



UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA

FACULTAD DE INGENIERÍA

PROYECTO DE INVESTIGACIÓN

MÉTODO PARA LA DETECCIÓN AUTOMATIZADA DE ESPONDILOLISTESIS BASADO EN

MODELO DE DETECCIÓN DE VERTEBRAS POR APRENDIZAJE PROFUNDO.

PREVIO A LA OBTENCIÓN DEL TÍTULO

INGENIERO EN BIOMÉDICA

PRESENTADO POR:

21941028 DIEGO RENE ESPINAL VALLE

41841027 CARLOS GUILLERMO MEJÍA GUARDADO

ASESOR: KARLA MIRIAM REYES LEIVA

CAMPUS: SAN PEDRO SULA

JUNIO, 2024

HOJA DE FIRMAS

Los abajo firmantes damos fe, en nuestra posición de miembro de Terna, Asesor y/o jefe 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 Reyes
Asesor UNITEC



Ing. José Luis Ordoñez



Ing. William Taylor



Dra. Esperanza Corea



Ing. Reyna Valle

Jefe Académico de Ingeniería en Biomédica | UNITEC

DEDICATORIA

Primero y, ante todo, quiero dedicar este trabajo a Dios, por su infinita sabiduría y guía constante en mi vida. Este logro no habría sido posible sin su gracia y sus bendiciones. A mis padres, Carlos Mejía y Sonia Hernández quienes siempre confiaron en mí y me brindaron su amor incondicional y apoyo constante. Gracias por enseñarme la importancia de la perseverancia y el esfuerzo, sin ellos no sería la persona que soy hoy. A mi abuela Martha Gonzales, por su amor y sus palabras de aliento que siempre estarán en mi corazón. A mis hermanos, quienes me apoyaron y me dieron ese empujón cuando más lo necesitaba. Su compañía y ánimo fueron fundamentales para superar los desafíos que encontré en el camino. A los demás integrantes de la familia, amigos y compañeros de carrera, quienes con su compañía, amistad y ánimo constante hicieron de este camino una experiencia más llevadera y enriquecedora. También quiero dedicar este trabajo a aquellos que ya no están con nosotros, pero que me impartieron su sabiduría y me acompañaron en otras etapas de este lindo camino. Su influencia y enseñanzas siguen presentes en mi vida, guiándome y dándome fuerza para seguir adelante.

- **CARLOS MEJÍA**

Primeramente, quisiera dedicar este trabajo a Dios, quien sin duda alguna estuvo presente en todo momento para guiarme a tomar las decisiones correctas en mi vida. A mis padres, Lincoln Espinal y Dunia Valle, quienes me inculcaron desde el primer día a nunca dejar de creer, a pesar de cualquier adversidad. Su esfuerzo constante para permitirme estudiar la carrera de mis sueños no fue en vano y este trabajo es producto de ello, por y para ustedes seguiré dando mi mayor esfuerzo en alcanzar mis metas. Por último, quisiera dedicar este trabajo a todos mis hermanos y a sus parejas, quienes me han apoyado de innumerables formas para poder seguir adelante. Todo el apoyo emocional que me brindaron en mis momentos más altos y bajos ha sido gran parte de quien soy actualmente.

- **DIEGO ESPINAL**

AGRADECIMIENTOS

En primer lugar, quiero agradecerle a Dios por la oportunidad que me brindó de poder estar concluyendo mi carrera universitaria. A mis padres, gracias infinitas por el esfuerzo y la lucha constante para que yo pudiese realizar mi sueño de estudiar la carrera que deseaba. Gracias por enseñarme a apreciar el valor en los detalles y a comprender que los malos momentos son solo eso: momentos, y nunca son eternos.

A todos los docentes que me impartieron clases, quiero expresar mi sincero agradecimiento. Cada uno de ustedes ha aportado un granito de arena con su sabiduría y sus valiosos consejos, acercándome un paso más a ser el profesional que aspiro ser.

A mis compañeros y amigos de la carrera, quienes estuvieron desde las primeras clases y a aquellos que se han ido sumando al grupo: con ustedes compartí innumerables risas, anécdotas, horas de estudio, desvelos y cafés. Jenifer, Fátima, Gloria, Valeria, Majo, Jose gracias por aguantarme y seguirme en todas mis ocurrencias.

A mi grupo de amigos de la universidad, que siempre alegran mis días con sus loqueras, que estuvieron para mí, me apoyaban en todo lo que hacía y siempre estaban pendientes de mí. Gracias por hacer de esta experiencia algo inolvidable.

A mis amigos de mi ciudad, que me han visto crecer y que son mis hermanos de otras madres: ya sea a la distancia o cuando vuelvo "al pueblo", siempre me brindan su apoyo incondicional. Su amistad ha sido un pilar fundamental en mi vida.

Por último, pero no menos importante, a mi compañero de tesis, que primero fue mi compañero de clase, luego de grupo y que, a través de los años, se ha convertido en un gran amigo. Ese compañero que, cuando menos confiaban en nosotros, más nos motivábamos para demostrar de qué somos capaces. Hoy es mi amigo y colega, pero el día de mañana espero que seamos socios en Colimed Biomedica.

- **CARLOS MEJÍA**

Quisiera expresar mi sincera gratitud a Dios por darme la sabiduría necesaria para poder completar este trabajo de investigación. A mis padres una vez más, por creer en mí en todo

momento y guiarme por el camino correcto, por demostrarme que uno es capaz de todo lo que se proponga y por luchar todos los días para que sus hijos tengan la mejor educación posible.

A mis compañeros de carrera y amigos que hice en este largo trayecto, creando anécdotas interminables a través de las distintas clases, a las compañeras Valeria, Fátima, Gloria, María José, Jenifer y todas las demás que aguantaron mis entregas cercanas a la hora límite pero que a pesar de todo estuvieron hasta el final, a los hermanos que hice en el deporte y al resto de mis amistades, estoy eternamente agradecido con ustedes por hacer de mi experiencia universitaria una bonita historia.

A mi compañero de proyectos en la mayoría de las clases, y ahora también mi compañero de tesis, por ser el amigo que me siguió desde la primera vez que le pedí hacer un proyecto complicado, gracias por siempre confiar en que a pesar de todo somos capaces de completar todo lo que nos propusiéramos.

Por último, pero no menos importante, a los Halcones. Esa segunda familia que me otorgó la vida, quienes estuvieron para mí en todo momento para escucharme y apoyarme en cualquier situación a través de los años, les agradezco por todo esto y más.

- **DIEGO ESPINAL**

Finalmente, queremos expresar nuestro más profundo agradecimiento a nuestra asesora de tesis, Dra. Karla Reyes, por su invaluable apoyo, orientación y paciencia a lo largo de este proceso. Su fe en nosotros y en el proyecto, incluso cuando nosotros mismos dudábamos, y su lucha constante por nosotros, son testimonio de su extraordinaria generosidad y profesionalismo. Gracias por creer en nosotros y por guiarnos con su experiencia y sabiduría. Sabemos que donde quiera que esté contamos con su apoyo, y que en la medida que sea posible también usted cuenta con el nuestro.

- **CARLOS MEJÍA Y DIEGO ESPINAL**

RESUMEN EJECUTIVO

La espondilolistesis es una afección prevalente de la columna vertebral caracterizada por el desplazamiento anterior de una vértebra sobre otra adyacente. El uso de métodos manuales para su diagnóstico implica una amplia preparación para el médico y puede considerarse una práctica muy compleja y que requiere mucho tiempo, especialmente cuando se trata de múltiples pacientes. Por lo tanto, se necesita un sistema para la detección automática de la espondilolistesis lumbar que reduzca el tiempo empleado en el diagnóstico y proporcione valores objetivos basados en mediciones. Este estudio presenta un método de detección automática de espondilolistesis utilizando un modelo de aprendizaje profundo para la detección vertebral. El objetivo fue desarrollar y validar un algoritmo capaz de identificar con precisión los límites vertebrales y medir la traslación vertebral utilizando el sistema de clasificación Meyerding. La metodología implicó el preprocesamiento de imágenes de rayos X, el entrenamiento de un modelo de detección de vértebras y la implementación de un algoritmo para calcular la traslación vertebral. Para el entrenamiento y la validación se utilizaron un total de 714 imágenes, incluidas 197 de pacientes hondureños y 517 de la base de datos LSPINE de la BUU. El modelo alcanzó una precisión media del 98% en la predicción de los límites y del 95% en la detección de los puntos clave. El rendimiento del algoritmo se validó frente a mediciones manuales realizadas por especialistas, demostrando una alta precisión y fiabilidad. Los resultados indican que el método desarrollado puede reducir significativamente la variabilidad diagnóstica y proporcionar mediciones objetivas, mejorando la toma de decisiones clínicas. Este sistema de detección automatizado ofrece una herramienta prometedora para mejorar el diagnóstico y la planificación del tratamiento de la espondilolistesis.

ABSTRACT

Spondylolisthesis is a prevalent spinal condition characterized by the anterior displacement of a vertebra over an adjacent one. The use of manual methods for its diagnosis involves extensive preparation for the physician and can be considered a highly complex and time-consuming practice, especially when dealing with multiple patients. Therefore, a system for the automatic detection of lumbar spondylolisthesis is needed to reduce the time spent in diagnosis and to provide objective values based on measurements. This study introduces an automatic detection method for spondylolisthesis using a deep learning model for vertebral detection. The objective was to develop and validate an algorithm capable of accurately identifying vertebral boundaries and measuring vertebral translation using the Meyerding classification system. The methodology involved pre-processing X-ray images, training a vertebrae detection model, and implementing an algorithm to calculate vertebral translation. A total of 714 images, including 197 from Honduran patients and 517 from the BUU LSPINE database, were used for training and validation. The model achieved an average accuracy of 98% for bounding box predictions and 95% for keypoint detection. The algorithm's performance was validated against manual measurements by specialists, demonstrating high accuracy and reliability. The findings indicate that the developed method can significantly reduce diagnostic variability and provide objective measurements, enhancing clinical decision-making. This automated detection system offers a promising tool for improving the diagnosis and treatment planning of spondylolisthesis.

ÍNDICE DE CONTENIDO

I.	Introducción	2
II.	Estado del Arte.....	4
	3.1 Antecedentes	4
	3.2 Problemática.....	20
	3.3 Imagen Integradora.....	22
	3.4 Tabla de Limitaciones	23
	3.5 Tabla de Metodologías.....	26
III.	Objetivos	34
	2.1 Objetivo General	34
	2.2 Objetivos Específicos	34
IV.	Método	35
	4.1 Enfoque	35
	4.2 Variables de Investigación.....	35
	4.2.1. Variables Dependientes	35
	4.2.2. Variables Independientes	36
	4.3 Técnicas e Instrumentos aplicados	37
	4.3.1. Roboflow.....	37
	4.3.2. Google Colaboratory.....	37
	4.3.3. Python.....	38
	4.3.4. ImageJ.....	38
	4.4 Metodología de Estudio	38
	4.4.1. Pre-Procesamiento de Imágenes.....	38

4.4.2. Modelo de Detección de Vértebras.....	39
4.4.3. Algoritmo de Medición para la Clasificación de Meyerding Propuesto.....	39
4.5 Metodología de Validación	43
4.5.1. Métricas de Evaluación de Desempeño	43
4.6 Cronograma de Actividades	45
V. Resultados y Análisis.....	46
5.1 Base de Datos	46
5.2 Reentrenamiento del Modelo de Detección de Vértebras.....	47
5.3 Inferencia	49
5.4 Módulo de Medición y Detección de Espondilolistesis	51
5.5 Validación	53
5.5.1. Verificación manual con ImageJ	53
5.5.2. Verificación manual por parte de expertos	55
VI. Discusión de Resultados.....	58
VII. Conclusiones.....	62
VIII. Recomendaciones.....	63
IX. Anexos.....	65
X. Referencias	109

ÍNDICE DE TABLAS

Tabla 1 métricas de rendimiento y precisión en el estudio de Varcin et al. (2019)	10
Tabla 2 métricas evaluadas en la investigación de Xuan et al., (2023).....	16
Tabla 3 comparación entre algoritmo y grupo de doctores en la investigación de Zhang et al., (2023)	16
Tabla 4 comparación entre diagnostico con asistencia y sin asistencia en la investigación de Zhang et al., (2023).....	17
Tabla 5 Resultados obtenidos en la investigación de Dshad Ahmed & Hawezi, (2023)...	17
Tabla 6 Resultados del total de imágenes clasificadas en la investigación de Varcin et al., (2019).....	19
Tabla 7 Total de imágenes clasificadas en la investigación de Cai et al., (2017).....	19
Tabla 8 Metodologías de Estudios previos.....	26
Tabla 10 Métricas de evaluación de rendimiento.....	43
Tabla 9 Bases de Datos a utilizar	46
Tabla 11 Comparativa entre resultados de entrenamiento de modelo de detección de vertebras con estudio previo	49
Tabla 12 Resultados de Concordancia en Clasificación.....	56
Tabla 13 Grados de Kappa de Fleiss	57
Tabla 14 Comparativa de modelo con investigaciones previas.....	58

ÍNDICE DE ILUSTRACIONES

Ilustración 1 Imagen Integradora.....	22
Ilustración 2 Variables Dependientes	36

Ilustración 3 Variables Independientes.....	37
Ilustración 4 Diagrama de Flujo de Algoritmo de Medición.....	40
Ilustración 5 Cronograma de Actividades.....	45
Ilustración 6 Ejemplo de detección de vertebras en entrenamiento.....	48
Ilustración 7 Ejemplo de puntaje de confianza.....	50
Ilustración 8 Ejemplo de inferencia	51
Ilustración 9 Medición de Traslación	52
Ilustración 10 Cálculo de Porcentaje de Traslación.....	53
Ilustración 11 Validación Manual con ImageJ.....	54
Ilustración 12 Gráfico de Dispersión de Mediciones Manuales vs Automáticas.....	54
Ilustración 13 Gráfico de Bland-Altman de Mediciones Automática y Manual.....	55

ÍNDICE DE ECUACIONES

Ecuación 1 Porcentaje de Traslación Vertebral.....	42
--	----

ÍNDICE DE ANEXOS

Anexo 1 Anotaciones manuales con Keypoints en Roboflow.....	66
Anexo 2 Anotaciones manuales con bounding boxes.....	67
Anexo 3 Métrica de precisión promedio del modelo con 100 épocas en bounding boxes	68
Anexo 4 Métrica de sensibilidad del modelo con 100 épocas en bounding boxes	68
Anexo 5 Métrica de precisión promedio del modelo con 100 épocas en Keypoints.....	69
Anexo 6 Métrica de sensibilidad del modelo con 100 épocas en Keypoints	69
<i>Anexo 7 Resultados del módulo de detección de espondilolistesis</i>	<i>70</i>

<i>Anexo 8 Resultados del módulo de detección de espondilolistesis</i>	<i>71</i>
<i>Anexo 9 Resultados del módulo de detección de espondilolistesis</i>	<i>72</i>
<i>Anexo 10 Resultados del módulo de detección de espondilolistesis.....</i>	<i>73</i>
<i>Anexo 11 Resultados del módulo de detección de espondilolistesis.....</i>	<i>74</i>
<i>Anexo 12 Resultados del módulo de detección de espondilolistesis.....</i>	<i>75</i>
<i>Anexo 13 Resultados del módulo de detección de espondilolistesis.....</i>	<i>76</i>
<i>Anexo 14 Resultados del módulo de detección de espondilolistesis.....</i>	<i>77</i>
<i>Anexo 15 Resultados del módulo de detección de espondilolistesis.....</i>	<i>78</i>
<i>Anexo 16 Resultados del módulo de detección de espondilolistesis.....</i>	<i>79</i>
<i>Anexo 17 Resultados del módulo de detección de espondilolistesis.....</i>	<i>80</i>
<i>Anexo 18 Resultados del módulo de detección de espondilolistesis.....</i>	<i>81</i>
<i>Anexo 19 Resultados del módulo de detección de espondilolistesis.....</i>	<i>82</i>
<i>Anexo 20 Resultados del módulo de detección de espondilolistesis.....</i>	<i>83</i>
<i>Anexo 21 Resultados del módulo de detección de espondilolistesis.....</i>	<i>84</i>
<i>Anexo 22 Resultados del módulo de detección de espondilolistesis.....</i>	<i>85</i>
<i>Anexo 23 Resultados del módulo de detección de espondilolistesis.....</i>	<i>86</i>
<i>Anexo 24 Medición con ImageJ (3729 L4-L5).....</i>	<i>87</i>
<i>Anexo 25 Medición con ImageJ (3729 L5-S1)</i>	<i>87</i>
<i>Anexo 26 Medición con ImageJ (3753 L5-S1)</i>	<i>88</i>
<i>Anexo 27 Medición con ImageJ (3808 L5-S1)</i>	<i>88</i>
<i>Anexo 28 Medición con ImageJ (3836 L3-L4).....</i>	<i>89</i>
<i>Anexo 29 Medición con ImageJ (3836 L4-L5).....</i>	<i>89</i>
<i>Anexo 30 Medición con ImageJ (3870 L5-S1)</i>	<i>90</i>

