstico. En la figura 9 se pueden observar los resultados de este proceso sobreimpresos en la imagen original.

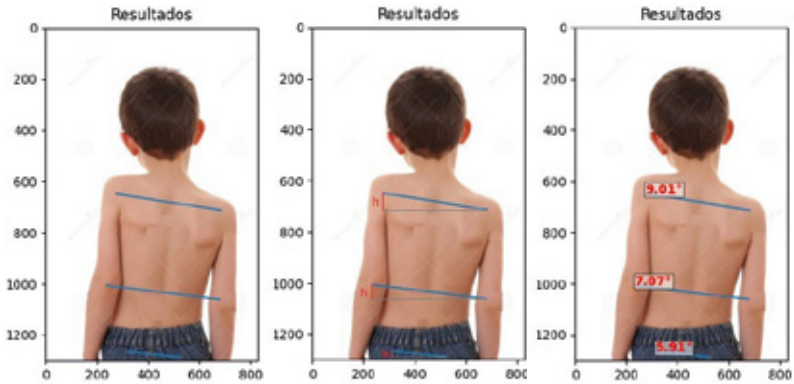


Fig. 10. Resultados visuales de los cálculos de los puntos de interés y ángulos

Con estos datos, el paso siguiente es simular el “método de la plomada” que utilizan los kinesiólogos para identificar si la postura del paciente está equilibrada o desequilibrada. Esta “plomada” es una línea central que se toma de referencia y es muy útil ya que el paciente puede tener una desviación que comience desde las vértebras cervicales y de esta manera sería fácil detectarla.

Para el sistema, la línea central comienza en el punto medio entre las dos orejas del paciente y finaliza en la cintura del mismo como se muestra en la figura 10. Se puede ver claramente donde está el centro de cada parte de interés (hombro, codos y cintura) a partir de la línea central que las atraviesa. Finalmente, esta imagen (junto con el análisis de los datos) es enviada a la aplicación móvil donde el usuario puede conocer su pre diagnóstico.

tico de manera inmediata y si lo desea puede descargar o imprimir un archivo PDF con el resultado.

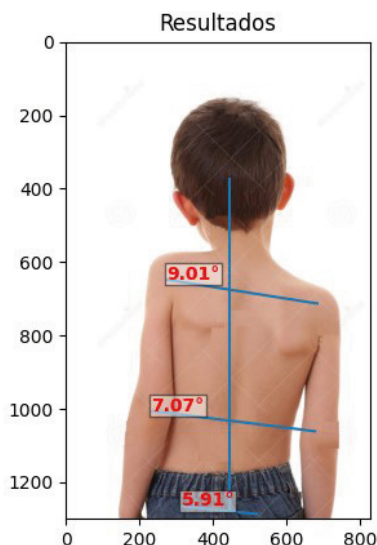


Fig. 11. Salida gráfica con datos y una línea de nivel que replica el método de la plomada

Es importante destacar que esta medición fue incorporada durante el desarrollo del sistema, y que inicialmente no se consideró que fuese necesaria una imagen donde las orejas sean visibles. Luego de las interacciones con los especialistas, se agregaron los puntos de las orejas ya que en el centro de estas se encuentra la horquilla occipital que es donde los kinesiólogos ponen la plomada para medir si está desequilibrada la postura.

Otro resultado importante, que fue incorporado durante el desarrollo del sistema, sale del análisis de la simetría de la persona y se brinda al usuario de forma gráfica. Para este análisis se compara la distancia entre los codos respecto al segmento de recta central y se determina de qué lado se forma el triángulo, siempre que las distancias codo-centro no sean iguales figura 11.

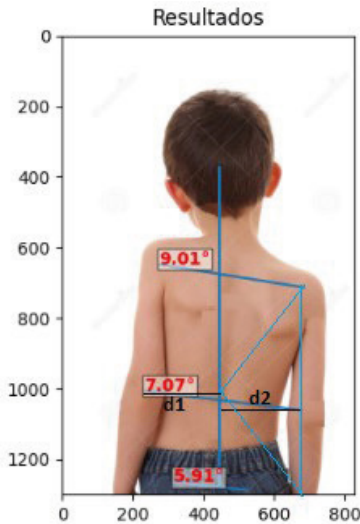


Fig. 12. Imagen que indica el triángulo de la talla

PRUEBAS Y RESULTADOS

En este punto, se encuentran dadas las condiciones para poner a prueba la funcionalidad de la aplicación desarrollada con el fin de analizar, comparar y validar los resultados obtenidos.

