



UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA

FACULTAD DE INGENIERÍA

PROYECTO DE INVESTIGACIÓN

PROCESAMIENTO DE IMÁGENES DE NEURORADIOLOGÍA UTILIZANDO

COMPUTER VISION PARA DETECCIÓN DE VERTEBRAS

PREVIO A LA OBTENCIÓN DEL TÍTULO

INGENIERO EN BIOMÉDICA

PRESENTADO POR:

21741367 YOSHIAKI ANTONIO SUGUIYAMA YANES

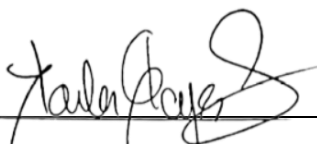
21851107 ALICIA JEANETH GARCIA REGALADO

ASESOR: KARLA MIRIAM REYES LEIVA

DICIEMBRE, 2023

HOJA DE FIRMAS

Los abajo firmantes damos fe, en nuestra posición de miembro de Terna, Asesor y/o Académico y en el marco de nuestras responsabilidades adquiridas, que el presente documento cumple con los lineamientos exigidos por la Facultad de Ingeniería y los requerimientos académicos que la Universidad dispone dentro de los procesos de graduación.



Ing. Karla Miriam Reyes Leiva
Asesor UNITEC



Ing. Manuel Adalid Gamero

Coordinador de terna



Ing. Alicia Reyes Duke

Miembro de terna



Ing. José Luis Ordoñez

Miembro de terna



Ing. Reyna Valle

Jefe Académico de Ingeniería en Biomédica | UNITEC

RESUMEN EJECUTIVO

La columna vertebral, compuesta por 24 vértebras en la etapa adulta, desempeña un papel fundamental en la estabilidad y movilidad del cuerpo humano. Sin embargo, está sujeta a deformidades o alteraciones que pueden manifestarse de diversas maneras en pacientes, dependiendo de factores como la edad, traumatismos y patologías. Entre estas condiciones, incluyendo espondilosis, tumores, hernia de disco, fracturas y la de nuestro interés espondilolistesis. Los métodos tradicionales de detección de vértebras se basan en técnicas de detección de bordes, las cuales pueden ser inexactas en imágenes de baja calidad o con presencia de ruido. En este artículo, proponemos un algoritmo de reconocimiento de vértebras que se fundamenta en la detección de puntos clave o keypoints utilizando lenguaje Python y la herramienta Roboflow.

En este contexto, la detección de vértebras a través de técnicas de visión computacional se ha convertido en un área de interés significativa. La capacidad de las CNN para discernir patrones anatómicos específicos contribuye a mejorar la eficacia del análisis, reduciendo la carga de trabajo manual y, al mismo tiempo, aumentando la consistencia y exactitud de los resultados. Nuestro algoritmo utiliza un conjunto predefinido de puntos clave para identificar las vértebras en la imagen, empleando el modelo KeyPoints RCNN para su detección. La evaluación del algoritmo se llevó a cabo en un conjunto de datos de imágenes de rayos X de la columna lumbar, específicamente en imágenes en plano sagital. Los resultados indican que el algoritmo logra una precisión promedio del 98% en el entrenamiento de la detección de vértebras, con una pérdida mínima en la detección de la región de interés. Esto facilita significativamente la detección de vértebras que tienen cierto grado de desplazamiento.

Estos resultados respaldan que los algoritmos de detección de vértebras basados en el uso de puntos clave representan una herramienta prometedora para la detección de deformidades en las curvaturas de la espina dorsal en imágenes de rayos X. Esta metodología ofrece una detección de vértebras, proporcionando información sobre la

gravedad del desplazamiento vertebral a través de la detección de los keypoints en las imágenes analizadas por el algoritmo.

Palabras claves: aprendizaje de transferencia profunda, detección de vértebras, inteligencia artificial, métricas de desempeño, procesamiento de imágenes, puntos clave, radiología.

ABSTRACT

The spine, composed of 24 vertebrae in adulthood, plays a fundamental role in the stability and mobility of the human body. However, it is susceptible to deformities or alterations that may manifest in various ways in patients, depending on factors such as age, trauma, and pathologies. Among these conditions, including spondylosis, tumors, herniated discs, fractures, and our focus, spondylolisthesis. Traditional methods of vertebral detection rely on edge detection techniques, which can be inaccurate in low-quality images or in the presence of noise. In this article, we propose a vertebral recognition algorithm based on the detection of key points or keypoints using python language and Roboflow tools.

In this context, vertebral detection through computer vision techniques has become a significant area of interest. The ability of CNNs to discern specific anatomical patterns contributes to improving the effectiveness of the analysis, reducing manual workload, and simultaneously increasing the consistency and accuracy of the results. Our algorithm uses a predefined set of keypoints to identify vertebrae in the image, employing the KeyPoints RCNN model for detection. Algorithm evaluation was carried out on a dataset of X-ray images of the lumbar spine, specifically in sagittal plane images. The results indicate that the algorithm achieves an average precision of 98% in vertebral detection training, with minimal loss in detecting the region of interest. This significantly facilitates the detection of vertebrae with a certain degree of displacement.

These results support the conclusion that vertebral detection algorithms based on the use of keypoints represent a promising tool for detecting deformities in the curvatures of the spine in X-ray images. This methodology provides vertebral detection, offering information on the severity of vertebral displacement through the detection of keypoints in the images analyzed by the algorithm.

Key words: artificial intelligence, deep transfer learning, image processing, key points, performance metrics, radiology, vertebrae detection.

ÍNDICE DE CONTENIDO

I.	Introducción	15
II.	Estado del Arte	17
	2.1 Antecedentes	17
	2.2 Problemática	23
	2.3 Imagen integradora	26
	2.4 Tabla de limitaciones	28
III.	Objetivo	32
	3.1. Objetivo General	32
	3.2. Objetivos específicos	33
IV.	Método	33
	4.5 Metodología de Estudio	42
	4.6 Metodología de Validación	48
	4.6.1 Métricas de Validación y Cálculos Numéricos.	48
	4.5 Cronograma de Actividades	51
	4.6 Matriz Metodológica	52
V.	Resultados y Análisis.....	54
	5.1. Resultados de primer entrenamiento.	54
	5.2. Resultados Primera Prueba	55
	5.3. Resultados de predicciones de modelo Propio	58
	5.4. Etapa final: Evaluación a través de comparativas	64
	5.5. Inferencia	65
VI.	Conclusiones.....	69

Anexos	71
Referencias Bibliográficas	99

ÍNDICE DE ILUSTRACIONES

Ilustración 1. Imagen integradora	26
Ilustración 2. Variables Dependientes.....	35
Ilustración 3. Variables Independientes.....	36
Ilustración 4. Variables de la investigación	37
Ilustración 5. Estructura de modelo KeyPoint RCNN	38
Ilustración 6. Referencia de imágenes de BUU-LSpine	41
Ilustración 8. Resumen de procedimiento para la detección de vértebras	47
Ilustración 8. Gráfica de entrenamiento 25 épocas	58
Ilustración 9. Gráfica de Resultados Validación 25 épocas	59
Ilustración 10. Gráfica Entrenamiento 100 épocas	60
Ilustración 11. Gráfica de Resultados de Validación 100 épocas	61
Ilustración 12. Resultados detección de vértebras en el etiquetado semi-supervisado ...	66
Ilustración 13. Ejemplos de inferencia	68

ÍNDICE DE TABLAS

Tabla 1. Investigación Krizhevsky, et al. (2012) ImageNet Classification with Deep Convolutional.....	18
Tabla 2. Investigación Kai Chen et al.,2021, El aprendizaje automático como revolución prometedora en la práctica clínica de la escoliosis.	19

Tabla 3. Resultados de Investigación de Klinwichit, et al., (2022), BUU-LSPINE.....	22
Tabla 4. Detección de área pulmonar con 100 épocas Dhart T., et al., (2021)	22
Tabla 5. Resultados de investigación Trinh, G. M., et al., (2023) LumbarNet.....	23
Tabla 6. Tabla de limitaciones.....	32
Tabla 7. Base de Datos de Imágenes a utilizar.....	40
Tabla 8. Base de Datos BUU-LSpine.....	41
Tabla 9. Cronograma de actividades.....	51
Tabla 10. Matriz metodológica	53
Tabla 11. Primeras pruebas con base 1	55
Tabla 12. Comparativa de resultados de investigaciones de algoritmos de detección espinal	65

ÍNDICE DE ECUACIONES

Ecuación 1: Fórmula para el cálculo de precisión.	49
Ecuación 2: Fórmula para Exhaustividad.	49
Ecuación 3: Fórmula para Coeficiente de Intersección sobre Unión.....	50

ÍNDICE DE ANEXOS

Anexo 1 Etiquetas de esquinas de cada vértebra para entrenamiento	71
Anexo 2 Precisión promedio de validación prueba 1	72
Anexo 3 Pérdida de caja de Regulación de validación prueba 1	72
Anexo 4 Precisión de Keypoints de validación prueba 1.....	73

Anexo 5 Pérdida de detección de Keypoints de validación prueba 1	73
Anexo 6 Validación Pérdida de Clasificador prueba 1- validación.....	74
Anexo 7 Pérdida de cajas reguladoras KeyPoint -Entrenamiento	74
Anexo 8 Pérdida de Keypoints modelo KeyPoint- Entrenamiento	75
Anexo 9 Pérdida de Clasificador modelo KeyPoint- Entrenamiento	75
Anexo 10 Pérdida de total modelo KeyPoint- Entrenamiento.....	76
Anexo 11 Presición Promedio 25 Épocas- validación.....	76
Anexo 12 Pérdida detección Bounding boxes 25 Épocas - validación	77
Anexo 13 Presición de detección de Keypoints 25 Épocas- validación	77
Anexo 14 Pérdida de Keypoints 25 Épocas- validación	78
Anexo 15 Pérdida de Clasificador KeyPoint 25 Épocas- validación.....	78
Anexo 16 Pérdida de cajas reguladoras KeyPoint -Entrenamiento.....	79
Anexo 17 Pérdida de Keypoints modelo KeyPoint- Entrenamiento.....	79
Anexo 18 Pérdida de Keypoints modelo KeyPoint- Entrenamiento.....	80
Anexo 19 Pérdida total de modelo KeyPoint- Entrenamiento.....	80
Anexo 20 Precisión Promedio KeyPoint- Validación	81
Anexo 21 Pérdida de cajas reguladoras KeyPoint - Validación.....	81
Anexo 22 Presición de Keypoints KeyPoint- Validación	82
Anexo 23 Pérdida de Keypoints modelo KeyPoint- Validación	82
Anexo 24 Pérdida de Clasificador modelo KeyPoint- Validación	83
Anexo 25 Pérdida total de modelo KeyPoint- Validación	83
Anexo 26 Pérdida exhaustividad de modelo KeyPoint- Validación	84
Anexo 27 Pérdida con 20% de base de datos- entrenamiento	84

Anexo 28 Pérdida de bounding boxes con 20% de base de datos entrenamiento ...	85
Anexo 29 Pérdida de Keypoints con 20% de base de datos- entrenamiento	85
Anexo 30 Pérdida de clasificador con 20% de base de datos- entrenamiento.....	86
Anexo 31 Precisión promedio de bounding boxes de modelo Keypoint- Validación	86
Anexo 32 Precisión promedio de Keypoints de modelo Keypoint- Validación	87
Anexo 33 Pérdida de bounding boxes de modelo Keypoint- Validación	87
Anexo 34 Pérdida de Keypoints modelo Keypoint- Validación	88
Anexo 35 Exhaustividad de Keypoints de modelo Keypoint- Validación	88
Anexo 36 Exhaustividad de bounding boxes de modelo Keypoint- Validación	89
Anexo 37 Pérdida con 50% de base de datos- entrenamiento	89
Anexo 38 Pérdida de bounding boxes con 50% de base de datos entrenamiento ...	90
Anexo 39 Pérdida de Keypoints con 50% de base de datos- entrenamiento	90
Anexo 40 Pérdida de clasificador con 50% de base de datos- entrenamiento.....	91
Anexo 41 Precisión promedio de bounding boxes con 50% de base de datos - Validación	91
Anexo 42 Precisión promedio de Keypoints con 50% de base de datos - Validación	92
Anexo 43 Pérdida de bounding boxes con 50% de base de datos - Validación	92
Anexo 44 Pérdida con 50% de base de datos modelo Keypoint- Validación.....	93
Anexo 45 Pérdida de clasificador con 50% de base de datos modelo Keypoint- Validación	93
Anexo 46 Pérdida con 100% de base de datos- entrenamiento.....	94
Anexo 47 Pérdida de bounding boxes con 100% de base de datos entrenamiento ..	94
Anexo 48 Pérdida de Keypoints con 100% de base de datos- entrenamiento	95
Anexo 49 Pérdida de clasificador con 100% de base de datos- entrenamiento	95

Anexo 50 Precisión promedio de bounding boxes con 100% de base de datos - Validación	96
Anexo 51 Precisión promedio de Keypoints con 100% de base de datos - Validación	96
Anexo 52 Pérdida de bounding boxes de modelo Keypoint- Validación	97
Anexo 53 Pérdida con 100% de base de datos modelo Keypoint- Validación	97
Anexo 54 Pérdida de clasificador con 100% de base de datos modelo Keypoint- Validación	98
Anexo 55 Exhaustividad de Keypoints de modelo Keypoint- Validación	98

LISTA DE SIGLAS

AP	Precisión Promedio.
AP view	Anterior view.
LA view	Lateral view.
CNN	Convolutional Neural Network o Red neuronal convolucional.
ED	Espondilolistesis Degenerativa.
IA	Inteligencia Artificial.
ISS	Índice de similitud estructural.
IoU	Intersección sobre Unión.
LBP	Patrones binarios locales.
RCNN	Region Based Convolutional Neural Networks o redes neuronales convolucionales basadas en regiones.
RELU	Rectified Lineal Unit o Rectificador.
SD	Similitud Dice.
VGG	Grupos visuales de geometría.

I. INTRODUCCIÓN

La columna vertebral está conformada en la etapa adulta por 24 vértebras, el conjunto de las mismas es esencial para la estabilidad y el movimiento del cuerpo humano (Serrano Rivera, 2019). Los huesos de la columna tienen la función de proteger la médula espinal y el sistema de raíces nerviosas adyacentes a la misma. La médula espinal es parte del sistema nervioso central y su labor es el intercambio de información sensorial y de movimiento con el cerebro y el resto del cuerpo (C.Hannaford, 2009). Esta parte del cuerpo humano pasa por muchos cambios ya sea en la etapa de crecimiento como deformarse dependiendo del trabajo que el ser humano hace o la genética que este tenga, siendo factores para la evolución de enfermedades en la columna vertebral que se pueden investigar para evitar y dar diagnósticos acertados para dar el tratamiento correcto dependiendo de la afección. La espina dorsal es propensa a sufrir deformidades o alteraciones que tienden a manifestarse en pacientes de diferentes formas, dependientes de un número de factores de riesgo como la edad, traumatismos y patologías.

La columna vertebral también puede presentar afecciones relacionadas al dolor lumbar una de estas es conocida como espondilolistesis, la cual se caracteriza por consistir de un desfase o un deslizamiento de una vértebra de la espina dorsal por sobre la otra. Este desfase es la transposición de una vértebra sobre la inferior, sin un evento de trauma como una ruptura o algún defecto en el anillo vertebral. La espondilolistesis causa inestabilidad en el segmento de la vértebra, y comúnmente se clasifica bajo la categoría de "degenerativa". Esta condición es común en pacientes que sobrepasan los 50 años de edad y se presenta de forma más frecuente en pacientes del sexo femenino (Fraiwan, et al. ,2022).

Como aporte hacia el diagnóstico de deformidades dorsales, se propone la implementación de un algoritmo de inteligencia artificial que utiliza técnicas de transferencia profunda de modo asistencial al diagnóstico llevado a cabo por profesionales de la medicina que se especialicen en el área de estudio. La literatura disponible en la actualidad y los numerosos experimentos realizados en la línea de estudio sugieren la existencia de un enorme potencial en la inteligencia artificial como complemento a este tipo de diagnósticos utilizando métodos de procesamiento de imágenes como segmentación, detección de objetos, y separación de contrastes. En

comparación con otras industrias o áreas de trabajo, la adopción de la inteligencia artificial por parte del sector de salud ha sido relativamente lenta. Ambrosio-Ambrosio, et al. (2023), consideran que este fenómeno es peculiar, en vista de que la enorme complejidad que conllevan los procesos vinculados a la salud es precisamente lo que le otorga un gran potencial de compatibilidad a los algoritmos de aprendizaje automático.

La visión por computadora es un campo de estudio que se centra en desarrollar algoritmos y sistemas que permiten a las máquinas interpretar y comprender la información visual del mundo que las rodea (Dhar, T., et al., 2021) La visión por computadora es un campo de la informática que se ocupa del procesamiento de imágenes y la extracción de información de ellas.

Las redes neuronales convolucionales (CNN) son una herramienta derivada de la inteligencia artificial. Consisten en numerosas capas convolucionales capaces de aprender características locales de las imágenes que le son otorgadas, y generar clasificaciones en base a ello. Las arquitecturas de CNN han demostrado en la literatura, su capacidad para mitigar problemas que involucren sombras indeseadas, u objetos interpuestos en los estudios (Hernández, 2018).

El presente proyecto propone la mejora de un algoritmo de detección de objetos en imágenes de radiografía con la capacidad de etiquetar vértebras en las imágenes de pacientes de nuestras bases de datos en imágenes en plano sagital. Se entrenará al algoritmo para poder mejorar el porcentaje de precisión promedio en la detección en imágenes de los pacientes con espondilolistesis, utilizando una base de datos con radiografías descritas con etiquetado manual utilizando la herramienta de Roboflow para el entrenamiento y luego evaluadas con etiquetación semiautomático. Además de proponer este algoritmo, se dará aumento a la base de datos de diferentes deformidades en plano sagital con más imágenes rayos x de la afección para mejor desempeño de resultados del proyecto a presentar y futuros trabajos que se deseen realizar.

II. ESTADO DEL ARTE

2.1 Antecedentes

Las afecciones de la columna vertebral son un área de estudio amplio a analizar, dando lugar a brindar una herramienta clasificatoria y de detección de vértebras en imágenes en plano sagital. Según Freppel et al. (2008), la espondilolistesis se caracteriza por el deslizamiento de una vértebra y de la columna adyacente respecto a las vértebras distales. La espondilolistesis es una deformidad que tiene el potencial de interrumpir o entorpecer cada una de funciones de la columna vertebral. El tratamiento quirúrgico de esta deformidad tiene como objetivo corregir el fallo de estas funciones, pero es infrecuente y únicamente 10-20% de las espondilolistesis diagnosticadas concluyen en tratamiento operatorio. Es por esto que es favorable una evaluación clínica de un paciente que aún no padece de síntomas sin necesitar acudir a una cirugía.

En su análisis narrativo de espondilolistesis degenerativa, García Ramos, et al. (2020) describen los métodos de diagnóstico convencionales para la espondilolistesis, siendo el primordial elemento de apoyo para el diagnóstico de la deformidad la radiografía. Las radiografías laterales lumbares se utilizan para identificar el grado de listesis de la deformidad. Existe un método de clasificación para determinar el grado de desfase de la espondilolistesis que es el mayormente utilizado a nivel universal. Este es el método de Meyerding, se utiliza en radiografías laterales de la columna lumbar para medir el diámetro antero-posterior de la plataforma superior de la vértebra inferior. Los algoritmos de inteligencia artificial pasaron de ser técnicas de ensayo y error, a técnicas manuales y estructuradas, y finalmente técnicas automatizadas de aprendizaje. Aún entre los años 2015-2017 un gran número de estudios que utilizaban inteligencia artificial dependían de modelos de aprendizajes con supervisión, uno de ellos siendo las redes neuronales convolucionales (CNN). Existen dos factores claves mostrados en los modelos de trabajo que utilizan redes neuronales convolucionales. El primero es la disponibilidad de bases de datos, lo cual aún sigue limitando el avance en algunas líneas de estudio. El segundo es el avance por la capacidad de hardware en unidades de procesamiento gráfico (GPU). El resultado de estas mejoras es una mejor capacidad para entrenar redes neuronales y, por ende, incrementar su aplicabilidad en el ámbito de la salud.

El inicio de las CNN para el análisis de imágenes médicas tomó realce como en el trabajo de Krizhevsky, et al. (2012) donde realizaron un estudio detallado sobre el uso de redes neuronales convolucionales para la clasificación de imágenes en el conjunto de datos en el reto de reconocimiento visual en gran escala de Imagenet o, "Imagenet Large Scale Visual Recognition Challenge" (ILSVRC). Su objetivo era entrenar una red neural capaz de reconocer y clasificar las 1.2 millones de imágenes en alta resolución en las 1000 diferentes clases del conjunto de datos, como podemos ver en la tabla 1 cuenta con una tasa de error de 15.3%, con su segundo modelo tenía una tasa de error de casi el doble en 26.2%.

Relación	Tasa de error	Pérdida	Modelo	Exactitud
Red Neuronal	15.3%	2%	ImageNet	47.1%
Red Neuronal	26.2%	16.6%	Sparse coding	28.2%

Tabla 1. Investigación Krizhevsky, et al. (2012) ImageNet Classification with Deep Convolutional

Las unidades lineales rectificadas (RELU) según Repetur, et al. (2019), tienen la propiedad de no requerir normalización de entrada para evitar la saturación, lo cual significa que, si al menos algunos ejemplos de entrenamiento producen una entrada positiva a una neurona RELU, esta neurona aprenderá durante el entrenamiento. Esta es la razón de explicar por qué las redes neuronales convolucionales han alcanzado una posición predominante en la empleabilidad en el procesamiento y análisis de imágenes médicas.

Otra investigación fue realizada por Chen, K., et al. (2021), donde describen los aspectos en los cuales el aprendizaje automático ha avanzado en el área de estudio de escoliosis. Se describen los avances del aprendizaje automático como su evolución en el diagnóstico de la escoliosis. Según los autores, el diagnóstico automático de la escoliosis podría detectarla en una etapa más temprana de lo que se detecta clínicamente en la actualidad. La detección temprana de la escoliosis puede ser controlada con menores costos para el paciente y mejor eficacia, además de evitar errores de interpretación de parte de los profesionales. La importancia de desarrollar un procedimiento automático proviene de las deficiencias inherentes al cálculo manual del ángulo de Cobb. El procedimiento puede consumir tiempo y existe un margen de error de

cálculo natural al ser aun un proceso humano. La tabla 2 muestra los resultados del algoritmo realizado por Chen, K., et al. (2021), mostrando que su exactitud es de 69%-85% y con una tasa de error de detección de vértebras de 15.3% y resultados negativos de la detección de 26.2%, siendo comparado con otros trabajos como el de Alharbi, R. H., et al., (2020) que también consiguió desarrollar un algoritmo que calcule el ángulo de Cobb.

Relación	Tasa de error	Pérdida	Modelo	Exactitud
Enfermedad vertebral	15.3%	2%	ImageNet	*
Algoritmo	26.2%	16.6%	Sparse coding	69%-85%

Tabla 2. Investigación Kai Chen et al.,2021, El aprendizaje automático como revolución prometedora en la práctica clínica de la escoliosis.

Alharbi, R. H., et al., (2020) en su aporte con respecto a un algoritmo basado en el aprendizaje profundo para la medición automática del ángulo de Cobb, se apoyan de redes neuronales convolucionales (CNN) para la detección de las vértebras en las imágenes. Bajo clasificación de imágenes con CNN, una imagen se clasifica bajo una categoría predeterminada. Las redes utilizan el reconocimiento de patrones en general para su aprendizaje a la hora de clasificar datos. No es necesario realizar ningún tipo de extracción de características de forma manual. Esos avances dependen de aprendizaje de transferencia o "transfer learning" que utilizan métodos de aprendizaje automático para desarrollar un modelo y entrenar el mismo para una tarea específica.

En la investigación de Alharbi, R. H., et al., (2020) se utilizó una red neuronal convolucional pre-entrenada de aprendizaje profundo residual llamada ResNet. ResNet50 en esta investigación fue utilizado para la detección de vértebras. La idea detrás de utilizar este tipo de red es aprovecharse de una "conexión atajo de identidad" que se ahorra el uso de una capa o más. ResNet-50 es una CNN pre-entrenada por millones de imágenes de la base de datos de ImageNet. La red tiene 50 capas de profundidad y es capaz de clasificar imágenes en 1000 categorías.

Continuando con la exploración de la literatura actualmente disponible, el estudio de Caesarendra W., et al. (2022), de cálculo automático el ángulo de Cobb en pacientes de escoliosis idiopática

adolescente explora a mayor detalle la etapa de pre-procesamiento orientado al estandarizado de dimensiones de imagen y protección de privacidad de los pacientes en las radiografías. Se utilizaron un tipo de descriptor visual llamado "clasificador de cascada" y patrones binarios locales (LBP) para estandarizar las imágenes radiográficas y recortar automáticamente las imágenes entre la cervical y el sacro. El histograma del código de LBP contiene información acerca de la distribución de pixel para los contornos u otras características dentro de la imagen. El autor decidió hacer uso de estos patrones binarios ya que es viable para obtener información de radiografías que se encuentran en escala de grises por utilizar patrones derivados para encontrar dirección a partir de un gradiente binario. Se establecen umbrales para limitar las imágenes de modo que se incluya la misma porción de columna vertebral en todas las imágenes con ayuda del clasificador de cascada que define el área de objetos detectados.

Una de las limitantes con mayor prevalencia en las investigaciones de esta línea de estudio revuelve alrededor de la privacidad del paciente, y la capacidad de anonimizar imágenes radiográficas sin alterar la imagen de modo que pueda perjudicarse su utilidad para estudio. El sistema propuesto por Caesarendra W., et al. (2022), no solo beneficia la investigación de modelos y algoritmos asistidos con inteligencia artificial con respecto a la ética médica y las implicaciones, la implementación de la inteligencia artificial para el procesamiento de imágenes se agrega a ella, cumpliendo con la tarea de estandarizar dimensionamientos de imágenes lo cual nutre a los estudios que se realicen utilizando bases de datos de gran tamaño o múltiples bases de datos de diferentes orígenes.

En lo que respecta a espondilolistesis, un estudio de Saravagi D., et al. (2023) exploran el uso de un modelo híbrido de aprendizaje profundo para la predicción de espondilolistesis lumbar. Se utiliza un modelo integrado que resulta de una combinación entre el modelo VGG16 y la red CapsNet. Los grupos visuales de geometría (VGG), son herramientas conocidas de CNN utilizadas para solventar categorización de imágenes médicas (Saravagi D., et al. ,2023). El trabajo de ellos es un modelo robusto que tiene una arquitectura bidimensional para detección de características en diferentes posiciones de imágenes objetivo. El modelo VGG16 fue de los modelos con mejor desempeño que salieron al público. A pesar de que el VGG16 se adapta de forma ideal para la categorización visual, tiene las siguientes desventajas que contaron como limitantes en su estudio

como ignorar algunas estructuras locales entre objetos sencillos perdiendo la información de posición de la entidad al hacer agrupamiento.

La solución propuesta de Saravagi D., et al. (2023), para mitigar estas desventajas es el uso de una red de capsula (CapsNet). El algoritmo de CapsNet utiliza vectores, lo cual lo hace mucho más sensible que el VGG16 previamente mencionado. Este concepto codifica los datos por capsulas o grupos de neuronas. Compensa la pérdida de información en las capas de agrupamiento en las CNN. Además de esto se utilizó un algoritmo dinámico de rutas, una opción más efectiva que el agrupamiento máximo otorgando soporte solido a las capsulas que alcanzan la salida principal.

En el estudio Saravagi D., et al. (2023), obtuvieron una base de datos privada de 466 radiografías de la espina dorsal. Las imágenes fueron organizadas entre las categorías de saludable y afectado. Posteriormente se realizó una división para asignar algunas imágenes a los datos de entrenamiento y algunas a los datos de experimento. Se entrenó el diseño neuronal con núcleos bidimensionales y con un tamaño normalizado de imagen. Se utilizaron métricas de calificación de desempeño. Se evaluó la exactitud donde el número de predicciones correctas se divide entre el número total de predicción. La precisión evalúa la exactitud de una predicción. El retiro es la métrica de la identificación de la verdad absoluta. El F1 es una métrica de la media armónica de la precisión y el retiro (García Maya, A. F.,2017). Por último, están las métricas de sensibilidad y especificidad. Se comparó el desempeño del VGG16 de forma autónoma y el modelo híbrido con CapsNet denominado S-VCNet.

Una de los investigación más recientes es el de Klinwichit, et al. (2022), ellos compararon diferentes modelos para la detección de vértebras como lo podemos ver en la tabla 3 descritos sus resultados, con el fin de identificar espondilolistesis a través de imágenes de rayos x de diferentes pacientes, la cantidad de imágenes de su base de datos que fueron analizadas con todos los modelos fueron un total de 3600 imágenes. Su trabajo "Un conjunto de datos tailandés abierto de columna lumbar para la detección de espondilolistesis", es un trabajo basado en clasificación dependiendo del grado de gravedad según el la clasificación Meyerding. Como podemos ver en la tabla 3 se desglosa los resultados de los diferentes modelos comparados junto el desempeño que estos tuvieron en las diferentes pruebas del estudio.

