

Implementation of a Model based on Deep Learning techniques applied to Computer Vision in the Classification of X-ray Images, for the support of the diagnosis of traumatological injuries of the Pelvic Structure

Marco Aedo López, Ing.¹, Eveling Castro Gutiérrez, Mg.¹

¹Universidad Nacional de San Agustín de Arequipa, Perú, maedol@unsa.edu.pe, ecastro@unsa.edu.pe

Abstract -- This article describes the implementation of a model based on Deep Learning techniques applied to computational vision in the activity of classifying X-ray images as support in the diagnosis of traumatic lesions of the pelvic structure, specifically the acetabulum of the pelvis. In the area of Medical Sciences, nowadays, it is essential to have automated tools that support medical diagnosis. For the construction of these tools it is necessary to analyze the different techniques or methods provided by Computing, specifically Deep Learning, for the processing and interpretation of images and potentialize them with the application of GPUs to accelerate the achievement of results.

Keywords – Deep Learning, Computer Vision, Traumatological injuries, Pelvic structure, Pelvis Acetabulum.

Digital Object Identifier (DOI):
<http://dx.doi.org/10.18687/LACCEI2019.1.1.395>
ISBN: 978-0-9993443-6-1 ISSN: 2414-6390

Implementación de un Modelo basado en técnicas de Deep Learning aplicado a la Visión Computacional en la Clasificación de Imágenes de Rayos X, para el soporte del diagnóstico de lesiones traumatológicas de la Estructura Pélvica

Marco Aedo López, Ing.¹, Eveling Castro Gutiérrez, Mg.¹

¹Universidad Nacional de San Agustín de Arequipa, Perú, maedol@unsa.edu.pe, ecastro@unsa.edu.pe

Abstract—El presente artículo describe la implementación de un Modelo basado en técnicas de Deep Learning aplicado a la visión computacional en la actividad de clasificar imágenes de Rayos X como soporte en el diagnóstico de lesiones traumatológicas de la estructura pélvica, específicamente el acetábulo de la pelvis. En el área de las Ciencias Médicas, hoy en día, es indispensable disponer con herramientas automatizadas que den soporte al diagnóstico médico. Para la construcción de estas herramientas es necesario analizar las diferentes técnicas o métodos que nos provee el área de la Computación, específicamente del Deep Learning, para el procesamiento e interpretación de imágenes y potencializarlas con la aplicación de GPUs para acelerar la obtención de resultados.

Keywords—Deep Learning, Visión Computacional, Lesiones Traumatológicas, Estructura Pélvica, Acetábulo de Pelvis.

I. INTRODUCCIÓN

Existen diversos estudios en el ámbito internacional orientados al procesamiento e interpretación de imágenes médicas, sin embargo, en Perú y en varios países de Latinoamérica es aún poco abordado.

Entre las imágenes médicas que se utilizan con mayor frecuencia se encuentran las Imágenes de Resonancia Magnética (MRI), Tomografías Computarizadas (CT) y Rayos X (Rx). De las antes mencionadas, las imágenes de Rx son inferiores en cuanto a la calidad de la imagen en comparación con las otras dos, sin embargo, a pesar de tener grandes desventajas tiene ventajas en costo, rapidez de obtención, equipo requerido no tan costoso y el efecto de la radiación.

Dichas imágenes permiten examinar diversas lesiones traumatológicas que pueda presentar una determinada estructura ósea, sin embargo, las fracturas ubicadas en la pelvis son las únicas fracturas en traumatología que pueden causar la muerte del paciente, de allí su importancia [1] [2].

Se utilizan técnicas de Deep Learning (DL), que se considera la última encarnación de las redes neuronales (NN), pero que aprovecha las ventajas del desarrollo tecnológico en hardware, de la existencia de mayor cantidad de datos disponibles para utilizar en la actualidad y de mejoras

importantes en los algoritmos utilizados [3]. El problema a resolver es de Visión Computacional (VC), los expertos humanos pueden resolver este problema sin mucho esfuerzo; pero es complicado para las computadoras, ya que deben entender el contenido de las imágenes y lo que el DL plantea es hacer que la computadora aprenda a partir de datos [4].

Nuestro trabajo presenta la experiencia de implementación de un Modelo basado en técnicas de DL aplicado a la VC, específicamente en la actividad de clasificación de Rx, que permita un buen soporte en el diagnóstico de lesiones traumatológicas de la estructura pélvica, puntualmente en el acetábulo de la pelvis; contribuyendo así, en el ámbito de clasificación de Rx pélvicos que son poco abordados en el ámbito nacional y regional, luchando así frontalmente contra una de las consecuencias más comunes que trae consigo los accidentes de tránsito (atropellos), un problema a nivel mundial.