<i>Anexo 31 Medición con ImageJ (4093 L4-L5)</i>	90
<i>Anexo 32 Medición con ImageJ (4093 L5-S1)</i>	91
<i>Anexo 33 Medición con ImageJ (4321 L4-L5)</i>	91
<i>Anexo 34 Medición con ImageJ (4321 L5-S1)</i>	92
<i>Anexo 35 Medición con ImageJ (4367 L4-L5)</i>	92
<i>Anexo 36 Medición con ImageJ (4426 L5-S1)</i>	93
<i>Anexo 37 Medición con ImageJ (4622 L4-L5)</i>	93
<i>Anexo 38 Medición con ImageJ (4622 L5-S1)</i>	94
<i>Anexo 39 Medición con ImageJ (4662 L5-S1)</i>	94
<i>Anexo 40 Medición con ImageJ (4698 L4-L5)</i>	95
<i>Anexo 41 Medición con ImageJ (4698 L5-S1)</i>	95
<i>Anexo 42 Métrica de pérdida del modelo con 100 épocas</i>	96
<i>Anexo 43 Métrica de pérdida de clasificación del modelo con 100 épocas</i>	96
<i>Anexo 44 Métrica de pérdida de bounding boxes del modelo con 100 épocas</i>	97
<i>Anexo 45 Métrica de pérdida de keypoints del modelo con 100 épocas</i>	97
<i>Anexo 46 Métrica de precisión promedio del modelo con 150 épocas en bounding boxes</i> . 98	
<i>Anexo 47 Métrica de sensibilidad del modelo con 150 épocas en bounding boxes</i>	98
<i>Anexo 48 Métrica de precisión promedio del modelo con 150 épocas en keypoints</i>	99
<i>Anexo 49 Métrica de sensibilidad del modelo con 150 épocas en Keypoints</i>	99
<i>Anexo 50 Métrica de pérdida del modelo con 150 épocas</i>	100
<i>Anexo 51 Métrica de pérdida de clasificación del modelo con 150 épocas</i>	100
<i>Anexo 52 Métrica de pérdida de bounding boxes del modelo con 150 épocas</i>	101
<i>Anexo 53 Métrica de pérdida de keypoints del modelo con 150 épocas</i>	101

<i>Anexo 54 Métrica de precisión promedio del modelo con 50 épocas en bounding boxes..</i>	102
<i>Anexo 55 Métrica de sensibilidad del modelo con 50 épocas en bounding boxes</i>	102
<i>Anexo 56 Métrica de precisión promedio del modelo con 50 épocas en keypoints</i>	103
<i>Anexo 57 Métrica de pérdida del modelo con 50 épocas.....</i>	103
<i>Anexo 58 Métrica de pérdida de clasificación del modelo con 50 épocas.....</i>	104
<i>Anexo 59 Métrica de pérdida de bounding boxes del modelo con 50 épocas</i>	104
<i>Anexo 60 Métrica de pérdida de keypoints del modelo con 50 épocas.....</i>	105
<i>Anexo 61 Mockup de Aplicación.....</i>	105
<i>Anexo 62 Mockup de Aplicación.....</i>	106
<i>Anexo 63 Mockup de Aplicación.....</i>	106
<i>Anexo 64 Mockup de Aplicación.....</i>	107
<i>Anexo 65 Mockup de Aplicación.....</i>	107
<i>Anexo 66 Mockup de Aplicación.....</i>	108
<i>Anexo 67 Mockup de Aplicación.....</i>	108

LISTA DE SIGLAS

TC	Tomografía Computarizada
IA	Inteligencia Artificial
CNN	Redes Neuronales Convolucionales
RCNN	Redes Neuronales Convolucionales basadas en regiones
YOLO	You Only Look Once
IoU	Intersección sobre la Union
mAP	Precisión Promedio
RM	Resonancia Magnética
DSC	Coficiente de similitud Dice
Mm	Milímetros
TP	Verdaderos positivos
FP	Falsos Positivos
TN	Verdaderos Negativos
FN	Falsos Negativos

I. INTRODUCCIÓN

El desplazamiento hacia adelante de una de las vértebras lumbares con respecto a la vertebra inferior se le conoce como espondilolistesis lumbar, y esta es una de las afecciones de la columna más comunes en los seres humanos. Se define como la progresión del deslizamiento después de la maduración esquelética y se asocia con la degeneración del disco según el grado de deslizamiento (Varçin et al., 2021). Esta afección puede causar una variedad de síntomas, que incluyen dolor de espalda, rigidez entumecimiento y debilidad en las piernas, dependiendo del grado de desplazamiento, el impacto sobre los nervios y las estructuras circundantes. En pacientes clínicamente sintomáticos, la espondilolistesis degenerativa es seis veces más común en mujeres que en hombres y de seis a nueve veces más común en la vertebra L4 que en regiones adyacentes. De acuerdo con Wiltse (1981), ocurre cuatro veces más a menudo cuando la quinta vértebra lumbar está sacralizada. La espondilolistesis se divide en 5 grados según el sistema de clasificación de Meyerding. Este sistema se basa en el porcentaje de deslizamiento de una vértebra en relación con la vertebra sobre la cual se desliza. Esta clasificación se divide en grado 1 (0-25%), grado 2 (26-50%), grado 3 (51-75%), grado 4(76-99%) y grado 5 (99% o más se define como espondiloptosis) (Gallagher et al., 2020).

Interpretar los trastornos degenerativos de la columna en múltiples niveles puede ser una tarea tediosa y propensa a una alta variabilidad de quien emite un diagnóstico. Esta tarea es adecuada para la inteligencia artificial (IA), que puede mejorar la fiabilidad de los diagnósticos en tareas repetitivas (Gorelik & Gyftopoulos, 2021). La clasificación de imágenes es un campo de aplicación de las redes neuronales convolucionales (CNN). Las CNN se han aplicado principalmente al análisis de imágenes, especialmente al reconocimiento de patrones, y también se pueden utilizar para resolver otros problemas de análisis de datos, como la clasificación de imágenes y detección de objetos (Hoque et al., 2021). Recientemente, se ha intentado combinar los algoritmos tradicionales de CNN para la detección de objetos y la segmentación semántica; ya que el propósito de la detección de objetos es determinar si un objeto existe en una categoría predefinida (Yang & Yu, 2021).

Con esta motivación, el presente proyecto de investigación plantea el desarrollo de un método que involucra el refinamiento y reentrenamiento de un modelo basado en CNN, para la

detección de vertebras en imágenes de rayos X en el plano sagital. Para esto se utilizará una base de datos que contiene imágenes de rayos X. La finalidad del método desarrollado es que fuese capaz de detectar casos de espondilolistesis automáticamente y, definir el grado de afección según el sistema de clasificación de Meyerding. Se refinó un modelo y se implementó como herramienta principal para la clasificación de las imágenes y una posterior validación de los diagnósticos automáticos de la mano de tres médicos especialistas en el área de radiología y neurocirugía.

El resto del informe se estructura de la siguiente manera: En el estado del arte se describen los trabajos previos a esta investigación, la problemática y las limitaciones de investigaciones previas. En el capítulo de métodos se describe de qué forma se realizó el proyecto, se resume el procedimiento realizado y se definen las métricas de evaluación. En los resultados y análisis se describen los hallazgos obtenidos a base de gráficos e imágenes. La discusión analiza e interpreta los resultados en un contexto teórico y explora sus implicaciones para el campo. En las conclusiones se resumen los principales hallazgos del estudio y sugiere posibles direcciones para futuras investigaciones basadas en los resultados obtenidos.

II. ESTADO DEL ARTE

En este capítulo se presenta el estado del arte relacionado con el tema del proyecto de investigación. El objetivo es proporcionar el contexto teórico y práctico en el que se basa este trabajo, identificando los antecedentes de mayor relevancia en el campo de estudio. Se revisaron estudios anteriores, sus principales hallazgos y limitaciones se resaltaron y presentaron en una tabla de limitaciones de estudios anteriores. Además, se proporciona una tabla comparativa de los métodos utilizados en estudios anteriores para proporcionar una comprensión detallada de los diferentes enfoques metodológicos utilizados hasta la fecha.

3.1 ANTECEDENTES

La columna vertebral es una estructura ósea fundamental para el cuerpo humano, brindando soporte estructural y protección a la médula espinal. Sin embargo, está sujeta a distintas afecciones que pueden afectar la calidad de vida de quienes las padezcan. Dentro de las afecciones de columna vertebral, existe una condición en la que ocurre una traslación de un cuerpo vertebral sobre otro respecto a las vértebras distales. Koslosky & Gendelberg (2020), definen la espondilolistesis como el deslizamiento o luxación de una vértebra sobre otra adyacente. Esta representa una incidencia de 4-6% en infantes y aumentando a 5-10% en adultos, principalmente siendo esta degenerativa y más frecuente en mujeres.

Wiltse, (1981) en su investigación exhaustiva acerca de la espondilolistesis, lo cataloga como una afección en la que una vértebra se desliza sobre otra. Estas se clasifican en varios tipos: Displasia (congénita), que ocurre debido a un defecto congénito del sacro o del arco neural y que muchas veces tiene un componente genético. Puede ser ístmico, causado por fracturas interarticulares por estrés, que son comunes en personas jóvenes, o lítico, similar al tipo lítico pero causado por el alargamiento indiviso de un área con un alargamiento interarticular intacto. Postraumático, es el resultado de una lesión grave que provoca un deslizamiento progresivo de las vértebras sin que se produzca una fractura aguda. y condiciones patológicas causadas por enfermedades óseas que brindan un apoyo insuficiente para la alineación de la columna. Las mediciones típicas se basan en la cantidad de desplazamiento de una vértebra con respecto a la otra y la inclinación del sacro.

En el estudio hecho por Hoque et al., (2021) se habla de las diferentes sistemas para la detección de objetos en 3D y detección de poses en 6D. se destaca la complejidad de una imagen ya que en ellos se encuentra mucha información, diferentes objetos a diferentes escalas y posiciones que aquí se propone la extracción de características para identificar objetos por sus características predominantes y la clasificación de objetos que ayuda a diferenciar lo que se encuentra en la imagen para enfocarse en un objeto a la vez. Los autores destacan varios de los modelos y algoritmos especializados en la rama de detección de objetos algunos de los más populares como YOLO o Faster RCNN Hoque et al., (2021). También distintas formas de detectar objetos, ya sea por detección de puntos clave los bounding boxes que son útiles para objetos de diferentes tamaños y profundidades. También se destacan las características de varios algoritmos, sus desventajas y sus ventajas para que cualquier lector pueda guiarse y elegir el algoritmo o modelo que mejor le convenga.

En un artículo por Gallagher et al. (2020) hacen mención a los diferentes métodos para la clasificación de espondilolistesis. La clasificación de Wiltse hace una distinción basándose en la etiología de la espondilolistesis, siendo Tipo I si es congénita, Tipo II si es ístmica y con más incidencia en infantes, Tipo III si es degenerativa, siendo más frecuente en adultos mayores, Tipo IV es traumática y Tipo V es patológica. Si bien la clasificación de Wiltse brinda información valiosa, no es útil para llevar un control de la progresión y es más descriptiva. La clasificación de Marchetti y Bartolozzi surge como un intento de mejorar la clasificación de Wiltse debido a inconformidades específicamente con el Tipo II, estableciendo cambios para diferenciar entre una afección evolutiva y una adquirida. La evolutiva se subdivide entre grado alto y grado bajo, basándose en el grado de displasia del segmento afectado. La adquirida representa los tipos III-V de Wiltse, subdividiéndose en las siguientes 4 categorías: traumática, postquirúrgica, patológica y degenerativa. El método de clasificación de Meyerding permite cuantificar la magnitud de la listesis, dependiendo del grado de desplazamiento de la vértebra en relación con la vértebra adyacente (Rajpurkar et al., 2018). Separándolo en grados de I a V, representados por una traslación de 0-25%, 25-50%, 50-75%, 75-100% y más de 100% respectivamente.

Los autores Gorelik & Gyftopoulos, (2021) descubrieron en su investigación descriptiva que el uso de IA ofrece avances en la detección de fracturas y la interpretación de imágenes de la

columna. Para las fracturas, varios estudios han demostrado la capacidad de la IA para detectar y clasificar fracturas esqueléticas axiales y apendiculares tanto en radiografías como en tomografías computarizadas. Se ha observado que los modelos de IA logran una precisión significativamente mayor, en varios casos llegan a superar a los lectores humanos. Sin embargo, señalan que los modelos de IA pueden tener problemas con fracturas menos obvias, como las fracturas no desplazadas del cuello femoral o del escafoides. Otro aspecto importante es que cuando se trata de interpretar imágenes de la columna, la IA ha demostrado ser eficaz para identificar y etiquetar los cuerpos vertebrales y los discos intervertebrales, así como para clasificar la patología que afecta a la columna. Los modelos de IA han logrado resultados similares a los de los radiólogos al evaluar la gravedad de las enfermedades degenerativas de la columna, pueden mejorar la coherencia y reducir la variabilidad en el criterio de los radiólogos para este tipo de interpretación radiológica Gorelik & Gyftopoulos, (2021).

Existe una gran variedad de imágenes médicas, como rayos X, resonancia magnética (RM) ultrasonidos, endoscopias, tomografías computarizadas (TC), entre otras. Yang & Yu, (2021), destacan el uso de la inteligencia artificial para la detección de objetos y la segmentación en las imágenes médicas. Ven imperativo la entrada de inteligencia artificial para ayudar al especialista a analizar las imágenes generadas y a resolver problemas en ciertos diagnósticos. Se proponen algoritmos para la detección de objetos y la segmentación semántica, todos son derivados de las redes neuronales convolucionales que podrían ser utilizados dependiendo de diferentes aspectos como la tarea a realizar o la información disponible para entrenar al algoritmo. Los autores destacan las similitudes y las limitantes de los diferentes modelos de detección de objetos y segmentación semántica.

En su investigación, Moon et al., (2022) se centraron en la detección de fracturas de huesos faciales en imágenes de TC basadas en el modelo YoloX. Para ello, los investigadores recopilaron datos de tomografías computarizadas de pacientes con narices rotas y otras afecciones en un hospital universitario de 2014 a 2020 y etiquetaron las imágenes con cuadros que indicaban la ubicación de la fractura. Luego, los datos se preprocesaron utilizando filtros de imagen para aumentar la visibilidad de los pómulos. Los autores eligieron el modelo YoloX por su rendimiento y estructura central separada en el conjunto de datos. También aplicaron la tecnología de

aumento de datos de YoloX al entorno de imágenes de TC, de esta forma cambian la manera en que se llenan los espacios en las imágenes aumentadas y se ajusta el grado de distorsión. Luego realizaron un estudio para evaluar el rendimiento del modelo utilizando métricas como precisión promedio, sensibilidad y especificidad para comparar diferentes configuraciones de pérdida y métodos de aumento de datos. Después de realizadas las pruebas, los autores comprobaron que el modelo con las configuraciones propuestas con la pérdida de la intersección sobre la unión (IoU) y el CTMixup que es una superposición de dos imágenes TC obtuvo mejores resultados con una precisión promedio de 0.69, sensibilidad del 100% y especificidad del 84% que significa que todos los pacientes que presentaron una lesión fueron identificados.

En un estudio realizado por Singh et al., (2022), propusieron un modelo a base de Deep Learning para la detección automática en radiografías musculoesqueléticas. Utilizaron el conjunto de datos de Mura Rajpurkar et al., (2018), uno de los conjuntos de datos más grandes disponibles públicamente, para la detección de anomalías en radiografías musculoesqueléticas de las extremidades superiores. Consta de 40.561, radiografías de 14.863 exámenes obtenidos del Archivo de Comunicaciones e Imágenes del Hospital Stanford. Los autores seleccionaron un conjunto de muestras de radiografías de dedos y las dividieron en conjuntos de datos de entrenamiento, validación y prueba. Luego implementaron un modelo de CNN utilizando la biblioteca Keras con el servidor TensorFlow y utilizar las CNN para el reconocimiento y clasificación de imágenes también, utilizaron la capa de activación ReLU y la capa de agrupación. El modelo fue entrenado y validado utilizando imágenes destinadas para los diferentes procesos, logrando una precisión de validación del 93,51 % y una pérdida de validación de 0,09 después de 50 épocas. Finalmente, el modelo se evaluó en un conjunto de pruebas para validar su capacidad de predicción de anomalías y logró una precisión de detección del 89,41 %, superando a los modelos existentes. Los radiólogos pueden utilizar el modelo ComDNet-512 propuesto para detectar automáticamente anomalías en el dedo para facilitar y agilizar los procesos de detección para las radiografías.

Por otro lado, el uso de las redes neuronales para la identificación automática de regiones de interés o puntos clave para el análisis de la columna vertebral ha ido en aumento desde hace ya más de diez años. En un estudio publicado por Jaehan Koh et al. (2012), se estudia la extracción

de bordes de manera automática para la columna utilizando imágenes de RM. Los autores propusieron métodos para la caracterización y el diagnóstico asistidos por computadora utilizando LumbarDiagnostics, para utilizarse como punto de referencia al marcar regiones de fondo y encontrar regiones de interés. Para la validación de los resultados comparan el método con los límites de referencia en términos de la distancia de Fréchet, mostrando una tasa de acierto del 99% a partir de 65 datos clínicos, sin embargo, contaba con ciertos fallos en su funcionamiento. Cuando el gradiente entre la vértebra y las regiones adyacentes es grande, el método de gradiente inicializado se separa del canal medular y tiende a desviarse si la región del límite es borrosa.