Relación	IoU [0.5]	Pérdida	Modelo	Exactitud
Detección de vértebras	95.83	5.56%	YOLOv5	95.82%
Detección de vértebras	89.48	7.29%	MobileNetV1	81.91%
Detección de vértebras	91.79	6.37%	MobileNetV2	83.52%
Detección de vértebras	91.03	8.51%	ResNet50V1	81.97%
Detección de vértebras	89.26	5.86%	ResNet101V1	83.20%
Detección de vértebras	92.76	5.83%	ResNet152V1	82.80%
Detección de vértebras	92.39	24.38%	EfficientDet D0	77.25%
Detección de vértebras	92.79	71.82%	EfficientDet D1	75.63%

Tabla 3. Resultados de Investigación de Klinwicht, et al., (2022), BUU-LSPINE

Otro trabajo que también utilizó el modelo ResNet y VGG16 fue el de Dhar, T., et al., (2021), teniendo un trabajo experimental basado en la detección de objetos. Este toma relevancia para nuestro proyecto por el uso de sus épocas. En su entrenamiento hacen hincapié en que el cambio de cantidad de épocas afectó sus resultados teniendo como valores más acertados cuando tenían 100 épocas para no perder exactitud ni sensibilidad. Otro punto a tomar en cuenta era la variante de tiempo que tuvieron dependiendo de su modelo, teniendo mejores valores con el VGG16 y menor tiempo para dar resultados

Relación	Tiempo	Épocas	Modelo	Exactitud
Detección de vértebras	150 min	100	VGG16	94.13%
Detección de vértebras	202 min	100	ResNet	91%

Tabla 4. Detección de área pulmonar con 100 épocas Dhart T., et al., (2021)

Por último, para detectar el deslizamiento lumbar en la investigación de Trinh, G. M., et al. (2023) trata sobre un método para detectar el deslizamiento lumbar en imágenes de rayos X. El método utiliza una red neuronal convolucional profunda llamada LumbarNet. LumbarNet consta de dos rutas principales: una ruta espacial y una ruta de reducción de muestreo. La ruta espacial conserva la información espacial y genera características de alta resolución. La ruta de reducción de muestreo extrae características robustas de la imagen. Dentro de su preprocesado y selección de imágenes se necesitaba homogenizar todas las imágenes contando con imágenes de 512 x 512 x 3 para la entrada, al igual para la selección de imágenes debe de tener las imágenes un mínimo de 5 vértebras. Como podemos ver en la Tabla 5 como resumen, la detección de su región de interés con el modelo LumbarNet logra una tasa de precisión del 88.83%, tomando 312 imágenes de rayos x y obteniendo 250 imágenes de salida con esa tasa de precisión.

Tasa de error	Pérdida	Modelo	Exactitud	Sensibilidad	IoU
12.7%	2%	LumbarNet	88.83	91.24	81

Tabla 5. Resultados de investigación Trinh, G. M., et al., (2023) LumbarNet

2.2 Problemática

La detección de las afecciones de la columna vertebral, como el diagnóstico preciso de las deformidades, son procesos clínicos cruciales, pero a menudo están sujetos a la interpretación subjetiva de imágenes de rayos X por parte de profesionales de la salud. Esta subjetividad puede resultar en diagnósticos imprecisos o en la falta de detección de afecciones en etapas tempranas (Jaiswal, A., et al., (2021). Además, el proceso de clasificación y detección manual es laborioso y consume una cantidad significativa de tiempo, lo que puede retrasar el inicio del tratamiento y aumentar los costos médicos. Estas afecciones afectan la salud de millones de personas en todo el mundo y su diagnóstico temprano puede marcar la diferencia en términos de resultados clínicos y calidad de vida.

Actualmente, el proceso de detección de estas condiciones se basa en gran medida en la interpretación manual de imágenes de rayos X por parte de radiólogos y especialistas en ortopedia.

El progreso tecnológico en el campo de la salud, se abre la puerta a soluciones innovadoras. La aplicación de técnicas de aprendizaje profundo y la automatización en la anotación de imágenes de rayos X representan una respuesta que optimiza el proceso. Aunque ha habido avances significativos en la detección de escoliosis, aún persisten desafíos en la detección de espondilolistesis. La mayoría de deformidades en la columna vertebral pueden causar dolor, limitaciones físicas y afectar la movilidad de las personas que las padecen. Si estas afecciones se diagnostican de manera tardía o incorrecta, los pacientes pueden experimentar un deterioro en su salud y un empeoramiento de los síntomas. Es por ello que la posibilidad de automatizar la predicción de estas deformidades puede ser un favorecedor avance tecnológico y apoyo en diagnóstico médico, Además, la automatización de este proceso puede reducir significativamente el tiempo necesario para el diagnóstico, lo que a su vez puede mejorar la eficiencia de los sistemas de atención médica y permitir un tratamiento más oportuno.

Los algoritmos de aprendizaje profundo pueden aprender patrones y características sutiles en las imágenes de rayos X que a veces pueden pasar desapercibidos para el ojo humano. La aplicación de técnicas de aprendizaje profundo y la automatización de la anotación de imágenes de rayos X pueden ayudar a superar los desafíos asociados con la interpretación manual. Al usar algoritmos de inteligencia artificial, se pueden identificar patrones sutiles en las imágenes de rayos X que pueden no detectar los radiólogos o especialistas en ortopedia. En la actualidad, el tema de la detección automatizada de escoliosis ha tenido mayor avance para el cálculo del ángulo de Cobb, siendo un apoyo para nuevas investigaciones de escoliosis, pero para lo que son enfermedades degenerativas donde su detección son imágenes de rayos x en plano sagital, aún no ha experimentado un progreso tecnológico.

Por consiguiente, es importante asegurar el desarrollo, implementación y efectividad de esta clase de algoritmo. La exactitud de este algoritmo es necesario para que los resultados derivados de su funcionamiento no sean inestimables en la detección temprana de las deformidades aludidas. Además, la continua optimización de dicho algoritmo sea para tener una

herramienta innovadora dentro del ámbito de la salud, lo cual redundaría en beneficio palpable para la población aquejada por estas afecciones de la columna vertebral. Su desarrollo y aplicación se perfilan como un recurso de asistencia y como una herramienta de detección para los profesionales especializados en neurocirugía dando una detección a través de puntos claves de las vértebras de radiografías en plano sagital.

El material científico relacionado a deformidades de la columna vertebral como la espondilolistesis en Honduras, se basa únicamente en la revisión de casos, brindando como contribución a la Sociedad de Neurocirujanos de Honduras ,una herramienta para facilitar el proceso de detección de vértebras, al igual de una de base de datos recolectada que cumple con el criterio de ser imágenes rayos x en plano sagital, para el desarrollo de un algoritmo en base de un modelo ya existente que nos pueda brindar la detección de las vértebras para la verificación de la afecciones de la columna vertebral.

2.3 Imagen integradora

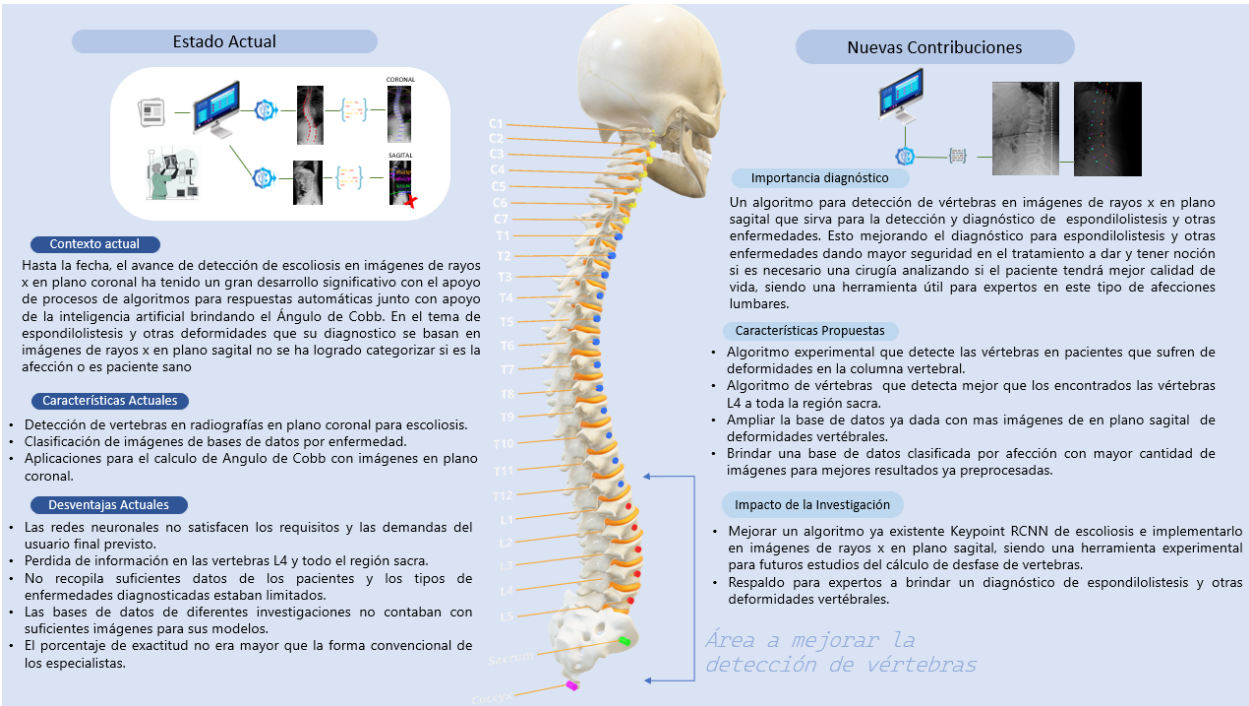


Ilustración 1. Imagen integradora

Fuente: Elaboración de Ilustración propia.

La imagen integradora se encuentra dividida en dos segmentos principales: el "Estado actual" que trata de cómo ha tenido alcance el tema de algoritmos que calculen el ángulo de Cobb para escoliosis y lo poco que hay del cálculo de desfase de vértebras en pacientes con espondilolistesis y otras enfermedades que se detectan con imágenes de rayos x en plano sagital. En las "Nuevas contribuciones" está en forma de resumen la investigación actual, buscando aportar en la detección de vértebras, enfocándonos en mejorar el cálculo de detección de estas a través de KeyPoints, es un área que necesita avances según la literatura, para obtener mejores resultados en imágenes de rayos x que están en plano sagital.

En el lado izquierdo se desglosa el contexto de algoritmos de cálculo de enfermedades vertebrales en la actualidad, donde se obtuvo mayor información acerca del cálculo de escoliosis. Se brindan los logros hasta la fecha, como el uso de la inteligencia artificial para detección de objetos y su porcentaje de exactitud, se mencionan algunas de las limitaciones que tuvieron otros investigadores en este tema como la ineficiencia en algoritmos que cumplieran con una exactitud

mayor en el cálculo que los expertos, esto debido a la escases de imágenes de acceso libre para entrenar mejor sus algoritmos. Se resalta que las soluciones actuales en escoliosis para el cálculo de Cobb lograron tener una aplicación mientras que en espondilolistesis no se ha tenido un avance notorio actualmente.

En el lado derecho se enfoca la importancia del diagnóstico correcto para afecciones de la columna vertebral y las mejoras propuestas por esta investigación, colocando como enfoque el desarrollo del algoritmo para detección de vértebras en radiografías en plano sagital. Se destaca la relevancia de una base de datos más abundante de imágenes de ambas enfermedades para mejorar la calidad de resultados al igual que optimizar un algoritmo que identificar las vértebras que son difíciles de detectar en otros algoritmos (de la L4 a las parte sacra). Las características propuestas incluyen un algoritmo para el procesado, una base de datos con mayor cantidad de imágenes de diferentes deformidades en plano sagital y ya pre procesada con el fin de tener un algoritmo que logre analizar las vértebras basándose en el algoritmo existente de cálculo de ángulo de Cobb, pero aplicándolo para la detección de vértebras en imágenes de rayos x en plano sagital.

2.4 Tabla de limitaciones

Autores	País del estudio	Enfoque de estudio	Tipo de limitación	Detalle de limitación
Fraiwan et al., 2022	Estados Unidos de América	Comparativa de redes neuronales clasificatorias.	Precisión, uso móvil y uso de usuario final.	Carencia en la implementación de redes neuronales en aplicaciones móviles destinadas a profesionales de la salud, lo que resulta en una falta de accesibilidad y facilidad de uso. En la actualidad, las redes neuronales no satisfacen los requisitos y las demandas del usuario final previsto.
Jain, B. Sharma et al., 2023	Madhya Pradesh, India	Propuesta de modelo combinando red neuronal convolucional y la red de cápsulas.	Detección de espondilolistesis	La incapacidad de clasificar con precisión algunos en la detección de todas las vértebras en casos de espondilolistesis como negativos o viceversa, lo que indica la necesidad de investigaciones adicionales en el diagnóstico de la espondilolistesis lumbar y mejorar la detección de vértebras.
Yongcheng Tu, et al., 2019.	No identificado	Propuesta de algoritmo automático de medición del ángulo de Cobb en la escoliosis.	Inestabilidad en la detección del ángulo.	La segmentación de imágenes de rayos X de la columna vertebral se complicó. El algoritmo propuesto utiliza una combinación de algoritmos de detección de objetivos y segmentación de imágenes llamada DU-Net, hubo

				problemas con la identificación de las vértebras.
Junbo Xuan, et al., 2023	Guilin, Guangxi, China	Desarrollo el software de diagnóstico de IA y proporcionar una cómoda interfaz de usuario.	Ejecución con menor exactitud	La investigación no recopila suficientes datos de los pacientes y los tipos de enfermedades diagnosticadas estaban limitados.
Justin Ker et al., 2018	Estados Unidos de América	Aplicar algoritmo a análisis de imágenes médicas, enfocándose en redes neuronales convolucionales.	Base de datos incompleta.	La cantidad de datos disponibles por cada condición patológica no es suficiente. Aún no se ha establecido un valor definitivo para el número óptimo de imágenes requeridas para lograr un entrenamiento exhaustivo de estos modelos. El desafío de la disponibilidad de imágenes debido a la aprehensión de algunos pacientes en permitir su uso
Nan Meng, et al., 2022	No identificado	Desarrollar y validar un modelo híbrido para la interpretar la alineación de las vértebras de	Identificación para alineación automática.	SpineHRNet+ utiliza algoritmos de detección y segmentación de landmarks para visualizar los puntos de referencia en las imágenes radiográficas de la columna vertebral. Si la imagen contaba con

		forma automática		problemas de contraste tenía mayor margen de error.
Ajay Jaiswal, et al.,2021	No identificado	Clasificación de informes de radiología.	Herramientas y computadora de alto desempeño.	Hablan de que BERT es un modelo de lenguaje grande y complejo que requiere una gran cantidad de recursos computacionales y memoria para su entrenamiento y uso, esto lo hace costoso y difícil de implementar en entornos con restricciones de recursos, si contamos con la base de datos de radiología, sus imágenes se necesitan procesar con el modelo.
Saechueng, et al., (2023, June).	Taiwan	Detección de vértebras	Desempeño de modelo arquitectura U-Net.	La arquitectura U-Net es que no admite ninguna conexión de capas. Además, sólo se utiliza la parte de verificación de cada convolución, lo que permite el mapa de segmentación para contener cada píxel de la imagen de entrada en su contexto completo.
Cai, Y., et al., (2017).	No identificado	Detección de espondilolistesis a partir de escaneos reales de resonancia magnética/tomografía	Modelo de detección de lumbar	El modelo no especificó la etiología y el grado de deslizamiento de la espondilolistesis. Los resultados finales son inconclusos, se logra solo hacer un test de validación con nueve imágenes MR +CT. De las nueve imágenes logran tener una

		computarizada (RM/TC).		sensibilidad de 91% en el entrenamiento.
Nguyen, T. P., et al., (2021).	Korea	Desarrollo de sistema de aprendizaje profundo en el que se entrena el modelo CNN suplementario para corregir los puntos clave ubicados en las esquinas de las vértebras.	Pérdidas respecto al rendimiento del método para la detección de puntos clave.	Los requisitos para construir el conjunto de datos que coincida con los requisitos de detección de cada modelo individual. Tiempo de entrenamiento para cada modelo. Algunos tipos de deformación de la escoliosis fueron menos explicados por nuestra base de datos. Falta de validación estadística del rendimiento del método propuesto con respecto a otras enfermedades lumbares La confiabilidad del rendimiento del programa con respecto a estas enfermedades mencionadas necesita ser mejorada y validada.
(Trinh, et al., 2023)	Taiwan	Detección de espondilolistesis	Desarrollo de modelo de detección de lumbar LumbarNet	El modelo no especificó la etiología y el grado de deslizamiento de la espondilolistesis. La presencia de osteoporosis afecta los resultados de la segmentación debido a la baja calidad de la imagen. No se logró identificar la quinta vértebra lumbar (L5) y el primer sacro (S1) en comparación con las otras vértebras lumbares debido a su superposición

con la pelvis o la presencia de sacralización lumbar. Además, los osteofitos vertebrales (espuelas óseas) introducían errores en el ajuste cuadrilateral ya que no representan las esquinas reales de las vértebras, si las esquinas no pueden detectarse correctamente, todas las evaluaciones fallan.

Tabla 6. *Tabla de limitaciones*

Fuente: Elaboración de tabla propia

III. OBJETIVO

3.1. Objetivo General

Efectuar un algoritmo de aprendizaje profundo computarizado capaz de automatizar la detección de vértebras en imágenes de rayos X tomadas en el plano sagital de pacientes con deformidades en la región de la espina dorsal.

3.2. Objetivos específicos

- Seleccionar y clasificar las imágenes para el entrenamiento del modelo y separar los sets de entrenamiento y validación.
- Etiquetar las imágenes de acuerdo al formato de Keypoint-RCNN para la detección de las vértebras.
- Modificar parámetros de los modelos existentes para la creación de un modelo entrenado capaz de realizar predicciones de vértebras en imágenes de rayos x en plano sagital.
- Evaluar el rendimiento del algoritmo de detección de vértebras, y su viabilidad al realizar inferencia de predicciones sobre imágenes desconocidas de la espina dorsal.

IV. MÉTODO

En el capítulo de metodología se exponen los métodos utilizados para el desarrollo de la presente investigación. En este capítulo se definen y establecen el enfoque de la investigación, su alcance, el diseño de este, y las variables dependientes e independientes de la investigación. De igual forma se exponen los métodos empleados para llegar a cabo el alcance de los objetivos.

4.1 ENFOQUE

El enfoque de la presente investigación es cuantitativo ya que se realizan en ella mediciones utilizando métricas numéricas de desempeño para evaluar la efectividad del entrenamiento del algoritmo y por consecuencia, la capacidad del mismo para hacer predicciones satisfactorias. Estas métricas incluyen la exhaustividad, la precisión promedio del coeficiente de intersección sobre unión, mejor conocido en inglés como Intersection over Union (IoU), y la pérdida del procesamiento de las imágenes brindadas para la detección de vértebras. La comparación final incluye los valores de dichas métricas como calificaciones de la validación del algoritmo posterior a su entrenamiento, por lo cual cae bajo la categoría de cuantificación de valores.

El diseño de esta investigación es un diseño experimental, provocando un cambio en las variables de salida mientras se sostiene un control total sobre las variables de entrada. Un algoritmo de aprendizaje profundo es entrenado y algunas de sus variables son modificadas para verificar si las modificaciones producen algún tipo de mejora en el modelo.

La presente investigación es de alcance exploratorio debido a la escasez de estudios de la misma línea de investigación en Honduras y en la región centroamericana. En el contexto nacional para Honduras, la inteligencia artificial aplicada al sector de la salud es un tema especialmente novedoso. La viabilidad de este tipo de modelos de inteligencia artificial en el ámbito médico es un tópico que se encuentra aún en una etapa de infancia. Esto resulta en el brote de incógnitas sin respuesta en la materia de estudio, lo que es considerado un efecto natural con temas con tal grado de novedad. Establecer esta aplicabilidad del aprendizaje profundo a la salud como línea para futuras investigaciones representaría un enorme paso para la medicina de Honduras en general, y también para el diagnóstico y detección de deformidades en la espina dorsal a nivel mundial.

4.2 VARIABES DE INVESTIGACION

Según Sampieri et al., (2014), variables son cualidades que son cambiantes y pueden observarse, monitorearse, o medirse. La relación que múltiples variables tienen entre sí dan forma a lo que es conocido como una teoría, lo cual asigna importancia a cada una de esas variables en lo que respecta a una investigación. Usualmente se les denomina constructos hipotéticos.

4.2.1 Variables Dependientes

Esta investigación toma las métricas de desempeño utilizadas, cuyo estado de variación será dependiente del entrenamiento del algoritmo de aprendizaje profundo orientado al procesamiento de imágenes y la detección de vértebras en radiografías de deformidades lumbares en plano sagital. Las métricas a utilizar en la investigación incluyen las siguientes:

- Precisión promedio del coeficiente IoU (Intersection over Union): Precisión se refiere a la fracción de instancias recuperadas que son relevantes. El parámetro medido en precisión promedio es el coeficiente IoU, una métrica utilizada comúnmente para evaluar modelos de aprendizaje profundo (Tu Y. et al., 2019). Es una relación entre la intersección y la unión entre un cuadro delimitador predicho en la detección de objetos, y el cuadro delimitador real del objeto. (Abdullah et al., 2018).
- Exhaustividad: Se refiere a la fracción de instancias relevantes que han sido recuperadas. (Buckland, M & Gey, F., 1994).
- Pérdida: evalúa la discrepancia entre las predicciones del modelo y las etiquetas (ground truth) reales en el conjunto de datos. El objetivo durante el entrenamiento del modelo es minimizar esta función de pérdida. Un valor bajo de pérdida indica que el modelo está haciendo predicciones cercanas a las etiquetas reales.



Ilustración 2. Variables Dependientes

Fuente: Elaboración de Ilustración propia

4.2.2 Variables independientes

Las variables independientes son las variables que no presentarán ningún cambio ni se verán afectadas, aun habiendo cambios en el algoritmo como tal. Estas por definición están compuestas por factores que en torno afectan a las variables dependientes. Las variables dependientes son de la investigación incluyen:

- Modelo pre-entrenado de aprendizaje profundo: La red neuronal de inteligencia artificial aplicada para realizar tareas de aprendizaje profundo. Esta red neuronal es entrenada previamente para ser capaz de cumplir objetivos como procesar imágenes, detectar elementos y realizar cálculos.
- Épocas: La época es la cantidad de veces que el algoritmo aprende de todos los datos de ejemplo durante la etapa de su entrenamiento.
- Tamaño de lote: Es la cantidad de ejemplos de entrenamiento que actualizan los pesos en el algoritmo en cada paso de la fase de entrenamiento.



Ilustración 3. Variables Independientes

Fuente: Elaboración de Ilustración Propia.

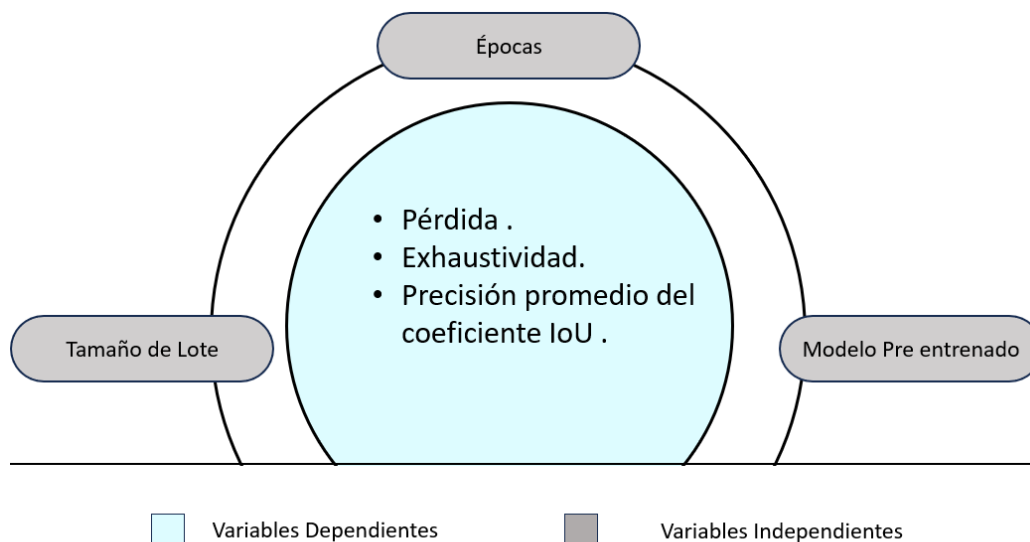


Ilustración 4. Variables de la investigación

Fuente: Elaboración de Ilustración Propia.

4.3 Técnicas e Instrumentos Aplicados

4.3.1 Técnicas

En el marco de la presente investigación, se aplicaron algoritmos donde se llevó a cabo la detección de objetos o regiones de interés en la forma de vértebras. Para poder detectar cada una de las vértebras de la parte lumbar, es necesario implementar un entrenamiento un modelo de detección de objetos, que le permita aprender a identificar las regiones de interés en imágenes de rayos x en plano sagital representando las vértebras, y los puntos específicos de interés que representarían las esquinas de cada vertebra. Este enfoque se apoya en la adaptación de un código previamente adquirido a través de la profunda exploración de la literatura con relación a la línea de estudio del presente proyecto. Dicho código establece un modelo de redes neuronales convolucionales basadas en regiones mejor conocida por sus siglas en inglés R-CNN, llamado Keypoint RCNN. Keypoint RCNN es un modelo pre-entrenado de aprendizaje profundo que utiliza R-CNN para llevar a cabo la detección de puntos clave y cuadros delimitadores de regiones de interés simultáneamente. El modelo fue desarrollado,

entrenado y validado utilizando un cuaderno Jupyter de Google Colaboratory. Para su diseño, se utilizaron las librerías de Pytorch, Pandas, Numpy, Matplotlib, OpenCV, Pycocotools, y Pillow en Google Colaboratory. La arquitectura del modelo está estructurada de la siguiente forma. El proceso consta del ingreso de una imagen de Rayos X de la espina dorsal a la base de datos. La imagen es sometida a una base de datos en Roboflow, en la cual se realiza un etiquetado manual de la verdad absoluta para las ubicaciones de los puntos clave que simbolizan las diferentes esquinas de las vértebras en la imagen. Estas etiquetas sirven para la detección de la ubicación de las vértebras completas y sus cajas delimitadoras que las encapsula al pasar las imágenes etiquetadas por el cuaderno de pre-procesamiento en Google Colab. Cada una de las imágenes a utilizar en la base de datos designada a entrenamiento necesita esta característica para ser procesada.

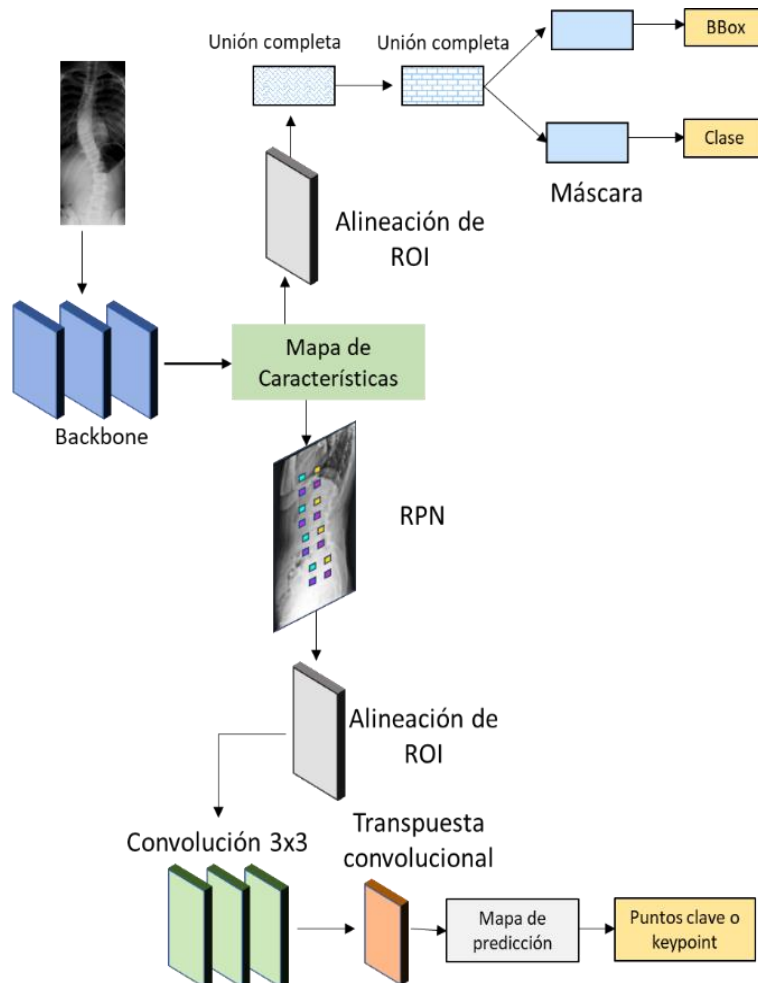


Ilustración 5. Estructura de modelo KeyPoint RCNN

Fuente: Basado en estructura explicada por Imran, et al., 2020.

4.3.2 Herramientas

- Roboflow

Dentro del conjunto de herramientas utilizadas en este proyecto se encuentran diversas plataformas empleadas para llevar a cabo el preprocesado para apoyo a las pruebas y validación del algoritmo. Entre estas herramientas, destaca Roboflow, una plataforma especializada en el ámbito del aprendizaje automático y la inteligencia artificial. Su misión principal es proporcionar ayuda a etiquetar las imágenes de un conjunto de datos de manera rápida e intuitiva y ayuda a generar el conjunto de datos en un formato que pueda ser leído por otras plataformas como TensorFlow.

- Colaboratory Python

Los algoritmos y modelos a utilizar están realizados en lenguaje Python. Google Colaboratory es una herramienta en línea que permite ejecutar código Python en la nube utilizando recursos de Google, lo que incluye acceso a unidades de procesamiento gráfico (GPU) y unidades de procesamiento tensorial (TPU) para acelerar tareas de aprendizaje automático. Es una elección popular para ejecutar modelos de detección de imágenes debido a su opción de uso de GPU's públicos compartidos en línea, permitiendo los beneficios que estos brindan a tareas de inteligencia artificial de forma accesible. Esta herramienta fue escogida para trabajo colaborativo y la optimización del entrenamiento ya que tanto las GPUs como las TPUs son unidades de procesamiento utilizadas para acelerar operaciones computacionales intensivas, pero las TPUs están diseñadas específicamente para el procesamiento de tensores y, en muchos casos, pueden superar a los GPUs en términos de velocidad y eficiencia en tareas de aprendizaje automático. La elección entre GPU y TPU dependerá de las necesidades específicas del proyecto y los recursos disponibles. Otro factor relevante para la elección de esta herramienta fue la facilidad de distribución de información, permitiendo facilidad al editar cuadernos de programación entre más de una persona.

4.4 Materiales

Los recursos fundamentales en el contexto de esta investigación consistieron en las bases de datos seleccionadas. Estas bases de datos albergan imágenes de rayos X de la columna vertebral, capturando tanto a pacientes que exhiben deformidades vertebrales, tales como espondilolistesis, como a aquellos individuos que mantienen la integridad de su columna vertebral, es decir, carecen de deformidades o patologías detectables en dicha región anatómica.