Para ello, se realizaron 30 pruebas con distintas imágenes y se compararon los resultados obtenidos con la respectiva validación del especialista kinesiólogo.

Antes de comenzar, cabe aclarar que en un primer momento se pretendía que las imágenes procesadas sean de pacientes reales que padecen de escoliosis idiopática adolescente. Es por ello que, se procedió a contactar con diferentes kinesiólogos de la ciudad de Posadas y de Buenos Aires a través de las redes, con el propósito de solicitarles la siguiente información de sus pacientes:

- Imágenes de la espalda de su paciente con escoliosis.
- Radiografía de la espalda de dicho paciente.
- Información sobre el paciente, como a la fecha que inició el tratamiento, si se realizó o se realizará una cirugía para tratar la escoliosis y la edad del paciente.
- Información referida al especialista: si el posee un sistema para realizar el seguimiento de su paciente.

De todos los kinesiólogos y fisioterapeutas contactados, sólo se obtuvo alguna información de sus experiencias y de cómo realizaban su labor de manera verbal. Sin embargo, no nos podían suministrar la información solicitada de los pacientes debido a muchos factores: no tenían registrados estos datos, no trabajaron con pacientes con escoliosis, trataron pacientes, pero ya le otorga-

ron el alta y no tenían sus datos, no registraban las imágenes de la forma que necesitábamos para este trabajo, entre otras. Con esta experiencia se evidenciaron las dificultades de los trabajos interdisciplinarios y de las diferentes lógicas/metodologías de trabajo, además se comprobó la problemática presentada en este trabajo sobre la falta de una herramienta útil para el registro, evaluación y seguimiento de los pacientes, así como también la dificultad de tener datos de los pacientes.

Con base en esto, se decidió extraer imágenes de prueba de internet y validarlas con un kinesiólogo, obteniendo así su visto bueno sobre el uso de esas imágenes. Además, con el fin de simular imágenes de pacientes reales se procedió a generar un set de imágenes propio. Para esto, nos tomamos distintas fotografías de espalda simulando las posturas que son congruentes con las de un paciente con escoliosis, este proceso fue guiado por un kinesiólogo que nos asistía mientras realizamos las posturas.

EXPERIMENTOS

En esta sección se presentan los resultados de analizar las imágenes que se recopilaron. El sistema reconoce la imagen, obtiene los puntos y realiza los cálculos para obtener los ángulos y el diagnóstico. Inicialmente, se consideraron algunas restricciones o características que debían poseer las imágenes como ser la ubicación de la cámara que se explican al final de la sección.

Como ya se comentó las imágenes fueron obtenidas de internet y validadas con un especialista. Es importante mencionar que éstas no son imágenes que presentan alta definición, buena exposición a la luz, correcto ángulo de toma de la fotografía, fondos homogéneos en la imagen, ni resoluciones uniformes, entre otros. Se evaluaron 219 imágenes en total de las cuales 211 corresponden a fotografías de pacientes con distintos grados de escoliosis obtenidas de las siguientes fuentes: <https://openi.nlm.nih.gov>, <https://www.hindawi.com>, <https://www.elsevier.es>, <https://www.shutterstock.com>, <https://www.barcelonaquiropactic.es>, <https://peakhealthcalgary.com>, y <https://www.chroniclive.co.uk>. Luego, se incorporaron ocho fotografías propias tomadas con el fin de evaluar distintos factores que se deben de tener en cuenta en la fotografía y que se describen en la sección de discusiones.

Si bien el presente trabajo está enfocado en la definición y desarrollo de una herramienta útil para el prediagnóstico, asistencia y seguimiento de desviaciones anormales de la columna vertebral, se incluye el análisis y la discusión de las limitaciones y condiciones adecuadas de funcionamiento. A continuación, se presentan algunos ejemplos de los resultados del sistema, se agruparon según el grado de desviación detectada: *leve*, *moderado* y *grave*.

RESULTADOS DE DESVIACIÓN LEVE

Estos casos se dan cuando se detecta cuando el mayor de los ángulos detectados está entre 0.50 y 1.50. En la figura 39 se presentan dos casos de desviaciones “leves”, en el caso de la imagen de la izquierda el ángulo de mayor valor es el superior que vale 1.1° , mientras que en la de la derecha el ángulo de mayor valor es el central y su valor es 1.28° .