El resto del artículo está organizado de la siguiente manera: la sección II describe el fundamento médico del problema a resolver y su importancia. En la sección III se describe al DL como subárea del Machine Learning (ML) y su aplicación a la VC. En la sección IV se presentan las Convolutional Neural Networks (CNN) que serán la técnica específica de Deep Learning aplicada al presente trabajo. En la sección V se presentan las generalidades de nuestra propuesta de modelo basado en técnicas de Deep Learning para el problema a resolver, en la sección VI se muestra la experimentación con diferentes modelos y los resultados obtenidos. Finalmente se presentan las conclusiones.

II. LESIONES TRAUMATOLÓGICAS DE LA ESTRUCTURA PÉLVICA

Es bastante claro que en el área de las Ciencias Médicas, hoy en día, es indispensable contar con herramientas automatizadas que den soporte al diagnóstico médico.

La Traumatología es una especialidad de la Medicina encargada del diagnóstico y tratamiento de las lesiones del aparato locomotor. Utiliza las imágenes médicas como un

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LACCEI2019.1.1.395>

ISBN: 978-0-9993443-6-1 ISSN: 2414-6390

medio para realizar su trabajo, estas imágenes principalmente son MRI, CT y Rx.

Entre dichas técnicas de radiología que han despertado gran atención por diversos autores se encuentran las imágenes de Rx, las cuales presentan un mayor desafío en comparación con las otras técnicas radiológicas, como es el caso de las CT; aunque estas últimas son superiores en cuanto a calidad de la imagen, resultan ser muy costosas y perjudiciales, por la radiación, para el paciente.

En el presente estudio nos centramos en los Rx que, aunque tiene desventajas, tiene las siguientes ventajas:

- Más accesibles en costo
- Más rápidas de obtener
- No requieren un equipamiento costoso
- El paciente no sufre exposición excesiva a radiación

Basados en dichas imágenes se pueden examinar diversas lesiones traumatológicas que pueda presentar una determinada estructura ósea, sin embargo, nos centramos en las fracturas ubicadas en la pelvis que son las únicas fracturas en traumatología que pueden causar incluso la muerte del paciente, de allí su importancia [2]. Específicamente nos centramos en las fracturas del acetábulo de la pelvis [2][5], que, aunque no producen la muerte del paciente, son de alto impacto, así, producen artrosis o daño del nervio ciático.

Según estadísticas del Instituto Nacional de Estadística e Informática (INEI) en [6] [7], en el Perú, el grupo con más riesgo de fracturas de acetábulo son personas entre las edades de 21 a 30 años (35%), 65% son de sexo masculino, y la causa más común fue el accidente de tránsito peatonal (atropello).

Analizando nuestra realidad nacional y local, el INEI [6] determinó lo siguiente:

- El tercer tipo de accidente de tránsito con mayor porcentaje de casos registrados durante el año 2016 es el atropello, después de la colisión y choque
- La segunda región con mayor incidencia de accidentes de tránsito en Perú es la región Arequipa con 7 mil 076 accidentes de tránsito que representa el 6,06%

Según los Centros para el Control y la Prevención de Enfermedades (CDC), las lesiones traumatológicas causan la muerte de más personas entre las edades de 1 y 44 años que cualquier otra enfermedad. Así, según [1] las lesiones traumatológicas presentes en la estructura pélvica constituyen uno de los problemas más complejos de las estructuras óseas y constituyen el 3% de las lesiones del esqueleto.

Hay que considerar que las fracturas localizadas en la estructura pélvica son de alta incidencia [8] y la mayoría de veces son causadas por accidentes de tránsito, caídas o accidentes deportivos. También debemos tener en cuenta que se producen junto con múltiples lesiones en otras localizaciones corporales, así se les asocia con altas tasas de mortalidad y discapacidad [2] [5].

Según [9] [10] las fracturas de acetábulo constituyen un tipo de lesiones de incidencia creciente y pueden ser incluso potencialmente letales a pesar de los avances en cuanto a su

tratamiento. Incluso en fracturas de alto impacto, hasta el 30% de los casos pueden producir lesiones neurológicas, generalmente parciales del nervio ciático, que afectan con más frecuencia a la división peronea que a la tibial [11] [12].

Tradicionalmente, el diagnóstico de fracturas de acetábulo en pacientes, se realiza clínicamente y radiológicamente mediante una vista anteroposterior (AP) de imágenes de Rx de la región pélvica (debido a la cantidad significativa de información que éstas proveen). Dicha vista permite a los especialistas tomar decisiones de diagnóstico y obtener información para la planificación del tratamiento que se debe seguir [13].