- Base de datos

El conjunto de datos para entrenamiento del modelo a utilizar es de la base de datos pacientes hondureños espondilolistesis, esta base solo consta con imágenes de esta afección siendo de diferentes pacientes y distinta cantidad de vértebras, el criterio a ser utilizado es que sean en plano sagital para cumplir con el objetivo de la detección de las vértebras lumbares en plano sagital, sin importar la cantidad de vértebras en la radiografía.

Base de datos	Año	Tipo de estudio	Rotulado	# Imágenes
BUU-LSPINE ¹	2019-2023	AP view y LA view	Vértebra Lumbar	308
Base de Datos pacientes hondureños espondilolistesis	Dif años	Clasificación de afección de vértebras	Posición de la vértebras	197
Dos bases de datos	Rango >2019	Columna Vértebral	Múltiples anomalidades	505

Tabla 7. Base de Datos de Imágenes a utilizar

¹ <https://services.informatics.buu.ac.th/spine/>

Fuente: Elaboración de tabla Propia

No.	Campo	Descripción
1	Imagen	Tipo de imagen: X-ray (PF)
2	Vista	AP y LA
3	Parte del cuerpo	Lumbar spine (LSPINE)
4	Total de pacientes	3600
5	Total de imágenes	7200
6	Pacientes con spondylolisthesis	621
7	Instancias de spondylolisthesis	788 (algunos pacientes incluyen múltiples instancias)
8	Tamaño del almacenamiento de los datasets	18.5 GB
9	Ground truth	Vertebral position, Spondylolisthesis diagnosis, LSTV
10	Objetivos de la posición vertebral lumbar	Las coordenadas de los puntos de esquina de L1 a L5
11	Objetivos del diagnóstico de spondylolisthesis	Left laterolisthesis
		Right laterolisthesis
		Anterolisthesis
		Retrolisthesis
12	Objetivos de LSTV	Sacralization, Lumbarization
13	Tipos de archivos	Ground truth: .csv, Imágenes: .jpg
14	Ubicaciones	Burapha University Hospital, Thailand
15	Período de años	2000-2021
16	Rango de edad	6-97 años
17	Dimensiones de la imagen	Altura: 1434-3408 pixels
		Ancho: 860-3040 pixels
		Espacio de píxeles: 0,125-0,175 mm
18	Parámetros de imagen de rayos X	AP: kV: 70-75, mAs: 32-40
		LA: kV: 80-90, mAs: 50-60

Tabla 8. Base de Datos BUU-LSpine

Fuente: DataSet informatics BUU-LSpine

El conjunto de imágenes del BUU-LSPINE consiste en una colección de imágenes de rayos X de la columna vertebral lumbar de 3600 pacientes anónimos, divididas en secciones como



(A) Anteroposterior

(B) Lateral

Ilustración 6. Referencia de imágenes de BUU-LSpine

vemos de forma detallada en la Tabla 8. Cada registro de paciente incluye dos vistas: anteroposterior (AP) y lateral (LA), presentadas en la Ilustración 6.

El conjunto de datos abarca 2700 imágenes. Es importante destacar que todas las imágenes laterales fueron invertidas para garantizar una orientación uniforme hacia el lado izquierdo. Cada imagen se diagnosticó con 4 trastornos de la columna que incluyen anterolistesis, retrolistesis, laterolistesis izquierda y laterolistesis derecha. El diagnóstico lo hicieron los médicos. Todas las imágenes de la vista LA en este conjunto de datos son vistas laterales izquierdas (el paciente mira hacia el lado izquierdo de la imagen) y todas las imágenes de las vistas laterales derechas (el paciente mira hacia el lado derecho de la imagen) están volteadas a la vista lateral izquierda.

4.5 Metodología de Estudio

Según el objetivo de la presente investigación, la metodología de la misma es de tipo exploratorio. Este tipo de metodología se utiliza para definir el alcance que puede tener el estudio. Cada etapa o avance implica la añadidura de descubrimientos cada vez más complejas al sistema, eventualmente aproximándolo a su etapa final de desarrollo. Se centró en explorar el potencial del uso de un algoritmo desarrollado para el entrenamiento de la visión de computadora para obtener la detección de vértebras. Se utilizará un modelo Keypoint-RCNN, un modelo de visión de computadora de detección de objetos que identifica puntos exactos relevantes en las regiones de interés denominados 'keypoints' y áreas de mayor espacio de interés denominados cuadros delimitadores o 'bounding boxes'. Este tipo de modelo se implementa con la finalidad de lograr la detección de cada vértebra en una

imagen de la espina dorsal de rayos X, lo cual puede visualizarse en el [Error! No se encuentra el origen de la referencia.](#) El proceso de entrenamiento, evaluación y consecuente mejora del algoritmo consta de varias etapas fundamentales, que se describen a continuación:

El primer paso de la parte exploratoria planteada consiste en ampliar la base de datos existente recopilada con ayuda de la Sociedad Hondureña de Neurocirugía, el Hospital Dr. Mario Catarino Rivas, el Hospital del Área de Puerto Cortés y el Hospital General El Progreso con la obtenido de BUU-LSpine.

Lo siguiente consiste en la prueba de datos y entrenamiento del algoritmo propuesto, esto incluye un método de pre-procesamiento para establecer la verdad absoluta de la información de las vértebras en las imágenes recolectadas. Este pre-procesamiento, consiste dos etapas: el etiquetado de los cuadros delimitadores y los puntos clave en las imágenes obtenidas en formato YoLoV5, y la extracción de dicha información en un formato compatible con el entrenamiento propuesto, siendo KeyPoint RCNN. El etiquetado se llevará a cabo en la aplicación Roboflow, facilitando la tarea y la extracción de las características de los puntos clave. La extracción de las anotaciones de puntos clave y cajas delimitadoras consiste en un cuaderno de programación en Google Colaboratory para interpretar las etiquetas realizadas en Roboflow y extraídas en formato YOLO V5, al formato KeyPoint RCNN para su uso en el entrenamiento. Posteriormente un entrenamiento es requerido para cumplir la función de ajuste fino del modelo de modo que este pueda en primera instancia aprender los detalles y características específicas de las muestras de pacientes con deformidades lumbares mediante la manipulación de las variables independientes. Al obtener esto se hace la inferencia del modelo, llevando a cabo la detección de vértebras en todas las imágenes posibles en plano sagital. Por medio de la manipulación de variables dependientes en el entrenamiento se pretende reflejar variaciones en las capacidades del algoritmo para realizar la detección, siendo posible optimizar su desempeño. El último paso consiste en la inferencia del algoritmo finalizado en imágenes de la base de datos que sean desconocidas al modelo en su etapa de entrenamiento. Al lograr una inferencia que pueda considerarse exitosa, se pretende realizar un etiquetado semi-supervisado dando anotaciones en todas las imágenes nuevas para el modelo de forma simultánea.

4.5.1 Criterios para pruebas

- Precisión de la detección de vértebras a través de la precisión promedio del coeficiente IoU, se presenta como una herramienta en la evaluación del desempeño de un modelo, al permitirnos visualizar la tasa de verdaderos positivos frente a la tasa de falsos positivos en diversos umbrales de decisión. A través de esta representación gráfica, se obtiene una apreciación detallada de cómo el modelo se desenvuelve en la tarea de clasificación.

La pérdida de clasificador, por su parte, se erige como un indicador de considerable relevancia que resume la capacidad de discriminación del modelo en un solo valor (Gurei, E. 2023). Esta medida sintetiza la eficacia del modelo en términos de distinguir entre las clases de interés, donde un valor más bajo de pérdida sugiere una mayor habilidad para realizar dicha discriminación.

- El tiempo de procesamiento, se debe de considerar que el procesamiento se torna imprescindible en la evaluación de un modelo, dado que este factor determina la eficiencia del sistema en la realización de predicciones. La eficiencia adquiere una dimensión crítica en aplicaciones clínicas que requieren respuestas en tiempo real, donde un modelo capaz de ofrecer resultados de manera expedita se convierte en una elección idónea.
- La capacidad de generalización representa una dimensión crucial de la evaluación de un modelo, esta faceta se refiere a la habilidad del modelo para extender sus predicciones a diferentes conjuntos de datos o escenarios clínicos. La evaluación de la generalización resulta esencial para verificar si el modelo puede aplicarse con éxito en diversas situaciones clínicas, lo que fundamenta su utilidad y robustez.
- En el caso para la aplicación clínica, es imperativo evaluar si el modelo posee la capacidad de mejora continua a medida que se le suministran nuevos datos y actualizaciones. Esta característica se traduce en una ventaja significativa en entornos clínicos dinámicos, donde la información evoluciona con el tiempo.
- Garantizar la seguridad y privacidad de los datos del paciente representa un principio cardinal en entornos clínicos. En consecuencia, es vital que se verifiquen

y cumplan rigurosamente las normativas y prácticas pertinentes (Fernández,2020). El escrutinio y la adopción de medidas adecuadas en torno a la seguridad y privacidad de los datos es necesario en la evaluación de modelos en aplicaciones clínicas, con el objetivo de preservar la integridad y confidencialidad de la información sensible del paciente.

4.5.2 Resumen de Procedimiento

El procedimiento del cálculo de estas imágenes va en base al código obtenido de Scoliovis, este código ya viene hasta con una interfaz de usuario para los expertos que necesiten calcular el ángulo de Cobb. Para la utilización y la implementación de pre procesamiento de esta ocupamos apoyo de la herramienta Roboflow para la etiquetación de las vértebras y la creación de keypoints así tener mejor estructurado lo que es nuestro ROI. Contamos con un total de 197 imágenes completas con el análisis realizado en el pre entrenamiento y los datos resultantes de estos obtenidos y tabulados

Para la utilización del modelo Keypoint RCNN que es capaz de realizar la detección de puntos clave y detección de objetos simultáneamente, es necesario hacer un pre procesamiento con la herramienta de Roboflow. En la herramienta Roboflow se colocaron cajas delimitadoras, con el fin de marcar la ubicación de las esquinas para cada vértebra. Al tener etiquetadas todas las imágenes designadas al entrenamiento, se extrajo de este programa el conjunto de datos en formato YoLoV5 donde se transformaría para lograr que pueda utilizarse en el entrenamiento del modelo de KeyPoint RCNN en Colaboratory. El cuaderno de pre-procesamiento cumple tres funciones fundamentales para hacer posible el uso de la base de datos. La primera función es interpretar las marcas realizadas en Roboflow como puntos clave. La aplicación Roboflow únicamente permite el etiquetado de cajas delimitadoras en las imágenes, por lo cual es necesario interpretar los centros de cada caja hecha, como puntos clave.

La segunda es calcular las coordenadas de las verdaderas cajas delimitadoras que encapsularían cada vértebra, en base a los nuevos puntos clave en la imagen. Se toman los valores de los cuatro puntos clave para cada una de las vértebras, y se agrupan por clase. Las

etiquetas realizadas en Roboflow fueron clasificadas, de modo que cada esquina para las vértebras fuera una clase, resultando en cuatro clases. De este modo, el pre-procesamiento tendría un parámetro para diferenciar cuales marcas representaban cada una de las esquinas. Al tener los puntos clave agrupados por clase, se organizaron según su coordenada en Y o en X , siendo el primer punto clave cuando la coordenada en Y y en X es la menor dentro de la imagen. El segundo, cuando X tiene el mayor valor, pero Y el menor valor, la tercera cuando Y es mayor y X es menor, siendo el cuarto punto clave cuando Y tiene el valor mayor en X y en Y . Esto se hizo con el fin de detectar la cantidad de vértebras en la imagen (habiendo un puesto por vértebra, deduciéndose como la cantidad de "esquinas" para cada clase en la imagen). Según esta jerarquía, los cuatro primeros lugares se interpretarían como las cuatro esquinas de la vértebra superior marcada en la imagen, los segundos lugares se interpretarían como la siguiente vértebra consecutiva, realizando esto consecutivamente para todas las instancias dentro de la imagen reagrupando así todos los puntos clave marcados correctamente según su vértebra. Luego se calculan los valores máximos y mínimos en las coordenadas X y Y para esos puntos como grupo, y se determinan así las coordenadas de las cajas delimitadoras que encapsulan cada vértebra.

La tercera tarea que hace el preprocesamiento es cambiar el formato de las anotaciones. Estas se encuentran en formato YOLO V5, y se transforman a formato KeyPoint RCNN, de modo que puedan utilizarse en el entrenamiento. El cuaderno de programación en Colaboratory, permite extraer esta información final en formato correcto, en archivos tipo ".json". La información guardada para la base de datos asigna nombres para los archivos de las anotaciones, que coinciden con los nombres de archivo de las imágenes. Es por esta coincidencia de nombres que el entrenamiento determina cuales anotaciones son para cada imagen. Una vez se cuenta con la verdad absoluta en ese formato, se ubica en las carpetas correspondientes en Google Drive, y es posible realizar el entrenamiento.

Se realiza el desarrollo y entrenamiento del modelo KeyPoint RCNN. El cuaderno de entrenamiento utilizado otorga métricas de validación para cada época de entrenamiento, dentro de las cuales están incluidas las variables a evaluar para esta investigación: precisión promedio del IoU, exhaustividad, y pérdida. Es posible extraer en texto la información para el

entrenamiento terminado, permitiendo la evaluación de su comportamiento para cada época. Los modelos entrenados con variaciones en el número de épocas utilizadas en el entrenamiento o variaciones en el tamaño de la base de datos utilizada para entrenar, resultan en cambios en el desempeño del modelo entrenado al hacer predicciones. Se pretende evaluar estos cambios y comparar las métricas finales para alcanzar un modelo capaz de hacer predicciones la mayor precisión posible a la hora de evaluar imágenes de la espina dorsal. En el aprendizaje automático, la inferencia es el proceso de utilizar un modelo entrenado para hacer predicciones sobre nuevos datos. Por medio de esta inferencia, es posible visualizar el rendimiento del modelo al hacer predicciones en imágenes que no están incluidas en el entrenamiento. A partir de esta inferencia, se crea un cuaderno de etiquetado "semi-supervisado", capaz de realizar anotaciones de vértebras en imágenes con el modelo, y la extracción en una carpeta con archivos de tipo ". json" siendo correspondientes a cada imagen.

4.5.3 Imagen de Procedimiento

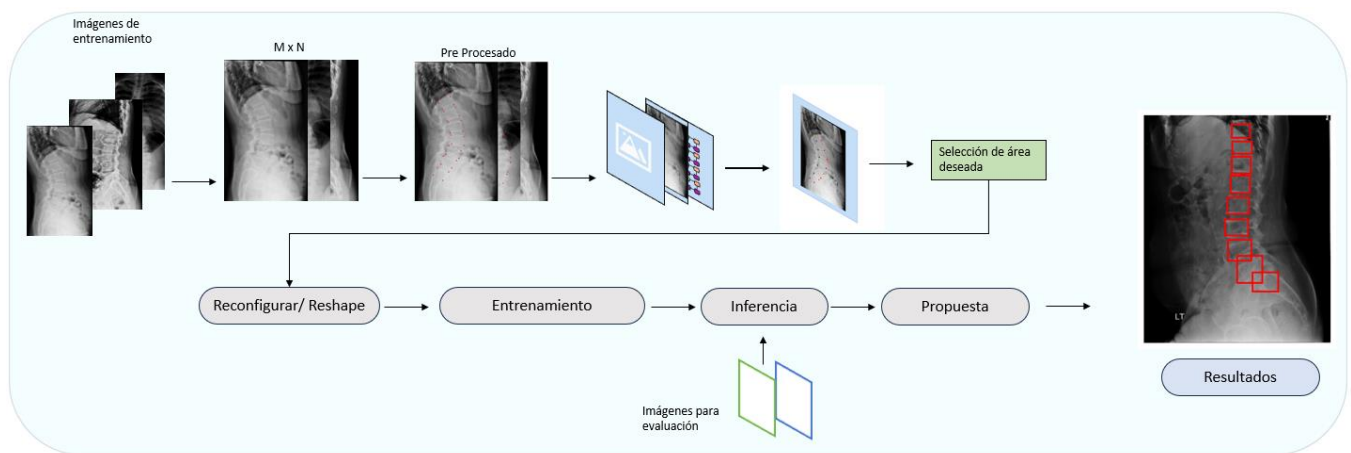


Ilustración 7. Resumen de procedimiento para la detección de vértebras

4.6 Metodología de Validación

Durante el proceso de entrenamiento de una red neuronal, se emplea la validación con el propósito de evaluar el desempeño del modelo en datos previamente no expuestos a él. Esta práctica se convierte en una salvaguarda efectiva contra el fenómeno de sobreajuste, el cual se manifiesta cuando un modelo se adapta excesivamente a los datos utilizados en su entrenamiento, limitando su capacidad para generalizar a conjuntos de datos inéditos. El proceso de validación implica la subdivisión del conjunto de datos de entrenamiento en dos componentes fundamentales: un conjunto de entrenamiento y un conjunto de validación. El primero se dedica a entrenar el modelo, mientras que el segundo se destina a la evaluación del desempeño del modelo. Una vez que el modelo ha completado su entrenamiento con el conjunto de entrenamiento, se aplica al conjunto de validación. La eficacia del modelo se cuantifica mediante métricas de evaluación, como el error cuadrático medio o la precisión. La validación representa un eslabón crítico en el proceso de entrenamiento de una red neuronal, ya que su inclusión sistemática permite contrarrestar el sobreajuste y asegura la capacidad del modelo para generalizar efectivamente en el contexto de nuevos datos.

4.6.1 Métricas de Validación y Cálculos Numéricos.

Posterior a la adquisición de datos procedentes de las diferentes etapas de entrenamiento, para hacer posible el análisis cuantitativo de los resultados, la discusión de los resultados obtenidos y las conclusiones de la investigación, se realizaron los cálculos de las métricas establecidas para definir el rendimiento del algoritmo frente a los diagnósticos reales por parte de profesionales. Para la parte del algoritmo que realiza las predicciones de los puntos de referencia en las imágenes, se midieron la precisión, y el IoU.

4.6.1.1. Precisión

La precisión se refiere a la fracción de instancias recuperadas que son relevantes. La precisión es, por tanto, una medida de pureza en el desempeño de la recuperación, una

medida de eficacia a la hora de excluir elementos no relevantes del conjunto recuperado. Es deseable una alta precisión, así como es deseable una alta recuperación. Lo ideal es alcanzar el 100% en ambos al mismo tiempo. (Buckland, M & Gey, F., 1994)

$$\%Precision = \frac{TP}{TP + FP} \times 100$$

Ecuación 1: Fórmula para el cálculo de precisión.

TP = verdaderos positivos, y FP = falsos positivos.

4.6.1.2. Exhaustividad

La exhaustividad se refiere a la fracción de instancias relevantes que han sido recuperadas. Esto la hace una medida de efectividad en la recuperación (o selección) del desempeño y puede verse como una medida de la eficacia a la hora de incluir elementos relevantes en el conjunto recuperado. Un 100% de exhaustividad siempre se puede lograr recuperando (examinando) toda la base de datos, pero esto anula el propósito de un sistema de recuperación. Sin embargo, la capacidad de lograr una alta recuperación (100% o cerca) de manera eficiente es claramente deseable. (Buckland, M & Gey, F., 1994)

$$\%Exhaustividad = \frac{TP}{TP + FN} \times 100$$

Ecuación 2: Fórmula para Exhaustividad.

TP = verdaderos positivos, FN = Falsos Negativos.

4.6.1.3 Coeficiente IoU

El coeficiente de IoU (Intersection over Union) es una métrica utilizada comúnmente para evaluar modelos de aprendizaje profundo, es una relación entre la intersección y la unión entre un cuadro delimitador predicho en la detección de objetos, y el cuadro delimitador real del objeto. Se recomienda aplicar un umbral mínimo de 0.3 durante entrenamiento para evitar que las cajas delimitantes predigan la misma vértebra más de 1 vez. Un IoU final superior al 0.5 se considera aceptable (Tu Y. et al., 2019).

$$IOU = \frac{A(o)}{A(u)}$$

Ecuación 3: *Fórmula para Coeficiente de Intersección sobre Unión.*

A(o) = área de superposición, A(u) = área de unión.

Fuente: (Taleon C. et al., 2023)

4.6 Matriz Metodológica

Título	Problema de Investigación	Objetivos	Variables	Metodologías y Herramientas
Algoritmos Procesamiento de Imágenes de Neuroradiología Utilizando Computer Visión	El diagnóstico preciso de afecciones de la columna vertebral, es crucial para un tratamiento efectivo y una mejor calidad de vida para los pacientes. Actualmente, la interpretación subjetiva de imágenes de rayos X por profesionales de la salud puede resultar en diagnósticos imprecisos y retrasos en el tratamiento. El uso de técnicas de aprendizaje profundo y la automatización de la anotación de imágenes de rayos X ofrecen una solución prometedora al identificar patrones sutiles que los humanos pueden pasar por alto,	General	Dependientes	Técnicas
		Efectuar un algoritmo de aprendizaje profundo computarizado capaz de automatizar la detección de vértebras en imágenes de rayos X tomadas en el plano sagital de pacientes con deformidades en la región de la espina dorsal.	<ul style="list-style-type: none"> • Precisión • Pérdida • Exhaustividad • Intersection over view 	<ul style="list-style-type: none"> • Etiquetado de imágenes en Roboflow • Modelo KeyPoint RCNN.
		Específicos	Independientes	Herramientas
		Modificar parámetros de los modelos existentes para la creación de la clasificación de la gravedad de las afecciones.	<ul style="list-style-type: none"> • Tamaño de Lote • Épocas • Modelo Pre Entrenado 	<ul style="list-style-type: none"> • Roboflow • Python • Colaboratory

	<p>mejorando la detección temprana, reduce el tiempo necesario para el diagnóstico y mejora la eficiencia en la atención médica. La capacitación y supervisión de estos algoritmos son fundamentales para garantizar su exactitud. En Honduras, no se cuenta con opciones de apoyo para diagnóstico de este tipo. Debido a esto se busca contribuir con la Sociedad de Neurocirujanos para mejorar la detección de vértebras de estas afecciones y la atención médica en el país.</p>	<p>Evaluar el rendimiento del sistema de detección en un conjunto de datos de prueba y compararlo con diagnósticos manuales realizados por expertos.</p>		
--	---	--	--	--

Tabla 10. Matriz metodológica

Fuente: Elaboración de tabla Propia

V. RESULTADOS Y ANÁLISIS.

El siguiente capítulo muestra los resultados obtenidos en cada una de las etapas de esta investigación experimental. Los resultados se presentan en base a las métricas obtenidas del modelo en las diferentes pruebas realizadas, específicamente se contemplan las métricas de precisión promedio de IoU, exhaustividad y pérdida además de los ponderados de cada una de estas.

5.1. Resultados de primer entrenamiento.

Los valores predichos se exhiben en contraste con los valores reales, acompañados de una evaluación exhaustiva de las métricas de entrenamiento y predicción. Estas métricas comprenden las precisiones promedio de IoU tanto para cajas delimitadoras como para puntos clave, exhaustividad para ambas, la pérdida para ambas, y la pérdida general del entrenamiento. Se realizaron ajustes en los parámetros del tamaño de lote, las épocas y el modelo pre-entrenado para evaluar la precisión en cada una de las pruebas efectuadas. Con el propósito de lograr una comparativa más precisa entre las épocas, se establecieron los mismos parámetros modificados en cada prueba. La cantidad de imágenes al comienzo de nuestro proyecto fue de 245 imágenes, estas imágenes ya habían sido clasificadas como espondilolistesis, aunque no eran todas aceptables para uso puesto que no se encontraban tomadas en plano sagital, un parámetro necesario predeterminado para tener visión las vértebras, dejando nuestras primeras pruebas con 197 fotos.

Después de llevar a cabo una revisión de la base de datos, se utilizaron 197 imágenes en la primera prueba, logrando totalizar una precisión de entrenamiento del 91.1% en la detección, un resultado considerado óptimo para los propósitos de la investigación, donde esto se puso a prueba con el comienzo de resultados de código. A través de este resultado se hizo un incremento en nuestra base de datos para obtener mejores análisis teniendo un total de 405 imágenes donde se dividen 70% en entrenamiento y 30% validación. Cada tabla exhibe las predicciones derivadas de las pruebas realizadas con el modelo entrenado, junto con sus métricas de evaluación correspondientes para cada variante. En total, se generó **1** predicción del código original por

parámetro durante 50 épocas. En lo que respecta al modelo KeyPoint RCNN, se ejecutaron **2** predicciones en total, distribuidas equitativamente entre cada parámetro, es decir, error, predicción y número de épocas, divididas en bloques de 25, 50, 75 y 100 épocas para el modelo específico. Se realizó de igual forma una comparativa con el modelo original de Scoliovis (2023) antes de editarlo para cumplir con nuestra detección de vértebras en imágenes de radiografías en plano sagital.

Pruebas	Tamaño de lote	Épocas	Precisión	Pérdida
KeyPoint	197	50	89.3%	2.69%
KeyPoint	197	100	91.1%	3%

Tabla 11. *Primeras pruebas con base 1*

Fuente: Elaboración de tabla Propia

5.2. Resultados Primera Prueba

La fase de entrenamiento y prueba del método de medición se llevó a cabo de manera independiente utilizando conjuntos de datos distintos para lograr una validación objetiva. En el caso de los modelos de aprendizaje automático supervisados, como las redes neuronales, las épocas adicionales permiten que el modelo aprenda en más instancias con los datos a su disposición. Esto puede conducir a una mejor generalización del modelo, lo que significa que el modelo funcionará mejor con datos que no ha visto antes resultando en una mejor agrupación de los datos, lo que puede ser útil para una variedad de tareas como la reducción de la dimensionalidad o la detección de anomalías.

Para la mejora de resultados de precisión, fue necesario un pre-procesado con el fin de mejorar la detección de puntos clave o keypoints. Este pre-procesado está representado en el [Anexo 1](#). Se obtuvieron las métricas de entrenamiento de cada prueba hecha al modelo para cada uno de los parámetros. Los parámetros en el modelo KeyPoint RCNN en la primera prueba se muestran en

, donde es posible visualizar el rendimiento del algoritmo de detección de objetos en términos de precisión promedio (AP) como función del umbral de intersección de regiones (IoU).

La precisión promedio de IoU cuantifica la medida de similitud entre el cuadro delimitador predicha por el algoritmo y el cuadro delimitador de la verdad absoluta anotada del objeto. En este caso, el algoritmo fue entrenado para detectar objetos con un IoU de 0,50, esto significa que el algoritmo debe predecir una caja delimitadora que se superponga al menos en un 50 % con la caja delimitadora real del objeto. Se logra denotar que el rendimiento del algoritmo es satisfactorio, con una precisión promedio superior a 0,90 para todos los valores de umbral de IoU. Esto significa que el algoritmo es capaz de detectar objetos con precisión en la mayoría de los casos. Sin embargo, el rendimiento del algoritmo disminuye ligeramente a medida que aumenta el umbral de IoU. Esto se debe a que es más difícil para el algoritmo predecir cajas delimitadoras que se superpongan exactamente con la caja delimitadora real del objeto.

Esta primera prueba obtuvo diferentes valores de nuestro interés, el [Anexo 3](#) Pérdida de caja de Regulación de validación es la gráfica que muestra la pérdida en la regresión de cuadros delimitadores. La métrica cuantifica las diferencias de distancia entre las coordenadas predichas para cada cuadro delimitador, y las coordenadas verdaderas para los cuadros anotados como verdad absoluta en la base de datos.

En el parámetro de precisión de Keypoints en el

[Anexo 4](#), muestra la precisión promedio del algoritmo de detección de puntos clave o Keypoints para un valor de umbral de IoU de 0,5. En este caso como se ha mencionado previamente, los puntos clave son puntos de referencia ubicado en las esquinas de cada una de las vértebras en la columna vertebral. Esto significa que el algoritmo debe predecir la ubicación de un punto clave que se encuentre al menos 50 % superpuesto con la ubicación real del punto clave. El entrenamiento devuelve una precisión promedio superior a 0,80 para todos los valores de umbral de IoU. Esto significa que el algoritmo es capaz de detectar puntos clave con precisión en la mayoría de los casos.

En la evaluación de la pérdida para la detección de keypoints, podemos ver sus resultados en [Anexo 5](#), donde se encuentra graficada la pérdida de puntos clave de vértebras a lo largo de las épocas en el entrenamiento. La gráfica muestra que la pérdida de puntos clave aumenta ostensiblemente en su comienzo, y se reduce y estabiliza a medida que el entrenamiento avanza

en sus épocas. Esto se debe a que el entrenamiento optimiza su detección, concluyendo en un valor de pérdida aceptable para llevar a cabo la detección de las esquinas de las vértebras.

En lo que respecta a la pérdida del clasificador, [Anexo 6](#), la pérdida del clasificador del validación del modelo a lo largo del tiempo. El clasificador del modelo es un elemento que se encarga de realizar predicciones sobre a cuál clase pertenecen los objetos que están siendo detectados. La gráfica muestra que la pérdida del clasificador disminuye a medida que aumenta el número de iteraciones. Esto se debe a que el clasificador aprende a ajustar el modelo para minimizar la función de pérdida. La pérdida del clasificador se estabiliza después de un cierto número de iteraciones ya que el clasificador ha aprendido todo lo que puede de los datos de entrenamiento. Se toma en cuenta que la pérdida del clasificador de preentrenamiento del modelo puede variar en función del tamaño del conjunto de datos de entrenamiento y la cantidad de clases a detectar. En este caso en particular se están detectando únicamente objetos bajo la clase de "vertebras".

Teniendo examinados los resultados de la primera prueba se logró hacer una comparativa entre la finalidad asignada al modelo en la presente investigación experimental y la que tenía originalmente el modelo de KeyPoint RCNN. Para mejorar resultados y así contrarrestar la calidad de nuestras imágenes a nuestros entrenamientos se les asignaron 100 épocas, mientras que al comienzo solamente 50 épocas y los resultados de los entrenamientos fueron exitosos y con un porcentaje alto de aceptabilidad. Tras la evaluación de los resultados obtenidos en la prueba inicial, se llevaron a cabo dos predicciones adicionales. La primera se basó en 25 épocas para su proceso de entrenamiento y validación que son de [¡Error! No se encuentra el origen de la referencia.](#) al [Anexo 15](#) comparándola con la segunda predicción que utilizó 100 épocas [Anexo 16 al Anexo 26](#). Ambas predicciones arrojaron resultados similares tanto en el conjunto de entrenamiento como en el de validación.