El resultado detallado del primer caso sería: ángulo superior = 1.1° , ángulo central = 0.14° y el inferior = 0.38° ; y las etiquetas serían: leve, no posee, no posee. En el segundo caso los resultados serían [0.45° , 1.28° , 0.03°] y [no posee, leve, no posee]. También pueden observarse las radiografías respectivas a cada caso, que indican la existencia de una curva leve que concuerda con el resultado de nuestra aplicación.



Fig. 13. Ejemplos de desviaciones leves

RESULTADOS DE DESVIACIÓN MODERADA

En el caso de una desviación moderada de la columna vertebral, el ángulo de mayor valor detectado suele encontrarse en un rango de 1.51 a 4 grados. Esto indica que hay una curvatura anormal en la columna vertebral, pero que no es lo suficientemente grave y puede ser visualizado claramente tanto en fotografías de espalda como en radiografías, ya que la curvatura de la columna vertebral se ve pronunciada en estos casos.

Además, podemos distinguir una inclinación en la postura del paciente (figura 14a y 14c), y se puede notar el triángulo de la talla que indica donde se encuentra la curva. Por otra parte, en el ejemplo de la figura 13b, este no es tan notorio, aunque existe un pliegue del lado derecho que es un signo claro de que existe una anomalía y el especialista debería de tenerlo en cuenta ya que el sistema actual no lo tiene en cuenta.



Fig. 14. Ejemplos de desviaciones moderadas

RESULTADOS DE DESVIACIÓN GRAVE

Se produce una desviación grave cuando el ángulo de mayor valor detectado supera los 4 grados, lo que muestra una curvatura anormal de gran consideración que nuestro sistema lo denomina grave. Esta curvatura es más pronunciada que la moderada, e incluso puede tener rotación de las vértebras, las desviaciones graves son aquellas que pueden necesitar un tratamiento quirúrgico dependiendo del caso.

En la figura 14a y b se pueden observar en las radiografías que las curvaturas son pronunciadas y genera que la persona no se encuentre equilibrada. Lo mismo ocurre con el ejemplo de la figura 15d, un caso proporcionado por el especialista en RPG Ricardo Rodríguez, donde la paciente posee una escoliosis lumbar de 37°. En este nivel de escoliosis habitualmente se necesita una cirugía, sin embargo, esta paciente logró disminuir su curvatura y evitar la operación realizando tratamiento de RPG.



Fig. 15. Ejemplos de desviaciones graves

DISCUSIÓN

Después de llevar a cabo un minucioso análisis de los resultados obtenidos a través de las pruebas realizadas anteriormente, se han identificado una serie de factores críticos para garantizar el correcto funcionamiento del sistema. Para ello, se han definido una serie de requisitos que deben ser tomados en cuenta en el momento de capturar las fotografías, estos incluyen tanto aspectos técnicos como físicos que son esenciales para garantizar la precisión y confiabilidad del sistema en su conjunto.

En primer lugar, es fundamental cuidar la perspectiva de la fotografía, situando al paciente de espaldas y procurando tomar la imagen lo más centrada posible. Esto podría guiarse agregando visualmente en la aplicación alguna referencia, como ser una silueta aproximada de un cuerpo humano para encuadrar la captura de la imagen. Además, es importante prestar atención a la vestimenta del paciente, asegurando que vista una remera ajustada o que directamente no lleve ninguna prenda superior.

A pesar de estas limitaciones, los resultados obtenidos son de gran importancia y brindan una valiosa información para el diagnóstico de desviaciones anormales en las imágenes médicas. Por lo tanto, es fundamental seguir trabajando en la mejora de estas técnicas y herramientas para lograr una mayor precisión y fiabilidad en los resultados.

Para validar la precisión del sistema, se utilizaron 213 imágenes de prueba, que fueron evaluadas por kinesiólogos, El sistema clasifica las desviaciones en base a tres etiquetas de severidad, las cuales son: *leve*, *moderada* y *grave*.

El kinesiólogo etiquetó cada una de las desviaciones en las imágenes para luego compararlas con la respuesta del sistema. Se contabilizaron las predicciones realizadas por los kinesiólogos y las del sistema para cada imagen. Los resultados se presentan en una tabla de valores que muestra la predicción manual realizada por los kinesiólogos y la predicción generada por el sistema para cada imagen (ver tabla 2).