Algunos de los principales los desafíos para la medición automática de desplazamiento lumbar radican en la detección precisa de las superficies superiores e inferiores de cada vértebra, y poder etiquetar dichas vértebras para brindar un significado semántico de la medición. Para tratar de resolver dichos desafíos, en un artículo publicado por Liao et al. (2016), se explora directamente la medición automática de espondilolistesis en imágenes de TC, basado en el etiquetado automático y la determinación de regiones críticas. Los autores utilizaron el método de etiquetado integrando el aprendizaje que integra tanto el aspecto de la imagen como la información geométrica de la columna para la detección de cada vértebra y propusieron un método jerárquico para el posicionamiento de los puntos anteriores y posteriores. El método logró una precisión del 94.4% (es decir, 776 de 821) en predicción de los grados de espondilolistesis, sin embargo, identificaron que había múltiples enfermedades adicionales que existían en pacientes como las fracturas vertebrales y escoliosis que afectaban en la clasificación. Adicionalmente, la mayor parte de la clasificación errónea se produjo cuando el porcentaje de desplazamiento real estaba muy cercano al rango de un nivel de clasificación adyacente. Dando validez a estudios futuros para la mejora de dichos modelos, para aumentar la precisión y posiblemente, aumentar las funciones y capacidades.

Tomando como referencia una de las limitantes del estudio realizado por Liao et al. (2016), se identifica una nueva publicación el año siguiente, en el cual los autores Jamaludin et al. (2017) de la universidad de Oxford, desarrollaron la estructura de una CNN que toma los volúmenes de los discos intervertebrales como datos de entrada para producir automáticamente clasificaciones

radiológicas de resonancias magnéticas lumbares y así realizar distintas graduaciones (Pfirrmann, espondilolistesis, estenosis de canal medular, entre otras) a la vez. Utilizando la caja de herramientas de MatConvNet para el entrenamiento de las redes neuronales, lograron proponer variantes de la arquitectura CNN utilizando convoluciones 3D, mostrando como este tipo de arquitectura permite el entrenamiento utilizando una función de pérdida multitarea que no requiere de etiquetado para segmentación. A pesar de solamente haber usado cortes sagitales en sus pruebas ya que ciertas clasificaciones pueden analizarse con más vistas, como en el caso de la estenosis medular utilizando cortes axiales, los autores lograron resultados cercanos al rendimiento humano en las ocho graduaciones, permitiendo de igual forma visualizar las pruebas para apoyar en el entendimiento de las predicciones.

Un artículo publicado para el *European Spine Journal*, se explora similarmente un análisis automatizado de deformaciones lumbares utilizando técnicas de aprendizaje profundo. Los autores Galbusera et al. (2019), idearon un método de extracción de parámetros anatómicos a partir de radiografías biplanares de la columna, para determinar su forma de manera automática y así calcular parámetros posturales en un amplio escenario de condiciones clínicas. Basándose en la transformación espacial a numérica diferenciable (DSNT, por sus siglas en inglés), combinando con una red completamente convolucional para convertir el mapa de calor espacial generado por la red en coordenadas numéricas. Dicho modelo puede ser utilizado en imágenes adquiridas de distintos sujetos, tanto adolescentes como adultos mayores, para tratar de eliminar la dependencia del operador en los resultados. Los autores obtuvieron parámetros altamente correlacionados, sin embargo, los errores estándar oscilaron entre 2.7° (para la inclinación pélvica) y 11.5° (para la lordosis L1-L5). Esto puede darse debido al tamaño de muestras utilizadas, ya que puede considerarse muy bajo en el contexto del aprendizaje automático (443 pares de imágenes).

Para poder evaluar correctamente el rendimiento de redes neuronales para el diagnóstico de espondilolistesis, existen investigaciones como la de Varcin et al. (2019), que comparan diferentes herramientas y evalúan sus resultados. Los autores identificaron principalmente que no existe una base de datos pública de imágenes radiológicas exclusivamente de espondilolistesis, siendo esta su principal limitante que abordaron montando así una base de datos para poder llevar a cabo el estudio. Parte de la problemática mencionada anteriormente es la cantidad

limitada de muestras en las bases de datos, utilizando *Transfer Learning* para tomar ventaja de la cantidad de datos de entrenamiento en un dominio de interés para facilitar el aprendizaje en otro dominio con pocos datos de entrenamiento, y *Data Augmentation* para producir datos artificiales utilizando métodos de procesamiento de imágenes, se logra así reducir la necesidad de bases de datos muy extensas. Los autores realizaron la comparativa entre dos redes previamente entrenadas, AlexNet y GoogLeNet, para implementar el modelo de aprendizaje de transferencia profunda, obteniendo resultados experimentales mostrados en la Tabla 1, que demostraron una leve ventaja de GoogLeNet sobre AlexNet en cuanto a su precisión y rendimiento.

Tabla 1 métricas de rendimiento y precisión en el estudio de Varcin et al. (2019)

Medición	Valor	
	AlexNet	GoogLeNet
Sensibilidad	95.45%	95.65%
Especificidad	88.46%	92.00%
Precisión	87.50%	91.67%
Valor predictivo negativo	95.83%	95.83%
Tasa de falsos positivos	11.54%	8%
Tasa de falsos descubrimientos	12.5%	8.33%
Tasa de falsos negativos	4.55%	4.35%
Exactitud	91.67%	93.75%
Puntuación F1	91.30%	93.62%

La implementación de redes neuronales para la detección de espondilolistesis es un tema aún por solidificar mediante más trabajos de investigación; la poca cantidad de trabajos relacionados abren la necesidad de recurrir por investigaciones de afecciones de columna similares, como lo es la medición del ángulo de Cobb en imágenes de rayos X para la detección de escoliosis Horng et al., (2019). Los autores proponen un sistema automático de medición de la curvatura de la columna vertebral mediante radiografías en vista anteroposterior, reconstruyendo una imagen segmentada completa y así calcular la curvatura basándose en el criterio del ángulo de Cobb. Utilizaron una red implementada basada en el marco de Tensorflow en Python, y obtuvieron resultados favorables que no varían significativamente con relación a registros manuales realizados por dos médicos diferentes. Mostrando así la validez de un método que proporciona una medición fiable del ángulo de Cobb para aplicaciones clínicas utilizando redes neuronales convolucionales.

Poder clasificar la espondilolistesis en imágenes de resonancia magnética puede representar un desafío debido a la dificultad en la localización de puntos de interés u objetos críticos en imágenes que poseen distintas características. Zhao et al., (2019) propusieron una red de reconocimiento adversarial rápida (FAR, por sus siglas en inglés), que permite realizar con precisión la graduación de la espondilolistesis en 3 niveles detectando de manera correcta las vértebras relevantes sin necesidad de localizar los puntos de referencia. Utilizando la red de reconocimiento multitarea, que es una red de supervisión híbrida compuesta por una red jerárquica de extracción de rasgos, una red de propuesta regional y un módulo de clasificación de detección multitarea, obtuvieron resultados en la clasificación de espondilolistesis (precisión de entrenamiento: 0.9883 ± 0.0094 , precisión de prueba: 0.8933 ± 0.0276) para imágenes de RM de diferentes modalidades. Esta precisión es comparable a la de los médicos y supera a la de otros métodos actuales. Este tipo de trabajos de investigación abren camino para la retro propagación de gradiente mediante la introducción de un tema de supervisión híbrida en la red de detección multietapas.

Siguiendo la línea de la medición automática del ángulo de Cobb para la detección de escoliosis, Tu et al. (2019), proponen de manera similar un algoritmo para calcular automáticamente el ángulo de Cobb sin la definición manual del médico, utilizando una red de segmentación y detección DU-Net para remover las regiones sin relación y poder segmentar el contorno de la columna en una imagen de rayos X. Utilizaron una CNN profunda basada en el marco PyTorch, implementada en Python para obtener resultados de medición de ángulo de Cobb con un error medio de 2.9° respecto al ángulo medido manualmente por un ortopeda especializado. El algoritmo de detección de este artículo ofrece una precisión media del 98.5% y una recuperación del 99.5%. Además, la DU-Net alcanza un coeficiente Dice medio con respecto a la segmentación de referencia del 90.28%, un IoU del 82.29% y una precisión del 86.30%. Con el fin de reducir la dependencia de la subjetividad que trae la experiencia de un médico, la propuesta de los autores logra mostrar buena consistencia entre la medición automática del ángulo de Cobb sobre la manual, dejando como futuro trabajo la optimización de la red de segmentación y así mejorar la precisión.

Otro trabajo publicado en el *European Spine Journal* realiza una comparativa entre la medición manual y automatizada del ángulo de Cobb, sin embargo, esta vez basado en la tecnología de aprendizaje profundo para la detección de puntos clave. Estudio en el cual los autores marcaron el nombre de cada vértebra junto con sus cuatro vértices, para poder luego medir el ángulo de Cobb de la curva principal y visualizar cualquier curva adyacente mayor a 10° por encima y por debajo de ella. Utilizando la tecnología de detección de objetos de CenterNet, que usa la estimación de puntos clave para encontrar puntos centrales y luego retrocede, midiendo el tamaño de los bordes, lograron obtener resultados en el ángulo de Cobb medio en los casos de prueba fue de $27.4^\circ \pm 19.2^\circ$ (intervalo 0.00-91.00°) con las mediciones manuales y de $26.4^\circ \pm 18.9^\circ$ (intervalo 0.00-88.00°) con las mediciones automatizadas. El método automatizado necesitó una media de 4.45 s para medir cada radiografía. Los autores lograron identificar que, cuando el ángulo de Cobb no excedía los 90°, los resultados tenían un grado alto de fiabilidad con respecto a las anotaciones realizadas por los médicos, sin embargo, cuando estas eran mayores a 90°, las mediciones se desviaban mucho. Dentro de las limitantes también mencionan el bajo número de imágenes del conjunto de pruebas, y que el sistema tenía mayor dificultad para identificar las cuatro esquinas de la vértebra L5. (Sun et al., 2022)

En un artículo publicado por Kim et al. (2020) se propone un sistema automatizado para la evaluación de la escoliosis a partir de radiografías de la columna vertebral, con el fin de resolver el inconveniente común con la medición manual, que estas pueden llegar a requerir mucho tiempo, son subjetivas dependiendo de la experiencia y el criterio del médico quien está realizando la evaluación y pueden volverse complejas. Utilizando PyTorch 1.1. en Python, realizaron el entrenamiento y la evaluación para poder localizar e identificar las vértebras tanto torácicas como lumbares, estimar la inclinación de dichas vértebras y medir automáticamente el ángulo de Cobb, identificando como problema principal la falta de un método que logre una medición precisa y visualmente explicable del ángulo de Cobb en términos de rendimiento clínico, el método de estimación directa a pesar de tener resultados muy precisos carece de interpretación, mientras que el de estimación indirecta depende de la calidad de la estimación de los puntos de referencia o de la segmentación de las vértebras, por lo cual la fiabilidad de sus resultados puede verse comprobada. Los autores integraron las ventajas del método de

estimación directa al método indirecto en el marco propuesto, obteniendo como resultado un error absoluto medio de 3.51 y un error porcentual absoluto medio de 7.84% para el ángulo de Cobb. Demostrando así que dicho método alcanza una estimación precisa del ángulo de Cobb mediante un sistema visualmente explicable.

Otro estudio realizado por Balzer et al. (2022) donde se evalúa la viabilidad, el desarrollo y la validación de una estructura de aprendizaje profundo basado en la combinación de CNN y arquitectura U-Net que abarca desde la entrada del estudio de resonancia magnética, hasta la segmentación y clasificación de los discos intervertebrales, con el fin de obtener como resultado cuantitativo distintos grados radiológicos de degeneración. El modelo propuesto fue entrenado para detectar de manera automática con base en los siguientes sistemas de clasificación:

- Clasificación Pfirmann
- *Endplate defects*
- Protrusión discal / herniación
- Nódulos de Schmorl
- Fisuras anulares
- Clasificación de Meyerding para espondilolistesis

Donde se demostró que, aun existiendo espacio para mejora del rendimiento de las clasificaciones individuales, el modelo propuesto brinda una manera objetiva de evaluar la columna vertebral y reduce o elimina la dependencia humana para su diagnóstico. A diferencia de otros trabajos de investigación relacionados, los autores aseguran que el modelo es completamente independiente de la interacción manual, mostrando resultados específicamente para la espondilolistesis en un $89.03\% \pm 2.42\%$ de coeficiente de similitud Dice (DSC) y $95.98\% \pm 1.82\%$ de área bajo la curva de precisión-sensibilidad (AUC-PR).

Continuando con una investigación realizada por Trinh et al., (2022) se pretendió utilizar el aprendizaje profundo para la detección de la espondilolistesis en imágenes de rayos X. Se utilizó el sistema de clasificación de meyerding para clasificar el grado de deslizamiento de una vértebra sobre la otra. El método utilizado fue el LumbarNet para el preprocesamiento de las imágenes de las vértebras en las imágenes de rayos X, este es un modelo creado por los autores que se distingue de los métodos tradicionales como UNet debido a una característica llamada feature fusion module (FFM) Fue diseñado para reorganizar la característica latente extraída por la ruta de codificación, de modo que se pueda obtener una característica más representativa para la

segmentación lumbar. Se utilizaron métodos de búsqueda de contornos para detectar la región de interés. En la investigación se utilizaron un total de 706 imágenes de rayos X en la zona lumbar donde 312 imágenes presentaban algún tipo de anomalía en dicha zona, se utilizaron 250 imágenes para el entrenamiento y 62 imágenes para las pruebas. Se compararon diferentes métodos para la detección de la espondilolistesis, el P-grade donde se definen ciertos puntos de las vértebras y se compara con una constante k para determinar el grado de afección, obtuvo un 88.05% de acierto, el Piecewise Slope Detection (PSD) aquí se detecta el ángulo de desplazamiento vertical y se utiliza para detectar directamente la variación del ángulo entre la región lumbar proximal y su espacio. El ángulo se calcula usando el teorema del coseno para la brecha entre los puntos superior e inferior y la vertical correspondiente, este obtuvo un 81.22% de acierto. Otro de los métodos fue el Dynamic Shift (DS), aquí se necesitaban 2 imágenes de rayos X, una era en la posición de extensión y otro en la posición de flexión para calcular el desplazamiento en las zonas lumbares, el método obtuvo un 81.42% de acierto. Por último, se juntaron dos métodos, el P-grade y el PSD para obtener un 88.83% de acierto.

Algunas especialidades médicas y enfermedades desarrollaron algoritmos basados en inteligencia artificial, como detectar fibrilación auricular, ictus y realizar diagnósticos basados en histopatología o imágenes médicas. Raraz-Vidal & Raraz-Vidal, (2022) explican en su investigación acerca de la inteligencia artificial en la medicina. Para ellos, la introducción de estas herramientas tecnológicas ha sido anticipada por los pacientes para una mayor autonomía y su capacidad de brindar atención personalizada, pero han encontrado resistencia por parte de los médicos que no están preparados para estos avances. Esto crea la necesidad de validar estas herramientas modernas a través de la práctica clínica y de laboratorio tradicional y de discutir la actualización del plan de estudios en las facultades de medicina con un enfoque en la medicina digital. La inteligencia artificial ha demostrado ser útil en neurología, oncología médica, radiología y patología clínica, contribuyendo a un diagnóstico más preciso, a la estratificación de enfermedades y a una mayor eficiencia en la práctica médica. Según Raraz-Vidal & Raraz-Vidal, (2022) es importante formarse en inteligencia artificial e integrar esta tecnología en la práctica clínica. Las universidades y las instituciones sanitarias deberían crear políticas sanitarias que

fomenten la implementación de la IA y promuevan la investigación científica en todas las especialidades médicas.

En otra investigación realizada por Fraiwan et al., (2022) se busca explorar y comparar las habilidades de 14 redes neuronales convolucionales pre entrenadas y saber si son aptas para detectar de manera automática si el paciente sufre de escoliosis, espondilolistesis o si no sufre de ninguna afección en la columna. El conjunto de imágenes incluye 338 sujetos donde 71 personas no sufren de afección alguna en la columna, 79 pacientes sufren de espondilolistesis y 188 sufren de escoliosis. Algunos de los modelos de redes neuronales son SqueezeNet, GoogleNet, Inception V3, DenseNet V2, MobileNet V2, etc. Los 14 modelos se han probado bajo diferentes circunstancias ya sea bajo las tres clases al mismo tiempo (normal vs escoliosis vs espondilolistesis) o usando dos de las tres clases (normal vs escoliosis, normal vs espondilolistesis, escoliosis vs espondilolistesis). El objetivo era probar cuál de ellos cumple con las métricas a tomar en cuenta como el promedio de precisión, precisión máxima, precisión mínima, desviación estándar, precisión, F1-score, sensibilidad, especificidad y es más preciso a la hora de detectar dichas afecciones. Después de haber realizado dichas pruebas, la red neuronal que obtuvo mayor puntaje fue el DensNet-201 con un promedio del 96%, una máxima del 99% y una mínima del 94%.

Xuan et al., (2023) en la investigación realizada, explica el desarrollo de un sistema de apoyo para el diagnóstico de enfermedades de la columna mediante RM y técnicas de aprendizaje profundo. Entre julio de 2016 y julio de 2021, se recopilaron imágenes de resonancia magnética de 604 pacientes con afecciones de la columna, incluidas hernia de disco lumbar y espondilolistesis. Se redujo la sensibilidad del conjunto de datos de imágenes para proteger la privacidad del paciente. Luego utilizaron una herramienta de anotación de imágenes para etiquetar imágenes de resonancia magnética de la columna para entrenar un algoritmo de aprendizaje profundo.