Durante el proceso de entrenamiento destinado a la detección de vértebras en radiografías de plano lateral, se procedió a evaluar la precisión del modelo KeyPoint RCNN entrenado. Nuestro objetivo consiste en superar los resultados obtenidos de trabajos anteriores relacionados con la detección de vertebras, para validar la mejora en la detección de nuestro código. Aunque nuestra

precisión promedio muestra valores elevados, otra métrica de interés es la evaluación de las pérdidas, con el propósito de cuantificar la magnitud de las mismas.

5.3. Resultados de predicciones de modelo Propio

5.3.1. Entrenamiento de 25 épocas

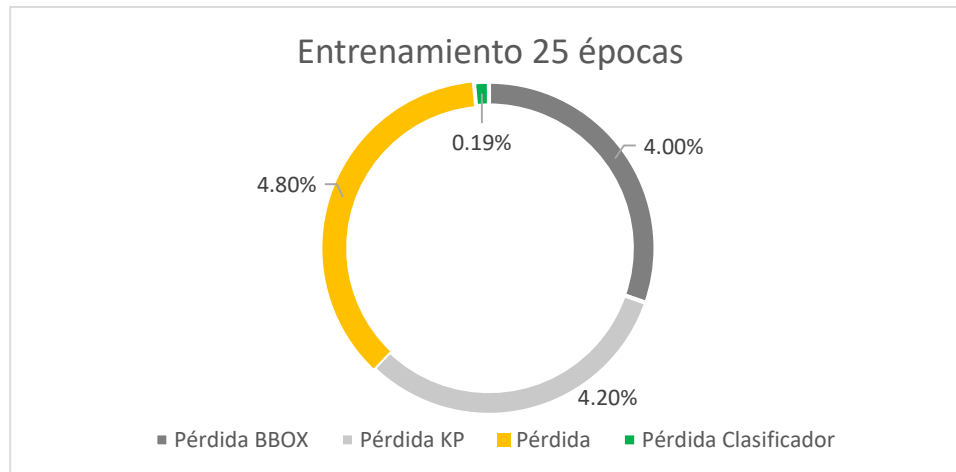


Ilustración 8. Gráfica de entrenamiento 25 épocas

Fuente: Elaboración de tabla Propia

Estos resultados de la gráfica están basados en los [Anexo 7](#) a [Anexo 10](#). La pérdida del modelo se desglosa durante el proceso de entrenamiento en dos componentes principales: la pérdida de cuadros delimitadores, que alcanzó un valor porcentual del 4.0%, aproximadamente similar a la pérdida total en el entrenamiento, que fue del 4.8%. En lo que respecta a la detección de las esquinas de las vértebras o puntos clave, se logró un 4.2% de pérdida, con una pérdida mínima del clasificador de 0.19%. En este contexto, la pérdida de clasificación se sitúa en alrededor de 0.02 después de 25 épocas, indicando que el modelo comete un error al clasificar incorrectamente el objeto detectado en aproximadamente el 2% de las instancias.

En términos generales, los resultados derivados de la gráfica son positivos, ya que la pérdida del modelo se mantiene baja después de 25 épocas de entrenamiento. Este hecho sugiere que el modelo exhibe una capacidad notable para detectar objetos con una elevada precisión. No obstante, se identifican áreas de mejora potencial, incluyendo la posibilidad de reducir la pérdida de regresión mediante la optimización de la capacidad del modelo para ubicar con precisión las cajas delimitadoras de los objetos.

5.3.2. Validación de 25 épocas

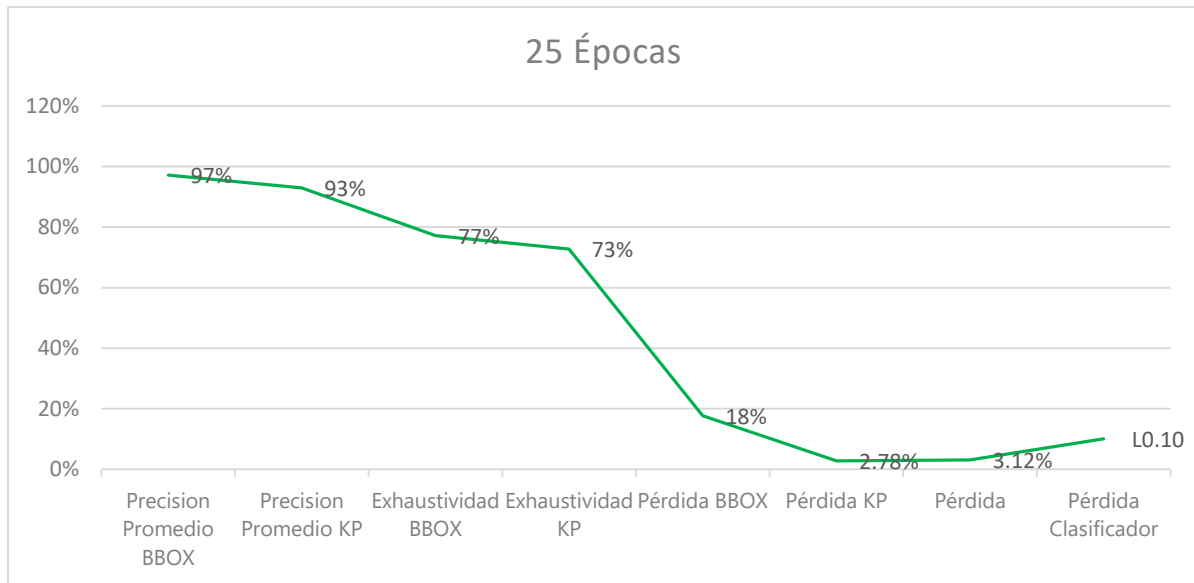


Ilustración 9. Gráfica de Resultados Validación 25 épocas

Fuente: Elaboración de tabla Propia

La representación gráfica está basada en los [Anexo 11](#) a [Anexo 15](#) proporcionando un resumen de los resultados obtenidos en después de 25 épocas de entrenamiento. En cuanto a la detección de vértebras mediante bounding boxes, alcanzamos un destacado porcentaje de precisión promedio de 97%, con una pérdida total a lo largo de todas las épocas del 18%. En el caso de la detección de KeyPoints, logramos un rendimiento del 93%, aunque con una pérdida del 2.70%. Es importante señalar que este último valor puede deberse a diversas situaciones, como la variabilidad en la cantidad de detecciones de vértebras en nuestras imágenes, así como el ruido y contraste presentes en las radiografías, factores que también pueden incidir en la exhaustividad, la cual fue del 77% para bounding boxes y del 73% para KeyPoints.

A pesar de estas consideraciones, estos números son aceptables en el contexto de nuestro entrenamiento, ya que la pérdida total se mantiene por debajo del 10%, específicamente en un 3.12%. Es relevante destacar que este último valor es el más alto, ya que representa la pérdida total del sistema

5.3.3. Entrenamiento de 100 épocas

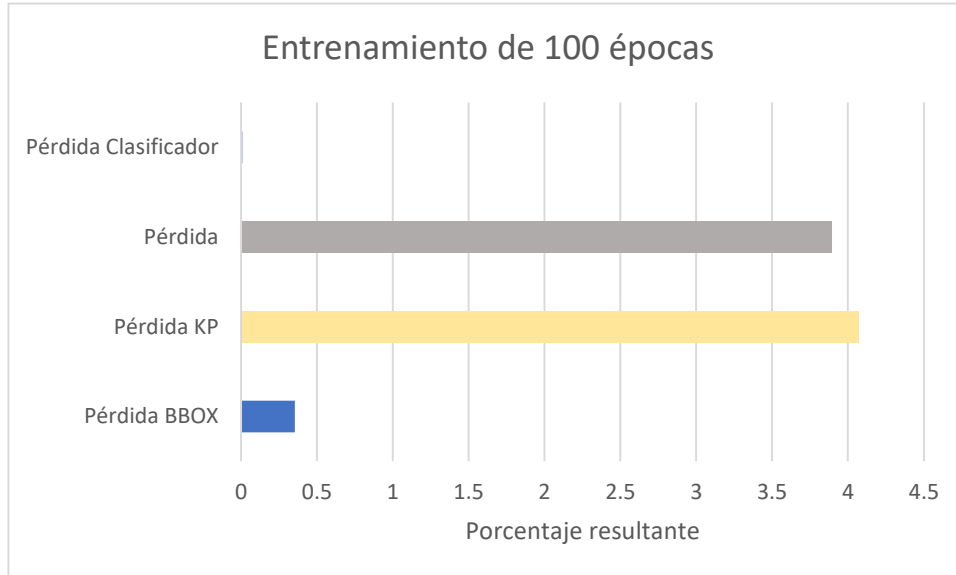


Ilustración 10. Gráfica Entrenamiento 100 épocas

Fuente: Elaboración de tabla Propia

Durante el proceso de entrenamiento, las pérdidas cumplen un papel fundamental al reflejar la efectividad de las predicciones. Los resultados expuestos en los anexos [Anexo 16](#) al [Anexo 19](#) arrojan luz sobre estas pérdidas, específicamente en relación con las bounding boxes. En nuestro entrenamiento de 100 épocas, logramos una pérdida inferior al 0.5%, demostrando una mejora con respecto a las 25 épocas anteriores. En cuanto a la pérdida asociada a los KeyPoints, se registró el valor más alto, alcanzando un 4.2%, un incremento de 0.2% en comparación con las 25 épocas previas. Vale la pena mencionar que, en contraste, la pérdida del clasificador fue prácticamente nula.

Estos resultados destacan la capacidad del modelo para minimizar las pérdidas en la detección de bounding boxes, evidenciando una mejora continua durante el entrenamiento de 100 épocas en comparación con el período de 25 épocas. Por otro lado, la ligeramente mayor pérdida en los KeyPoints podría sugerir áreas de oportunidad para optimizar la precisión en la detección de puntos clave. Es fundamental comprender estas variaciones en las pérdidas para perfeccionar el rendimiento general del modelo.

5.3.4. Validación de 100 épocas

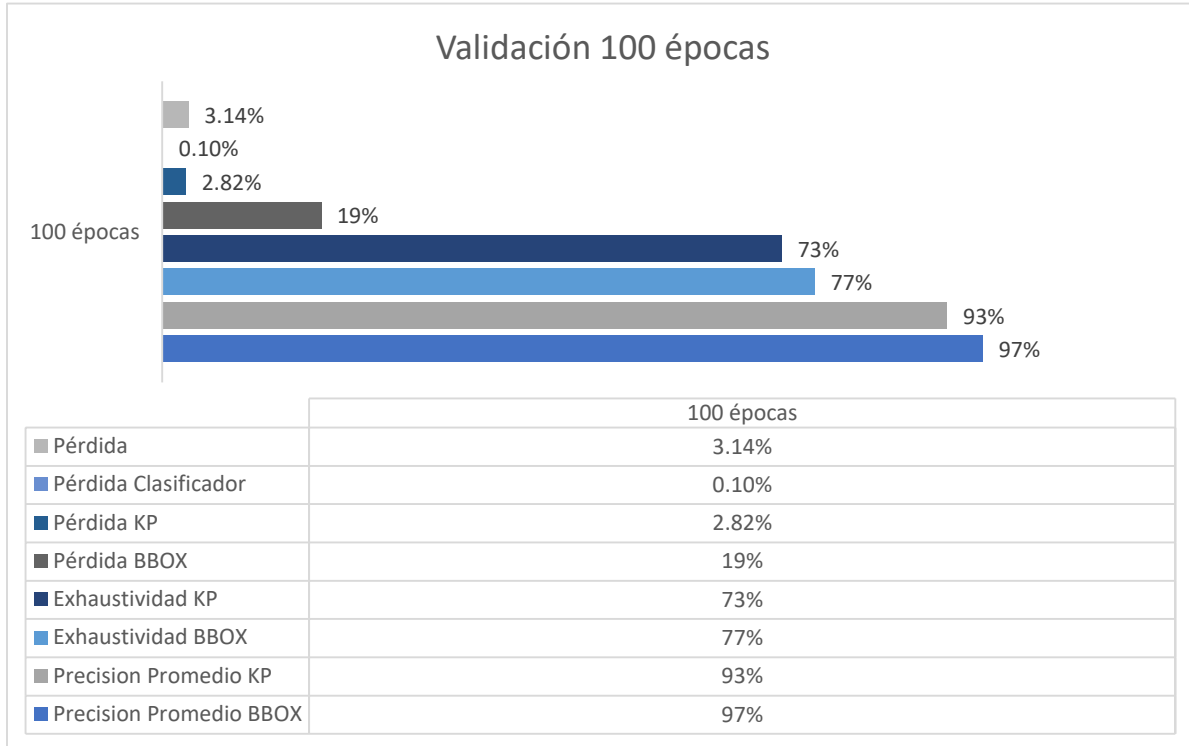


Ilustración 11. Gráfica de Resultados de Validación 100 épocas

Los valores obtenidos en los [Anexo 20](#) a [Anexo 26](#) ofrecen un análisis exhaustivo de los porcentajes durante las 100 épocas de predicción. La representación gráfica de estos datos proporciona una visión detallada de la evolución del rendimiento del modelo de aprendizaje automático a lo largo de las distintas épocas. En este marco, es relevante destacar que se logró una precisión promedio del 97% para las bounding boxes y del 93% para los KeyPoints. Este alto nivel de precisión se refleja en los valores de exhaustividad, con un 77% para bounding boxes y un 73% para KeyPoints. Es fundamental entender que, al considerar las pérdidas a lo largo de las 100 épocas, si bien no tienen un impacto significativo en los KeyPoints, sí afectan las bounding boxes, con una disminución del 17%, aunque este valor es menor que el observado con 25 épocas.

Estos resultados indican una mejora progresiva en el rendimiento del modelo a medida que transcurre el tiempo de entrenamiento. La disminución en la pérdida sugiere una reducción en las predicciones incorrectas del modelo. Simultáneamente, el incremento en la exhaustividad y precisión apunta a una identificación más precisa y completa de los objetos en cuestión. Este patrón continuo de mejora respalda la eficacia del aprendizaje automático aplicado en este

contexto específico, subrayando la robustez y capacidad de adaptación del modelo a lo largo de las 100 épocas de entrenamiento.

5.3.5. Resultados de 100 épocas variando cantidad de conjunto de datos

En términos generales, el modelo exhibe una capacidad notable para detectar objetos con una elevada precisión. No obstante, se identifican áreas de mejora, incluyendo la posibilidad de reducir la pérdida de regresión mediante la optimización de la capacidad del modelo para ubicar con precisión las cajas delimitadoras de los objetos. Otras de las pruebas realizadas para poder ver si sufrían anomalías con épocas fue entrenar el modelo y validarlo por épocas teniendo ya en cuenta que la segunda prueba que se hizo tuvo una precisión de 72.9% para KeyPoints, nuestro objetivo es tener un valor mejor al de otros investigadores utilizando este tipo de detección.

En general, los resultados de las pruebas del modelo de detección de vértebras muestran que el modelo es capaz de detectar vértebras con una precisión promedio de 97,2%. Sin embargo, hay algunas diferencias mínimas en los resultados entre las configuraciones de 25 y 100 épocas, dando paso a la comparativa de resultados variando la cantidad de imágenes que son sometidas en el algoritmo. La base de datos fue incrementada a 505 imágenes en esta última etapa de entrenamientos, en búsqueda de diferencias más significativas en la mejora de las métricas de desempeño. Se sometió la variación en el conjunto de datos, se muestra las pérdidas de entrenamiento y validación de detección de cajas delimitadoras y puntos clave que producen resultados comparables o mejores en fracciones del nuevo conjunto de datos del 20%, 50%, 80% y 100%, respectivamente. Como primeros resultados de las últimas pruebas comparamos la precisión promedio siendo afectada por el tamaño de conjunto de datos y manteniendo 100 épocas para estos resultados finales. Se muestra el rendimiento del algoritmo de detección de objetos en términos de precisión promedio por sus siglas en inglés AP, como función del umbral de intersección de regiones (IoU). El IoU es la medida de similitud entre la caja delimitadora predicha por el algoritmo y la caja delimitadora real del objeto. Dentro del conjunto de resultados, focalizamos la atención en la precisión promedio de las bounding boxes y KeyPoints.

En cuanto a la precisión de las cajas delimitadoras, como se evidencia en [Anexo 27](#) al [Anexo 36](#), donde se registra un valor notablemente bajo con la menor cantidad de conjunto de

datos, que corresponde al 20%. Es importante señalar que, para estas últimas pruebas, se empleó un total de 505 imágenes, y el 20% de este conjunto equivale a 101 imágenes.

Los resultados obtenidos fueron significativamente inferiores a las expectativas para esta cantidad de imágenes, con una precisión del 68% para las cajas delimitadoras y del 42.7% para la precisión promedio de puntos clave (KeyPoints). En contraste, para las otras tres pruebas que representan el 50%, 80% y 100% del conjunto de datos, con una precisión promedio del 97.5% para las cajas delimitadoras y del 93.8% para KeyPoints o puntos clave.

En lo que respecta a la pérdida de objetos, se tuvo una mejoría considerable que se puede visualizar en los [Anexo 37](#) al [Anexo 45](#) con el 50% de la base de datos, teniendo una pérdida mayor cuando hay menor cantidad de datos para verdad absoluta de bounding boxes, con este resultado se puede diferir que el aumento de imágenes hizo el trabajo de evitar un valor de pérdida mayor al que teníamos al comienzo en las primeras pruebas. Por otro lado, en lo que respecta a la pérdida de detección de keypoints la mayor pérdida es para la prueba con el 20% de conjunto de datos con un 3.75% y cuando se hizo con 100% también tuvo una pérdida notable de 3.06%.

Para la pérdida total del algoritmo se convalida el planteamiento de la necesidad de tener más datos para hacer la detección y de esta forma se entrena mejor el algoritmo. En términos más técnicos, con menos datos disponibles o una cantidad demasiado elevada de épocas de entrenamiento, el modelo puede enfrentar problemas de sobreajuste, donde se ajusta demasiado a los datos de entrenamiento específicos y no logra generalizar bien a nuevas instancias. Es decir que el modelo de detección de objetos alcanza un punto en el cual este comienza a memorizar los datos propios a la base de datos que le es entregada, en lugar de aprender de los datos. Además, la falta de datos puede hacer que el modelo sea más sensible a variaciones pequeñas y ruidos en el conjunto de entrenamiento, lo que afecta negativamente su capacidad para realizar predicciones precisas en nuevos datos, como es en este caso, cuando solo se le brindó 20% de las imágenes de la base de datos obteniendo la mayor pérdida registrada de 4.2289% y de 3.335% cuando es mayor cantidad de imágenes. Se busca evitar la falta de representatividad y diversidad en el conjunto de datos mediante el uso de conjuntos más grandes, lo que generalmente contribuye a un rendimiento mejorado y una pérdida reducida en tareas de detección de objetos y KeyPoints.

Se refleja desde [Anexo 46](#) al [Anexo 55](#). Los entrenamientos constan con el 100% de la base de datos, otorgando resultados de mayor exhaustividad implicando que en este aspecto de los entrenamientos, una base de datos con mayor cantidad de imágenes tiene un mayor rendimiento al identificar y recuperar de forma completa la información de los objetos relevantes detectados en las imágenes. Nuevamente, el modelo se encontró considerablemente afectado al entrenar utilizando el menor porcentaje de la base de datos, resultando en una notable diferencia en los valores finales. Esta tendencia observable sugiere una alta importancia entre el rango de imágenes utilizadas entre el 20% de la base de datos y el 50% de ella, y los efectos que esto tiene sobre la capacidad final del entrenamiento para hacer inferencia precisa sobre imágenes nuevas.

5.4. Etapa final: Evaluación a través de comparativas

Hasta donde alcanza nuestro conocimiento, no se ha diseñado ningún algoritmo para la detección de vértebras con imágenes de rayos X de la región lumbar en plano sagital mediante el uso combinado de cajas delimitadoras (bounding boxes) y puntos clave (KeyPoints). Nuestro enfoque se desarrolló con el propósito de detectar todas las vértebras de manera precisa, eliminando la necesidad de establecer límites para incluirlas en la base de datos. Motivados por requisitos clínicos específicos, nuestro método se especializa en la detección de vértebras en plano sagital, lo que justifica la comparación con otros métodos relacionados con la detección de vértebras.

En este estudio, comparamos nuestro modelo KeyPoint RCNN con investigaciones anteriores centradas en la detección de patologías espinales, como la utilización del aprendizaje profundo de transferencia por Fraiwan et al.,(2022) que logró una alta precisión del 96.34% para clasificar la espondilolistesis con radiografías frontales. Scoliovis por Taleon et al., (2023), informaron una precisión del 93.3% para la detección de escoliosis utilizando el mismo modelo que nuestra investigación. Además, se comparó nuestro trabajo con los resultados de Trinh et al., (2023), cuyo modelo LumbarNet alcanzó una precisión del 89.93% para la detección de espondilolistesis y su grado de desfase en radiografías con más de 7 vértebras. P. Klinwichit et al. (2023) presentaron resultados similares a nuestro estudio utilizando YoLoV5, con una precisión del 95.83%. Sin embargo, destacamos que, a pesar de las similitudes, ninguno de estos trabajos implementó una detección por KeyPoints eficiente para validación, lo que resultó en la pérdida

de información en las esquinas de las vértebras inferiores y errores en la clasificación del grado de deslizamiento vertebral.

Aunque nuestro estudio enfrenta limitaciones en la obtención de radiografías para la expansión de la base de datos y la complejidad en la programación del algoritmo, con horas pérdidas y resultados erróneos durante el entrenamiento y la validación por inestabilidad de redes de internet, logramos aprender de estos desafíos. El resultado final fue un algoritmo con una destacada precisión del 98% en la detección de vértebras en plano sagital, considerándose un logro satisfactorio.

No.	Estudio	Modelo	Precisión	Pérdida
1.	Trinh, et al. (2023)	LumbarNet	89.93	N/A
2.	C. A. T. Taleon, et al. (2023)	Keypoint RCNN	93.3%	18.15%
3.	P. Klinwichit, et al. (2023)	YoLoV5	95.83%	12.11%
4.	L. Fraiwan, et al. (2022)	DensNet-201	96.34%	10.50%
5.	Este trabajo (2023)	Keypoint RCNN	98%	17.64%

Tabla 12. Comparativa de resultados de investigaciones de algoritmos de detección espinal

5.5. Inferencia

La evaluación del sistema se basa en su rendimiento promedio, centrado en poner a prueba la capacidad del modelo para detectar vértebras. Durante el proceso de entrenamiento, el modelo ha sido instruido para aprender patrones y características relevantes que permiten la detección de vértebras en las imágenes médicas. Este entrenamiento realizado con un conjunto de datos específicos diseñados para fortalecer la capacidad del modelo para reconocer los KeyPoints asociados a las bounding boxes que delimitan las vértebras en las imágenes. En la fase de prueba, se aplicó un conjuntos de datos adicionales que no se integraron en el proceso de entrenamiento, esto garantiza una evaluación objetiva y realista del desempeño del modelo en situaciones no vistas previamente.

Al centrarse en el rendimiento promedio, se busca validar la consistencia y eficacia del modelo en la detección de vértebras en diversas instancias. La evaluación no solo se centra en la capacidad del modelo para identificar las vértebras correctamente, sino también en la capacidad para localizar de manera precisa las esquinas de estas vértebras. Aunque aún hay mejoras por hacer en el modelo, el porcentaje de exactitud obtenido es de 91.1% siendo uno de los más altos sin la cantidad de imágenes que otros proyectos han hecho, ni con imágenes de la mejor calidad. Las bounding boxes están siendo estructuradas correctamente excepto por una que se sitúa en una de las apófisis transversal.

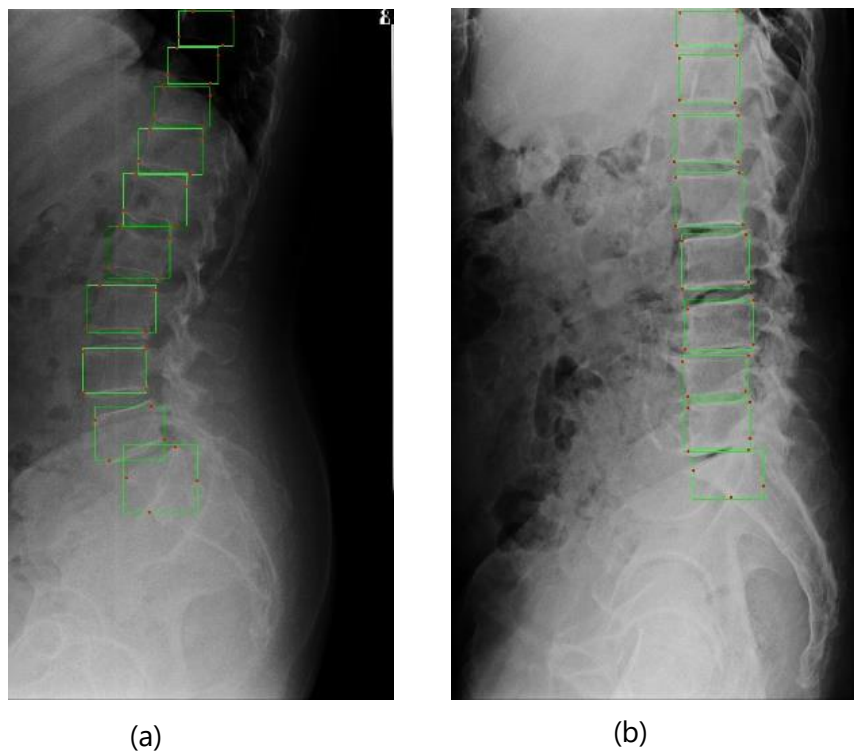


Ilustración 12. Resultados detección de vértebras en el etiquetado semi-supervisado

La Ilustración 12 denota las bounding boxes en los últimos resultados. Dentro de los errores encontrados en la selección se encuentra un problema repetitivo en otras investigaciones, la detección de bounding boxes cuando se sobreponen entre ellas, otro error mostrado en nuestro entrenamiento es que la visión de computadora aún sigue aprendiendo a detectar vértebras.

Las bounding boxes están siendo estructuradas correctamente. Posteriormente se desarrolla un cuaderno de inferencia en Google Colaboratory, que permite utilizar el modelo para hacer predicciones de anotaciones en una carpeta de múltiples imágenes, y las devuelve en un formato compatible con los modelos KeyPoint RCNN.

En la fase de prueba, utilizamos conjuntos de datos adicionales que no se habían empleado en el proceso de entrenamiento, esto garantizó una evaluación objetiva y realista del desempeño del modelo en situaciones no vistas previamente. Al centrarnos en el rendimiento promedio, buscamos validar la consistencia y eficacia del modelo en la detección de vértebras en diversas instancias. La evaluación no solo se centró en la capacidad del modelo para identificar las bounding boxes correctamente, sino también en su capacidad para localizar de manera precisa los KeyPoints dentro de estas cajas delimitadoras que representan las esquinas de las vértebras.

Este enfoque integral proporcionó una evaluación más completa de la capacidad del modelo para llevar a cabo la tarea específica de detección de vértebras en imágenes médicas. Aunque aún hay mejoras por hacer en el modelo, el porcentaje de exactitud obtenido es de 90% siendo uno de los más altos sin la cantidad de imágenes que otros proyectos han hecho, ni con imágenes de la mejor calidad. Como es visible en la Ilustración 13, las bounding boxes están siendo estructuradas correctamente. Posteriormente se desarrolló un cuaderno de inferencia en Google Colaboratory, que permite utilizar el modelo para hacer predicciones de anotaciones en una carpeta de múltiples imágenes, y las devuelve en un formato compatible con los modelos KeyPoint RCNN.

Este tipo de inferencia fue denominado etiquetado semi-supervisado. El objetivo de esta fase de resultados era aumentar las capacidades de inferencia del modelo. Las anotaciones en formato KeyPoint RCNN se guardan en archivos tipo ".json" y se relacionan con sus respectivas imágenes porque comparten el mismo nombre de archivo. El etiquetado semi-supervisado fue probado con una carpeta de 801 imágenes nuevas al modelo, originarias de la base de datos BUU-LSPINE

(2023). El cuaderno logro un etiquetado de inferencia exitoso, exportando 801 archivos tipo “.json” a la carpeta designada a su destino.

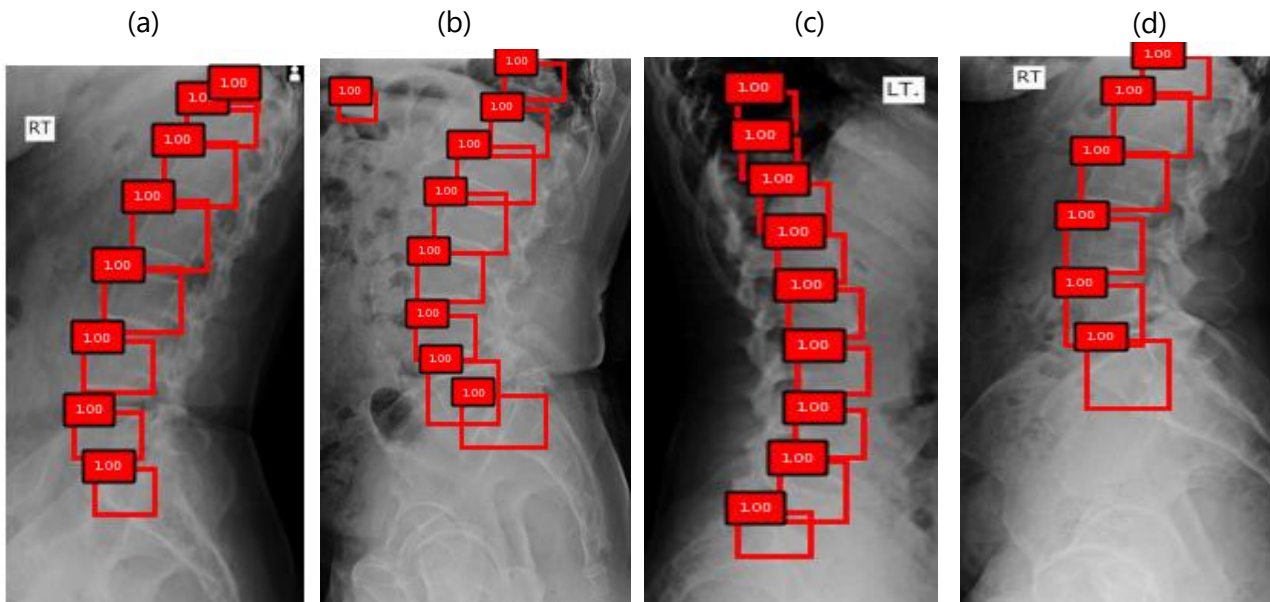


Ilustración 13. Ejemplos de inferencia

El propósito de la investigación fue evaluar la efectividad de los modelos pre entrenados, después de la adaptación para detección de vértebras, teniendo la identificación a partir de imágenes de rayos X, en este caso, radiografías de espondilolistesis. Además, dado que los algoritmos de aprendizaje profundo incurren en un alto costo computacional, se registró el tiempo utilizado tanto en el entrenamiento como en las pruebas llegando a durar de 15 minutos los entrenamientos cortos y los primeros para que la inteligencia artificial comenzara a detectar fue de 4 horas.