Severidad	Software	Kinesiólogo
Leve	26	57
Moderada	121	93
Grave	64	61

Tabla 2. Comparativa de predicciones del kinesiólogo y el software

La misma información se presenta de forma gráfica en la figura 16, donde se puede apreciar que el sistema de clasificación desarrollado tiene una tendencia a confundir muchos de los casos “Leves” con “Moderados”. Sin embargo, es importante destacar que se aproxima bastante bien en los casos graves, lo cual sería muy importante.

Comparativa global de predicciones por severidad

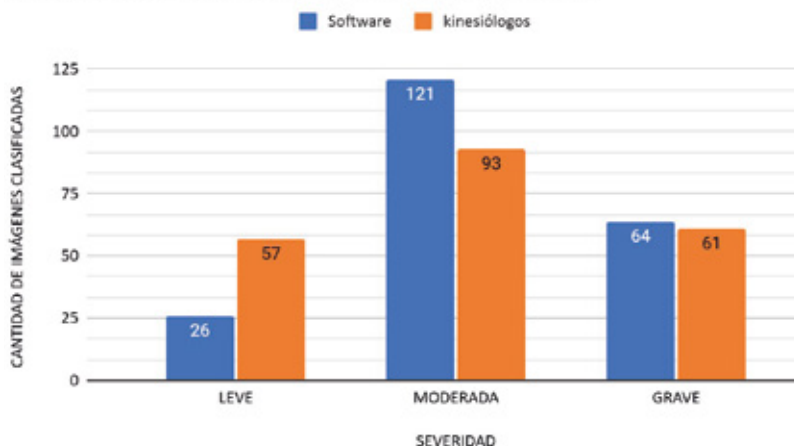


Fig. 16. Comparación global de las predicciones del sistema

Luego de esta primera apreciación de los resultados generales, se calculó la precisión efectiva del sistema, a partir de la obtención de los aciertos totales tanto del software como del kinesiólogo, que se visualizan en la tabla 3, con el fin de identificar la proporción de respuestas que fueron clasificadas correctamente, para cada una de las severidades, para hallar la precisión de cada severidad se utilizó la siguiente fórmula:

Precisión(i) = número de predicciones correctas(i) / número total(i)

Precisión (Leve) = 19 / 57 = 0.33 o 33%

Precisión (Moderada) = 67 / 93 = 0.72 o 72%

Precisión (Grave) = 39 / 61 = 0.64 o 64%

Severidad	Total Aciertos sistema	Total Respuestas Kinesiólogo
Leve	19	57
Moderado	67	93
Grave	39	61
Total	125	211

Tabla 3. Comparativa de aciertos del sistema

Los resultados muestran que el sistema tiene una precisión aceptable en la clasificación de imágenes con lesiones moderadas y graves. Demuestran que el sistema es capaz de responder de manera efectiva a este tipo de imágenes, lo que puede tener aplicaciones prácticas en la evaluación y diagnóstico de desviaciones en la columna vertebral. Es importante destacar que, si bien estos resultados son alentadores, aún es necesario realizar más pruebas y evaluaciones para determinar la eficacia del sistema en diferentes situaciones y casos de lesiones musculoesqueléticas de distinta gravedad. Además, es posible que la función de estimación de gravedad necesite ser ajustada para alcanzar mejores resultados, inclusive agregando más variables.

CONCLUSIONES

Se analizaron trabajos existentes relacionados a la investigación de métodos automatizados para el proceso de detección o diagnóstico. Se identificaron avances en el uso de fotografías, las redes neuronales artificiales profundas y la implementación de gráficos o líneas que representaban curvaturas o resultados aparentes del análisis a los pacientes. A partir de esto se propuso un método novedoso para aproximar diagnósticos en base a cálculos realizados sobre señaladores anatómicos obtenidos con una red neuronal profunda.

Se consultaron muchos especialistas kinesiólogos, pero no fue posible recopilar una base de datos provista por ellos, por lo que fue necesario el desarrollo de una base de datos de imágenes de internet con más de 200 imágenes que se validaron con los profesionales que nos asistieron en este trabajo.

Durante el desarrollo del trabajo fue posible estudiar y aprender temas y conceptos relacionados a la medicina y la kinesiología. Gracias a las interacciones con el especialista también fue posible conocer situaciones inherentes a la práctica cotidiana de su profesión, tanto de lo metodológico como de los pacientes. Se identificaron situaciones no deseadas por pacientes referidas a realizarse muchos análisis o el exceso de radiografías, y desde la metodología, se entendieron los lineamientos que pueden ser

abordados para identificar otros padecimientos en futuras evoluciones de nuestro prototipo.