Para el desarrollo de un sistema de soporte al diagnóstico, los autores eligieron el algoritmo de detección de objetos YOLO, es decir, sus versiones YOLOv3, YOLOv5 y PP-YOLOv2. utilizaron el aprendizaje por transferencia para aplicar estos modelos a la tarea de reconocer y diagnosticar imágenes de resonancia magnética de la médula espinal. Luego evaluaron el rendimiento de estos modelos utilizando métricas como precisión, sensibilidad, especificidad y

precisión promedio (mAP). Como resultado, en comparación con YOLOv3 y YOLOv5, se descubrió que el modelo PP-YOLOv2 tiene la mayor precisión en el diagnóstico de hernia, espondilolistesis y columna normal. Después se realizó una prueba donde se comparaba el diagnóstico de tres doctores con diferentes años de experiencia, esta prueba consiste en diagnosticar 50 imágenes de RM.

Tabla 2 métricas evaluadas en la investigación de Xuan et al., (2023).

	Sensibilidad	Especificidad	Precisión
Doctor 1	74.75%	93.78%	84.14%
Doctor 2	76.77%	92.75%	84.65%
Doctor 3	90.91%	94.30%	92.58%
Sistema de IA	97.98%	98.45%	98.21%

Como se puede observar en la tabla anterior, aunque la diferencia de los resultados entre los especialistas y la inteligencia artificial no es significativa, si se puede observar una ligera superioridad. Los autores proponen el uso de esta inteligencia artificial como asistente para los doctores ya que ayudaría a mejorar los diagnósticos.

En el artículo realizado por Zhang et al., (2023) se hace un algoritmo de aprendizaje profundo para la detección de la espondilolistesis basado en la clasificación de Meyerding. Aquí se utilizó un conjunto de imágenes de más de 1500 radiografías laterales de la zona lumbar donde se utilizaron 1117 para entrenamiento, 240 para validación y 239 para realizar pruebas. Utilizaron dos modelos, Faster R-CNN y RetinaNet, ambos algoritmos son especializados para la detección de objetos y la extracción de características. El objetivo de la investigación era comparar los resultados obtenidos del algoritmo con un grupo de doctores expertos en el diagnóstico de imágenes médicas.

Tabla 3 comparación entre algoritmo y grupo de doctores en la investigación de Zhang et al., (2023)

	Faster R-CNN	Doctores sin IA	Valor-P
Precisión	93.5%	92.7%	0.126
Sensibilidad	93.5%	89.2%	< 0.001
F1	93.5%	91%	< 0.001

Tabla 4 comparación entre diagnóstico con asistencia y sin asistencia en la investigación de Zhang et al., (2023)

	Doctores sin IA	Doctores con IA	Valor-P
Precisión	92.7%	97.5%	0.004
Sensibilidad	89.2%	97.4%	0.004
F1	91%	97.4%	0.004
Tiempo promedio	25.45min	18.313min	0.004

Las pruebas que se realizaron, se comparó primero la inteligencia artificial contra un grupo de doctores expertos en imágenes médicas donde en todos los aspectos a medir, la IA ha dominado los resultados. Después se realizó una comparativa de doctores con asistencia de la IA contra doctores los resultados de estos mismos doctores sin IA. Los doctores con IA consiguieron un mayor puntaje en todas las métricas, y mejoraron sus propios resultados.

Dlshad Ahmed & Hawezi, (2023) compararon varias técnicas para detectar fracturas en imágenes de rayos X. El sistema consta de cuatro módulos principales: preprocesamiento, detección de bordes, extracción de características y clasificación. El preprocesamiento convierte la imagen de RGB a escala de grises y utiliza un filtro gaussiano para eliminar el ruido. Luego aplicaron la detección de bordes utilizando el método Canny con ajustes que incluyen un histograma adaptativo para mejorar el contraste. La extracción de características se realiza utilizando una matriz de coproductos de nivel de grises (GLCM), que extrae características de textura como contraste, correlación, homogeneidad, energía e inconsistencia. Finalmente, la clasificación se realiza mediante técnicas como Gaussian Naive Bayes, Decision Tree, Nearest Neighbors, Support Vector Machines (SVM) y Random Forests. El sistema se evaluó utilizando conjuntos de datos radiográficos, utilizaron 60 imágenes de pacientes fracturados y 210 imágenes de pacientes no fracturados y se obtuvieron resultados satisfactorios.

Tabla 5 Resultados obtenidos en la investigación de Dlshad Ahmed & Hawezi, (2023)

ML Algorithm	Precisión	Sensibilidad	Exactitud
Naive Bayes	88%	65%	64%
Decision Tree	91%	85%	80%
Nearest Neighbors	84%	100%	84%
Random Forests	88%	96%	86%
SVM	94%	98%	93%

Como se puede observar en la Tabla 5 Resultados obtenidos en la investigación de Dlshad Ahmed & Hawezi, (2023), las diferentes técnicas han obtenido resultados positivos en la detección

de fracturas de hueso y cabe destacar que el algoritmo de Support Vector Machines fue el que mejores resultados obtuvo durante toda su investigación.

En su estudio centrado en la espondilolistesis Cai et al., (2017) desarrollaron un método para detectar automáticamente la espondilolistesis en imágenes de la columna y medir automáticamente el grado de dicha afección. Este problema surgió porque la medición manual de la gravedad de la espondilolistesis es difícil y poco confiable, lo que puede afectar el tratamiento de los pacientes. Para solucionar este problema se propuso un método que consta de tres pasos principales: síntesis del modelo, entrenamiento y prueba. En la etapa de síntesis, se creó un modelo de espondilolistesis utilizando métodos de dibujo y deformación articular. Estas muestras se utilizaron para entrenar un detector de espondilolistesis basado en aprendizaje supervisado utilizando una red profunda multimodal. La red fue entrenada para detectar y clasificar la presencia y gravedad de la espondilolistesis en imágenes de resonancia magnética (RM) y tomografía computarizada (TC). Los resultados mostraron una alta sensibilidad y especificidad en la detección de espondilolistesis, así como una precisión aceptable en la clasificación del grado de espondilolistesis.

En otro estudio realizado por Varçın et al., (2021) se desarrolló un diagnóstico computarizado de la espondilolistesis utilizando imágenes de rayos X de la zona lumbar del paciente. Para ello se utilizó un conjunto de imágenes de 2746 rayos x lumbares dentro de los cuales 1407 son rayos X normales y 1339 contienen espondilolistesis. Debido a que muchas de las imágenes variaban en tamaño y brillos, se tuvieron que aplicar diferentes operaciones como magnificación, voltear la imagen y cambios de brillo. El diagnóstico automatizado para la espondilolistesis lumbar se dividió en dos etapas. Primero, se utilizó YoloV3 para conseguir la región de interés (ROI) que en este caso son las imágenes que contienen las vértebras L4, L5 y S1. Luego se utilizó MobileNet para analizar el rendimiento de la clasificación y analizó métricas como la precisión y la sensibilidad. Aquí se evaluó la capacidad del algoritmo de clasificar las imágenes.

Tabla 6 Resultados del total de imágenes clasificadas en la investigación de Varcin et al., (2021)

Normal	Verdaderos Positivos	317
	Verdaderos Negativos	274
Espondilolistesis	Falsos Positivos	1
	Falsos Negativos	6

Como podemos observar en la Tabla 6 Resultados del total de imágenes clasificadas en la investigación de Varcin et al., (2021), los resultados obtenidos demuestran el excelente rendimiento del algoritmo con la menor tasa de error. El algoritmo ha demostrado una excelente capacidad para proporcionar diagnósticos precisos para los pacientes con espondilolistesis como los pacientes sin alguna afección.

Tabla 7 Total de imágenes clasificadas en la investigación de Cai et al., (2017)

Grado	# de casos	Tasa de error de detección	Tasa de error de clasificación
Grado 0 (real)	10 RM + 10TC	10%	11%
Grado 1 (real)	7RM + 6TC	15%	18%
Grado 2 (real)	3RM + 3TC	17%	20%
Grado 3 (sintético)	3RM + 3TC	0	16%
Grado 4 (sintético)	3RM + 3TC	0	0
Grado 5 (sintético)	3RM + 3TC	0	0

También se realizó una validación cruzada utilizando conjuntos de datos reales y sintéticos para demostrar la eficacia del método propuesto incluso cuando se utiliza información sintética para el aprendizaje. En conclusión, los autores han desarrollado un enfoque integral para la detección y medición automática de la espondilolistesis que supera las limitaciones de los métodos manuales existentes y proporciona una herramienta objetiva para la evaluación clínica de esta afección.

Qi et al., (2020) realizaron un estudio centrado en el análisis de radiografías de fémur. Se excluyeron pacientes con fracturas de cuello femoral o intertrocanterías, tumores u otras fracturas patológicas. Además, se excluyeron los pacientes con antecedentes de cirugía y radiografías sobreexpuestas o poco claras. Sin embargo, se incluyeron pacientes cuyas películas no correspondían completamente a las posiciones frontal y lateral estándar si esto no afectaba la calidad de la interpretación. Se recogieron un total de 2.333 radiografías de fémur, incluidas imágenes normales y de fractura. Estas imágenes se clasificaron en tres grupos de fracturas: fracturas simples, fracturas en cuña y fracturas complejas. La clasificación de las fracturas fue

evaluada por tres cirujanos ortopédicos participantes y un radiólogo musculoesquelético utilizando el sistema AO-Müller/Orthopaedic Trauma Association (AO/OTA). Las imágenes se etiquetaron manualmente para entrenar un detector de objetos para detectar y clasificar fracturas de fémur en imágenes de rayos X según el sistema AO. Se utilizó un modelo de detección RCNN más rápido con ResNet-50 y FPN como columna vertebral, y se evaluaron métricas como la precisión promedio, la precisión y la recuperación para diferentes tipos de fracturas. También se comparó el rendimiento del modelo de detección con un modelo de clasificación puro creado con la red neuronal convolucional ResNet50 y con tres doctores especialistas. Los resultados muestran que el modelo de detección con un porcentaje de precisión del 71.5% supera al modelo de clasificación con un 45% y a los cirujanos ortopédicos con un 54.8%, 50.5% y 50.7% de precisión, esto demuestra la efectividad del enfoque propuesto por los autores.

3.2 PROBLEMÁTICA

El diagnóstico de la espondilolistesis utilizando el sistema de clasificación de Meyerding suele considerarse como una práctica compleja para los médicos, pues, cuantificar de manera adecuada el grado de afección consistentemente requiere de mucha preparación y experiencia previa, además, la calibración y medición de las características utilizando imágenes de rayos X impresas consume una cantidad significativa de tiempo. Volviendo este método convencional subjetivo e impráctico para la evaluación de múltiples pacientes en un corto lapso. Las afecciones de la columna vertebral como la espondilolistesis generalmente no presentan síntomas de dolor, y, únicamente se puede detectar mediante estudios radiológicos, haciendo de su diagnóstico prematuro un desafío para el personal sanitario, influyendo en la condición de vida del paciente (Nguyen et al., 2021).

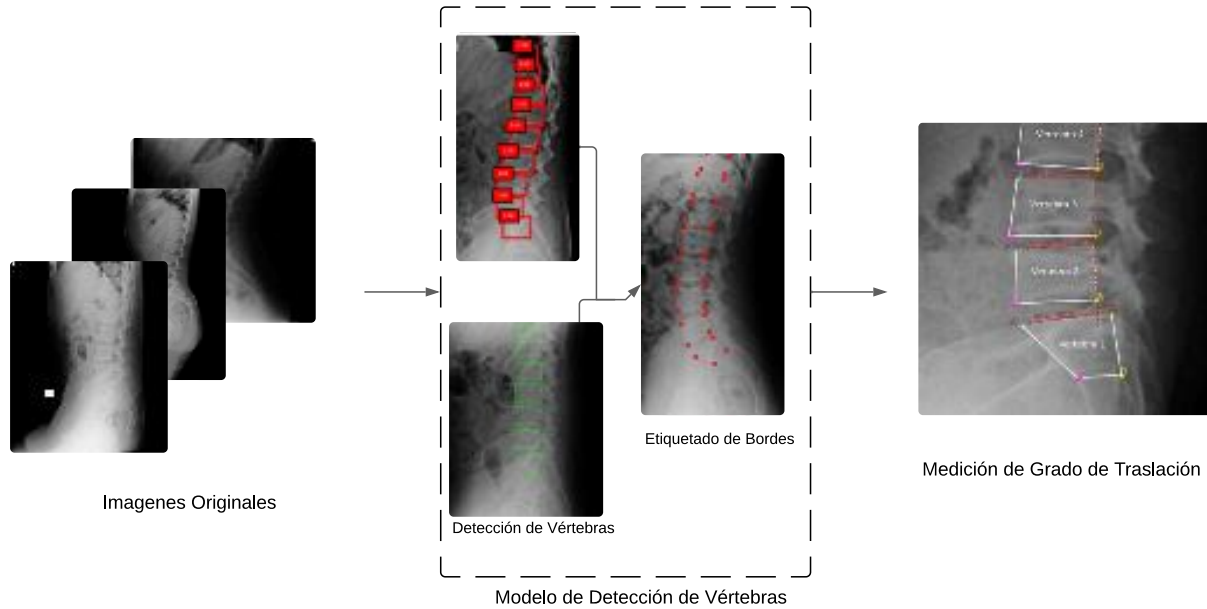
Actualmente, las radiografías siguen siendo la herramienta de imagen más utilizada para el diagnóstico inicial de la espondilolistesis debido a su accesibilidad y bajo costo. Las radiografías laterales de pie son especialmente eficaces, ya que pueden identificar deslizamientos vertebrales en adultos mayores con dolor lumbar crónico, y son más sensibles en comparación con la RM en posición supina para detectar espondilolistesis degenerativa (Mohile et al., 2022). Además, las radiografías dinámicas de flexión y extensión son el estándar de oro para detectar inestabilidad

vertebral, con una sensibilidad significativa en la identificación de traslaciones vertebrales superiores a 3mm (Vanti et al., 2021).

El campo de la salud está en constante evolución, con la llegada de la inteligencia artificial y los modelos de aprendizaje profundo, se logra establecer cada vez más un desarrollo tecnológico para los procesos de diagnóstico ya existentes. El uso de modelos de aprendizaje profundo permite agilizar procesos, reducir tiempos de espera brindando resultados fiables y cada vez más cercanos a los resultados propuestos por los humanos, mejorando así la calidad de vida de los pacientes. Es por ello por lo que el desarrollo de algoritmos para la medición automática de espondilolistesis representa un avance tecnológico relevante en el desarrollo clínico.

El desarrollo de un modelo de aprendizaje profundo para la detección de espondilolistesis presenta distintos desafíos que comprometen su rendimiento, factores como la visibilidad de las imágenes, la escasez de bases de datos extensas que contengan imágenes exclusivamente de espondilolistesis, las enfermedades adicionales que pueden padecer los pacientes de las imágenes estudiadas como fracturas, degeneración ósea, etc. Incluso considerándose más compleja la segmentación automática en imágenes de rayos X que en estudios de tomografía computarizada o resonancia magnética, debido a las sombras sobrepuestas de estructuras toracoabdominales (Kim et al., 2020). Razón por lo cual es un tema aún en desarrollo y con mucho espacio para mejoras.

3.3 IMAGEN INTEGRADORA



Fuente: Elaboración Propia

Ilustración 1 Imagen Integradora.

En la Ilustración 1 Imagen Integradora, la imagen de rayos X en vista lateral que se presenta muestra la columna vertebral de un paciente con el propósito de detectar y clasificar la espondilolistesis. Esta imagen es procesada mediante una red neuronal diseñada para identificar las vértebras. La red neuronal predice los puntos clave en cada esquina de las vértebras y dibuja una caja delimitadora alrededor de cada una. La salida generada por esta red neuronal se utiliza como entrada para un algoritmo de clasificación de espondilolistesis. Este algoritmo emplea los puntos clave para trazar líneas de referencia, medir y calcular el porcentaje de traslación, lo que permite evaluar la gravedad de la espondilolistesis en el paciente.