VI. CONCLUSIONES

Esta investigación desarrolló y evaluó un modelo de aprendizaje profundo para detectar vértebras en imágenes de rayos X en el plano sagital. La metodología destacó por un innovador pre-procesamiento que transformó puntos medios de cajas delimitadoras a puntos clave, facilitando el etiquetado en Roboflow. El modelo tiene potencial para diagnosticar deformidades espinales en aplicaciones futuras.

Se logró establecer una base de datos de 505 imágenes comparable en magnitud a las bases de datos comúnmente utilizadas en modelos de detección de vértebras, utilizando imágenes de la espalda lumbar tomadas en el plano sagital de pacientes hondureños e imágenes adquiridas en una base de datos publica llamada BUU-LSPINE diseñado para realizar entrenamientos de detección de objetos.

Se etiquetaron las imágenes de la base de datos de forma manual en la aplicación Roboflow, pasando posteriormente las imágenes a un cuaderno de pre-procesamiento el cual hizo posible extraer la información de los puntos clave marcados de forma manual, e interpretar los límites de cada conjunto de puntos clave como un cuadro delimitador. La información de la verdad absoluta fue extraída en formato Keypoint RCNN el cual es compatible con el cuaderno de entrenamiento. Las imágenes son organizadas en el pre-procesamiento de modo que se encuentren preparadas para el entrenamiento.

Se identificaron los cambios resultantes de modificar parámetros como el número de épocas y el número de lote a lo largo del entrenamiento, observando de igual forma que factores como la normalización de la base de datos, influyen sobre los resultados de desempeño del entrenamiento pese a ser difíciles de cuantificar. Se lograron realizar mejoras en métricas de desempeño de validación en comparación a lo disponible en la literatura, con una precisión promedio para el coeficiente de intersección sobre unión en detección de cuadros delimitadores de 98%.

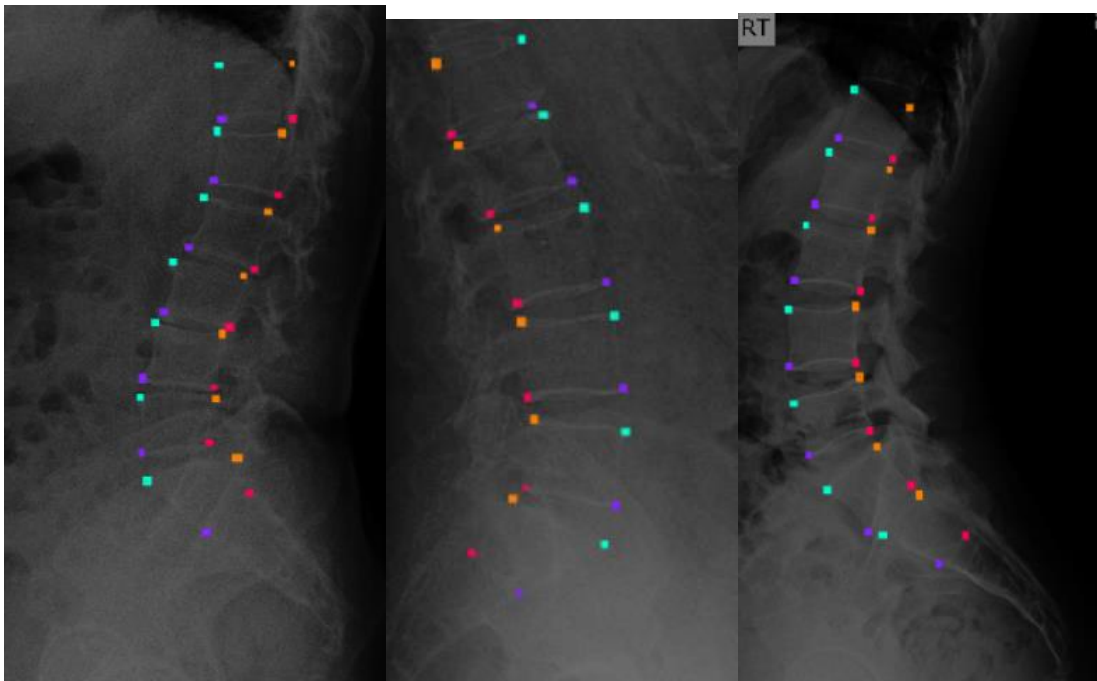
Se utilizaron imágenes resultantes de la base de datos de BUU-LSPINE que no fueron incluidas en el entrenamiento debido a límites de tiempo, para evaluar el modelo en su inferencia y comprobar si este lograba la detección de vértebras y las filtra en base a confiabilidad. El modelo

pre-entrenado de forma supervisada devolvió mediante su inferencia la misma cantidad de vértebras detectadas que el ojo humano en el 90% de las pruebas. Posteriormente, se utilizó el modelo pre-entrenado para realizar un etiquetado semi-supervisado en formato Keypoint RCNN para una carpeta de 801 imágenes y relacionar las anotaciones con las imágenes por medio del nombre de los archivos, probando la capacidad del modelo para etiquetar múltiples imágenes de rayos X de forma eficiente.

ANEXOS

- **Resultados Primeras Prueba -Validación**
- Anexo 2- Anexo 6
- **Entrenamiento 25 épocas** Anexo 7-Anexo 10
- **Validación 25 épocas** Anexo 11-Anexo 15
- **Entrenamiento 100 épocas** Anexo 16-Anexo 19
- **Validación 100 épocas** Anexo 20-Anexo 26
- **Resultados Pruebas 20% de base de datos** Anexo 21-Anexo 36
- **Resultados Pruebas 50% de base de datos** Anexo 37-Anexo 45
- **Resultados Pruebas 100% de base de datos** Anexo 46-Anexo 55

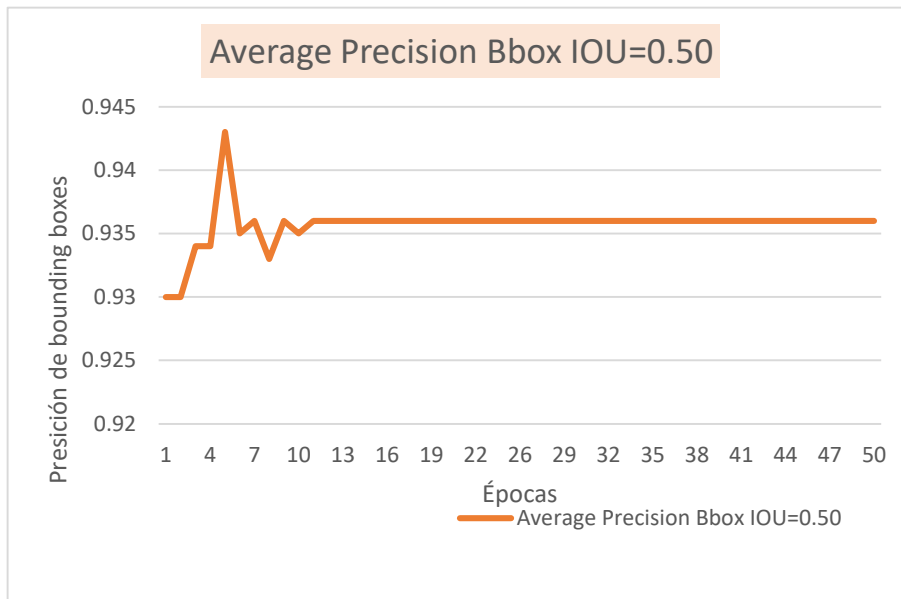
Anexo 1 Etiquetas de esquinas de cada vértebra para entrenamiento



Fuente: Elaboración propia obtenida por Roboflow

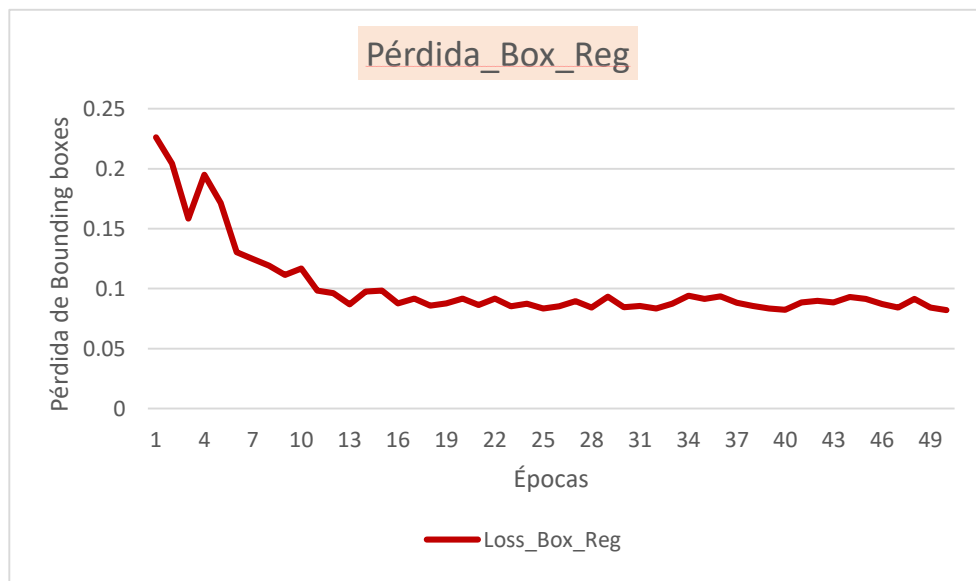
- **Resultados Primeras Pruebas-Validación**

Anexo 2 Precisión promedio de validación prueba 1



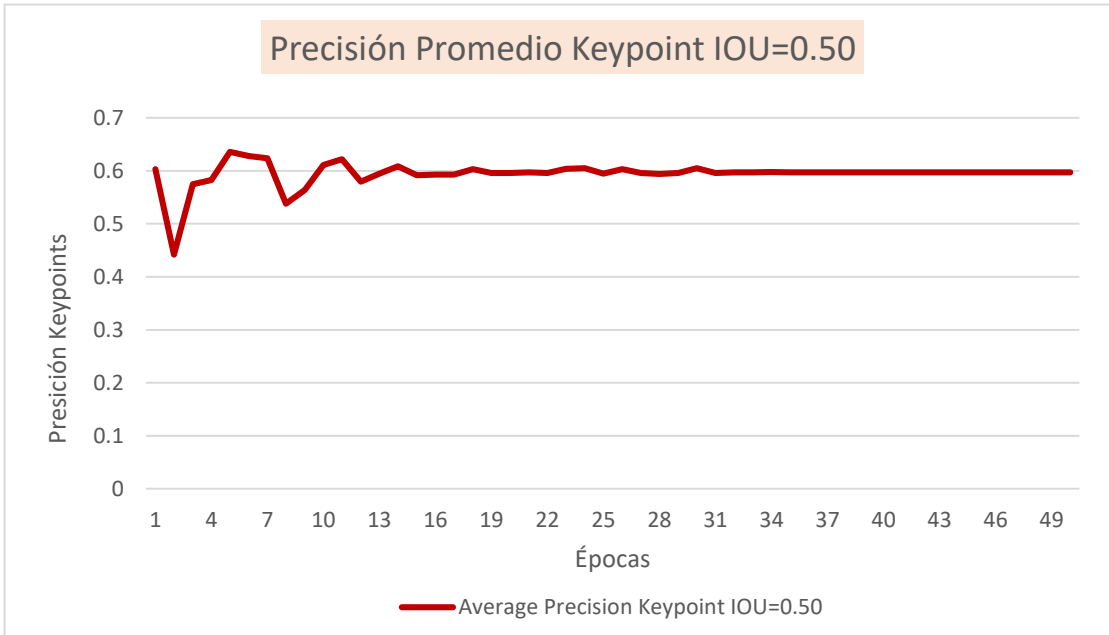
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 3 Pérdida de caja de Regulación de validación prueba 1



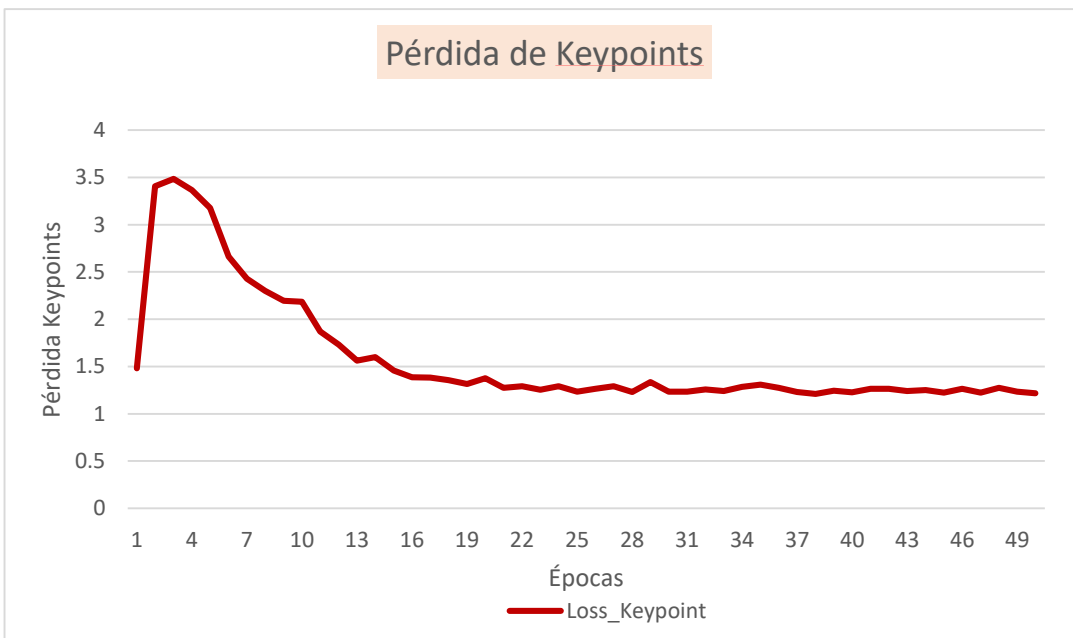
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 4 Precisión de Keypoints de validación prueba 1



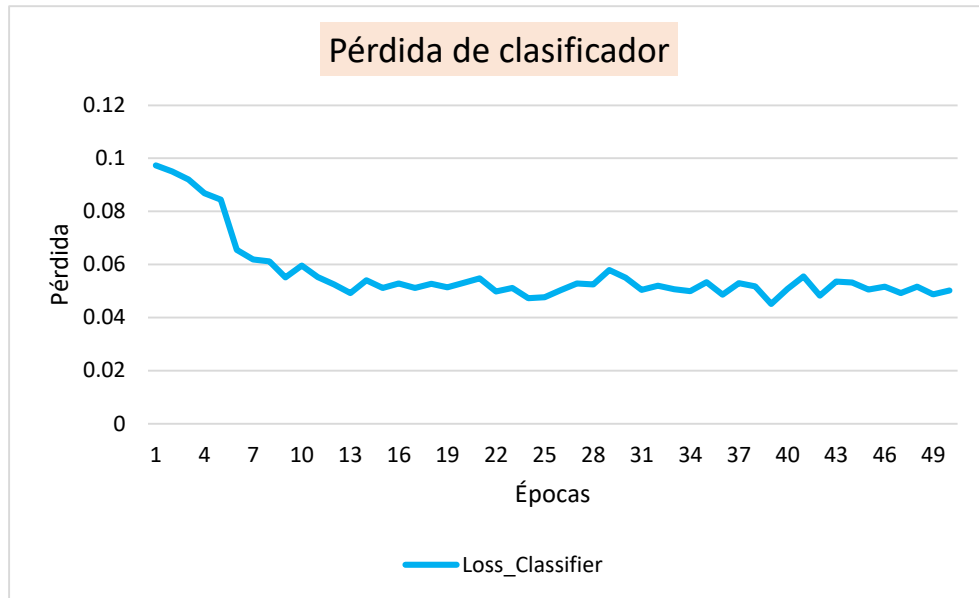
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 5 Pérdida de detección de Keypoints de validación prueba 1



Fuente: Elaboración propia obtenida por código de Google Colab.

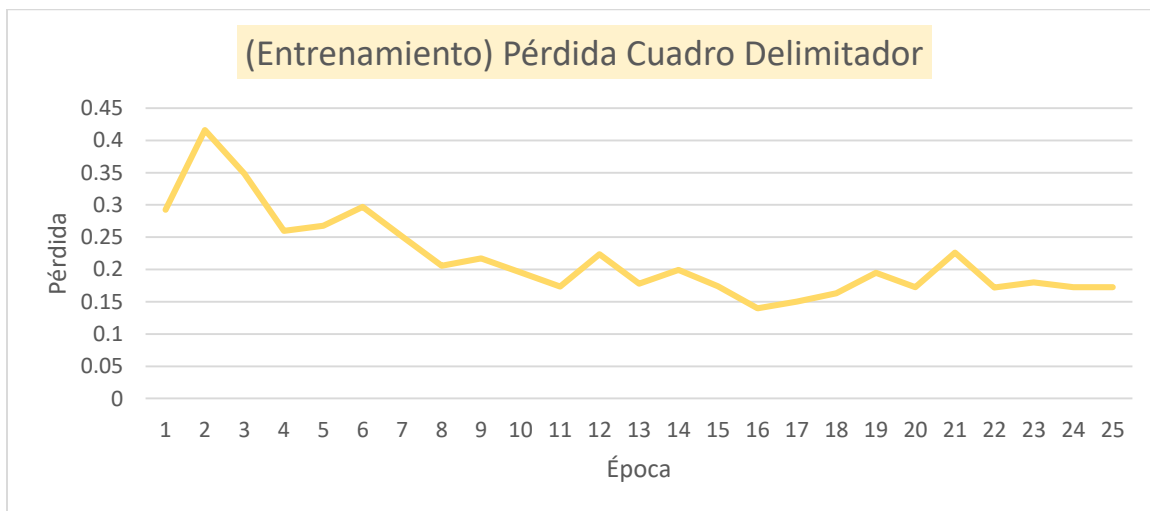
Anexo 6 Validación Pérdida de Clasificador prueba 1- validación



Fuente: Elaboración propia obtenida por código de Google Colab.

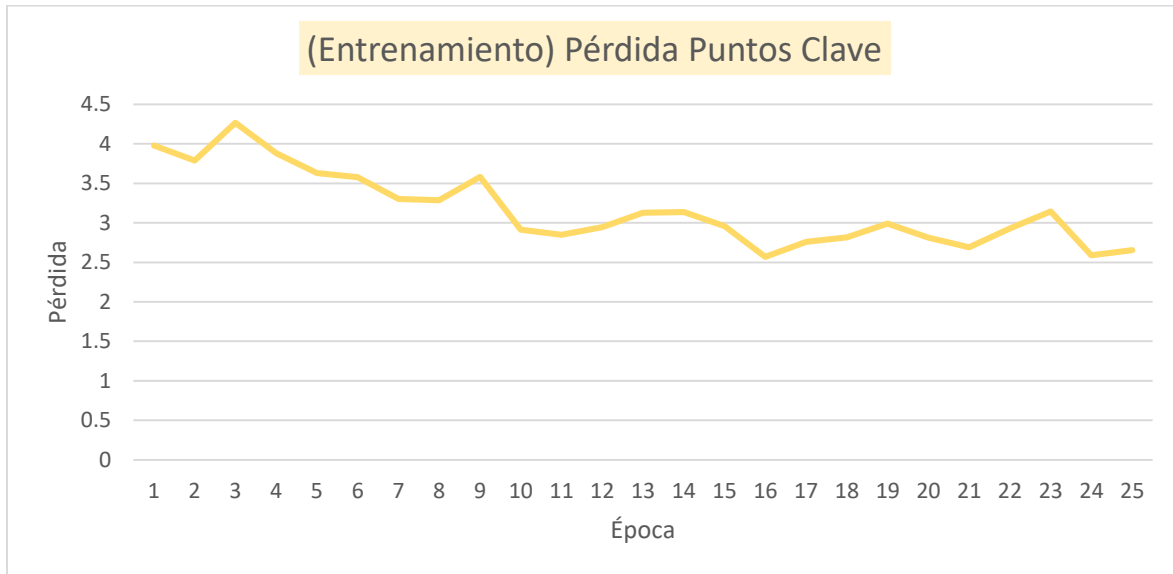
- **Entrenamiento 25 épocas**

Anexo 7 Pérdida de cajas reguladoras KeyPoint -Entrenamiento



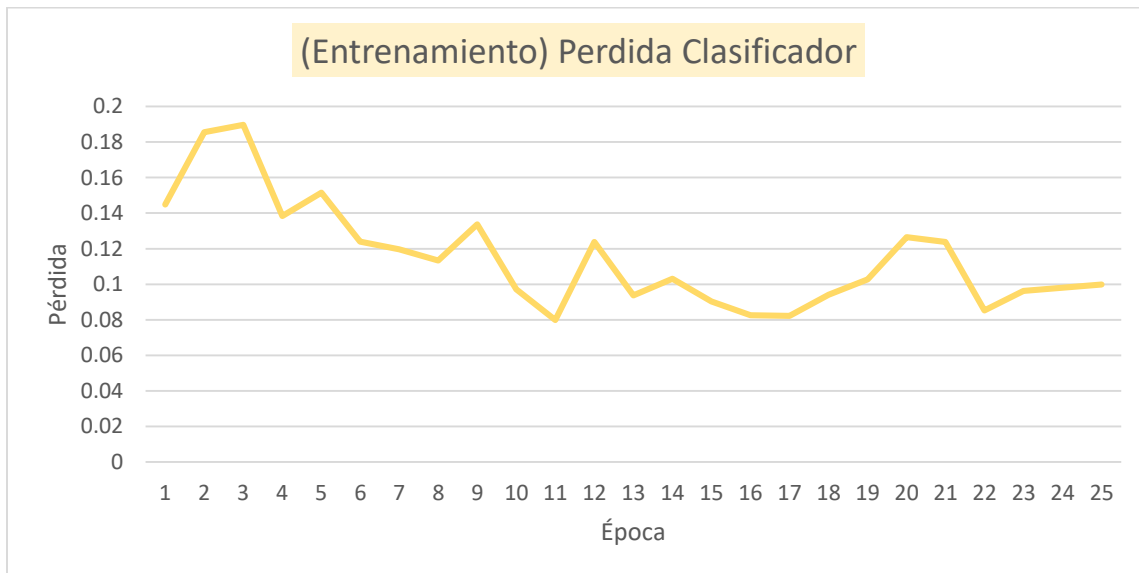
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 8 Pérdida de Keypoints modelo KeyPoint- Entrenamiento



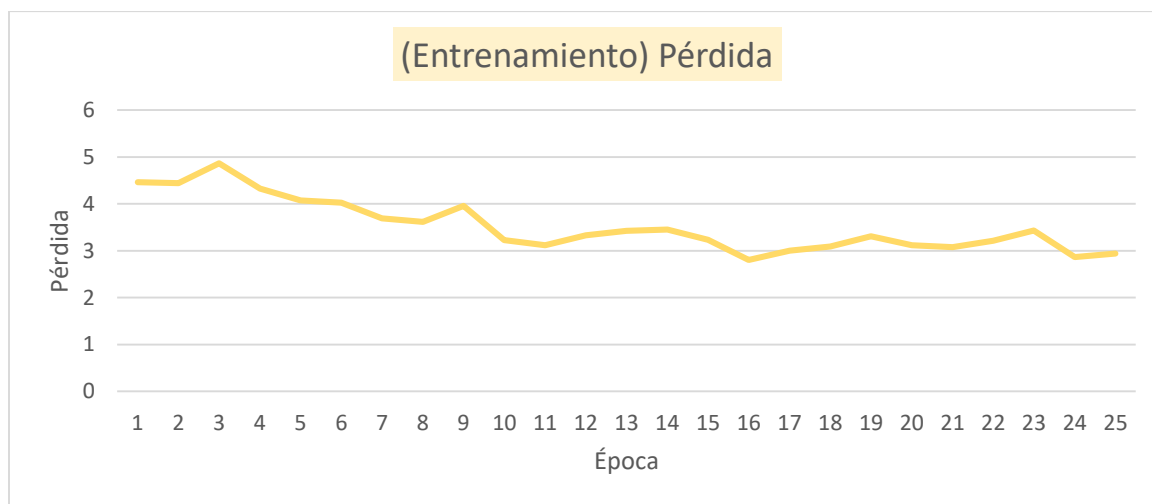
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 9 Pérdida de Clasificador modelo KeyPoint- Entrenamiento



Fuente: Elaboración propia obtenida por código de Google Colab.

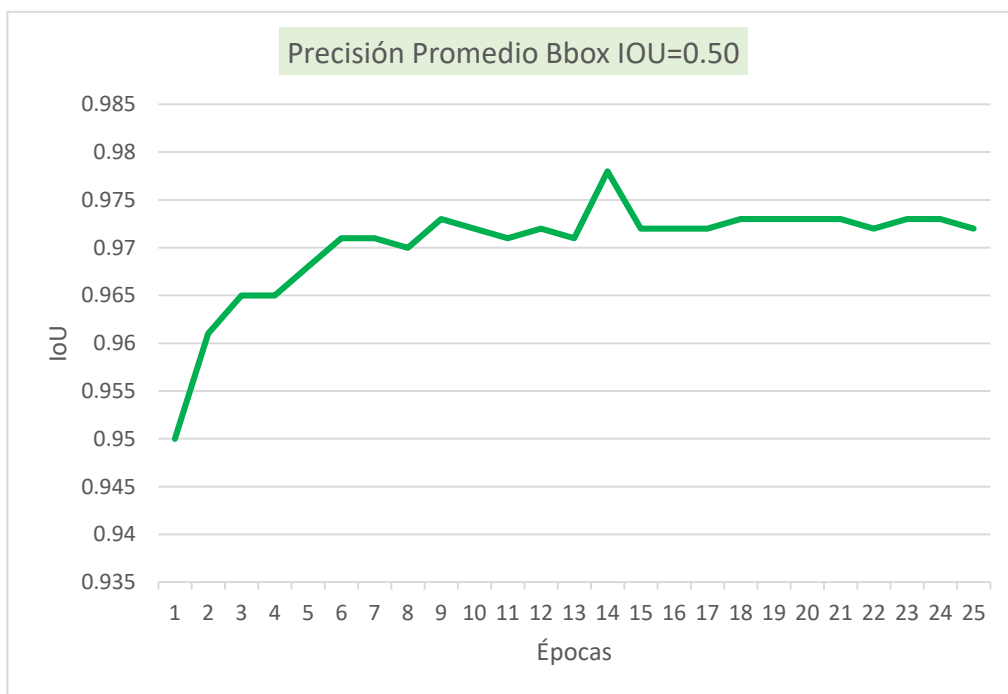
Anexo 10 Pérdida de total modelo KeyPoint- Entrenamiento



Fuente: Elaboración propia obtenida por código de Google Colab.