Pero en la actualidad, apoyarse de la detección automatizada de fractura de acetábulo en lesiones pélvicas politraumatizadas puede ayudar a los médicos en el análisis de imágenes de Rx para detectar la gravedad de las lesiones en un corto periodo de tiempo.

La tarea de detección de la fractura puede ser difícil debido a la baja resolución de las imágenes de Rx, solapamiento de huesos, variaciones en la forma ósea y diferencia entre cada paciente [14]. Considerando además que la presencia de ruido, efectos de bajo contraste y nitidez de las imágenes de Rx, hacen que la tarea de detección de fracturas de acetábulo de forma automatizada sea desafiante, debido a que contiene información que no se puede analizar a fondo con la inspección visual.

Teniendo en cuenta lo anterior, la clasificación precisa de las imágenes de Rx para determinar fracturas de acetábulo será fundamental para la interpretación de las mismas. Así, el proceso de clasificación automatizada es fundamental para el soporte al diagnóstico de lesiones localizadas en el acetábulo; determinando así, en base a imágenes de Rx, si un paciente presenta o no una fractura de acetábulo. Ver Fig. 1.

La clasificación con precisión de la fractura de acetábulo de la estructura pélvica es un paso fundamental para la interpretación de imágenes médicas, siendo el aspecto fundamental a considerar en el presente artículo, así, la interpretación que se obtenga a partir de la imagen médica será reflejada en el diagnóstico emitido por un médico especialista.

Aplicamos técnicas del DL para alcanzar dicho objetivo y también consideramos que es importante reducir el tiempo de respuesta en la emisión de un diagnóstico, el cual se alcanzará mediante el uso de Graphics Processing Units (GPU).



Fig. 1 Algunas imágenes utilizadas como datos para entrenar el Modelo

III. DEEP LEARNING

A. Generalidades

Aunque la Inteligencia Artificial (IA) nació en los años 50, es en los últimos años que se ha visto su resurgimiento de mano del ML y específicamente del DL. Se considera al ML como una subárea de la IA y a su vez el DL una subárea del ML [15], ver Fig. 2.

Aunque el DL es una subárea ya antigua del ML, su surgimiento estelar es reciente, a partir de los primeros años de la presente década; pero ha ocasionado una revolución en el área de la Computación.

A diferencia de otras técnicas, el DL tiene un enfoque diferente. En lugar de definir a mano un conjunto de técnicas y algoritmos con el objetivo de extraer características de una imagen, dichas características se aprenden automáticamente usando un proceso de entrenamiento.

Se dice que aprende representaciones a partir de los datos y que aprende capas sucesivas de representaciones más significativas incrementalmente. En DL se utiliza una jerarquía de conceptos, donde cada concepto se construye a partir de otros conceptos de las capas más bajas. A diferencia de otras aproximaciones del ML que sólo tienen 1 ó 2 capas (shallow learning), el DL tiene más capas.

En DL las representaciones en las capas son aprendidas por medio de modelos llamados redes neuronales [3]. Cada capa transforma la imagen de entrada en representaciones que son diferentes de dicha imagen, pero que se van construyendo a partir de las anteriores. Ver Fig. 3.

En síntesis, el objetivo del DL es encontrar representaciones ordenadas para datos complejos, toma la aproximación de descomponer incrementalmente una transformación geométrica complicada y enredada en una cadena de pasos elementales. Cada capa aplica una transformación que desenreda los datos un poco, haciendo que una secuencia de capas haga tratable un proceso de desenredo extremadamente difícil.

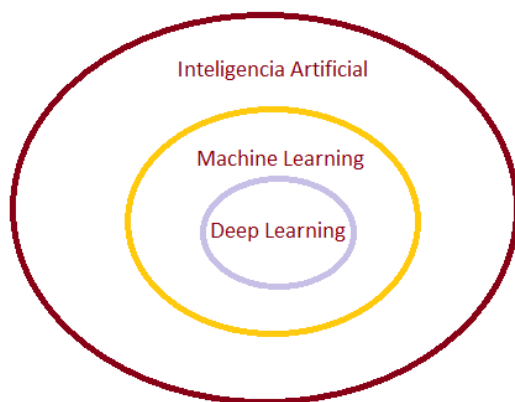


Fig. 2 La relación de inclusión entre la Inteligencia Artificial, el Machine Learning y el Deep Learning

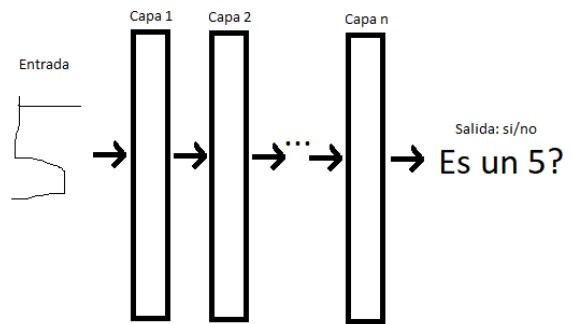


Fig. 3 Modelo con varias capas de profundidad que reconoce si la imagen es un 5 o no lo es (clasificación binaria)

B. Campos en los que ha alcanzado avances

Tiene resultados muy destacables en problemas de percepción que involucran habilidades naturales e intuitivas para los seres humanos, tales como visión y audición, pero que tradicionalmente han sido muy esquivas a las computadoras.