3.4 TABLA DE LIMITACIONES

Tabla 8 Tabla de Limitaciones

Autores	País del estudio	Enfoque de estudio	Tipo de limitación	Detalle de limitación
(Moon et al., 2022)	República de Corea	Desarrollo de modelo de detección de fracturas de huesos faciales	Falta de datos	Pocas imágenes para entrenamiento de otras fracturas que no fuesen fracturas nasales.
(Singh et al., 2022)	República de Corea	Desarrollo de modelo de detección automática de fracturas en partes superiores del cuerpo	Tamaño limitado de conjunto de datos	El uso único de radiografías de dedo en puede afectar la generalización del modelo a otros tipos de estudio s o radiografías de otras partes del cuerpo.
(Xuan et al., 2023)	China	Comparación de modelos de detección automática de espondilolistesis con doctores especialistas	falta de exploración de otros modelos	Se utilizaron diferentes modelos de aprendizaje profundo para el análisis de imágenes de como YOLOv3, YOLOv5 y PP-YOLOv2. No se exploraron otros modelos de inteligencia artificial que podrían ofrecer resultados iguales o superiores.
(Varçın et al., 2021)	Turquía	Desarrollo de un diagnóstico computarizado de espondilolistesis.	Limitación de base de datos	Se tuvieron que descartar rayos x de pacientes con dolores provocados por otras afecciones diferentes a la espondilolistesis, lo que acorto el conjunto de imágenes para entrenar al algoritmo.
Irina Balzer, et al.	Suiza	Evalúa la viabilidad, el desarrollo y la validación de una línea de aprendizaje automático que incluye todas las tareas desde la entrada de la IRM hasta la segmentación y clasificación de los discos intervertebrales de la columna lumbar	Limitante de datos de imágenes	Los casos extremos de escoliosis lumbar fueron excluidos. Por lo tanto, este marco no se aplica a estos casos. Además, el proceso se basa en la evaluación de cortes sagitales solamente. Para obtener una mayor relevancia clínica, el marco debería ampliarse con cortes transversales

Autores	País del estudio	Enfoque de estudio	Tipo de limitación	Detalle de limitación
Yu Sun, et al.	China	Comparativa entre medición manual y automatizada del ángulo de Cobb en la escoliosis idiopática basada en la tecnología de detección de puntos clave de aprendizaje profundo.	Dificultades con las mediciones de puntos y ángulos	Las mediciones realizadas por el sistema se desviaban mucho de las mediciones manuales cuando la curva era superior a 90°. El sistema tiene dificultades para identificar las cuatro esquinas de la vértebra L5.
Thong Phi Nguyen, et al.	República de Corea	el desarrollo de un sistema de aprendizaje profundo con el apoyo de la red neuronal convolucional (CNN)	Falta de tiempo para entrenar el modelo	La detección de múltiples pasos no sólo establece los requisitos para la construcción del conjunto de datos para que coincida con los requisitos de detección de cada modelo individual, sino que también toma tiempo para entrenar exhaustivamente cada modelo.
Kang Cheol Kim, et al	Corea del Sur	Técnica de segmentación automática de imágenes de rayos X que combina métodos de aprendizaje profundo y de conjunto de niveles	Dificultades para capturar los puntos clave	La pared posterior del cuerpo vertebral tiene límites poco claros debido a la superposición de dos pedículos en la vista lateral de la radiografía lumbar, es difícil capturar el límite de la pared posterior utilizando el conjunto de niveles.
Kang Cheol Kim, et al.	Corea del Sur	Un sistema automatizado y visualmente explicable para la evaluación de la escoliosis a partir de radiografías de la columna vertebral	Falta de recursos previos existentes	Ningún método existente ha logrado una medición precisa y visualmente explicable del ángulo de Cobb en términos de rendimiento clínico.
Fatih Varcin, et al.	Turquía	Comparativa de dos redes neuronales AlexNet y GoogLeNet	Falta de bases de datos de la afección a estudiar	No hay base de datos pública de imágenes radiológicas de espondilolistesis

Autores	País del estudio	Enfoque de estudio	Tipo de limitación	Detalle de limitación
Fabio Galbusera, et al.	Alemania	Un método automatizado de extracción de parámetros anatómicos a partir de radiografías biplanares de la columna vertebral	Base de datos limitada	el tamaño utilizado en el presente estudio (443 pares de imágenes) debe considerarse muy bajo en el contexto de la investigación sobre aprendizaje automático, en particular para el entrenamiento de redes neuronales profundas
Shu Liao, et al.	n/a	Método automático de estimación del grado de espondilolistesis lumbar. Etiquetado automático de la columna vertebral y determinación de regiones críticas	Se vieron afectadas las mediciones por otras enfermedades	Las enfermedades adicionales que existían en algunos pacientes, como la degeneración, fracturas vertebrales y escoliosis impedían utilizar esas imágenes para la base de datos.

3.5 TABLA DE METODOLOGÍAS

Tabla 9 Metodologías de Estudios previos

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	Modelo utilizado	Métricas	Resultados
End-To-End Computerized Diagnosis of Spondylolisthesis Using Only Lumbar X-rays. (Varçin et al., 2021)	Dataset de 2746 rayos x lumbares dentro de los cuales 1407 son rayos X normales y 1339 se diagnosticaron con espondilolistesis	Se utilizó augmentator, este es un software para generar artificialmente datos de imágenes para problemas de aprendizaje automático. Gracias al software, las imágenes se pueden rotar, recortar, cambiar de tamaño, etc. Se crean procedimientos aumentados que se vinculen a las siguientes acciones: Las tareas se aplican de forma estocástica, con parámetros asignados a cada tarea de forma aleatoria dentro de un intervalo predefinido.	Se utilizó YOLO V3 que utiliza DarkNet-53 para la detección de las vértebras y la región de interés (ROI). Para la detección de espondilolistesis se afinó un modelo de MobileNet, este afinamiento se llevó a cabo con el algoritmo de optimización de Adam, entropía transversal, tamaño de batch de 32. ritmo de aprendizaje de 0.001 y 25 épocas.	exactitud, sensibilidad, especificidad, AUC, precisión, puntuación F1 y matriz de confusión	Exactitud 0.99, Sensibilidad 0.98, Especificidad 0.99, AUC 0.99, Precisión 0.99, Puntuación F1 0.98. Para la matriz de confusión, se registraron 317 imágenes detectadas correctamente como normales, 1 detectado falsamente como espondilolistesis, 6 detectados erróneamente como espondilolistesis y 274 correctamente como espondilolistesis
Deep learning system assisted detection and localization of lumbar	se utilizó un conjunto de imágenes de más de 1596 radiografías laterales de la zona lumbar donde se	Las imágenes importadas se escalaron utilizando el tamaño de borde más corto y establecieron	Para esta investigación utilizaron modelos de aprendizaje	esta se dos de precisión, sensibilidad y puntuación F1	Se compararon primero los resultados obtenidos entre la red neuronal y un grupo de especialistas, después se compararon resultados

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	Modelo utilizado	Métricas	Resultados
spondylolisthesis. (Zhang et al., 2023)	utilizaron 1117 para entrenamiento, 240 para validación y 239 para realizar pruebas	los pixeles en 800, con el borde más largo escalado de igual forma. Además, existió la probabilidad del 50% de que una imagen se voltee de forma horizontal aumentando la riqueza de las imágenes. se realizó una normalización de las imágenes, esta normalización de imágenes es el proceso de convertir una imagen sin procesar a su propio formato normalizado mediante una serie de transformaciones para eliminar la influencia de otras funciones de transformación en la imagen.	profundo, Faster R-CNN y RetinaNet. El faster R-CNN se utilizó como modelo para la detección de objetos ya que posee características como capas convolucionales, red de propuestas regionales (RPN), clasificadores y una gran velocidad para la detección de objetos, poco visto hasta ahora. Se utilizo RetinaNet que utiliza la red de pirámide de características (FPN) y utiliza dos subredes para la clasificación de cada nivel para lograr alta precisión en el reconocimiento de objetos.		entre doctores con asistencia de inteligencia artificial y doctores sin asistencia
Detection of Lumbar Spondylolisthesis from X-ray Images Using Deep Learning Network. (Trinh et al., 2022)	En la investigación se utilizaron un total de 706 imágenes de rayos X en la zona lumbar donde 312 imágenes	El método utilizado fue el LumbarNet para el preprocesamiento de las imágenes de las vértebras en las imágenes de rayos X,	Se utilizaron métodos de búsqueda de contornos para detectar la región de interés. Se compararon	Exactitud, Sensibilidad, Especificidad, falsos positivos y falsos negativos.	el P-grade obtuvo un 88.05% de acierto, el Piecewise Slope Detection obtuvo un 81.22% de acierto. Otro de los métodos fue el Dynamic Shift (DS),

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	Modelo utilizado	Métricas	Resultados
	<p>presentaban algún tipo de anomalía en dicha zona, se utilizaron 250 imágenes para el entrenamiento y 62 imágenes para las pruebas.</p>	<p>este es un modelo creado por los autores que se distingue de los métodos tradicionales como UNet debido a una característica llamada feature fusión module (FFM) Fue diseñado para reorganizar la característica latente extraída por la ruta de codificación, de modo que se pueda obtener una característica más representativa para la segmentación lumbar</p>	<p>diferentes métodos para la detección de la espondilolistesis, el P-grade donde se definen ciertos puntos de las vértebras y se compara con una constante k para determinar el grado de afección, el Piecewise Slope Detection (PSD) aquí se detecta el ángulo de desplazamiento vertical y se utiliza para detectar directamente la variación del ángulo entre la región lumbar proximal y su espacio. El ángulo se calcula usando el teorema del coseno para la brecha entre los puntos superior e inferior y la vertical correspondiente. Otro de los métodos fue el Dynamic Shift (DS), aquí se necesitaban 2</p>		<p>obtuvo un 81.42% de acierto. Por ultimo los dos métodos juntos, el P-grade y el PSD obtuvieron un 88.83% de acierto.</p>

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	de	Modelo utilizado	Métricas	Resultados
				imágenes de rayos x, una era en la posición de extensión y otro en la posición de flexión para calcular el desplazamiento en las zonas lumbares. Por último, se juntaron dos métodos, el P-grade y el PSD.		
Direct Spondylolisthesis Identification and Measurement in MR/CT using Detectors Trained by Articulated Parameterized Spine Model. (Cai et al., 2017)	El conjunto de datos obtenidos de SpineWeb contenía 15 imágenes de cada tipo de estudio (RM+TC) con diferentes grados de problemas de espondilolistesis (principalmente Grado I y II). El conjunto de datos xVertSeg CT contenía 15 exploraciones lumbares y/o torácicas y el conjunto de datos Siegen MR contenía 17 imágenes de RM. Los conjuntos de	Como lo datasets de imágenes de espondilolistesis eran limitados, a los autores se les ocurrió crear imágenes de MRI y CT de pacientes con espondilolistesis a base de otras imágenes obtenidas utilizando spineWeb. Utilizaron las medidas de las bounding boxes para poder formar nuevas vertebras, cambiando parámetros como altura, anchura y largo, se realizaron movimientos de traslación y rotación y		El dataset con muestras de espondilolistesis reales y sintéticas se utilizó para entrenar un detector de espondilolistesis basado en aprendizaje supervisado utilizando una red profunda multimodal. La red fue entrenada para detectar y clasificar la presencia y gravedad (grado I a III) de la espondilolistesis en imágenes de resonancia magnética	Sensibilidad, especificidad, tasa de error en la detección y tasa de error en la graduación.	El método identificó con éxito 34 de 39 (RM+CT) casos reales de espondilolistesis y 9 de 9 casos de validación sintetizados en el experimento de prueba. La sensibilidad para los casos reales fue del 91,8% y la especificidad del 90,0%. La graduación estimada de la espondilolistesis en casos reales (Grado 0, 1, 2) alcanzó una tasa de éxito del 85,3% tanto para imágenes de RM como de TC.

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	de	Modelo utilizado	Métricas	Resultados
	datos xVertSeg y Siegen no contenían ningún caso de espondilolistesis.	se alteraron orientaciones ubicaciones determinadas vertebras.	las y de	(RM) y tomografía computarizada (TC).		
Using deep transfer learning to detect scoliosis and spondylolisthesis from x-ray images. (Fraiwan et al., 2022)	El conjunto de imágenes incluye 338 sujetos donde 71 personas no sufren de afección alguna en la columna, 79 pacientes sufren de espondilolistesis y 188 sufren de escoliosis.			Se compararon 14 modelos de redes neuronales convolucionales, entre ellos están: SqueezeNet, GoogleNet, Inception V3, DenseNet V2, MobileNet V2, ResNet-50, ResNet-101, ResNet-18, Xception, Inception-ResNet-v2, shuffleNet, NasNet-Mobile, DarkNet53, EfficientNet-b0	promedio de precisión, precisión máxima, precisión mínima, desviación estándar, precisión, F1-score, sensibilidad, especificidad	la red neuronal que obtuvo mayor puntaje fue el DensNet-201 con un promedio del 96%, una máxima del 99% y una mínima del 94%.
Hybrid Deep Learning Approach for Automatic Detection in Musculoskeletal Radiographs. (Singh et al., 2022)	Utilizaron el conjunto de datos de Mura, uno de los conjuntos de datos más grandes disponibles públicamente, para la detección de anomalías en radiografías	Se aplicaron múltiples funciones de procesamiento, como rotación, desplazamiento y volteo para obtener un mejor rendimiento, que luego pasó por compresión para		En este artículo se propone una arquitectura híbrida basada en redes neuronales profundas llamada ComDNet-512 para detectar las anomalías en la radiografía	Exactitud, Precisión, Recuperación, Puntuación F1, Valor Kappa, Verdaderos positivos, falsos negativos, verdaderos	El modelo entrenado obtuvo los siguientes resultados: Exactitud de 89.41%, Precisión de 0.82, Recuperación de 0.97, Puntuación F1 de 0.89, Valor Kappa de 0.74. se obtuvieron 80 verdaderos positivos, 5 falsos negativos,

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	Modelo utilizado	Métricas	Resultados
	<p>musculoesqueléticas de las extremidades superiores. Consta de 40.561, radiografías de 14.863 exámenes obtenidos del Archivo de Comunicaciones e Imágenes del Hospital Stanford. se utilizaron un total de 1702 imágenes normales y anormales y 300 radiografías para la validación.</p>	<p>eliminar la redundancia estadística. Se aplicó a todas las imágenes clínicas la compresión sin pérdida de fidelidad de los datos originales. La redundancia estadística de la imagen se eliminó aplicando la compresión.</p>	<p>musculoesquelética de un paciente. ComDNet-512 consta de tres fases: compresión, entrenamiento de la red neuronal densa y redimensionamiento progresivo. El modelo se entrenó con muestras de radiografías de dedos para realizar una predicción binaria, es decir, huesos normales o anormales.</p>	<p>positivos, falsos positivos</p>	<p>72 verdaderos positivos, 13 falsos positivos</p>
<p>Detection of bone fracture based on machine learning techniques. (Dlshad Ahmed & Hawezi, 2023)</p>	<p>se utilizó un conjunto de datos de imágenes de rayos X obtenidas del Hospital de Urgencias de Erbil, este contiene 60 imágenes fracturadas y 210 no fracturadas de huesos de la parte inferior de la pierna.</p>	<p>El preprocesamiento convierte la imagen de RGB a escala de grises y utiliza un filtro gaussiano para eliminar el ruido. Luego aplicaron la detección de bordes utilizando el método Canny con ajustes que incluyen un histograma adaptativo para mejorar el contraste. La extracción de características se</p>	<p>La clasificación se realizó mediante técnicas como Gaussian Naive Bayes, Decision Tree, Nearest Neighbors, Support Vector Machines (SVM) y Random Forests.</p>	<p>Exactitud, Precisión, Recuperación</p>	<p>Las diferentes técnicas han obtenido resultados positivos en la detección de fracturas de hueso y cabe destacar que el algoritmo de Support Vector Machines (SVM) fue el que mejores resultados obtuvo durante toda su investigación con una exactitud del 93%. recuperación de 0.98 y precisión de 94%.</p>

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	de	Modelo utilizado	Métricas	Resultados
		realiza utilizando una matriz de coproductos de nivel de grises (GLCM), que extrae características de textura como contraste, correlación, homogeneidad, energía e inconsistencia				
Automatic Lumbar Spondylolisthesis Measurement in CT Images. (Liao et al., 2016)	Utilizaron un conjunto de imágenes de CT para 258 pacientes, con resolución de 512x512	n/a		Método CARP («Critical Anatomy Region Propagation»), basado en el principio de propagación media no local multiatlasa.	Distancia euclidiana, inter-rater, intra-rater differences, matriz de confusión	El método logra una precisión de estimación del 94.4% (es decir, 776 de 821) de los grados de espondilolistesis de pares de vértebras se determinan correctamente.
Deep learning system for Meyerding classification and segmental motion measurement in diagnosis of lumbar spondylolisthesis. (Nguyen et al., 2021)	El conjunto de datos de entrenamiento fue construido con 1000 imágenes de rayos X laterales (382 hombres de edades 60.6 ± 15.5 años; 618 mujeres de edades 63.5 ± 16.1 años) obtenidas desde enero 2018 al diciembre 2019 del	todas las imágenes de rayos X del conjunto de datos se escalaron en el intervalo de 0-50 píxeles tanto horizontal como verticalmente y también se ajustaron con brillo y contraste múltiples		El primer modelo de regresión CNN basado en la arquitectura VGG-net se entrena para localizar la posición del punto clave en cada vértebra de T12 a S1 basándose en las imágenes de rayos X el segundo modelo	desviación de ángulos, matriz de confusión	La precisión del sistema de aprendizaje profundo propuesto muestra una desviación media de 1,76° entre la detección automática y la medición directa de un médico en 20 casos de prueba, pero también para la clasificación de Meyerding, que se valida de forma evidente basándose en los resultados

Titulo y Autor(es)	Dataset	Preprocesado de imágenes	Modelo utilizado	Métricas	Resultados
	International St. Mary's Hospital, Catholic Kwandong University College of Medicine, Incheon, Republic of Korea)		de regresión CNN se entrena para extraer los bordes izquierdo y derecho del cuerpo vertebral		de clasificación precisos para 49/50 casos experimentales.
Diagnosis of Lumbar Spondylolisthesis via Convolutional Neural Networks. (Varcin et al., 2019)	El conjunto de datos consta de 272 radiografías en total. 136 de estas imágenes pertenecen a pacientes que padecen espondilolistesis, y 136 imágenes sin espondilolistesis. Utilizando ImageNet	cada imagen se redimensionó a 227 × 227 × 3D.	Modelos pre-entrenados. AlexNet: Arquitectura "Traditional Sequential Network" con profundidad (8) y 25 capas. GoogleNet: arquitectura "Inception" con profundidad (22) y 144 capas.	Sensibilidad, especificidad, precisión, valor de predicción de negativo, tasa de falsos positivos, tasa de falsos negativos, tasa de descubrimiento de falsos, exactitud, F1 score, matriz de confusión	Los resultados experimentales demuestran que GoogleLeNet tiene una precisión del 93,87% y rinde ligeramente mejor que AlexNet, con una precisión del 91,67%.