Para dicho prototipo se realizó el diseño de una arquitectura que funciona de manera cliente-servidor, a partir de lo cual es posible generar clientes en cualquier dispositivo o sistema operativo móvil, los cuales se conectan a un servidor hospedado en la red disponible para consulta y evaluación de imágenes recibidas desde el cliente.

Los resultados de diagnóstico son prometedores ya que se comprobó que se aproximan a las evaluaciones que realizan los especialistas sobre sus pacientes. Incluso podrían justificar la necesidad de radiografías, dar soporte a los diagnósticos del profesional.

BIBLIOGRAFÍA CITADA

Bagnato, J. I. (2020). *Convolutional Neural Networks: La Teoría explicada en Español*.

Cobb, J. et al (1948). *Outline for the Study of Scoliosis*. Recuperado de <https://www.scienceopen.com/book?-1.ILinkListener-header-action~bar-public~button~container-bookmark~button-bookmark~button&vid=76a12f1e-c7ef-4cc2-8aec-0904b520cd98>.

El Aakel El Ouelkade, H. (2014). *Escoliosis idiopática*. Recuperada de <https://diposit.ub.edu/dspace/bitstream/2445/60463/1/60463.pdf>.

Esparza Olcina, M. J. (2015). ¿Hay que hacer cribado de la escoliosis idiopática del adolescente? *Formación Activa en Pediatría de Aten-*

- ción Primaria, 8(3), 157-160. Recuperado de https://fapap.es/files/639-1278-RUTA/10_FAPAP_3_2015_Cribado_escoliosis.pdf.
- IBM. (2021). *El modelo de redes neuronales*.
- López Briega, R. E. (2017). Introducción al Deep Learning. Recuperado de <https://relopezbriega.github.io/blog/2017/06/13/introduccion-al-deep-learning/>.
- Matich, D. J. (2001). *Redes Neuronales: conceptos básicos y aplicaciones*. Recuperado de https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monograias/matich-redesneuronales.pdf.
- Ramos-Álvarez, J. J.; Segovia-Martínez, J. C.; López-Silvarrey Varela, F. J. (2023). *La escoliosis. Manual de ejercicios correctivos en alteraciones estructurales y posturales de la columna*. Recuperado de https://www.ucjc.edu/pdfs/universidad/libro_escoliosis.pdf.
- Romero Rodríguez, C. V. (2019). *Reconocimiento de imágenes para detección temprana de alteraciones posturales*. Recuperado de <https://redicedia.edu.ec/document/434045>.
- Yang, J. Z. et al (2019). Development and validation of deep learning algorithms for scoliosis screening using back images. *Communications Biology*, 2, e390. Recuperado de <https://www.nature.com/articles/s42003-019-0635-8>.
- Yi, J.; Wu, P.; Huang, Q.; Qu, H.; Metaxas, D. N. (2020). Vertebra-Focused Landmark Detection for Scoliosis Assessment. arXiv:2001.03187v1. Recuperado de <https://arxiv.org/pdf/2001.03187>.

Sobre los autores



DIEGO ALBERTO GODOY

Es Doctor en Tecnologías de la Información y Comunicaciones por la Universidad de Vigo, España (Calificación Máxima con mención Honorífica “Cum Laude”), Magíster y Especialista en Ingeniería del Software por la Universidad Nacional de La Plata, Diploma de Estudios Avanzados en Ingeniería Telemática por la Universidad de Vigo, España e Ingeniero y Profesor Universitario en Informática en la Universidad Gastón Dachary (UGD). Es Decano del Departamento de Ingeniería y Ciencias de la Producción y fue Director de la Carrera de Ingeniería en Informática. Director del CITIC. Coordina proyectos de investigación en UGD relacionados con las TICs desde 2006.

ENRIQUE MARCELO ALBORNOZ

Actualmente es investigador adjunto del CONICET, Ingeniero en Informática y Doctor en Ingeniería en la Universidad Nacional del Litoral (UNL), con una mención en Inteligencia Computacional, Señales y Sistemas. Es también docente en el Departamento de Informática de la Facultad de Ingeniería y Ciencias Hídricas de la UNL desde 2007 y trabaja en el Instituto de investigación en Señales e Inteligencia Computacional. Es también investigador del CITIC.