Los campos en que DL ha alcanzado destacados avances son [3]:

- 1) *Clasificación de imágenes*: A nivel cercano a humano
- 2) *Reconocimiento de voz*: A nivel cercano a humano
- 3) *Transcripción de escritura a mano*: A nivel cercano a humano
- 4) *Traducción automática mejorada*
- 5) *Conversión de texto a voz mejorada*
- 6) *Asistentes digitales*: Como Google Now y Amazon Alexa
- 7) *Manejo autónomo de autos*: A nivel cercano a humano
- 8) *Búsquedas*: Mejora de resultados de búsqueda en web
- 9) *Habilidad para responder preguntas*: En lenguaje natural
- 10) *Videojuegos*

Inclusive ya existen estudios sobre su aplicación al razonamiento formal, que podrían aplicarse en la creación de asistentes inteligentes en diferentes áreas del conocimiento humano.

C. Las razones de su éxito

Los fundamentos teóricos del DL para la visión computacional existen desde fines de los años '80s y otros tópicos relacionados vienen de los años '90s, sin embargo, es a partir del 2012 que el DL hace su aparición con fuerza debido a 3 razones.

1) *El hardware*: Hace 20 años hubiera sido intratable ejecutar modelos de DL en el hardware de esa época. En la presente década empresas de hardware, básicamente NVIDIA [16] han invertido millones de dólares en el desarrollo de los GPUs (originalmente pensado sólo en videojuegos más reales). El 2007 NVIDIA lanzó Compute Unified Device Architecture (CUDA) [17] para la comunidad mundial, es una interface de programación para sus GPUs. Debemos considerar que un GPU puede reemplazar a un conjunto de CPUs en aplicaciones paralelizables y el DL es muy paralelizable y fue natural su adopción.

2) *Los datos*: Son el combustible, la materia prima que hace que el DL suceda. El internet ha permitido que se puedan coleccionar muchos datos de diferente tipo, tales como imágenes, en datasets (colección de datos ordenados). Un dataset en particular ha ayudado mucho, es el ImageNet [18], que contiene millones de imágenes categorizadas.

3) *Los algoritmos*: Antes las NN sólo poseían 1 ó 2 capas de representaciones y no podían competir con otros métodos como Support Vector Machines (SVM). En el 2010 se mejoraron los siguientes algoritmos:

- Funciones de activación mejoradas
- Esquemas de inicialización de pesos mejoradas
- Esquemas de optimización mejoradas, tales como RMSProp y Adam

Por el 2015 aparecieron nuevos algoritmos como:

- Batch Normalization
- Residual Connections
- Depthwise Separable Convolutions

En la actualidad se puede entrenar modelos de cientos de capas de profundidad.

IV. CONVOLUTIONAL NEURAL NETWORKS

A. Generalidades

Las CNN, también conocidas como “convnets” son muy utilizadas en aplicaciones de visión computacional y son incluso aplicables a problemas de clasificación de imágenes que cuentan con poca cantidad de datos para su entrenamiento.

Las CNN aprenden patrones locales que encuentran en las imágenes, los patrones se encuentran en pequeñas ventanas de entrada 2D.

B. Propiedades

1) *Los patrones que aprenden son invariantes de traslación*: Si aprendió un patrón en el extremo superior derecho de la imagen, lo reconocerá en cualquier otra parte de la misma. A diferencia de una red neuronal densamente conectada que tendría que aprender el patrón de nuevo si aparece en una nueva localización. Su uso es óptimo para procesamiento de imágenes (mundo visual es invariante de traslación). Necesitan menor cantidad de entrenamiento para aprender representaciones que tienen poder de generalización.

2) *Pueden aprender jerarquías de patrones espaciales*: la primera capa convolucional aprenderá patrones locales pequeños, tales como bordes, la segunda capa aprenderá sobre patrones aprendidos por la primera capa y así sucesivamente. Pueden aprender eficientemente conceptos visuales abstractos ante una complejidad en aumento, ya que el mundo visual es jerárquico espacialmente.