Fuente: Elaboración Propia

III. OBJETIVOS

En esta sección definen los objetivos que guiaron el proyecto de investigación. Estos se establecieron para proporcionar una dirección clara y precisa para los esfuerzos de investigación, asegurando que cada fase del trabajo cumpliera con los objetivos generales y específicos de la investigación.

2.1 OBJETIVO GENERAL

- Desarrollar un método para la detección automatizada de la espondilolistesis en imágenes de rayos X de la columna vertebral mediante la clasificación de Meyerding, logrando una precisión del 95% en la detección de vertebras aplicando el refinamiento de un modelo de aprendizaje profundo de detección de puntos clave.

2.2 OBJETIVOS ESPECÍFICOS

- Clasificar imágenes de espondilolistesis en la base de datos pública e interna para el reentrenamiento del modelo de aprendizaje profundo y etiquetar imágenes incorporadas para ampliar el conjunto a más de 700 imágenes.
- Seleccionar parámetros de refinamiento del modelo y realizar entrenamiento con la variación determinando las capacidades del modelo con al menos 5 métricas de evaluación para obtener puntos de interés con exactitud.
- Desarrollar e implementar el algoritmo de cálculo para la clasificación de Meyerding automática utilizando los puntos de interés.
- Validar el método en la clasificación de Meyerding y la fiabilidad para la detección de la espondilolistesis en imágenes de la columna vertebral, con soporte de 3 médicos especialistas.

IV. MÉTODO

En el capítulo de la metodología se detallan los procedimientos aplicados para llevar a cabo la presente investigación. Se definen el enfoque de la investigación, su alcance, las variables dependientes e independientes, las técnicas y materiales, y la arquitectura delimitada para alcanzar los objetivos planteados.

4.1 ENFOQUE

La investigación emplea un enfoque cuantitativo ya que en la primera parte se evalúan distintas métricas numéricas como la precisión, exactitud, desviación estándar y la sensibilidad, para analizar la capacidad del refinamiento del modelo para la detección de vertebras. En la segunda parte, se lleva a cabo una comparativa de los resultados de mediciones brindados por el modelo con mediciones manuales realizadas por especialistas en el área mediante una matriz de confusión.

El diseño de esta investigación es de tipo experimental, debido al refinamiento realizado al modelo existente y al desarrollo del algoritmo para obtener los resultados esperados en forma de porcentaje de traslación de vértebras.

4.2 VARIABLES DE INVESTIGACIÓN

Se consideraron variables de investigación solamente en la parte experimental del estudio (primera fase), referente al refinamiento del modelo de aprendizaje profundo.

4.2.1. VARIABLES DEPENDIENTES

En esta investigación se toma como variable dependientes las métricas de rendimiento del modelo que varían dependiendo de los parámetros para la extracción de puntos de interés (vertebras). Las métricas utilizadas son las siguientes:

- Precisión: Se define como la proporción de clasificaciones correctas realizadas en relación con el total de las instancias.
- Puntaje F1 es la media armónica de la sensibilidad y la precisión, y, expresa la exactitud de la clasificación en conjuntos de datos desequilibrados.
- Sensibilidad: La sensibilidad es la medición de todos los casos positivos clasificados adecuadamente, donde idealmente debe de acercarse a la unidad (Fraiwan et al., 2022).

- Exactitud: Hace referencia al porcentaje de todos los resultados correctos obtenidos de las muestras utilizadas en pruebas de validación (Abdullah et al., 2018).

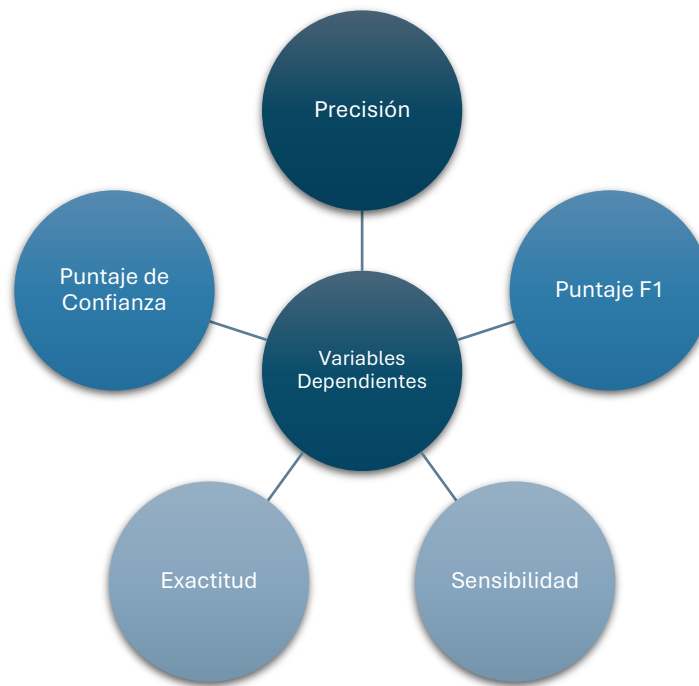


Ilustración 2 Variables Dependientes

Fuente: Elaboración Propia

4.2.2. VARIABLES INDEPENDIENTES

Las variables independientes no se ven afectadas por cambios realizados en el algoritmo, y poseen factores que si afecta las variables dependientes. La variable independiente de la investigación es la siguiente:

- Épocas: Se define como época cada iteración completa del conjunto de datos del algoritmo durante la etapa de entrenamiento. Aumentar la cantidad de épocas permite al modelo obtener mejor rendimiento, a cambio de mayor consumo de recursos se consideraron 10, 50, 100 y 150 épocas.

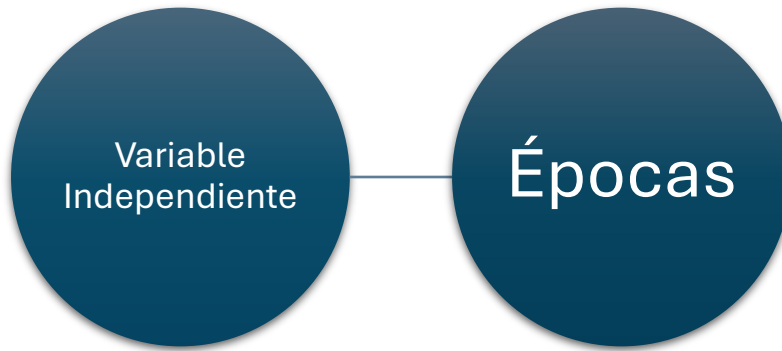


Ilustración 3 Variables Independientes

Fuente: Elaboración Propia

4.3 TÉCNICAS E INSTRUMENTOS APLICADOS

Para llevar a cabo el desarrollo del algoritmo para la detección automática de espondilolistesis se utilizaron los siguientes instrumentos en las etapas de refinamiento y entrenamiento del modelo.

4.3.1. ROBOFLOW

Roboflow es una herramienta multiusos para el desarrollo de modelos de visión de computadoras. Siendo eficaz para la anotación manual de puntos clave y cajas delimitadoras, haciéndolo útil para la creación de modelos de detección de vértebras mediante aprendizaje profundo. Se ha utilizado para la generación del dataset adicionado a la base original, con las imágenes etiquetadas de manera manual utilizando las herramientas de anotaciones.

4.3.2. GOOGLE COLABORATORY

También conocido como "Colab", es un entorno colaborativo multiusuario que permite a cualquier persona con acceso a Google escribir y ejecutar código Python de forma arbitraria a través de su navegador, lo que resulta especialmente útil para proyectos colaborativos de aprendizaje automático y análisis de imágenes médicas (Werth et al., 2022). Esta plataforma en la nube basada en cuadernos Jupyter se ha utilizado ampliamente en la enseñanza de aprendizaje automático al permitir la escritura de explicaciones de texto y códigos Python de manera conjunta Canesche et al., (2021). Además, Google Colab ha demostrado ser una herramienta efectiva para el aprendizaje profundo al ofrecer recursos computacionales basados en GPU de forma gratuita

(Balaniuk et al., 2020). Colab permitió tener un acceso directo al código de manera simultánea y facilitó el preprocesamiento y las modificaciones realizadas para el entrenamiento y ejecución del modelo de manera rápida y eficiente.

4.3.3. PYTHON

Python se presenta como un lenguaje de programación semi-interpretado, el cual procede a compilar el código con el fin de generar un pseudo código máquina, el cual es ejecutado por el intérprete respectivo. Este lenguaje de nivel elevado destaca por su accesibilidad gratuita, su sintaxis transparente y adaptable, y por su capacidad para soportar el paradigma funcional. Python es un lenguaje de programación ampliamente utilizado en el análisis de imágenes médicas y en el aprendizaje profundo. Se ha implementado en entornos como Google Colab para acelerar la ejecución de aplicaciones de aprendizaje profundo Dickson-Karn & Orosz, (2021).

4.3.4. IMAGEJ

ImageJ es una plataforma de dominio público y accesible para todo el mundo. Su abanico de opciones de uso es amplio porque los usuarios han creado complementos para el procesamiento, medición, análisis y escaneo de imágenes. La aplicación es sumamente útil para el preprocesamiento de las imágenes, eliminar ruidos y cambiar el contraste una imagen. Se ha implementado de igual forma para la segmentación automática o manual de algún objeto de interés en la imagen y su posterior seguimiento en un dataset de imágenes para su análisis volumétrico y comportamiento Schindelin et al., (2015).

4.4 METODOLOGÍA DE ESTUDIO

4.4.1. PRE-PROCESAMIENTO DE IMÁGENES

El preprocesamiento de las imágenes de las bases de datos se orienta a adecuar las imágenes para cumplir con los requisitos mínimos del algoritmo propuesto, centrado en el análisis de imágenes de rayos X. Primero, se corregirá la orientación de las imágenes para garantizar coherencia en su presentación. Se someten a una operación de inversión a las imágenes orientadas a la derecha, para que su orientación final sea hacia el lado izquierdo, asegurando coherencia al momento de la visualización de las imágenes. Posteriormente, se procede con la extracción y agrupación de puntos clave, se utiliza Roboflow para realizar esta dicha tarea. Estos puntos clave son organizados en conjuntos coherentes para facilitar su posterior procesamiento.

Posteriormente, se genera un cuadro delimitador alrededor de cada conjunto de puntos clave agrupados, definiendo así regiones de interés en las imágenes. Estos cuadros delimitadores se calculan con base en los límites mínimos y máximos de las coordenadas de los puntos clave, asegurando una captura precisa del área de interés. Finalmente, tanto las imágenes como las anotaciones preprocesadas son redimensionadas a un tamaño específico de 640x640 píxeles, compatible con el modelo propuesto. Este dimensionamiento está orientado a garantizar consistencia y eficiencia en el procesamiento de las imágenes durante la fase de entrenamiento y evaluación del modelo.

4.4.2. MODELO DE DETECCIÓN DE VÉRTEBRAS

Como siguiente paso es necesario contar con un modelo fuera capaz de identificar los bordes de las vértebras para así realizar las mediciones de distancias y llevar a cabo los cálculos necesarios para ponderar la traslación de las vértebras. Por lo cual se tomó como punto de partida un modelo de detección de vértebras previamente trabajado en esta misma línea de investigación por colegas de la Universidad Tecnológica Centroamericana, quienes desarrollaron un modelo basado en CNN capaz de transformar puntos medios de cajas delimitadoras (de vértebras) en puntos clave, para luego extraer los datos en formato Keypoint RCNN. El modelo de detección de vértebras permitió identificar los bordes de las vértebras en las imágenes de rayos X de pacientes con espondilolistesis. Este modelo se utilizó como base, sin embargo se llevó a cabo ciertas modificaciones en el refinamiento de este, como ser como el aumento de las épocas, el tamaño del tamaño de lote agregando más imágenes a la base de datos para refinar el modelo y obtener mejores resultados en la detección de vértebras y así facilitar la siguiente etapa de la medición de traslación de las vértebras. Es por eso, que este modelo se volvió a entrenar para obtener con mayor exactitud la detección de puntos de interés.

4.4.3. ALGORITMO DE MEDICIÓN PARA LA CLASIFICACIÓN DE MEYERDING PROPUESTO

Tomando como valores de entrada los resultados obtenidos por modelo de detección de vértebras en formato JSON, se lleva a cabo el procedimiento propuesto para calcular de manera automática el porcentaje de traslación de una vértebra sobre la otra, mostrado en la Ilustración 4. Como el modelo de detección de vértebras crea etiquetas para las esquinas de todas las vértebras

presentes en las imágenes resulta conveniente enfocarse únicamente en las vértebras L3-S1, pues, son los cuerpos vertebrales más propensos al deslizamiento (Koslosky & Gendelberg, 2020).

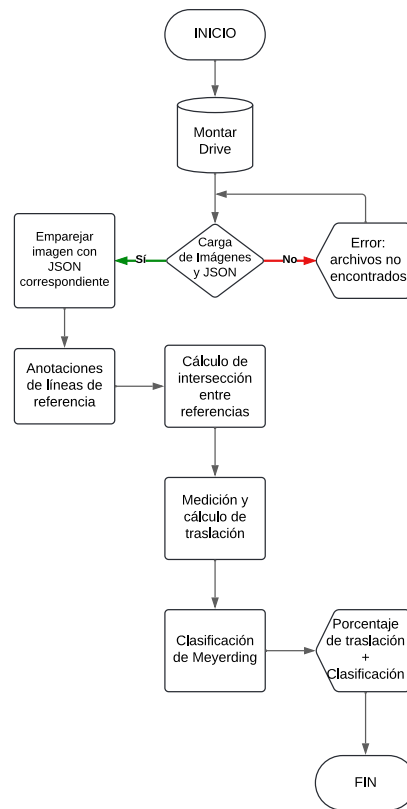


Ilustración 4 Diagrama de Flujo de Algoritmo de Medición

Fuente: Elaboración Propia

El proceso comienza con la importación y montaje de Google Drive, lo que permite acceder a los archivos almacenados. Se montan las unidades de Google Drive para acceder a las imágenes de rayos X y sus correspondientes archivos JSON que contienen las coordenadas de los puntos clave de las vértebras. Luego, se verifica si las imágenes y sus archivos JSON correspondientes están disponibles; si ambos archivos están presentes, se procede a emparejar la imagen con su archivo JSON correspondiente, cargando los puntos clave desde el archivo JSON; si falta alguno de los archivos, se genera un error indicando que los archivos no se encontraron, y el proceso finaliza.

Con los puntos clave cargados, se generan y anotan líneas de referencia en la imagen extrayendo las coordenadas x e y de sus extremos superior e inferior derecho, de la mano de la

función plt.plot. Similarmente se dibuja la segunda línea de medición entre los extremos superior izquierdo y derecho de la vértebra inferior, para luego realizar la división de la línea en cuatro cuadrantes, anotando los porcentajes de 25% a 100% para facilitar la visualización. Las líneas anotadas son necesarias para calcular las intersecciones y medir los desplazamientos vertebrales.