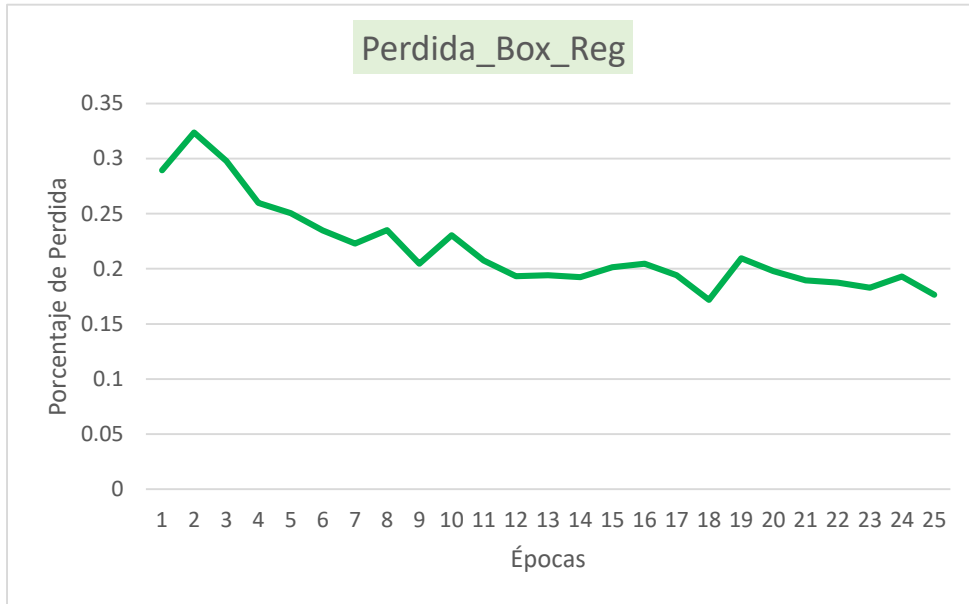
- **Validación 25 épocas**

Anexo 11 Precisión Promedio 25 Épocas- validación



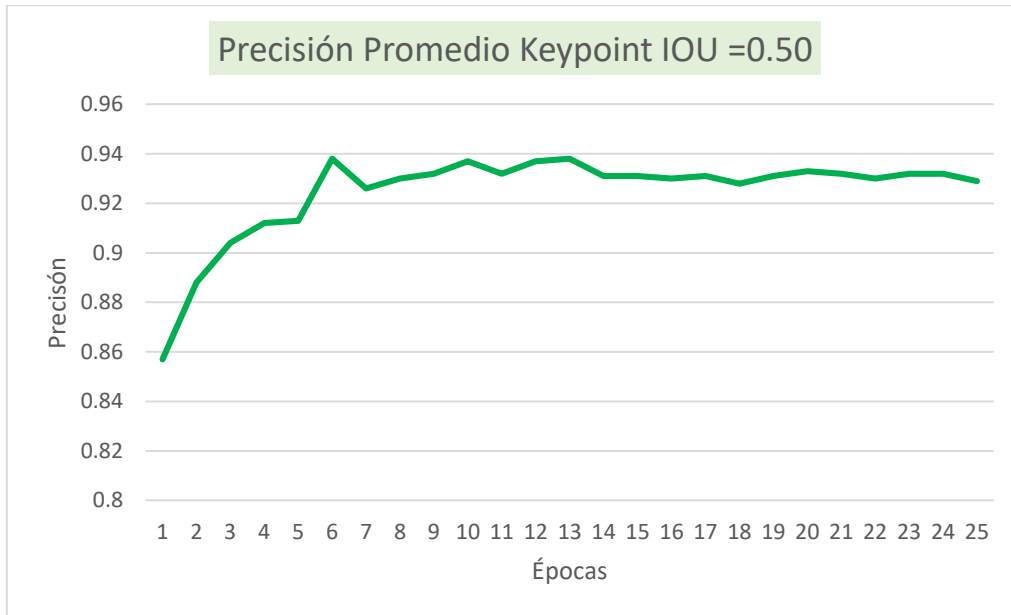
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 12 Pérdida detección Bounding boxes 25 Épocas - validación



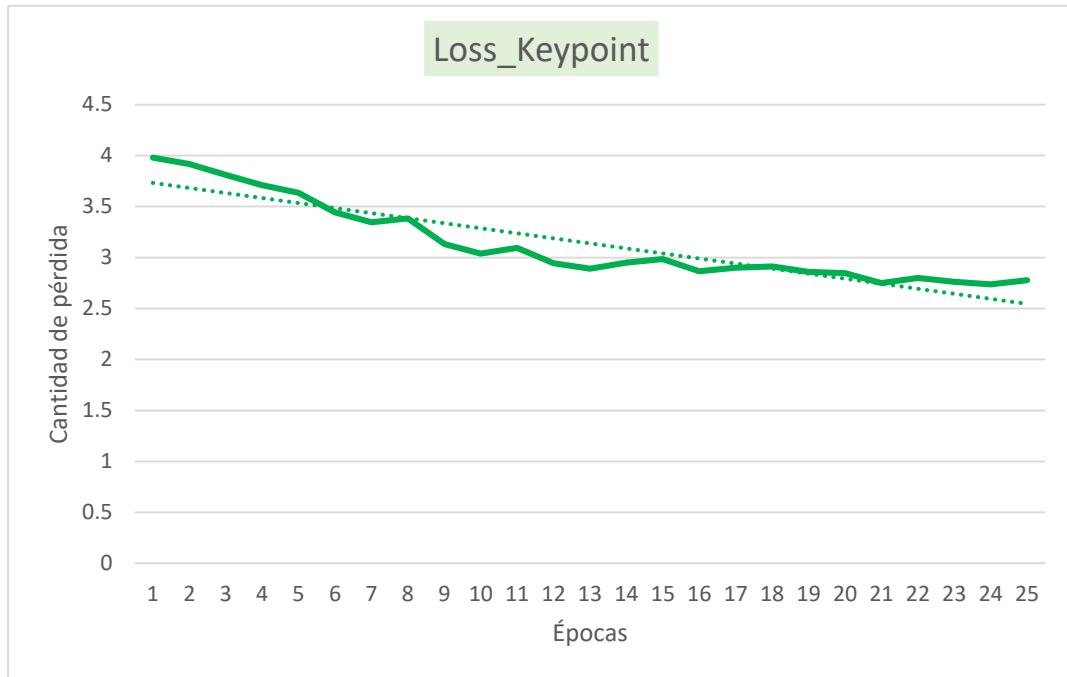
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 13 Precisión de detección de Keypoints 25 Épocas- validación



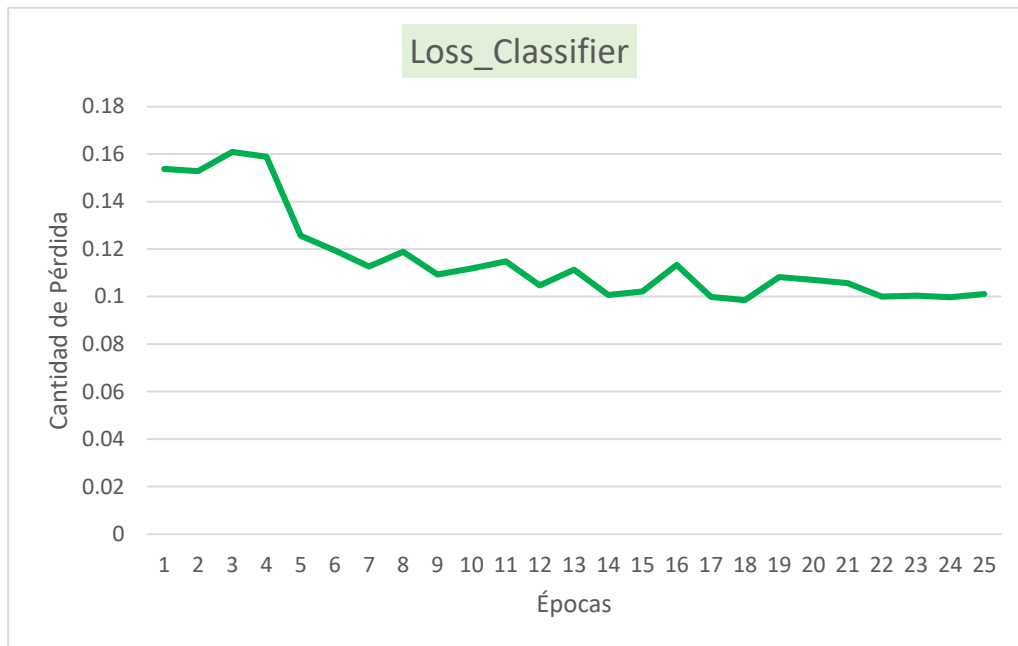
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 14 Pérdida de Keypoints 25 Épocas- validación



Fuente: Elaboración propia obtenida por código de Google Colab.

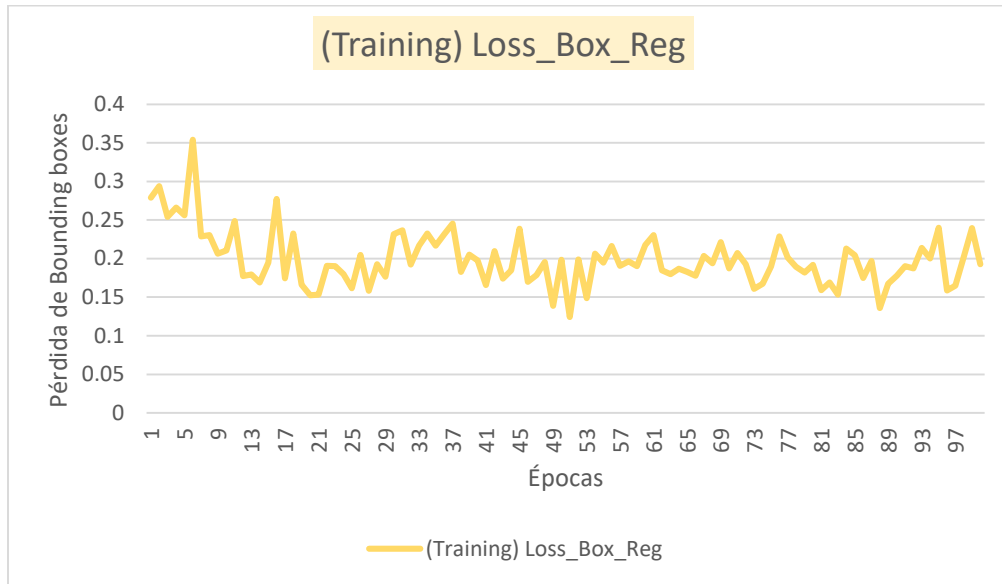
Anexo 15 Pérdida de Clasificador KeyPoint 25 Épocas- validación



Fuente: Elaboración propia obtenida por código de Google Colab.

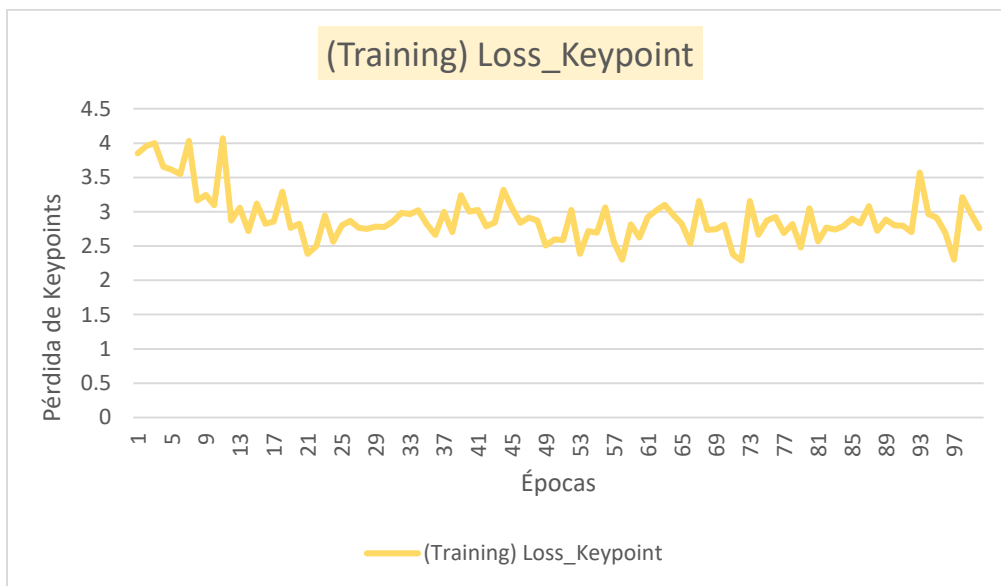
- **Entrenamiento 100 épocas**

Anexo 16 Pérdida de cajas reguladoras KeyPoint -Entrenamiento



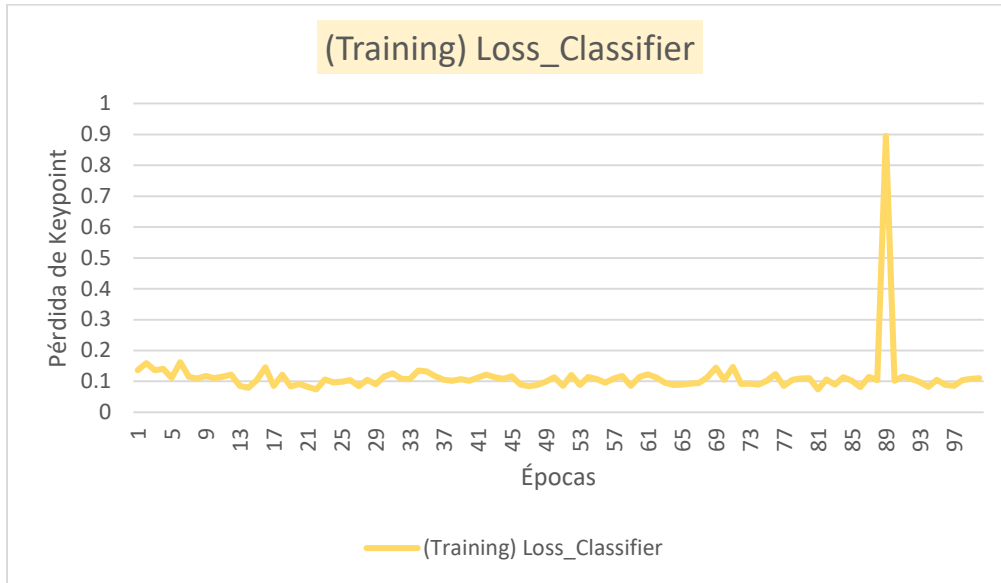
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 17 Pérdida de Keypoints modelo KeyPoint- Entrenamiento



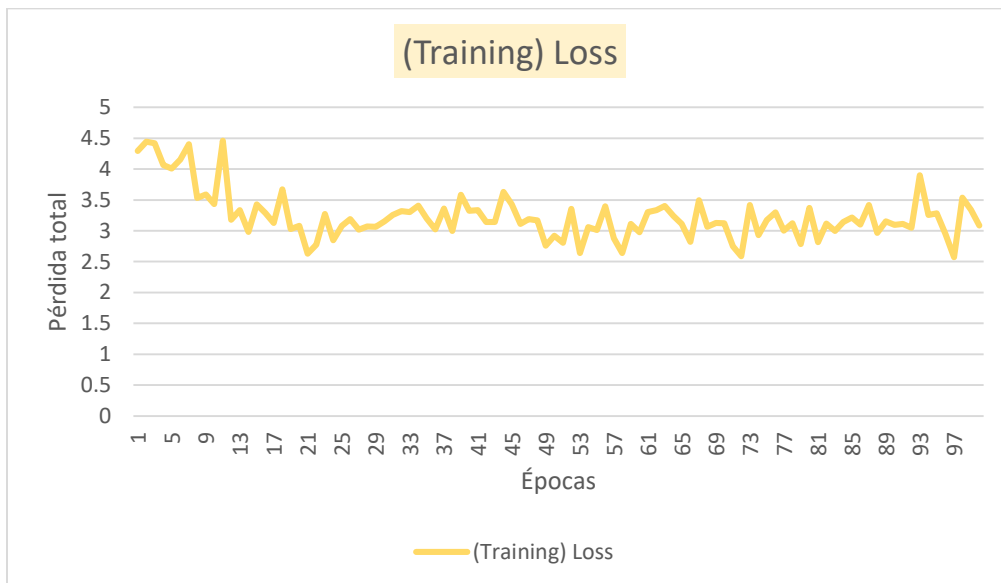
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 18 Pérdida de Keypoints modelo KeyPoint- Entrenamiento



Fuente: Elaboración propia obtenida por código de Google Colab.

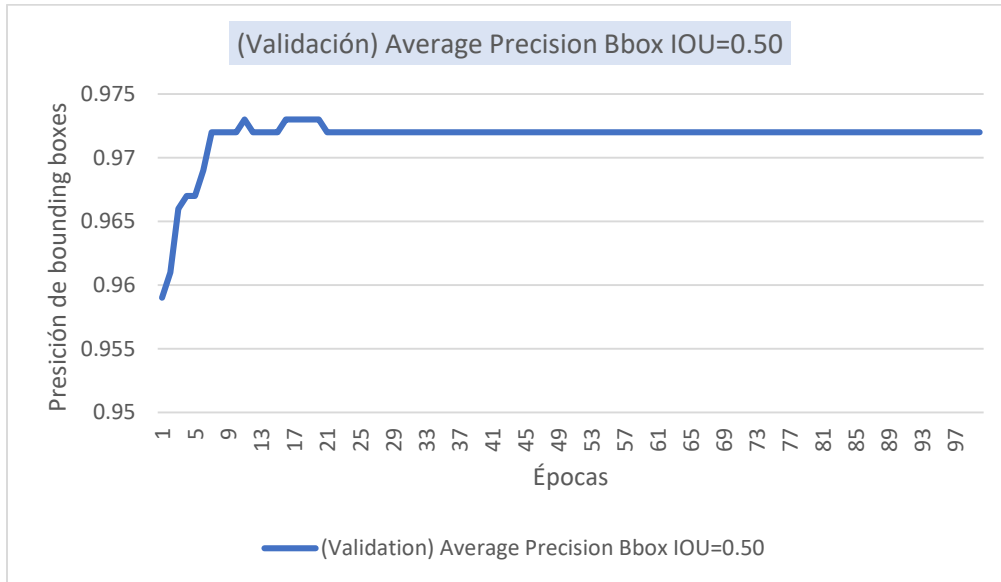
Anexo 19 Pérdida total de modelo KeyPoint- Entrenamiento



Fuente: Elaboración propia obtenida por código de Google Colab.

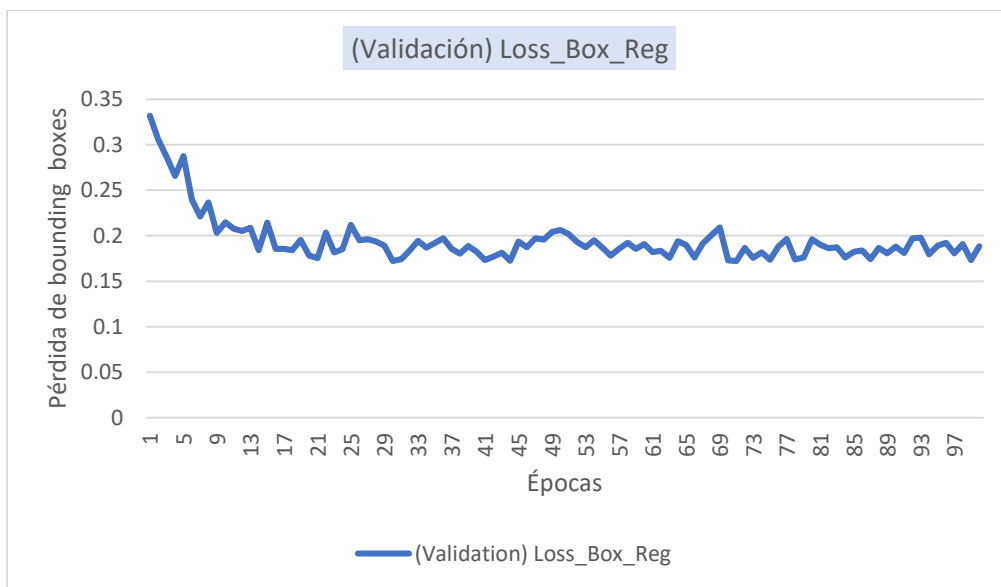
- **Validación 100 épocas**

Anexo 20 Precisión Promedio KeyPoint- Validación



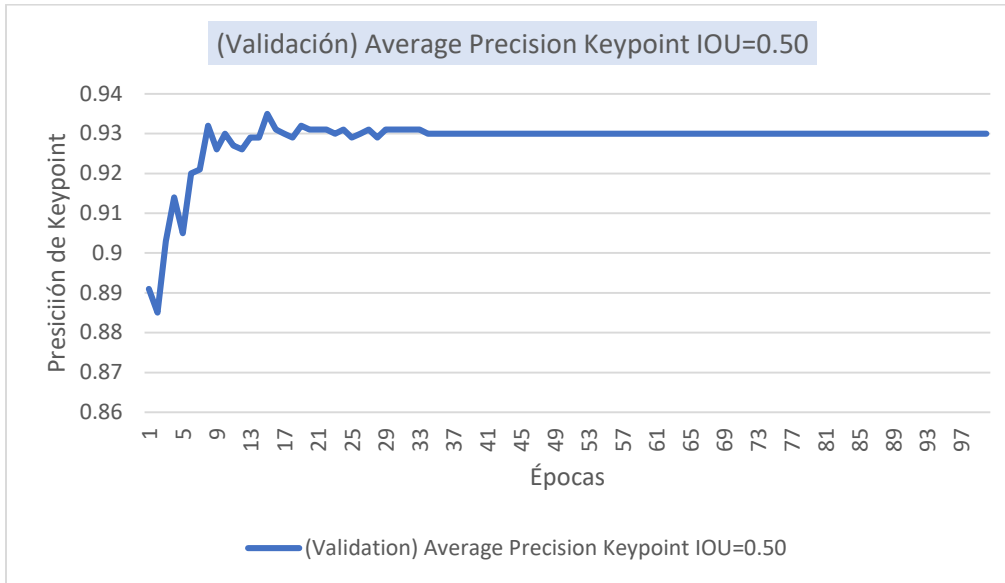
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 21 Pérdida de cajas reguladoras KeyPoint - Validación



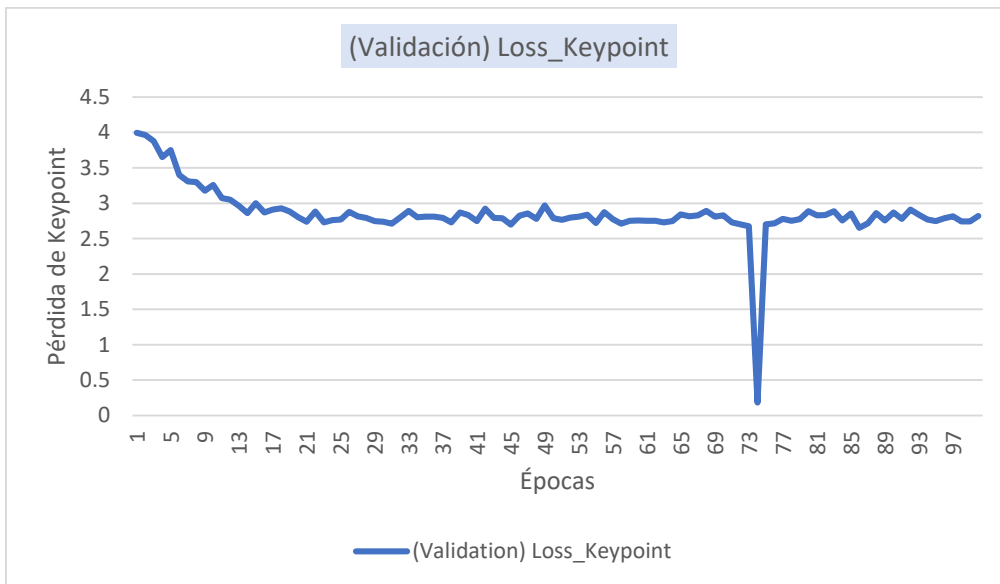
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 22 Precisión de Keypoints KeyPoint- Validación



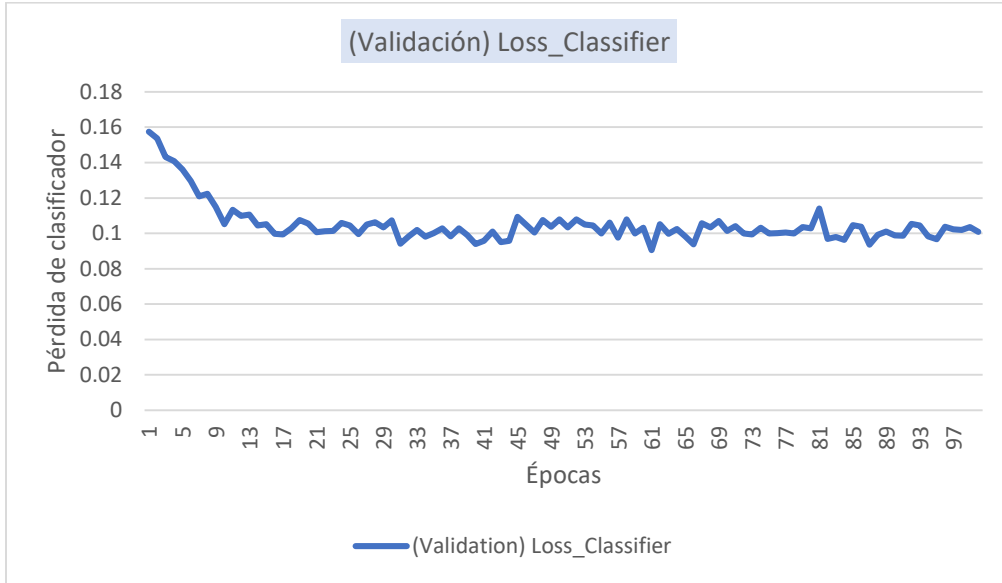
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 23 Pérdida de Keypoints modelo KeyPoint- Validación



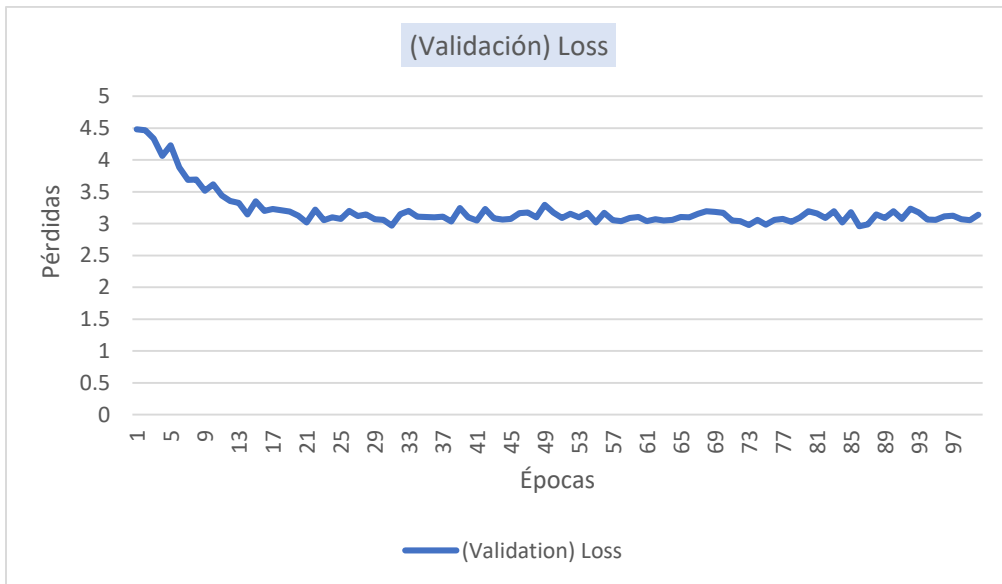
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 24 Pérdida de Clasificador modelo Keypoint- Validación



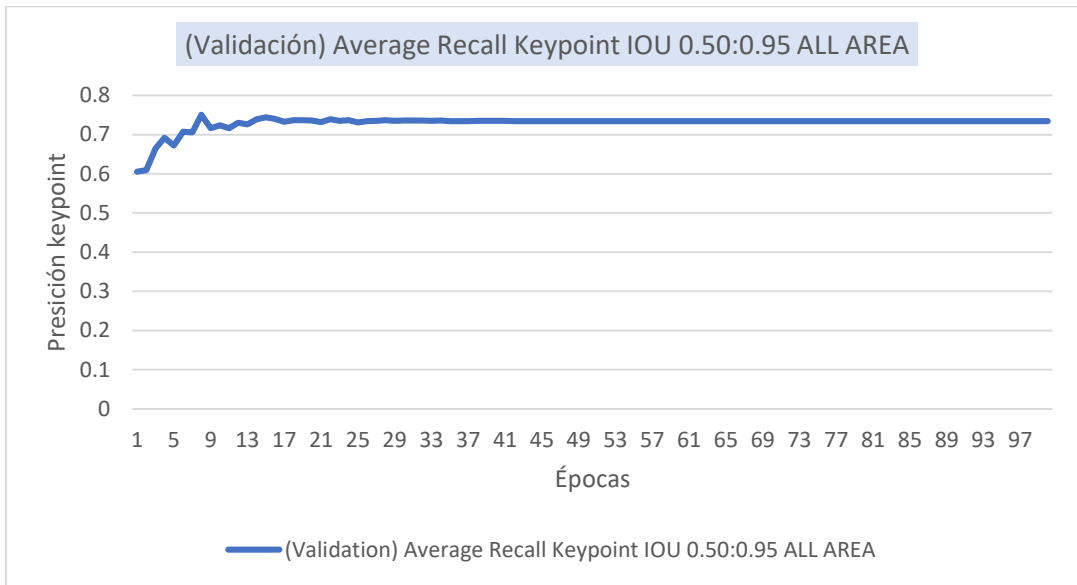
Fuente: Elaboración propia obtenida por código de Google Colab.

Anexo 25 Pérdida total de modelo Keypoint- Validación



Fuente: Elaboración propia obtenida por código de Google Colab.