C. Overfitting

Cuando se entrena un modelo de DL puede ser que con los datos de validación se empiece a estancar su precisión después de algunas épocas y luego incluso, puede llegar a degradarse. Esto significa que el modelo rápidamente empieza a hacer overfitting en el entrenamiento.

Hay que saber tratar con él, considerando que lo principal es la tensión existente entre optimización y generalización

1) *Optimización*: ajuste de modelo para conseguir la mejor performance posible en el entrenamiento.

2) *Generalización*: cuán bien el modelo entrenado se comporta con datos nunca antes vistos.

Se debe buscar una buena generalización, pero sólo podemos ajustarla basados en el entrenamiento. En el inicio ambos conceptos están correlacionados, pero eso puede variar al pasar de las épocas.

Así, después de algunas épocas, la generalización para de mejorar, las métricas sobre los datos de validación se estancan o degradan, decimos en ese caso que hay overfitting.

Se observa que el modelo aprende para los datos de entrenamiento, pero no mejora en nada con los casos nunca vistos o datos de validación. Debemos considerar que el desafío real es la generalización, más que la optimización.

D. Técnicas aplicadas en los experimentos propuestos

1) Dropout

Es una técnica sencilla y poderosa de regularización para NN y modelos de DL. Fue propuesta por Srivastava [19] y su objetivo es conseguir una mejor generalización y disminuir la probabilidad de overfitting.

En esta técnica se seleccionan neuronas al azar y se ignoran durante el entrenamiento. Ellas son “abandonadas” al azar. Esto significa que su contribución a la activación de las neuronas posteriores se elimina temporalmente en el paso forward y las actualizaciones de peso no se aplican a la neurona en el paso backward.

2) Data Augmentation

Los modelos del DL necesitan una adecuada cantidad de datos de entrenamiento para lograr un buen rendimiento. Cuando tenemos muy pocos datos de entrenamiento y deseamos crear un potente clasificador de imágenes, generalmente se requiere aplicar data augmentation para mejorar el rendimiento. Consiste en crear artificialmente más imágenes de entrenamiento a través de diferentes formas de procesamiento o combinación de procesamiento múltiple, como rotaciones, traslados, cambios, cortes, hacer zoom, etc.

El objetivo es que, al entrenar, el modelo nunca verá la misma imagen exacta dos veces y a más datos, mejor generalización.

3) Feature Extraction

Consiste en usar representaciones aprendidas por una red preentrenada para extraer características interesantes sobre nuevos datos. Luego, estas características se ejecutan a través de un nuevo clasificador que es entrenado desde cero.

En los CNN consiste en tomar la base convolucional de una red entrenada previamente (capas iniciales) y luego aplicar los datos nuevos a través de ella y entrenar un clasificador nuevo.

Al tomar sólo la base convolucional, se utilizarán las representaciones más genéricas y por ende más reusables, las cuales probablemente sean útiles independientemente del problema de visión computacional que se tenga que resolver.

V. PROPUESTA

A. Generalidades

Es de conocimiento general que el DL sólo funciona con una buena cantidad de datos disponibles. Esta afirmación es relativa, está claro que con pocas decenas de datos sería imposible, pero con varios cientos de datos es posible solucionar el problema si es una clasificación binaria de imágenes, por ejemplo.

También nos favoreceremos de la propiedad que los modelos de DL son reutilizables. Utilizaremos un modelo llamado VGG19 [20] ya preentrenado con un dataset a gran escala y lo utilizaremos aplicándolo a nuestro problema, realizando sólo cambios puntuales.

Todo el trabajo será realizado evitando el overfitting en lo posible, para tal fin aplicamos diferentes técnicas como data augmentation, dropout y feature extraction.

B. Entorno de desarrollo a utilizar

Utilizaremos el siguiente software:

1) *Tensorflow*: Biblioteca de código abierto desarrollada por Google y que se utiliza, entre otras áreas, para aplicaciones del ML como redes neuronales que permiten detectar patrones y correlaciones [21].

2) *Keras*: API de alto nivel para Tensorflow, fue diseñada para hacer experimentos usando menor tiempo [22].

3) *Jupiter Notebook*: Aplicación web que permite crear y compartir documentos que contienen “live code” y visualización de resultados. Muy buena para desarrollo incremental [23].

4) *CUDA*: Plataforma de procesamiento paralelo de NVIDIA que aprovecha los GPUs.

5) *cuDNN*: Biblioteca de redes neuronales profundas, permite aceleración de alta performance del GPU, ya que maneja primitivas para DL.

6) *Sistema Operativo*: puede ser Microsoft Windows 7/8/10 o cualquier versión de Ubuntu actual.