Utilizando las coordenadas de los puntos clave, se calculan las intersecciones entre las líneas de referencia y las líneas extendidas a partir de los bordes de las vértebras. Cada vértebra está definida por cuatro puntos en las coordenadas (x, y). Los puntos se almacenan en una lista "keypoints_list", y cada vértebra tiene un conjunto de cuatro coordenadas:

- (x1, y1) - Esquina superior izquierda
- (x2, y2) - Esquina superior derecha
- (x3, y3) - Esquina inferior derecha
- (x4, y4) - Esquina inferior izquierda

Las líneas se definen por dos puntos cada una:

- L1: (x1, y1) a (x2, y2) (de la vértebra inferior)
- L2: (x2, y2) a (x3, y3) (de la vértebra superior)

La intersección de dos líneas se calcula usando las coordenadas de los puntos finales de cada línea. La fórmula para encontrar la intersección en (x, y) de dos líneas se define en las ecuaciones Ecuación 1 y Ecuación 2. Si el denominador es cero, significa que las líneas son paralelas y no se interceptan, si el denominador no es 0, las coordenadas de la intersección se calculan como:

$$intx = \frac{(x1 \cdot y2 - y1 \cdot x2) \cdot (x3 - x4) - (x1 - x2) \cdot (x3 \cdot y4 - y3 \cdot x4)}{(x1 - x2) \cdot (y3 - y4) - (y1 - y2) \cdot (x3 - x4)}$$

Ecuación 1 Intersección en X entre líneas de referencia

Fuente: Elaboración Propia

$$inty = \frac{(x1 \cdot y2 - y1 \cdot x2) \cdot (y3 - y4) - (y1 - y2) \cdot (x3 \cdot y4 - y3 \cdot x4)}{(x1 - x2) \cdot (y3 - y4) - (y1 - y2) \cdot (x3 - x4)}$$

Ecuación 2 Intersección en Y entre líneas de referencia

Fuente: Elaboración Propia

Utilizando la fórmula de la distancia euclidiana para calcular el desplazamiento entre las vértebras se mide desde la intersección calculada hasta la esquina superior derecha de la vértebra inferior (x_2, y_2) se obtiene la Ecuación 3

$$d = \sqrt{(intx - x_2)^2 + (inty - y_2)^2}$$

Ecuación 3 Desplazamiento vertebral

Fuente: Elaboración Propia

De la misma forma, el ancho de la vértebra se calcula como la distancia entre la esquina superior izquierda (x_1, y_1) y la esquina superior derecha (x_2, y_2), mostrado en la Ecuación 4:

$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Ecuación 4 Ancho de vértebra

Fuente: Elaboración Propia

La Ecuación 5 establece la fórmula para calcular el porcentaje de traslación de una vértebra, tomando como referencia la longitud total de la vértebra (d) y la distancia del deslizamiento (D) sobre la vértebra adyacente. Con el porcentaje de traslación obtenido, se clasifica el grado de espondilolistesis según la clasificación de Meyerding, que varía de normal a grados I a V dependiendo de la severidad del desplazamiento. Siguiendo a exactitud la clasificación de Meyerding descrita por Koslosky & Gendelberg, 2020), se establecen los criterios para que el modelo establezca como datos de salida tanto el porcentaje de traslación como el Grado en el que corresponde (Grados I, II, III, IV).

$$\%Traslación = \frac{d (mm)}{D (mm)} \times 100$$

Ecuación 5 Porcentaje de Traslación Vertebral

Fuente: (Koller et al., 2019)

Finalmente, se presenta el porcentaje de traslación y la clasificación obtenida, proporcionando un diagnóstico visual y cuantitativo, concluyendo el proceso y mostrando los

resultados en la imagen para una interpretación rápida y precisa del diagnóstico. Este flujo de trabajo automatiza y optimiza el diagnóstico de la espondilolistesis, facilitando la visualización y clasificación de manera eficiente y precisa, reduciendo la subjetividad y el tiempo requerido en métodos convencionales.

4.5 METODOLOGÍA DE VALIDACIÓN

La metodología de validación representa un paso fundamental para esta investigación, ya que permite verificar la aplicabilidad del algoritmo propuesto en el diagnóstico de la espondilolistesis. Se emplea la validación inicialmente evaluando el desempeño del modelo dividiendo el conjunto de datos en un conjunto de entrenamiento y un conjunto de validación. El conjunto de entrenamiento, como su nombre lo indica, sirve para entrenar el modelo, mientras que el conjunto de validación es para evaluar el rendimiento utilizando las métricas establecidas como las variables dependientes de la investigación. Posteriormente se plantea una comparativa de los cálculos de traslación de las vértebras obtenidos por el algoritmo con mediciones manuales realizadas por tres especialistas, donde uno de ellos es neurocirujano, la Dra. Carla Chong con 10 años de experiencia y dos radiólogos el primero es el Dr. Marcelino López con 8 años de experiencia y el segundo es el Dr. Juan Paz, permitiendo contrastar los resultados del modelo con un método ya empleado de manera manual.

4.5.1. MÉTRICAS DE EVALUACIÓN DE DESEMPEÑO

Durante la recopilación de datos obtenidos de las distintas fases del algoritmo, se llevaron a cabo los cálculos necesarios de métricas específicas de desempeño, detalladas en la Tabla 10. Este proceso garantizó la obtención de resultados confiables para la investigación. Después, se analizó cuantitativamente los datos obtenidos, para discutir los hallazgos y llegar a conclusiones relevantes. Estos cálculos métricos fueron esenciales para evaluar el desempeño del algoritmo en diferentes escenarios y conjuntos de datos, contrastándolo con diagnósticos profesionales reales.

Tabla 10 Métricas de evaluación de rendimiento

Métrica	Fórmula	Descripción
Precisión	$\% \text{ Precisión} = \frac{TP}{TP + FP} \times 100$	La proporción de clasificaciones correctas realizadas en relación con el total de las instancias

Puntaje F1	$\% F1 = \frac{\text{Sensibilidad} \times \text{Precisión}}{\text{Sensibilidad} + \text{Precisión}} \times 100$	La media armónica de la sensibilidad y la precisión, y, expresa la exactitud de la clasificación en conjuntos de datos desequilibrados
Sensibilidad	$\% \text{Sensibilidad} = \frac{TP}{TP + FN} \times 100$	La medición de todos los casos positivos clasificados adecuadamente.
Exactitud	$\% \text{Exactitud} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \times 100$	El porcentaje de todos los resultados correctos obtenidos de las muestras utilizadas en pruebas de validación

Fuente: Elaboración Propia

4.6 CRONOGRAMA DE ACTIVIDADES

En esta sección se presenta el cronograma de actividades que determinará la ejecución del proyecto de investigación. Este cronograma se ha elaborado para planificar y organizar cada etapa del trabajo, asegurando que se alcancen los objetivos en el plazo establecido. Se tomaron en cuenta todas las fases del proyecto, desde la investigación inicial y la recopilación de datos hasta el análisis y la redacción del informe final.

	Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Semana 6	Semana 7	Semana 8	Semana 9	Semana 10
<i>Revisión del estado del arte</i>										
<i>Ajustes a base de datos</i>										
<i>Refinamiento de Modelo de Detección de Vértebras</i>										
<i>Evaluación de métricas del modelo</i>										
<i>Desarrollo de algoritmo de medición y cálculo de traslación vertebral</i>										
<i>Elaboración de mediciones manuales para validación</i>										
<i>Validación manual por parte de especialistas</i>										
<i>Comparativa de resultados de validación</i>										
<i>Análisis y discusión de resultados</i>										

Ilustración 5 Cronograma de Actividades

Fuente: Elaboración Propia

V. RESULTADOS Y ANÁLISIS

En el capítulo de resultados y análisis se muestran los datos obtenidos a través de las diferentes etapas de esta investigación siguiendo la estructura metodológica propuesta. Partiendo desde el entrenamiento del modelo de detección de vértebras, se presentan las métricas evaluadas para el modelo, luego los resultados obtenidos del código, y por último la comparativa con las mediciones manuales realizadas por los especialistas.

5.1 BASE DE DATOS

Los principales materiales utilizados en esta investigación son las bases de datos seleccionadas para el entrenamiento y validación del algoritmo. Estas bases de datos se han compilado cuidadosamente para poder conseguir imágenes de rayos X de la columna vertebral en la zona lumbar y en el plano sagital (lateral). La Tabla 11 muestra las características principales de las bases de datos utilizadas.

Tabla 11 Bases de Datos a utilizar

Base de datos	Año	# de Imágenes	Dimensiones
Base de Datos Hondureña	Diferentes años	197	Altura 224 pixeles Anchura 224 pixeles
BUU LSpine	2019-2023	517	Altura 1434-3408 pixeles. Anchura 860-3040 pixeles
Combinación	>2019	714	Diferentes dimensiones

Fuente: Elaboración Propia

Se seleccionaron un total de 714 imágenes, donde 197 provienen de la base de datos de pacientes hondureños que sufren de espondilolistesis tanto de mujeres como de hombres, se descartó cualquier imagen que no sea en el plano sagital. Se utilizaron de igual forma 517 imágenes obtenidas de la base de datos de BUU-LSPINE que cuenta con un total de 2700 imágenes de afecciones en la columna, donde igual existían pacientes masculinos y femeninos, esta base de datos fue recolectada entre 2019 y 2023. Dichas imágenes se utilizaron para desarrollar el método de diagnóstico de la espondilolistesis de manera automática, esto implica reentrenamiento de un modelo de detección de vertebras y puntos de interés en las vértebras,

así como el cálculo del grado de afección, mediante el algoritmo desarrollado específicamente para este propósito, como de manera manual. A demás se utilizaron para la evaluación de especialistas médicos. Esta comparación entre el diagnóstico automatizado y la evaluación manual por parte de profesionales de la salud es crucial para validar las métricas como la efectividad y precisión del algoritmo en la detección de esta patología.

5.2 REENTRENAMIENTO DEL MODELO DE DETECCIÓN DE VÉRTEBRAS

El etiquetado se llevó a cabo en Roboflow, donde una parte del etiquetado utiliza el punto medio de las bounding boxes para ser tomadas como keypoints en cada esquina de cada vertebra este etiquetado se llevó a cabo en una investigación previa de detección de vertebras realizada por compañeros de la Universidad Tecnológica Centroamericana, mostrado en el Anexo 2. Para esta investigación se realizó una ampliación de la base de datos previamente existente con el fin de refinar el modelo y mejorar todavía más los resultados obtenidos en los entrenamientos previos que ya de por si eran satisfactorios, se utilizaron bounding boxes para cada vertebra y keypoints para cada esquina de cada vertebra Anexo 1. La unificación de la base de datos a pesar de haber utilizado diferentes métodos de etiquetado se basa en el formato Keypoints RCNN donde las anotaciones se guardan como coordenadas en el plano X-Y para que sean entendibles por el cuaderno de entrenamiento.

En la fase de entrenamiento, el modelo fue entrenado por primera vez utilizando solamente 10 épocas y utilizando la totalidad de la base de datos de las imágenes previamente etiquetadas para generar la primera versión, este actúa como base para el modelo definitivo. Después de generar el modelo base, este se vuelve a llamar para realizar el reentrenamiento con 50, 100 y 150 épocas siendo este el modelo final para la detección de vertebras.

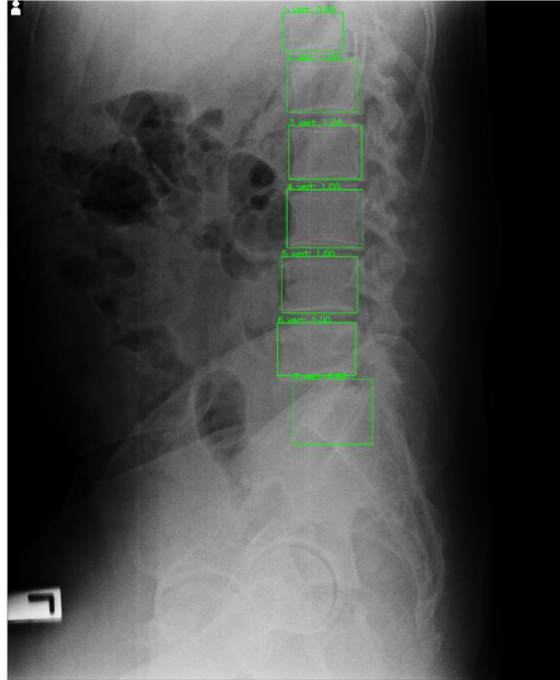


Ilustración 6 Ejemplo de detección de vertebras en entrenamiento

Fuente: Elaboración propia

Como se puede observar en la Ilustración 6, el modelo previamente entrenado mostró capacidad plena para detectar vertebras superando significativamente las expectativas en términos de precisión y confiabilidad. Los resultados con base en base a las métricas de evaluación mostraron ser más que excepcionales, el modelo obtuvo una precisión promedio de 98% para la predicción los bounding boxes, mostrado en el Anexo 3. Esto indica que casi todas las vértebras detectadas corresponden a la vertebras reales en las imágenes. Alcanzó un 95% para los keypoints, como se puede observar en el Anexo 5 detectando las esquinas de cada vertebra, el modelo obtuvo un 92% para puntaje F1 reflejando un balance sobresaliente entre precisión y sensibilidad. Finalmente, una exactitud del 91% que asegura que la gran mayoría de las predicciones son acertadas. Estos resultados dan la certeza y la confianza de que el modelo podrá detectar la posición de cada vértebra en la imagen de rayos x proporcionada.

Tabla 12 Comparativa entre resultados de entrenamiento de modelo de detección de vertebras con estudio previo

Investigación	Épocas	Precisión Promedio Bounding Boxes	Sensibilidad Bounding Boxes	Precisión Promedio Keypoints	Sensibilidad Keypoints	Exactitud	Puntaje F1
Investigación Previa	100	97%	77%	93%	73%	--	84%
Investigación actual	10	98%	76%	94%	96%	89%	90%
Investigación actual	50	98%	80%	94%	96%	90%	92%
Investigación actual	100	98%	80%	95%	96%	91%	93%
Investigación actual	150	98%	80%	94%	96%	91%	93%

Fuente: Elaboración Propia

En la Tabla 12 Comparativa entre resultados de entrenamiento de modelo de detección de vertebras con estudio previo se compararon los resultados del entrenamiento con el estudio anterior de nuestros compañeros de la Universidad Tecnológica Centroamericana, detallando métricas clave como la precisión promedio de las puntuaciones de los puntos clave y los cuadros delimitadores, así como la sensibilidad de los puntos clave y los cuadros delimitadores. Al aumentar la cantidad de imágenes anotadas manualmente y utilizar 100 épocas durante el entrenamiento, se observaron mejoras significativas en todas las métricas evaluadas. La precisión promedio de los puntos clave aumentó significativamente, lo que indica una mayor precisión en la identificación de los keypoints. De manera similar, mejoró la precisión de los bounding boxes, lo que refleja una representación más precisa de las estructuras vertebrales.

5.3 INFERENCIA

Después del entrenamiento del modelo de detección de vértebras, se genera un archivo que contiene los pesos del modelo, lo que permite que este pueda realizar predicciones. El enfoque del entrenamiento se basa en dos tipos de etiquetado, el primero fue el etiquetado utilizando el punto medio de bounding boxes que se tomarían como keypoints en cada esquina

de la vertebra, el otro método de etiquetado es el uso de bounding boxes para delimitar las vértebras y el uso de keypoints colocados en las esquinas de cada vertebra. El modelo se evalúa a través de las métricas proporcionadas por el entrenamiento y por su desempeño para detectar las vértebras en imágenes proporcionadas. Utilizando el cuaderno de inferencia desarrollado en Google Colaboratory, el archivo que contiene el modelo puede ser cargado y el modelo se utiliza para su posterior evaluación con imágenes que no han sido previamente etiquetadas. Se seleccionaron 20 imágenes al azar de la base de datos de BUU LSPINE para realizar la inferencia del modelo y realizar las pruebas pertinentes de detección de vertebras.

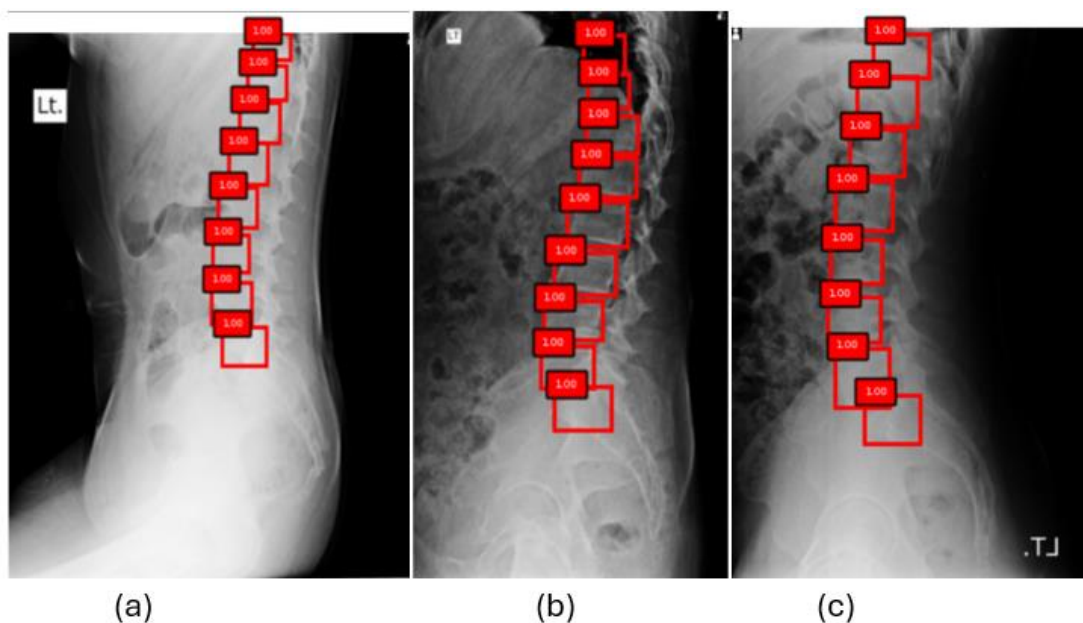


Ilustración 7 Ejemplo de puntaje de confianza

Fuente: Elaboración Propia

Como se puede observar en la Ilustración 7, el modelo tiene una alta puntuación de confianza a la hora de realizar la detección de vertebras. Esta puntuación define cuan seguro está el modelo de que la predicción que realizó es la correcta. Entre mayor sea el puntaje, como es el caso de la detección de las vértebras, indica que existe una mayor confianza para la detección de objetos.