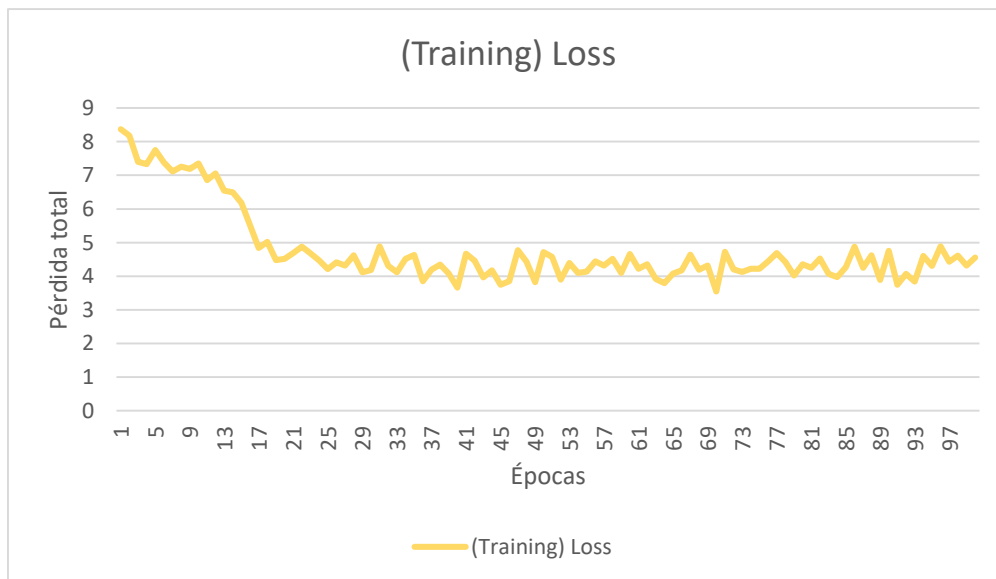
Anexo 26 Pérdida exhaustividad de modelo KeyPoint- Validación



Fuente: Elaboración propia obtenida por código de Google Colab

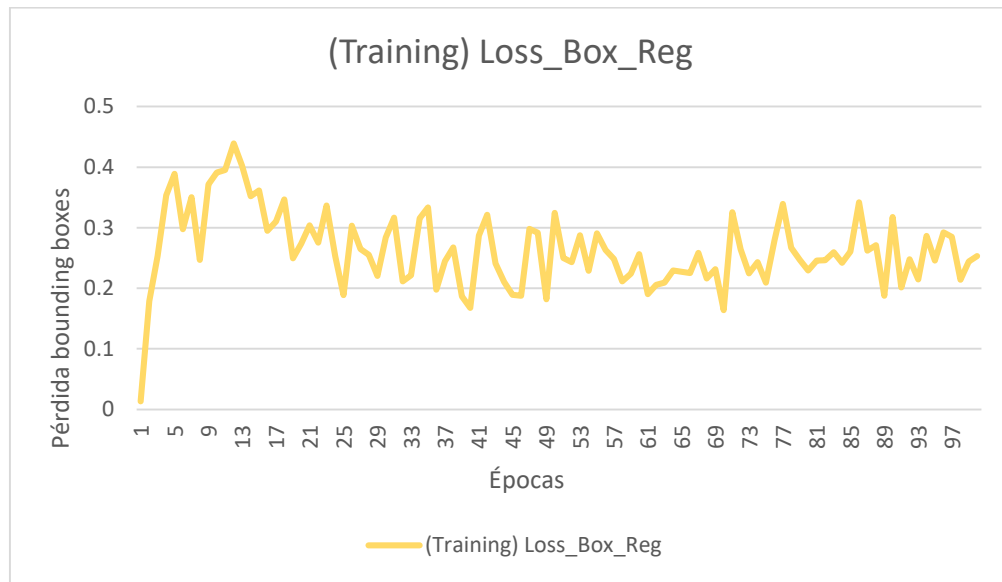
- **Resultados Pruebas porcentajes de base de datos -Validación**

Anexo 27 Pérdida con 20% de base de datos- entrenamiento



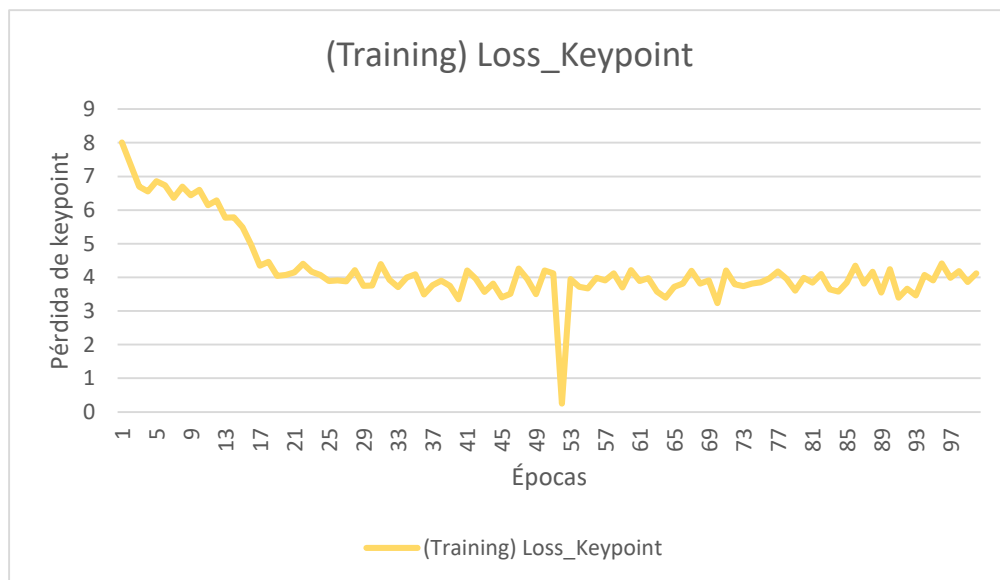
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 28 Pérdida de bounding boxes con 20% de base de datos entrenamiento



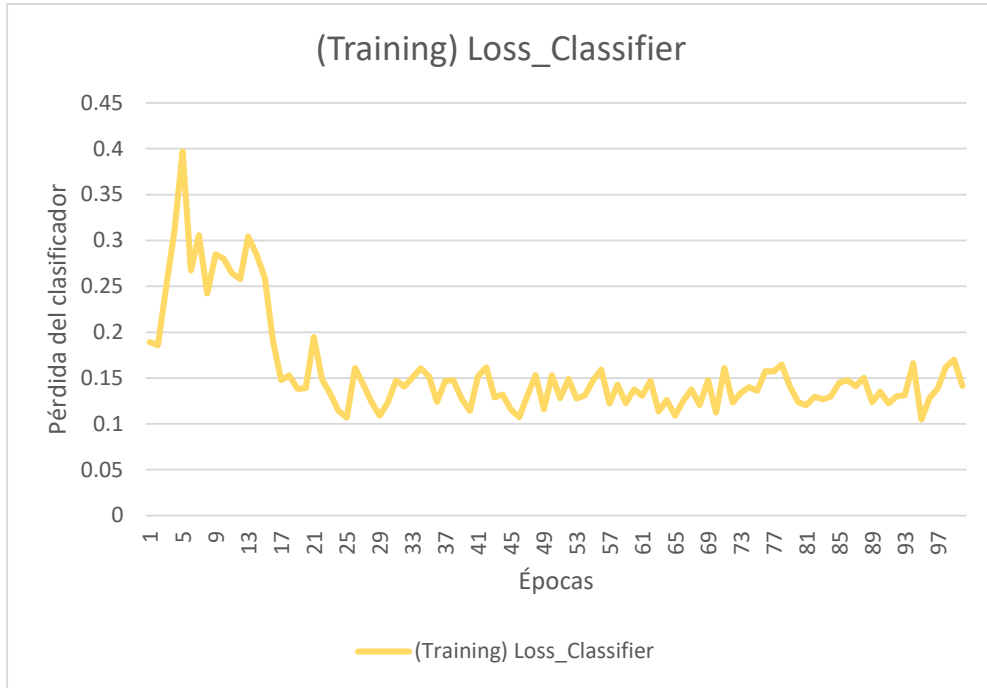
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 29 Pérdida de KeyPoints con 20% de base de datos- entrenamiento



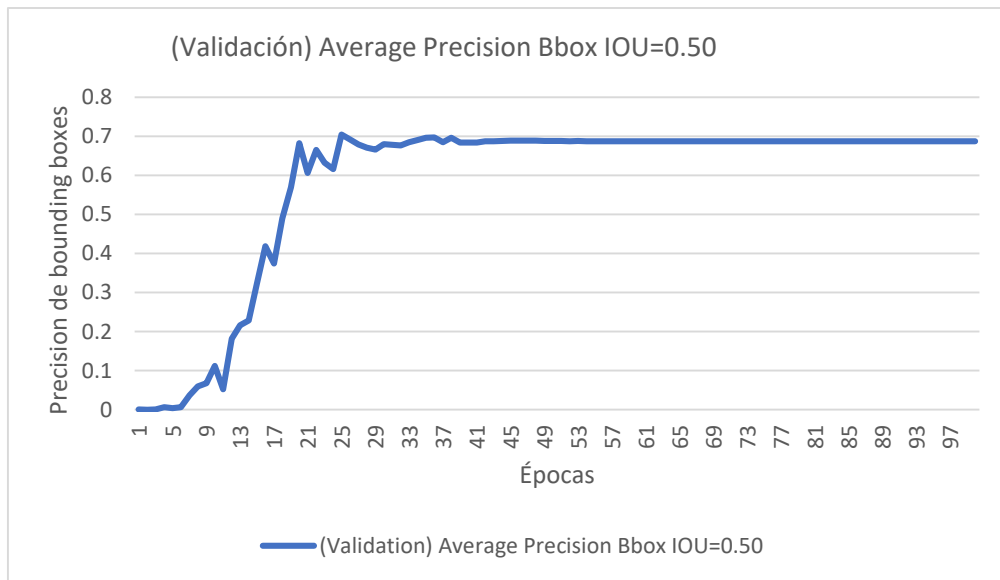
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 30 Pérdida de clasificador con 20% de base de datos- entrenamiento



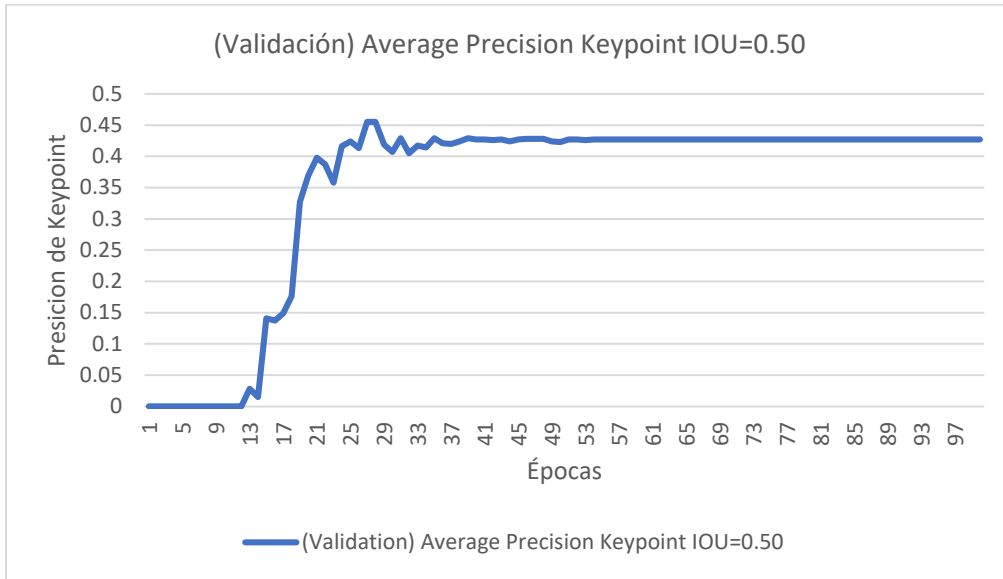
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 31 Precisión promedio de bounding boxes de modelo KeyPoint- Validación



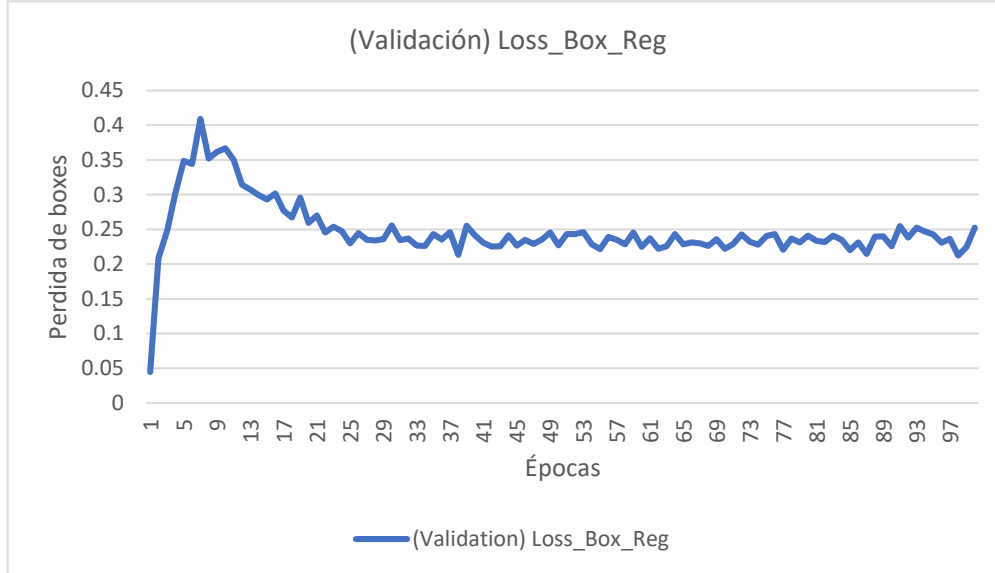
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 32 Precisión promedio de Keypoints de modelo Keypoint- Validación



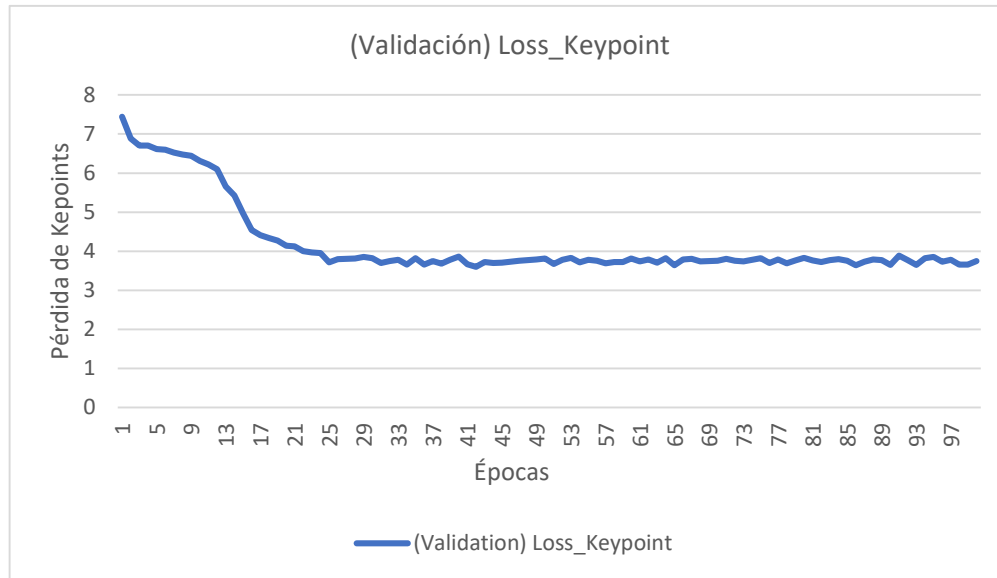
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 33 Pérdida de bounding boxes de modelo Keypoint- Validación



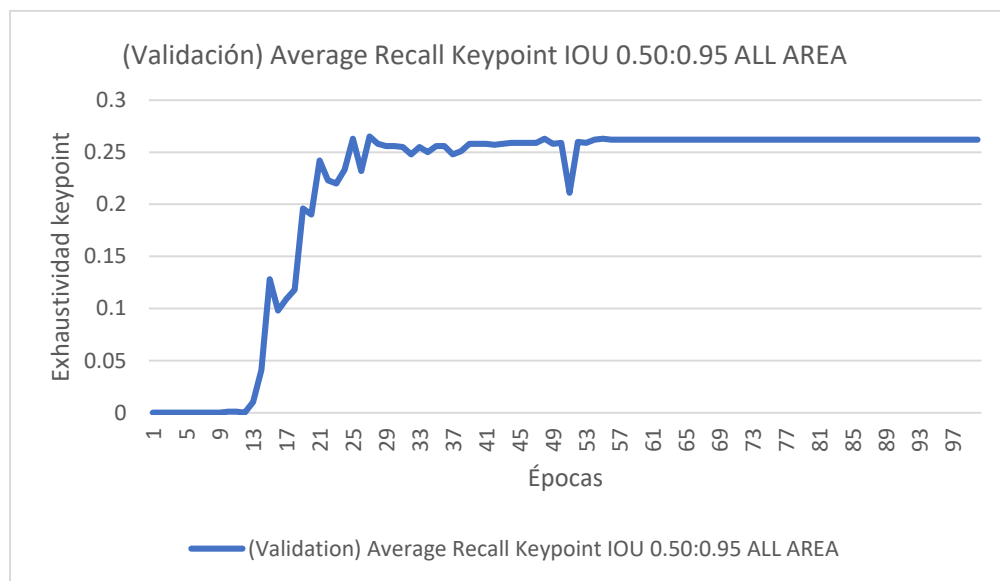
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 34 Pérdida de Keypoints modelo Keypoint- Validación



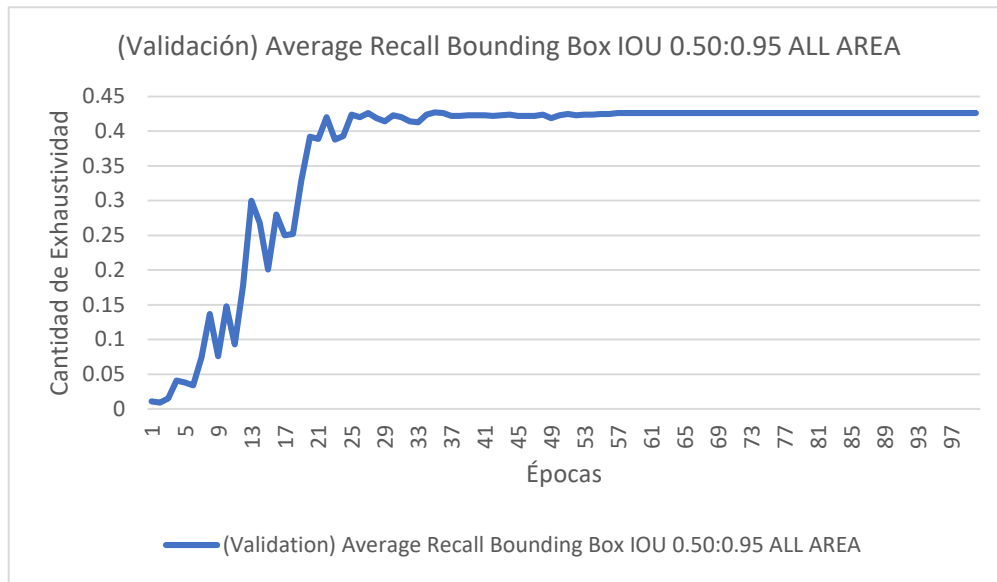
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 35 Exhaustividad de Keypoints de modelo Keypoint- Validación



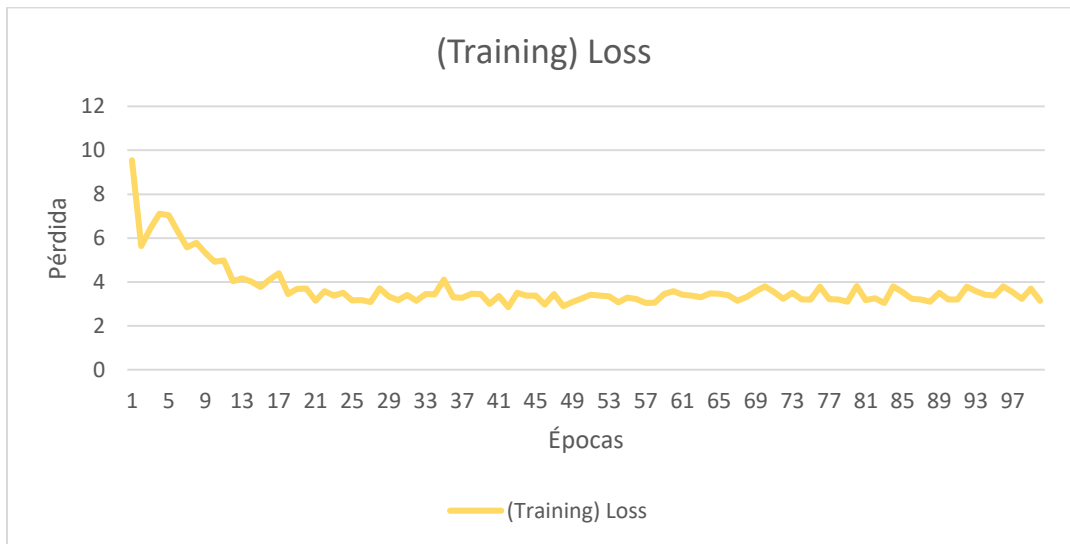
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 36 Exhaustividad de bounding boxes de modelo KeyPoint- Validación



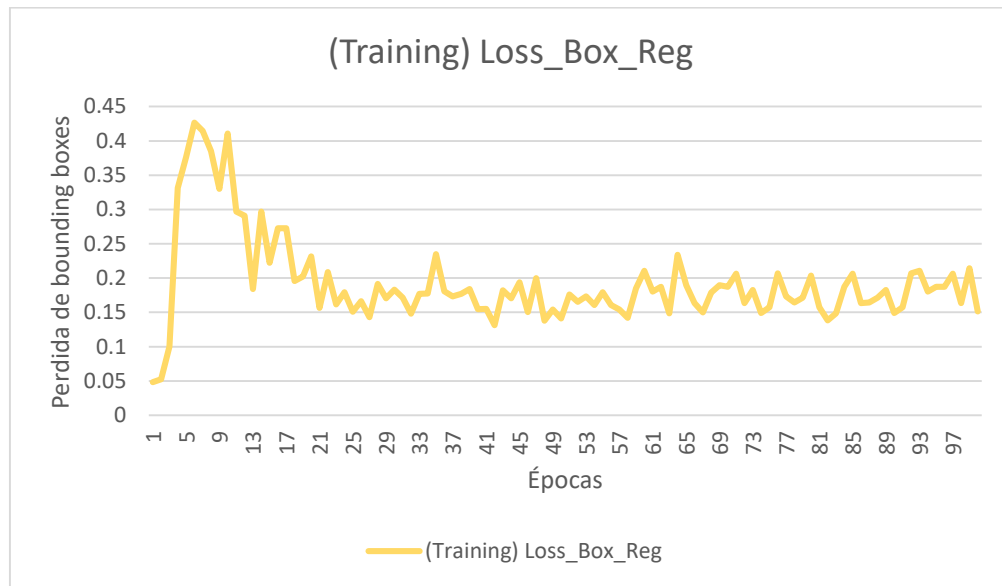
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 37 Pérdida con 50% de base de datos- entrenamiento



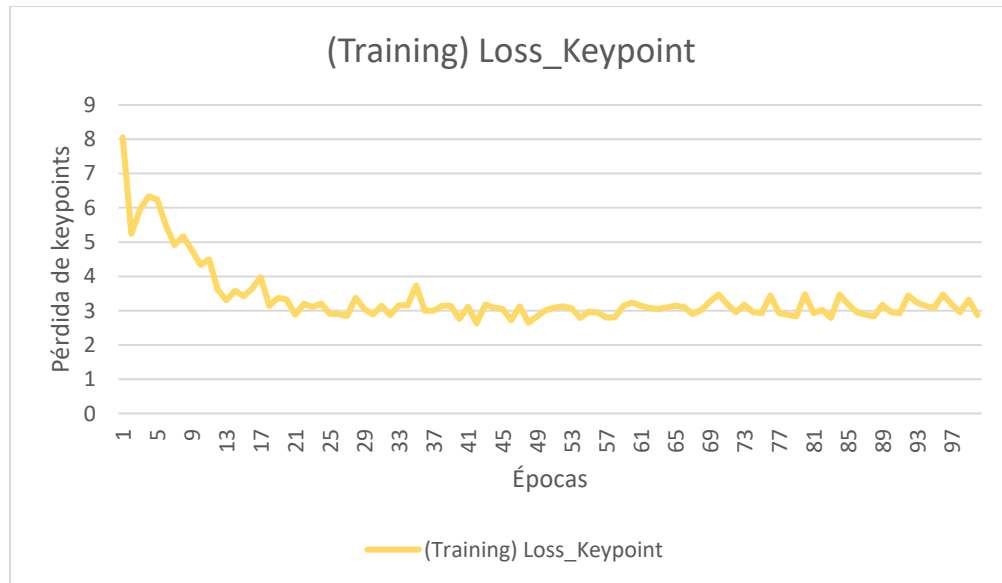
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 38 Pérdida de bounding boxes con 50% de base de datos entrenamiento



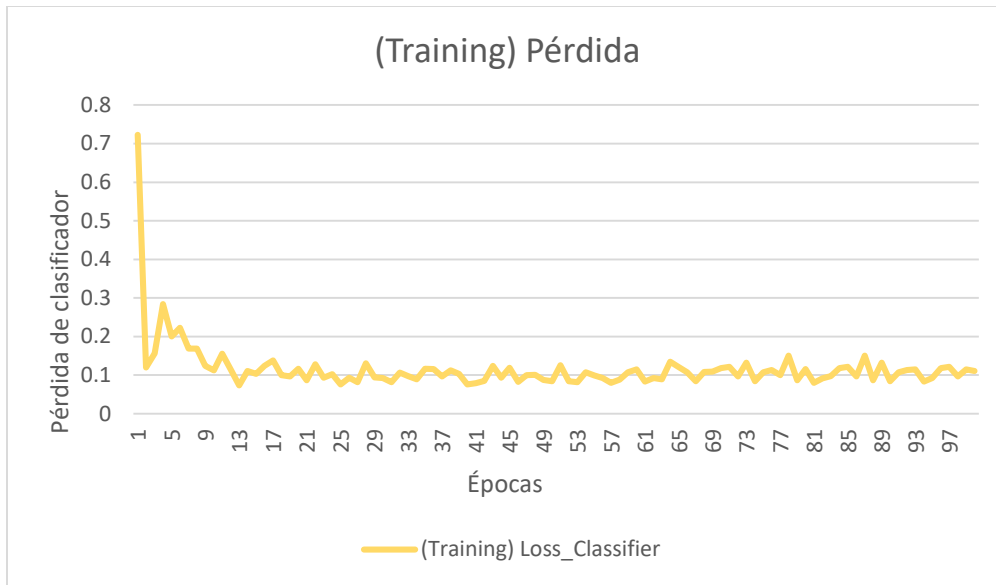
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 39 Pérdida de KeyPoints con 50% de base de datos- entrenamiento



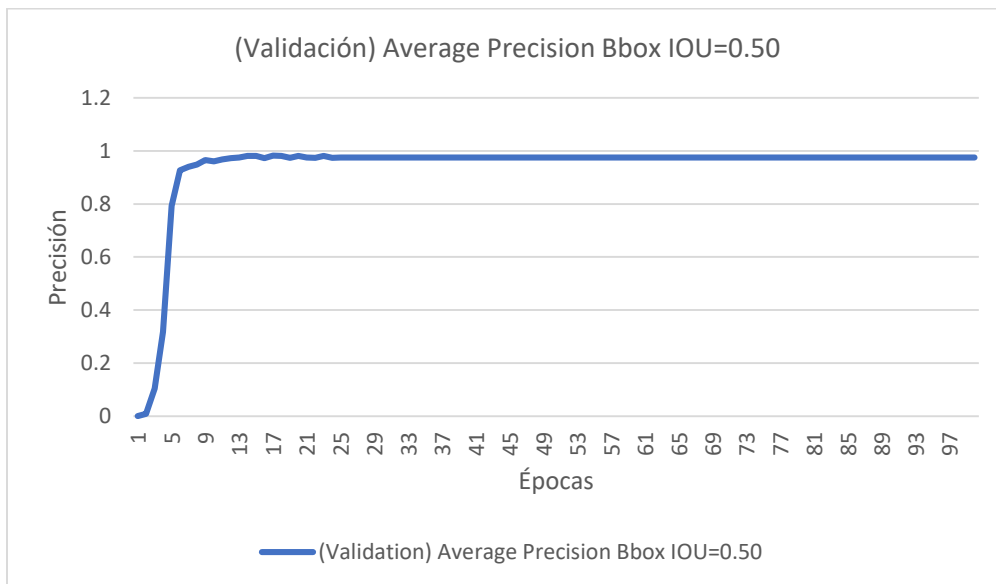
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 40 Pérdida de clasificador con 50% de base de datos- entrenamiento



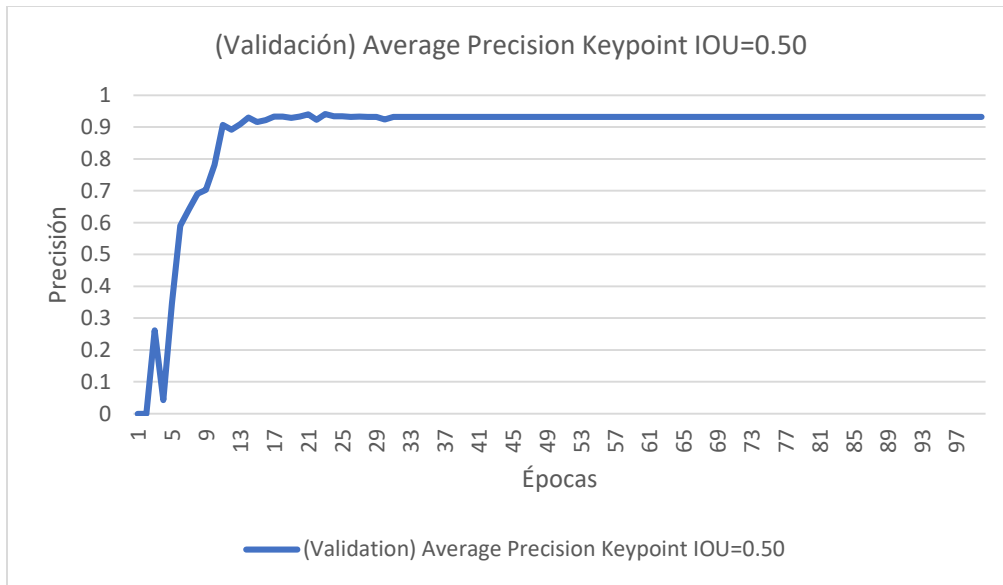
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 41 Precisión promedio de bounding boxes con 50% de base de datos - Validación



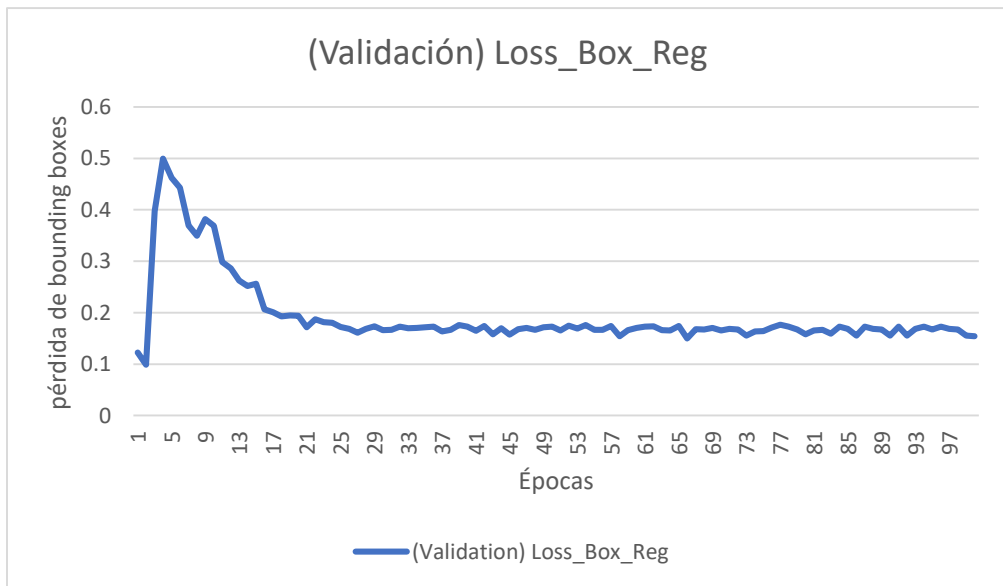
Elaboración propia obtenida por código de Google Colab