Utilizaremos el siguiente hardware:

1) *Tarjeta de Video*: NVIDIA GTX 1060 (GPU)

2) *Computadora*: con CPU Core-i5 8400 2.80 GHz, 16 GB de RAM, Disco duro de 2 TB

Utilizaremos el siguiente dataset:

1) *Datos*: 500 imágenes de pelvis (entre sanas y con fractura). Se usarán 250 para entrenamiento, 125 para validación y 125 para testing.

En la Fig. 4 se puede observar la pila de herramientas utilizada en la experimentación.

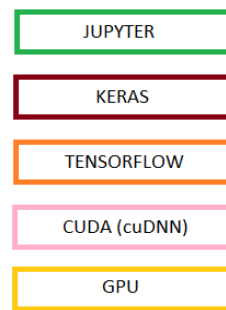


Fig. 4 Software y Hardware utilizado en experimentos. Pila de herramientas

C. Descripción general de los modelos a probar

Debemos ser conscientes que no existe una fórmula mágica para determinar el número de capas y el tamaño de cada capa. Se debe evaluar un conjunto de diferentes arquitecturas sobre los datos de validación, para encontrar el tamaño correcto del modelo para nuestros datos.

El flujo de trabajo para encontrar un tamaño de modelo apropiado es empezar con pocas capas y parámetros e ir aumentando el tamaño de las capas o añadir nuevas capas hasta que se vea que no hay estancamiento en la precisión de los datos de validación.

Se describirá la experimentación realizada hasta llegar al modelo propuesto. Para cada experimento se usarán 30 épocas o iteraciones (eje x) y se medirá el porcentaje de precisión del dataset de entrenamiento y de validación (eje y).

VI. RESULTADOS

A. Modelo inicial con CNN

Características: lo constituye una secuencia de 8 capas alternadas entre capas convolucionales (Conv2D) y capas pooling (MaxPooling2D). Seguidas secuencialmente con 2 capas densamente conectadas. Utiliza dropout.

Análisis: el modelo aprende, pero se observa que a partir de aproximadamente la iteración 8, se produce el overfitting (memoriza las imágenes de entrenamiento, pero no mejora sustancialmente con las imágenes de validación). El resultado para las imágenes de validación es muy pobre y debe ser mejorado. Ver Fig. 5.

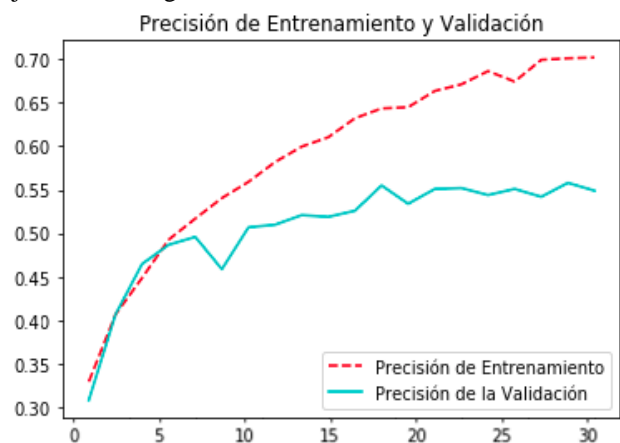


Fig. 5 Resultados de la precisión del conjunto de datos de entrenamiento y validación

B. Modelo CNN con Data Augmentation

Características: el mismo modelo anterior, pero aplicando la técnica de data augmentation para evitar el overfitting.

Análisis: el modelo aprende, no se produce el overfitting; sin embargo, los niveles alcanzados no son lo suficientemente óptimos. Ver Fig. 6.

C. Modelo que usa un modelo preentrenado VGG19 y aplicando Feature Extraction

Características: se utilizó un modelo preentrenado VGG19 congelado y seguido secuencialmente con 2 capas densamente conectadas. Utiliza dropout y feature extraction (no usa data augmentation).

Análisis: el modelo aprende, aunque se produce un claro overfitting a partir de la época 10. Ver Fig. 7.

D. Modelo que usa un modelo preentrenado VGG19, aplicando Feature Extraction y Data Augmentation

Características: se utilizó un modelo preentrenado VGG19 congelado y seguido secuencialmente con 2 capas densamente conectadas. Se descongela el bloque 5 y se utiliza feature extraction y data augmentation.

Análisis: el modelo aprende, hay muy poco overfitting; los resultados mejoran considerablemente en comparación a los modelos anteriores. Ver Fig. 8.