LUCAS GABRIEL KUCUK

Magíster en Tecnologías de la Información (Universidad Nacional de Misiones), Profesor Universitario en Informática (Universidad Gastón Dachary), actualmente profesor titular de las cátedras Proyectos Informáticos y Sistemas Operativos de la Carrera Ingeniería en Informática de la Universidad, Director de Proyectos de Investigación en el marco del CITIC. Con más de diez años de experiencia dedicados al desarrollo de investigaciones vinculadas a Educación, Formación Superior, Robótica y Tecnologías Inmersivas.

CARLOS KORNUA

Doctorando en Matemática Computacional e Industrial en la Universidad Nacional del Centro de la Provincia de Buenos Aires. Becario de CONICET e investigador en el Instituto de Inves-

tigación en Tecnología Informática Avanzada (INTIA, UNICEN). Licenciado en Sistemas de Información (Universidad Nacional de Misiones), Profesor Universitario en Informática por UGD. Director de la Carrera y docente de Ingeniería en Informática de la Universidad Gaston Dachary, donde también se desempeña como investigador del CITIC.

CÉSAR GALLARDO

Es Ingeniero en Informática por la UGD y actualmente trabaja en el área de Business Intelligence. Fue disertante en el Congreso Iberoamericano ICSC-CITIES 2022 y autor en la editorial Springer Nature. Fue disertante en el Congreso Iberoamericano ICSC-CITIES 2022 y desarrolló su trabajo final de Carrera en el CITIC.

RICARDO SELVA

Es Ingeniero en Informática por UGD. Fue disertante en el Congreso Iberoamericano ICSC-CITIES 2022 y autor en la editorial Springer Nature. Desarrolló su trabajo final de Carrera en el CITIC.

NICOLÁS IBARRA

Es Ingeniero en Informática por UGD y actualmente trabaja como *Backend Developer*. Desarrollador en lenguajes tanto de *frontend* como de *backend*, como JavaScript (React) o C# (.Net). Fue

disertante en el Congreso Iberoamericano ICSC-CITIES 2022 y autor en la editorial Springer Nature. Fue Auxiliar de Cátedra en la materia Análisis y Diseño de Sistemas. Desarrolló su trabajo final de Carrera en el CITIC.

ANGELES BELÉN PIOTROSKI

Es egresada de Ingeniería en Informática de la Universidad Gastón Dachary y actualmente trabaja como desarrolladora semi senior en Laravel. Fue disertante en 14° Congreso Argentino de Informática y Salud, enmarcado en las 52° Jornadas Argentinas de Informática (52 JAIIO). Desarrolló su trabajo final de Carrera en el CITIC.

DAMIÁN HORACIO STETSON

Ingeniero en Informática egresado de la Universidad Gastón Dachary, residente en Posadas, Misiones, Argentina. Fue disertante en 14° Congreso Argentino de Informática y Salud, enmarcado en las 52° Jornadas Argentinas de Informática (52 JAIIO). Actualmente trabaja cómo desarrollador *mobile frontend*. Desarrolló su trabajo final de Carrera en el CITIC.

CITIC-UGD

Impulsando el conocimiento y la formación de investigadores e ingenieros en Misiones

Debemos asumir el desafío que supone para los estudiantes de nuestra provincia cursar una carrera con un perfil productivo e industrial en un contexto geográfico y económico como el nuestro. La provincia de Misiones, aunque privilegiada por su belleza natural, se encuentra alejada de los principales centros tecnológicos y productivos del país, lo que genera en muchos estudiantes la sensación de estar lejos de donde pasan las cosas. Este sentimiento persiste incluso en un mundo globalizado y conectado, donde es posible acceder a grandes oportunidades *sin salir de casa*.

Quienes editan y publican este libro buscan ofrecer una mirada de la importancia estratégica del CITIC en la Universidad Gastón Dachary y, en particular, dentro de la carrera de Ingeniería Informática. Pero podríamos agregar que, más importante aún, esta publicación coincide con la celebración de un espacio de integración de una comunidad académica que mira hacia la sociedad y sus problemas, en un ejercicio permanente de diálogo con la misma, sin por ello ir en detrimento de la exploración científica como fin en sí mismo.

[Extracto del Prólogo]



UNIVERSIDAD
Gastón Dachary



CITIC
Centro de Investigación en
Tecnologías de la Información
y Comunicaciones