Anexo 42 Precisión promedio de Keypoints con 50% de base de datos - Validación



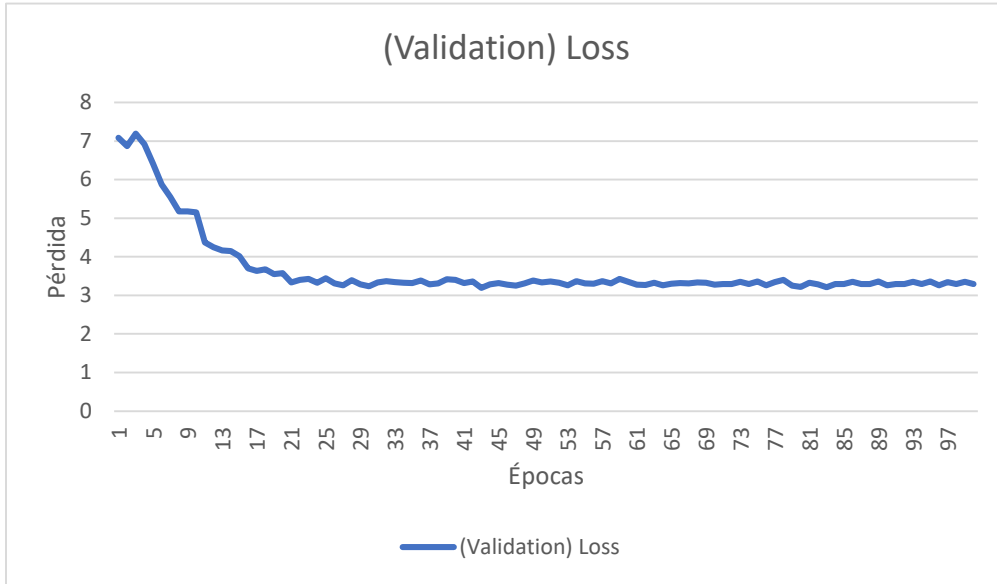
Elaboración propia obtenida por código de Google Colab

Anexo 43 Pérdida de bounding boxes con 50% de base de datos - Validación



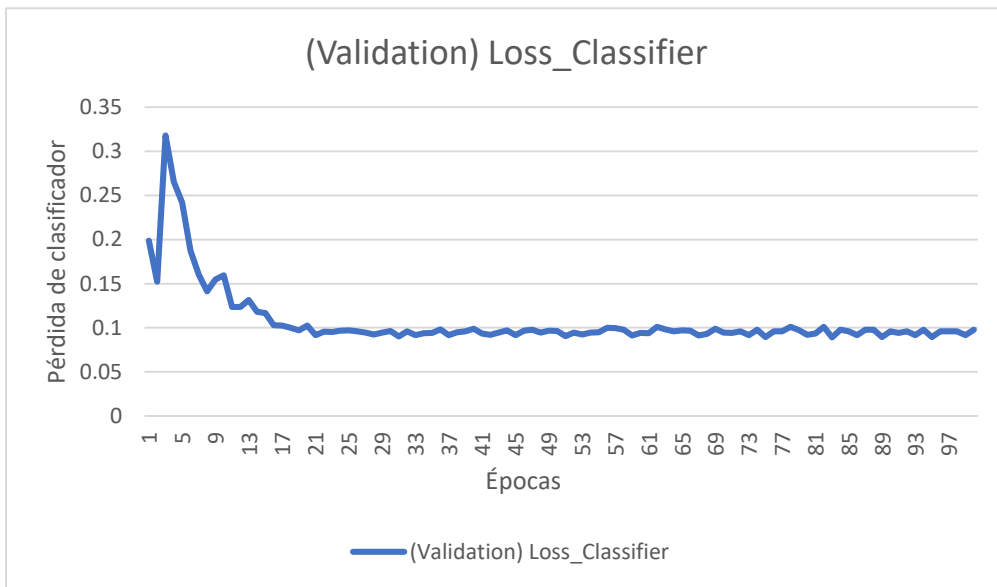
Elaboración propia obtenida por código de Google Colab

Anexo 44 Pérdida con 50% de base de datos modelo KeyPoint- Validación



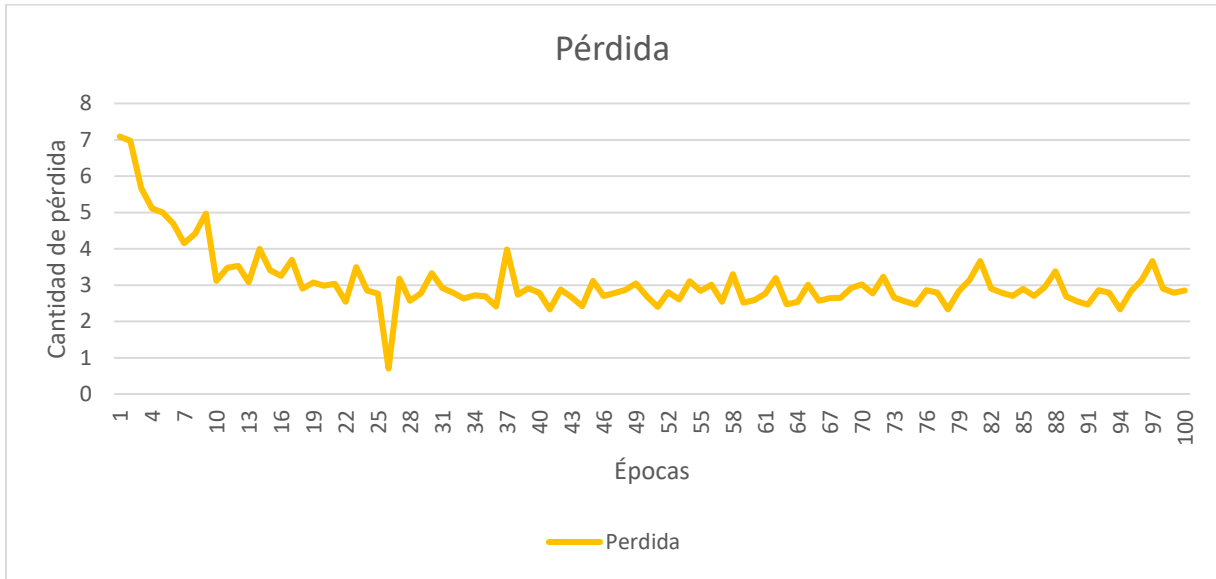
Elaboración propia obtenida por código de Google Colab

Anexo 45 Pérdida de clasificador con 50% de base de datos modelo KeyPoint- Validación



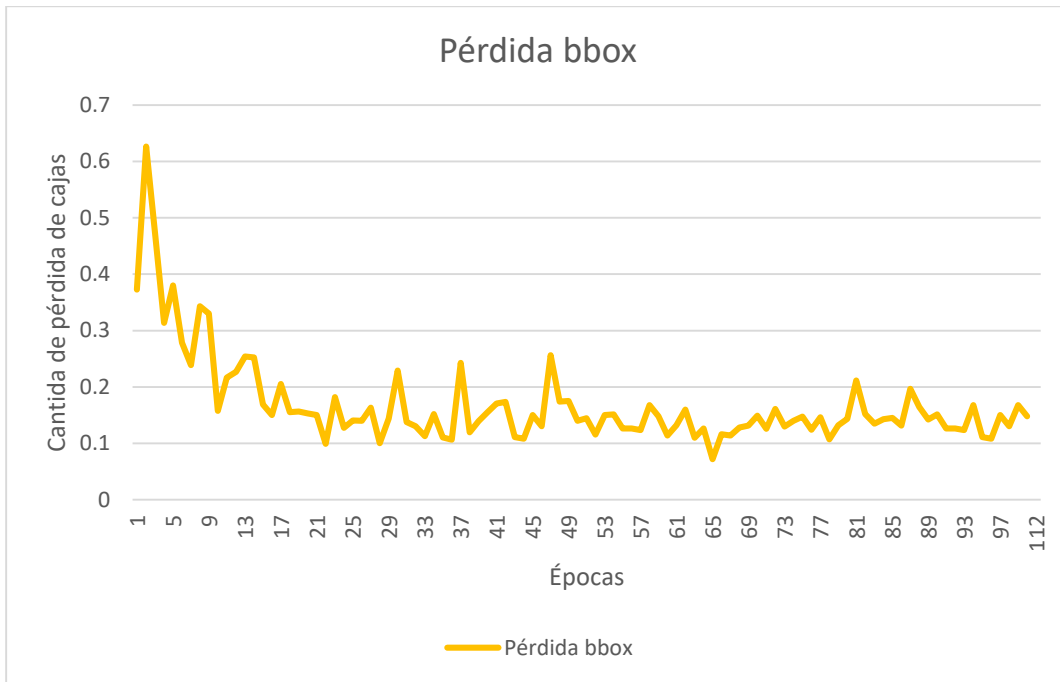
Elaboración propia obtenida por código de Google Colab

Anexo 46 Pérdida con 100% de base de datos- entrenamiento



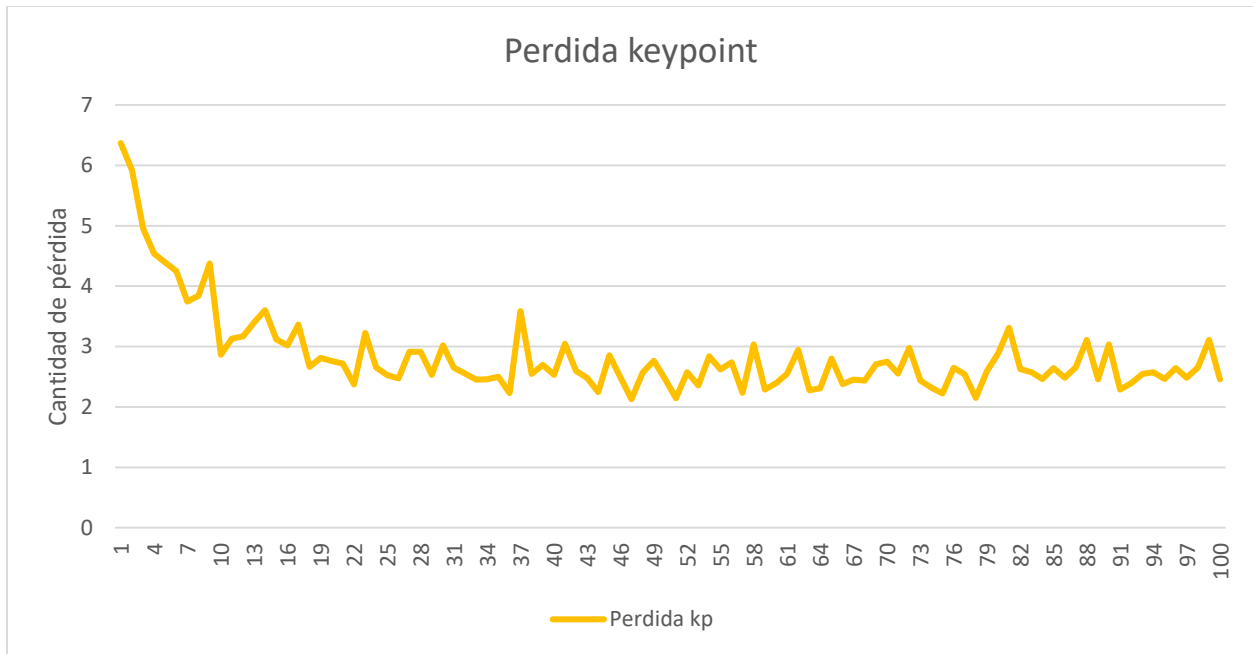
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 47 Pérdida de bounding boxes con 100% de base de datos entrenamiento



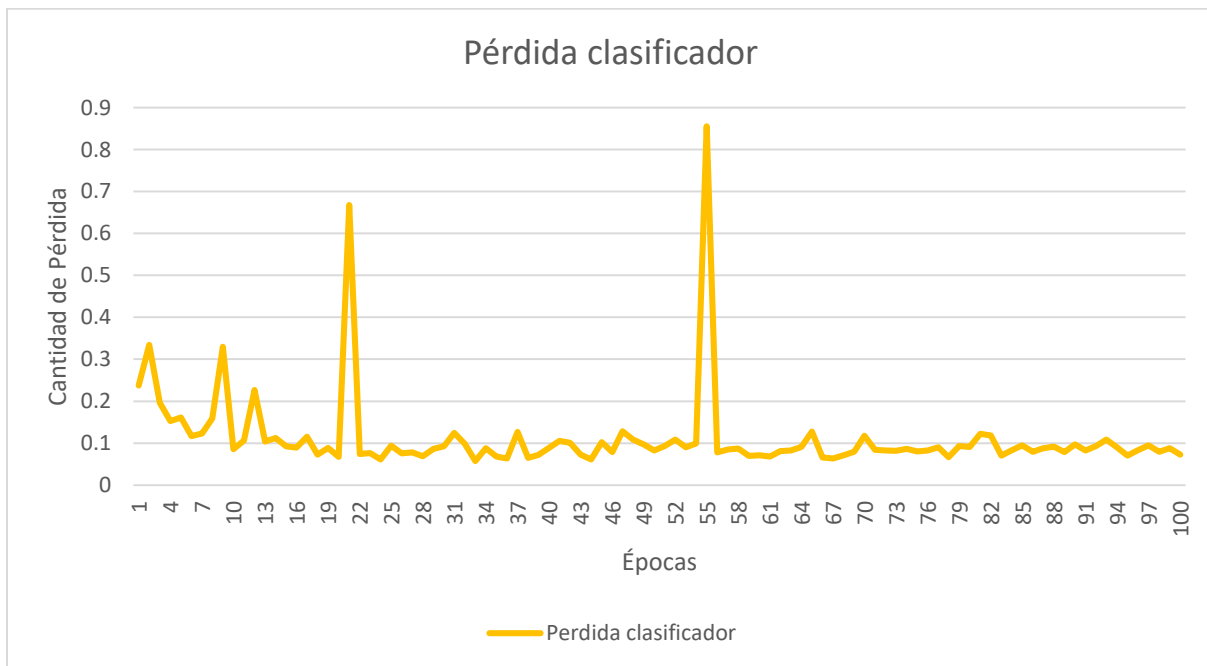
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 48 Pérdida de KeyPoints con 100% de base de datos- entrenamiento



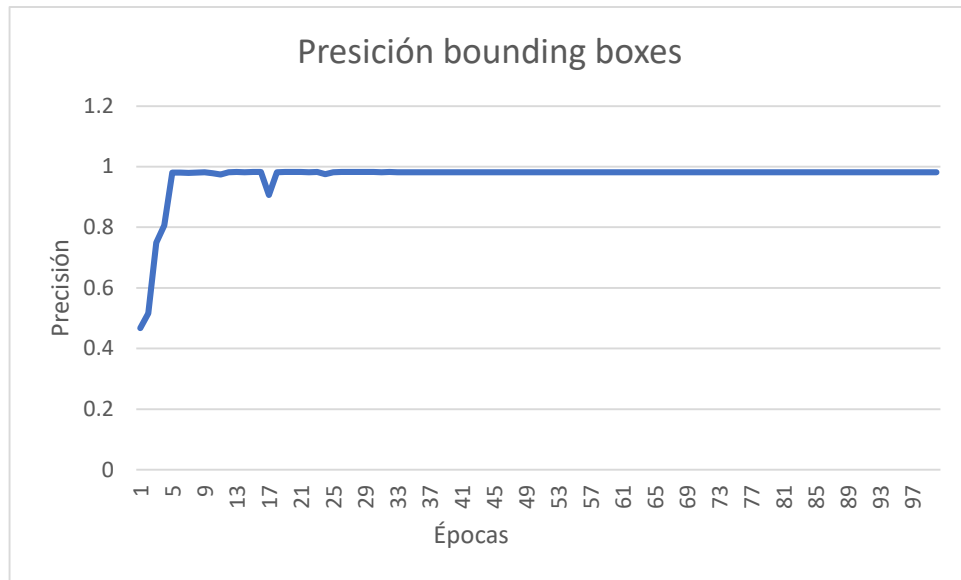
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 49 Pérdida de clasificador con 100% de base de datos- entrenamiento



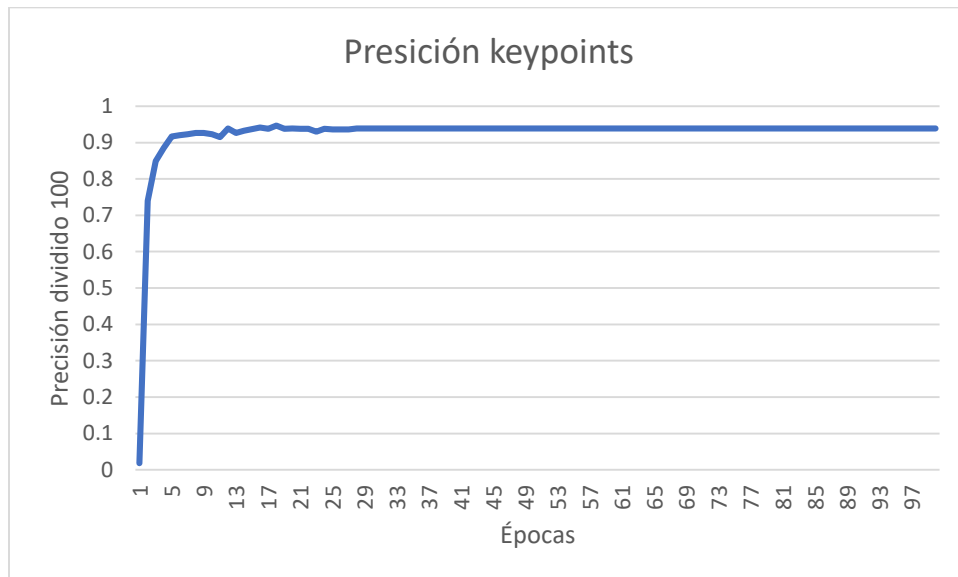
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 50 Precisión promedio de bounding boxes con 100% de base de datos - Validación



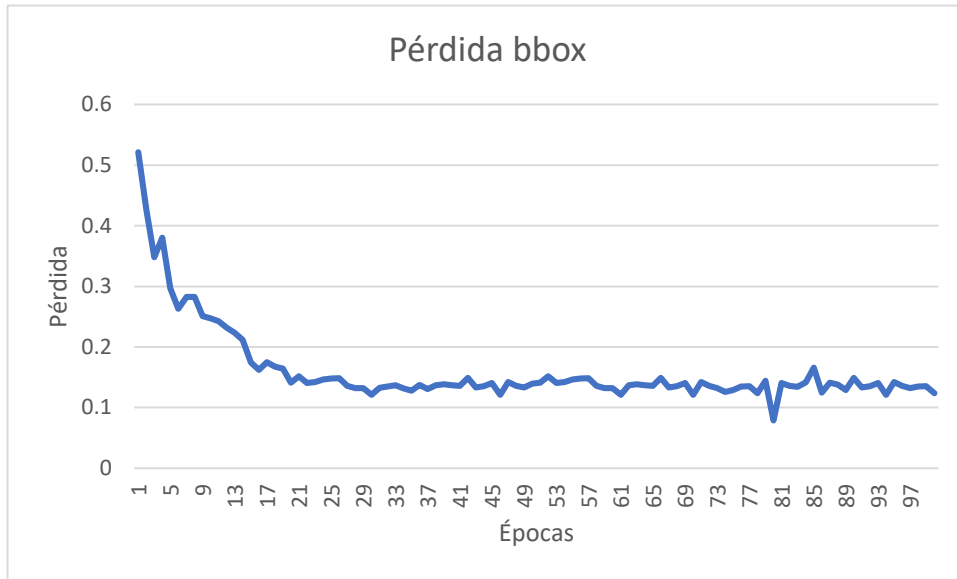
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 51 Precisión promedio de KeyPoints con 100% de base de datos - Validación



Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 52 Pérdida de bounding boxes de modelo KeyPoint- Validación



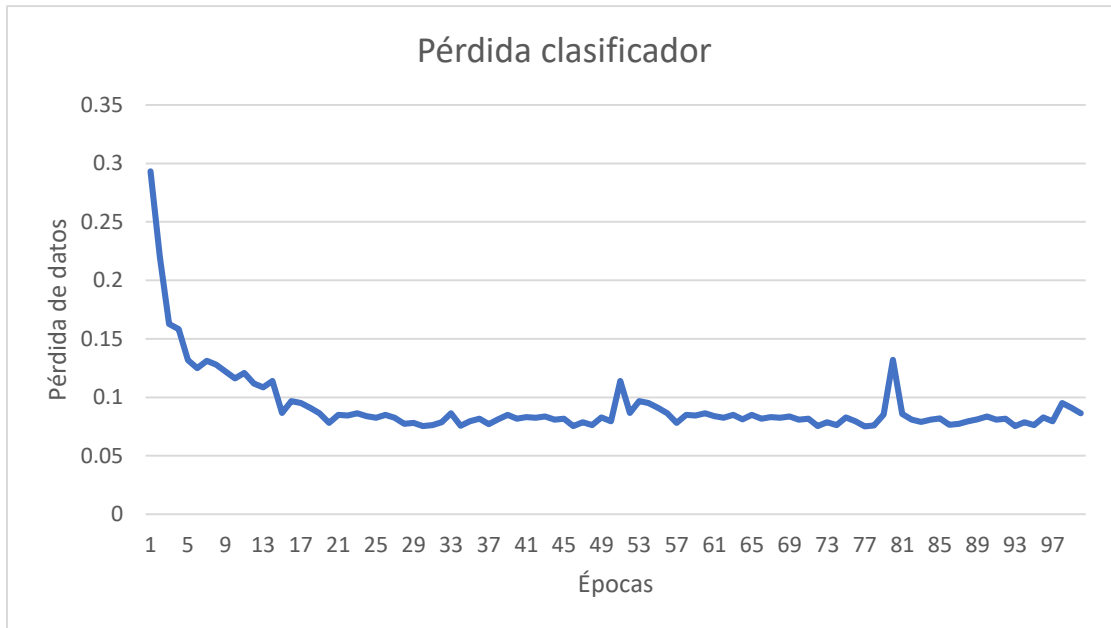
Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 53 Pérdida con 100% de base de datos modelo KeyPoint- Validación



Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 54 Pérdida de clasificador con 100% de base de datos modelo KeyPoint- Validación



Fuente: Elaboración propia obtenida por código de Google Colab

Anexo 55 Exhaustividad de KeyPoints de modelo KeyPoint- Validación



Fuente: Elaboración propia obtenida por código de Google Colab

REFERENCIAS BIBLIOGRÁFICAS

1. Fraiwan, M., Audat, Z., Fraiwan, L., & Manasreh, T. (2022). Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images. *Plos one*, 17(5), e0267851.
2. Ambrosio-Ambrosio, J. P., & González-Camacho, J. M. (2023). Algoritmos de aprendizaje automático para clasificar zonas de inundación a partir de imágenes de radar de apertura sintética. *Tecnología y ciencias del agua*, 14(4), 107-154.
3. Hernández, Y. J. C., Santana, M. R., Jorge, L. D. C. V., & Moreno, V. F. (2018) Aplicación de las redes neuronales convolucionales para la detección del cáncer de pulmón.
4. Berrones Reyes, M. C. (2019). *Clasificación de mamografías mediante redes neuronales convolucionales* (Doctoral dissertation, Universidad Autónoma de Nuevo León).
5. Ker, J., Wang, L., Rao, J., & Lim, T. (2017). Deep learning applications in medical image analysis. *IEEE Access*, 6, 9375-9389.
6. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
7. Chen, K., Zhai, X., Sun, K., Wang, H., Yang, C., & Li, M. (2021). A narrative review of machine learning as promising revolution in clinical practice of scoliosis. *Annals of Translational Medicine*, 9(1).
8. Repetur, A. E. (2019). *Redes neuronales artificiales*. Buenos Aires: Universidad Nacional del Centro de la Provincia de Buenos Aires.
9. Tan, Z., Yang, K., Sun, Y., Wu, B., Tao, H., Hu, Y., & Zhang, J. (2018, December). An automatic scoliosis diagnosis and measurement system based on deep learning. In 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO) (pp. 439-443). IEEE
10. Imran, A.-A.-Z., Huang, C., Tang, H., Fan, W., Cheung, K., To, M., Qian, Z., & Terzopoulos, D. (2020). Fully-Automated Analysis of Scoliosis from Spinal X-Ray Images. 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), 114–119. <https://doi.org/10.1109/CBMS49503.2020.00029>
11. Alharbi, R. H., Alshaye, M. B., Alkanhal, M. M., Alharbi, N. M., Alzahrani, M. A., & Alrehaili, O. A. (2020, March). Deep learning based algorithm for automatic scoliosis angle

- measurement. In 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS) (pp. 1-5). IEEE.
12. Gurei, E. (2023). *Predicción del rendimiento de los estudiantes en el aprendizaje basado en juegos* (Bachelor's thesis, Universitat Politècnica de Catalunya).
 13. Caesarendra, W., Rahmani, W., Mathew, J., & Thien, A. (2022). Automated Cobb angle measurement for adolescent idiopathic scoliosis using convolutional neural network. *Diagnostics*, 12(2), 396.
 14. Saravagi, D., Agrawal, S., Saravagi, M., Jain, S. K., Sharma, B., Mehboodniya, A., ... & Webber, J. L. Predicting Lumbar Spondylolisthesis: A Hybrid Deep Learning Approach.
 15. Tu, Y., Wang, N., Tong, F., & Chen, H. (2019, April). Automatic measurement algorithm of scoliosis Cobb angle based on deep learning. In *Journal of Physics: Conference Series* (Vol. 1187, No. 4, p. 042100). IOP Publishing.
 16. Xuan, J., Ke, B., Ma, W., Liang, Y., & Hu, W. (2023). Spinal disease diagnosis assistant based on MRI images using deep transfer learning methods. *Frontiers in Public Health*, 11, 1044525.
 17. Singha, A., Thakur, R. S., & Patel, T. (2021). Deep learning applications in medical image analysis. *Biomedical Data Mining for Information Retrieval: Methodologies, Techniques and Applications*, 293-350.
 18. Meng, N., Cheung, J. P., Wong, K. Y. K., Dokos, S., Li, S., Choy, R. W., ... & Zhang, T. (2022). An artificial intelligence powered platform for auto-analyses of spine alignment irrespective of image quality with prospective validation. *EClinicalMedicine*, 43.
 19. Jaiswal, A., Tang, L., Ghosh, M., Rousseau, J. F., Peng, Y., & Ding, Y. (2021, November). Radbert-cl: Factually-aware contrastive learning for radiology report classification. In *Machine Learning for Health* (pp. 196-208). PMLR.
 20. Abdullah, A. A., Yaakob, A., & Ibrahim, Z. (2018). Prediction of Spinal Abnormalities Using Machine Learning Techniques. 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA), 1-6. <https://doi.org/10.1109/ICASSDA.2018.8477622>

21. Fernández, L. M., García, E. D., & Riestra, S. G. (2020, May). Las responsabilidades derivadas del uso de las tecnologías de la información y comunicación en el ejercicio de las profesiones sanitarias. In *Anales de Pediatría* (Vol. 92, No. 5, pp. 307-e1). Elsevier Doyma.
 22. Wu, H., Bailey, C., Rasoulinejad, P., & Li, S. (2017). Automatic landmark estimation for adolescent idiopathic scoliosis assessment using BoostNet. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20* (pp. 127-135). Springer International Publishing.
 23. Buckland, M., & Gey, F. (1994). The relationship between recall and precision. *Journal of the American society for information science*, 45(1), 12-19.
 24. Taleon, C., Elizalde, G., & Rubinos, C. (2023). ScolioVis: Automated Cobb Angle Measurement on Anterior-Posterior Spine X-Rays using Multi-Instance Keypoint Detection with Keypoint RCNN. West Visayas State University College of Information and Communications Technology.
 25. AASCE. (2019). Task - Grand Challenge. <https://aasce19.grand-challenge.org/Task/>
- C.Hannaford. (2009). *Aprender moviendo el cuerpo*. México : Pax México .
- Fraiwan, L., Manasreh, T., & Audat, Z. (2022). Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images. *BMC Medical Informatics and Decision Making*.
- Imran, A.-A.-Z., Huang, C., Hui Tang, W. F., Cheung, K., To, M., Qian, Z., & Terzopoulos, D. (2020). Fully-Automated Analysis of Scoliosis from Spinal X-Ray Images. doi:<https://doi.org/10.1109/CBMS49503.2020.00029>
- Informatics, S. (2023). *BUU' DATASETS*. Retrieved from <https://services.informatics.buu.ac.th/spine/>
- Klinwichit, P., Yookwan, W., Limchareon, S., Chinnasarn, K., Jun-Su, J., & Athita, O. (2023). A Thai Open Lumbar Spine Dataset for. *Multidisciplinary Digital Publishing Institute*.
- Pantoja, T. S., & Chamorro, M. (2015). Escoliosis en niños y adolescentes. *Revista Médica Clínica*, 99-108.

- Serrano Rivera, D. (2019). *Ejercicios de Charriere en la cervicalgia mecánica del adulto*. Ecuador : BCBG Riobamba.
- Taleon, C. A., Elizalde, G., & Rubinos, C. (2023, Marzo). SCOLIOVIS: Automated Cobb Angle Measurement on AnteriorPosterior Spine X-Rays using Multi-Instance Keypoint. *West Visayas State University*, 117. Retrieved from <https://www.scoliovis.app/paper>
- Trinh, G. M., Hao-Chiang, S., Li-Chun Hsieh, K., Lee, C.-Y., Liu, H.-W., Lai, C.-W., . . . Huang, T.-J. (2023). *Detection of Lumbar Spondylolisthesis from X-ray Images*. Obtenido de MDPI: <https://www.mdpi.com/2077-0383/11/18/5450>
32. Castillo, H. M. (2013). *Escoliosis idiopática del adolescente: caracterización de variables semiológicas, de laboratorio y actividad física* (Doctoral dissertation, Universidad Católica de Córdoba).
33. Mora-de Sambricio, A., & Garrido-Stratenwerth, E. (2014). Espondilolisis y espondilolistesis en niños y adolescentes. *Revista Española de Cirugía Ortopédica y Traumatología*, 58(6), 395-406.
34. García Maya, A. F. Estudio comparativo de técnicas de clasificación binaria con múltiples anotadores.
35. Nguyen, T. P., Chae, D. S., Park, S. J., Kang, K. Y., & Yoon, J. (2021). Deep learning system for Meyerding classification and segmental motion measurement in diagnosis of lumbar spondylolisthesis. *Biomedical Signal Processing and Control*, 65, 102371.
36. Dhar, T., Adhikari, G., & Chaudhuri, S. S. (2021, September). An Improved Classification of Chest X-ray Images Using Adaptive Activation Function. In 2021 5th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech) (pp. 1-6). IEEE.