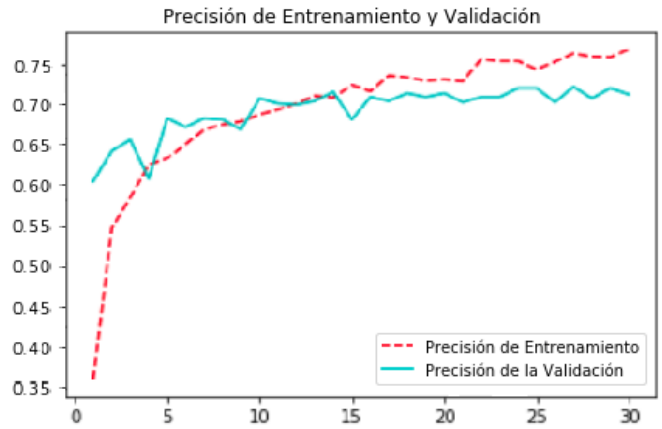


Fig. 8 Resultados de la precisión del conjunto de datos de entrenamiento y validación

En la Fig. 9 se puede ver detalladamente la arquitectura del modelo propuesto que dio los mejores resultados y sus respectivos parámetros de entrenamiento.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 150, 150, 3)	0
block1_conv1 (Conv2D)	(None, 150, 150, 64)	1792
block1_conv2 (Conv2D)	(None, 150, 150, 64)	36928
block1_pool1 (MaxPooling2D)	(None, 75, 75, 64)	0
block2_conv1 (Conv2D)	(None, 75, 75, 128)	73856
block2_conv2 (Conv2D)	(None, 75, 75, 128)	147584
block2_pool1 (MaxPooling2D)	(None, 37, 37, 128)	0
block3_conv1 (Conv2D)	(None, 37, 37, 256)	295168
block3_conv2 (Conv2D)	(None, 37, 37, 256)	590080
block3_conv3 (Conv2D)	(None, 37, 37, 256)	590080
block3_conv4 (Conv2D)	(None, 37, 37, 256)	590080
block3_pool1 (MaxPooling2D)	(None, 18, 18, 256)	0
block4_conv1 (Conv2D)	(None, 18, 18, 512)	1180160
block4_conv2 (Conv2D)	(None, 18, 18, 512)	2359808
block4_conv3 (Conv2D)	(None, 18, 18, 512)	2359808
block4_conv4 (Conv2D)	(None, 18, 18, 512)	2359808
block4_pool1 (MaxPooling2D)	(None, 9, 9, 512)	0
block5_conv1 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv2 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv3 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv4 (Conv2D)	(None, 9, 9, 512)	2359808
block5_pool1 (MaxPooling2D)	(None, 4, 4, 512)	0
flatten_1 (Flatten)	(None, 8192)	0
dense_1 (Dense)	(None, 256)	2097408
dense_2 (Dense)	(None, 1)	257
Total params: 22,122,049		
Trainable params: 22,122,049		
Non-trainable params: 0		

Fig. 9 Arquitectura del modelo que dio mejores resultados y sus parámetros de entrenamiento

Precisión de Entrenamiento y Validación

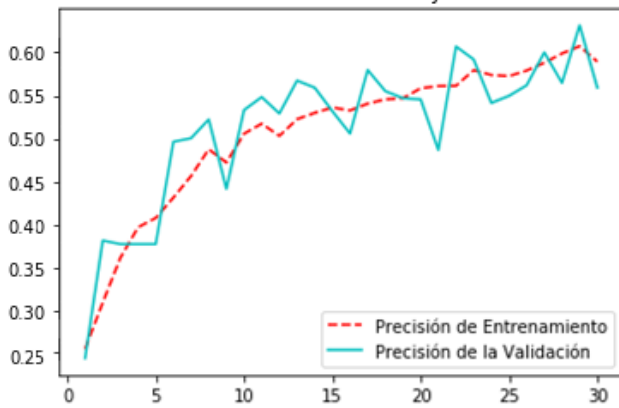


Fig. 6 Resultados de la precisión del conjunto de datos de entrenamiento y validación

Precisión de Entrenamiento y Validación

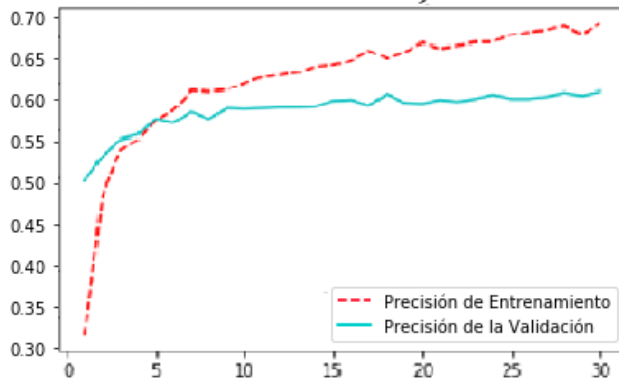


Fig. 7 Resultados de la precisión del conjunto de datos de entrenamiento y validación

CONCLUSIONES

En el presente trabajo se presentó la experiencia de implementación de un modelo basado en técnicas de DL.

Se realizaron 4 modelos de forma iterativa, hasta llegar al modelo propuesto.

Cada modelo utiliza técnicas planteadas en el anterior, pero va mejorando en los resultados que alcanza.

Trabajar con relativamente pocos datos en DL es complicado, pero hay técnicas que nos pueden ayudar a alcanzar nuestros objetivos.

Se espera que con un dataset mayor, los resultados obtenidos serán mucho mejores.

Existen modelos preentrenados más modernos que podrían darnos incluso mejores resultados.

Es indispensable el uso de GPUs, hacer el entrenamiento sólo con CPUs (aunque fueran de última generación) se hace intratable por el tiempo consumido.

REFERENCIAS

- [1] F. Coccolini, P. F. Stahel, G. Montori, W. Biffi, T. M. Horner, F. Catena, Y. Kluger, E. E. Moore, A. B. Peitzman, R. Ivatury, et al., "Pelvic trauma: Wses classification and guidelines," *World Journal of Emergency Surgery*, vol. 12, no. 1, 2017.
- [2] J.V. Yzaguirre, "Guías de práctica clínica de la especialidad de ortopedia-traumatología y guías de práctica clínica de la especialidad de medicina física y rehabilitación," Departamento de traumatología del Hospital de Emergencias José Casimiro Ulloa - Resolución Directorial del Ministerio de Salud, 2011.
- [3] N. Buduma, "Fundamentals of Deep Learning", 1st ed., O'Reilly Media, Inc., 2017.
- [4] A. Rosebrock, "Deep Learning for Computer Vision with Python", 1st ed., PyImageSearch, 2017.
- [5] E. J. Karadimas, T. Nicolson, D. D. Kakagia, S. J. Matthews, P. J. Richards, and P. V. Giannoudis, "Angiographic embolisation of pelvic ring injuries. treatment algorithm and review of the literature," *International orthopaedics*, vol. 35, no. 9, 2011.
- [6] "Análisis de los accidentes de tránsito ocurridos en el 2016- iii. Perú: Iv censo nacional de comisarías 2015. resultados definitivos," Instituto Nacional de Estadística e Informática - INEI, 2017.
- [7] "Perú: Enfermedades no trasmisibles y transmisibles, 2016," Instituto Nacional de Estadística e Informática - INEI, 2017.
- [8] E. S. Sánchez, Estudio de la precisión de la imagen térmica infrarroja en el diagnóstico de lesiones traumatológicas en la edad pediátrica. PhD thesis, Universitat de València, 2014.
- [9] F. Granell-Escobar, A. Montiel-Giménez, S. Gallardo-Villares, and A. Coll-Bosch, "Complicaciones de las fracturas de acetábulo," *Revista Española de Cirugía Ortopédica y Traumatología*, vol. 50, no. 2, 2006.
- [10] J. Moreta, X. Foruria, and F. Labayru, "Presentación tardía de lesión del nervio ciático secundaria a compresión por placa de osteosíntesis acetabular," *Revista Española de Cirugía Ortopédica y Traumatología*, vol. 60, no. 4, 2016.
- [11] D. A. Lawrence, K. Menn, M. Baumgaertner, and A. H. Haims, "Acetabular fractures: anatomic and clinical considerations," *American Journal of Roentgenology*, vol. 201, no. 3, 2013.
- [12] P. R. Mardones and B. C. Larrain, "Artrosis de cadera: tratamiento no protésico y alternativas de manejo con células madres mesenquimáticas," *Revista Médica Clínica Las Condes*, vol. 25, no. 5, 2014.
- [13] J. Jeuthe, "Automatic tissue segmentation of volumetric ct data of the pelvic region," 2017.
- [14] W. Xie, J. Franke, C. Chen, P. Gruetzner, S. Schumann, L.-P. Nolte, and G. Zheng, "A complete pelvis segmentation framework for image-free total hip arthroplasty (tha): methodology and clinical study," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 11, 2014.
- [15] F. Chollet, "Deep Learning with Python", 1st ed., Manning Publications Co., 2018.
- [16] NVIDIA. <https://www.nvidia.com/en-us/>
- [17] CUDA. <https://developer.nvidia.com/cuda-zone>
- [18] IMAGENET. <http://www.image-net.org/>
- [19] N. Srivastava, G. Hinton, A. Krizhevsky and I. Sutskever, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, 2014.
- [20] VGG19. <https://www.kaggle.com/keras/vgg19/home>
- [21] TENSORFLOW. <https://www.tensorflow.org>
- [22] KERAS. <https://keras.io/>
- [23] JUPYTER NOTEBOOK. <https://jupyter.org/>