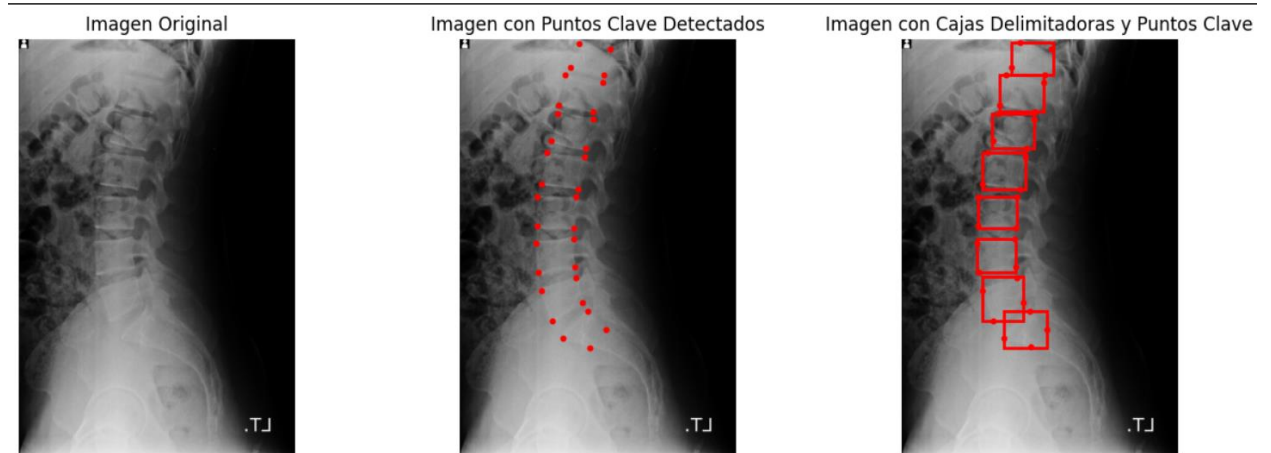


Ilustración 8 Ejemplo de inferencia

Fuente: Elaboración Propia

En la Ilustración 8 Ejemplo de inferencia se puede observar un ejemplo de cómo el cuaderno de inferencia pone a prueba el modelo de detección de vértebras, primero se carga una imagen sin anotaciones previas y de vista lateral orientada hacia la izquierda, después el propio modelo realiza la predicción de los puntos clave en las esquinas de cada vértebra y de las cajas delimitadoras por cada vértebra.

5.4 MÓDULO DE MEDICIÓN Y DETECCIÓN DE ESPONDILOLISTESIS

Habiendo refinado previamente el modelo de detección de vértebras, obteniendo resultados elevados de precisión en las predicciones, se implementa el módulo donde se realiza la medición del porcentaje de traslación vertebral para la detección de la espondilolistesis. El módulo consta del código que, tomando como variable de entrada los archivos JSON generados por el modelo de detección de vértebras y la imagen correspondiente, realiza las anotaciones necesarias para tomar las mediciones, y hace el cálculo del porcentaje de traslación para así definir el grado de espondilolistesis presente en dichas vértebras.

El código elaborado se divide en tres etapas, la primera es para la carga de los archivos de entrada, en la cual, el usuario selecciona la imagen a analizar y el programa la relaciona con el archivo JSON generado. Luego la etapa de visualización de imagen y anotaciones despliega la

imagen seleccionada con los puntos clave identificados por el modelo y las anotaciones realizadas para la medición de la traslación vertebral, como es mostrado en la Ilustración 9 Medición de Traslación

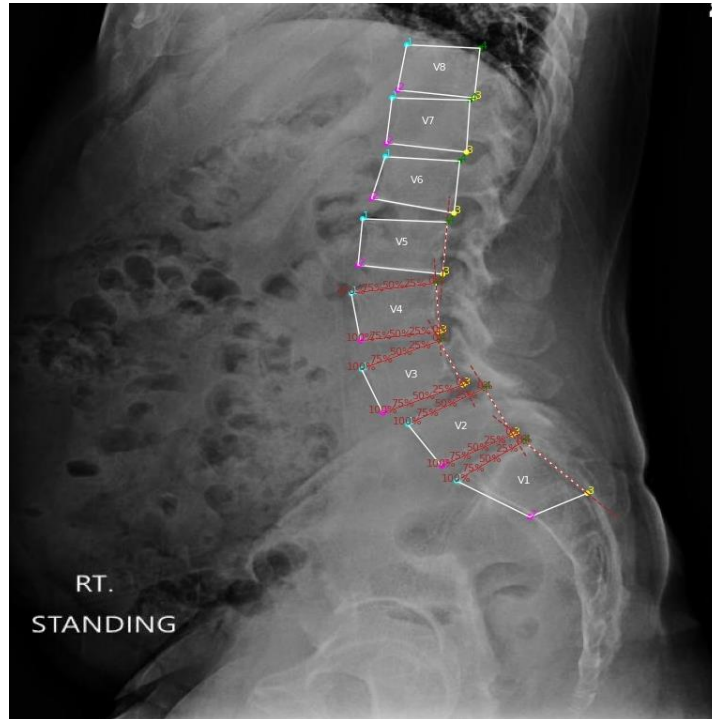


Ilustración 9 Medición de Traslación

Fuente: Elaboración Propia

La tercera etapa realiza el cálculo de la traslación vertebral, tomando las coordenadas de las vértebras identificadas por el modelo de detección de vértebras, anotando líneas punteadas que se extienden a lo largo de la vértebra, coincidentes con las dos esquinas donde se encuentran los puntos clave. Se lleva a cabo el cálculo utilizando la Ecuación 5, donde se mide el desplazamiento horizontal de una vértebra con respecto a la adyacente por debajo de esta, y se divide entre la longitud horizontal del platillo terminal de la vértebra, demostrado en la Ilustración 10. Dentro de la tercera etapa se establecen las condiciones para determinar el grado de espondilolistesis, siguiendo a exactitud la lógica propuesta en la clasificación de Meyerding.

```
Intersección para V1: (454.32, 391.73)
Entre V1 y V2: Desplazamiento = 5.42 píxeles, Porcentaje de deslizamiento = 7.80%, Clasificación = Grado I
Intersección para V2: (410.13, 348.53)
Entre V2 y V3: Desplazamiento = 15.33 píxeles, Porcentaje de deslizamiento = 20.40%, Clasificación = Grado I
Intersección para V3: (384.54, 299.8)
Entre V3 y V4: Desplazamiento = 0.58 píxeles, Porcentaje de deslizamiento = 0.78%, Clasificación = Grado I
Intersección para V4: (386.36, 248.36)
Entre V4 y V5: Desplazamiento = 5.40 píxeles, Porcentaje de deslizamiento = 7.15%, Clasificación = Grado I
```

Ilustración 10 Cálculo de Porcentaje de Traslación

Fuente: Elaboración Propia

5.5 VALIDACIÓN

Se realizaron distintos métodos para verificar los resultados obtenidos del modelo. Inicialmente, se observaron las métricas de evaluación en la fase de entrenamiento para asegurar que el modelo funcionara de acuerdo con los parámetros y criterios requeridos. Luego se realizó una validación cruzada utilizando conjuntos de datos independientes para evaluar la precisión y generalización del modelo. Además, se hacen comparaciones con investigaciones anteriores, donde se requería aumentar la precisión y la exhaustividad del modelo. También hubo diagnósticos y consultas con expertos en el campo que brindaron perspectivas y métodos de detección de la espondilolistesis de forma tradicional. Estas actividades de verificación son necesarias para garantizar la solidez y confiabilidad de los resultados.

5.5.1. VERIFICACIÓN MANUAL CON IMAGEJ

En el presente estudio de medición y clasificación automatizadas de espondilolistesis utilizando el sistema Meyerding, se utilizó ImageJ para realizar mediciones manuales precisas entre puntos de interés en imágenes de rayos X. Estas mediciones manuales se utilizan para verificar los resultados obtenidos por el algoritmo automático propuesto. Utilizando ImageJ, se identificó y midió manualmente distancias entre vértebras y puntos de referencia para determinar el grado de espondilolistesis.



Ilustración 11 Validación Manual con ImageJ

Fuente: Elaboración Propia

Los datos obtenidos fueron analizados utilizando gráficos de dispersión y un gráfico de Bland-Altman para facilitar la interpretación de los resultados. Las mediciones de porcentaje de traslación vertebral se mostraron altamente consistentes con las mediciones manuales, como se observa en el gráfico de dispersión (Ilustración 12) Este gráfico demuestra que la mayoría de los puntos están cercanos a la línea de referencia $y = x$, lo que indica una alta precisión y confiabilidad del algoritmo desarrollado.

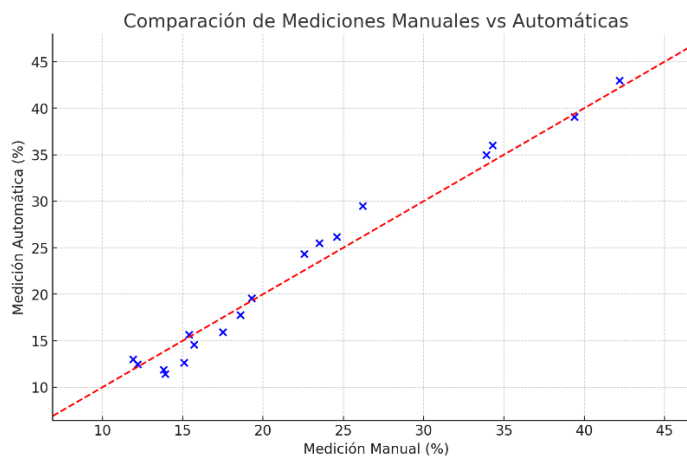


Ilustración 12 Gráfico de Dispersión de Mediciones Manuales vs Automáticas

Fuente: Elaboración Propia

El gráfico de Bland-Altman presentado en la Ilustración 13, mostró una variabilidad aceptable sin sesgo sistemático significativo. La media de las diferencias entre las mediciones automáticas y manuales fue de 0.20%, y los límites de acuerdo (± 1.96 SD) estuvieron entre -2.94% y 3.34%. Esto sugiere que las diferencias entre las mediciones están dentro de un rango aceptable y que el algoritmo automático es consistente en comparación con las mediciones manuales.

Al comparar las mediciones manuales y automáticas, se encontró que la variación entre ellas es mínima, como lo evidencia la proximidad de los puntos en el gráfico de dispersión a la línea de referencia y = x. Esto confirma la precisión y confiabilidad del algoritmo desarrollado. Este proceso de validación es crucial para garantizar el uso efectivo del sistema automatizado en la práctica clínica, asegurando resultados consistentes y precisos al evaluar la traslación vertebral.

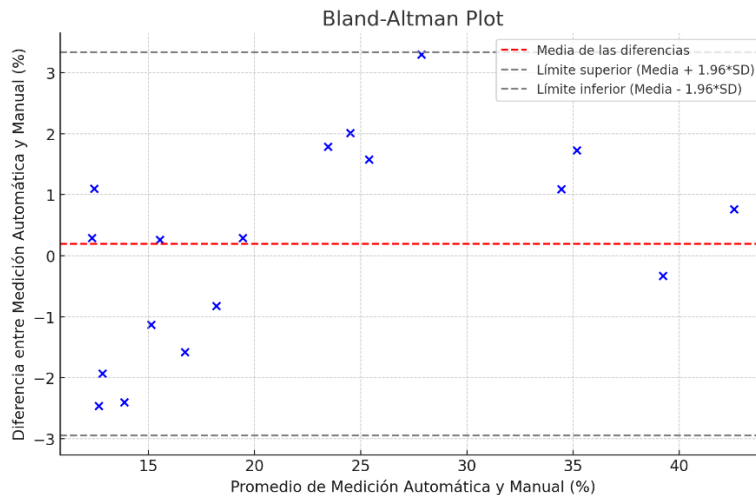


Ilustración 13 Gráfico de Bland-Altman de Mediciones Automática y Manual

Fuente: Elaboración Propia

5.5.2. VERIFICACIÓN MANUAL POR PARTE DE EXPERTOS

El segundo método de validación se basa en la medición del grado de traslación de una vértebra respecto a la siguiente, realizada por expertos en el campo de la detección de espondilolistesis. Dos médicos radiólogos y una neurocirujana formaron este grupo de especialistas, que aportaron su conocimiento y años de experiencia para medir y clasificar esta afección en la columna vertebral. Para asegurar la fiabilidad de los resultados, se realizaron reuniones individuales con cada especialista. Durante estas sesiones, cada experto evaluó las

imágenes de manera independiente y proporcionó su diagnóstico basado en su criterio clínico. Los tres especialistas concordaron en que las mediciones se hacen mediante estimaciones y visualmente sin usar objetos de medición. La neurocirujana, por su parte, aportó una perspectiva quirúrgica, enfatizando las implicaciones clínicas de los grados de traslación observados y su relevancia para el tratamiento de los pacientes. Cada diagnóstico fue registrado y comparado para identificar concordancias y discrepancias en las mediciones realizadas. Este enfoque permitió obtener una validación del algoritmo desarrollado, dado que los expertos utilizaron sus métodos tradicionales y experiencia para confirmar o refutar los resultados obtenidos automáticamente.

Para evaluar la consistencia y validez de las clasificaciones de espondilolistesis de Meyerding realizadas por los especialistas y compararlas con los resultados obtenidos por nuestro algoritmo, se utilizaron los métodos estadísticos del coeficiente Kappa de Cohen para la evaluación de cada par de evaluadores combinando las matrices de confusión generada entre las mediciones realizadas y el coeficiente Kappa de Fleiss, que relaciona para más de dos evaluadores el grado de concordancia obtenido, se utilizó este último para relacionar las mediciones realizadas por los evaluadores con las mediciones realizadas por el algoritmo. (Albakkosh, 2024).

Tabla 13 Resultados de Concordancia en Clasificación

Nivel Vertebral	Especialista 1 vs Especialista 2	Especialista 1 vs Especialista 3	Especialista 2 vs Especialista 3	Fleiss Kappa entre Especialistas	Fleiss Kappa Algoritmo vs Especialistas
L5-S1	0.603	0.856	0.434	0.628	0.357
L4-L5	0.642	0.775	0.642	0.686	0.636

Fuente: Elaboración Propia

Los resultados de la concordancia mostrados en la Tabla 13, relacionan las mediciones realizadas por cada uno de los especialistas, haciendo notar la variación que existe entre diagnósticos debido a la subjetividad del método empleado. Un coeficiente Kappa de Fleiss de 0.628 y 0.686 en los pares L5-S1 y L4-L5, respectivamente, demuestran concordancia substancial según la categorización de Fleiss' Kappa (Tabla 14). Sin embargo, una vez se comparan los resultados de los 3 especialistas adicionando el algoritmo propuesto, se reduce notablemente el coeficiente de concordancia en L5-S1, llegando a 0.357 que entra dentro de la interpretación de

concordancia débil, mientras que en L4-L5 se mantiene dentro de la concordancia sustancial. Esto es causado por la variación en mediciones subjetivas versus las mediciones objetivas realizadas por el algoritmo, analizando individualmente cada caso se observó que los especialistas identificaron como Grado 1 a imágenes que eran Grado 2 pero estaban muy cercanos al límite de 25%, causando la imprecisión en los resultados.

Tabla 14 Grados de Kappa de Fleiss

Kappa	Nivel de Concordancia
<0.00	Sin concordancia
0.00 – 0.20	Concordancia pobre
0.21 – 0.40	Concordancia débil
0.41 – 0.60	Concordancia moderada
0.61 – 0.80	Concordancia sustancial
0.81 – 1.00	Concordancia casi perfecta

Fuente: (Gordillo & Rodríguez, 2009)

VI. DISCUSIÓN DE RESULTADOS

La extracción manual de mediciones cuantitativas del porcentaje de traslación vertebral representa una inversión de tiempo significativa, por lo cual los especialistas tienden a optar por diagnósticos basados en su interpretación subjetiva de la imagen, motivada por la experiencia acumulada a través de los años. Por lo tanto, el uso de métodos para obtener mediciones objetivas es de gran interés ya que automáticamente presenta los resultados concretos para reducir así el margen de error ocasionado por la subjetividad y por consiguiente apoyar a la toma de decisiones. Este trabajo presenta un método para la detección de espondilolistesis combinando un modelo de aprendizaje profundo para detección de vértebras y extracción de puntos de interés, con un algoritmo de procesamiento, medición de traslación vertebral y clasificación de Meyerding.

El reentrenamiento del modelo de detección de vertebras se realizó con el fin de mejorar las salidas del modelo en cuanto a las métricas como la precisión y sensibilidad (por consiguiente, F1), para asegurar resultados más exactos en la medición de traslación vertebral, pues, las operaciones matemáticas empleadas en el algoritmo dependen directamente de la precisión de los puntos clave que predice el modelo de detección de vértebras. Se realizó el reentrenamiento con 10, 50, 100 y 150 épocas ampliando la base de datos a 714 imágenes, obteniendo las mejoras más significativas con la configuración de 50 épocas, en relación con el tiempo empleado.

Tabla 15 Comparativa de modelo con investigaciones previas

Autor	Modelo	Épocas	Cantidad de imágenes	Precisión promedio	Sensibilidad	Exactitud	Puntaje F1
Varcin et al 2021	Mobile-Net	25	2746	99%	98%	99%	98%
Varcin et al 2019	AlexNet	N/A	272	87.50%	95%	92%	91%
Varcin et al 2019	GoogleNet	N/A	272	91.70%	96%	94%	94%
Zhang et al 2023	Faster RCNN	N/A	1500	93.50%	93.50%	n/a	93.50%
Esta investigación	Faster RCNN	100	707	96%	88%	91%	93%

Fuente: Elaboración propia

En la Tabla 15 Comparativa de modelo con investigaciones previas, se puede observar los resultados y el desempeño de estudios anteriores comparados con el modelo propuesto de detección de vertebras. La tabla muestra que, en general, el modelo presenta una mejora significativa o está entre parámetros en métricas clave como ser la exactitud, la precisión y la puntuación F1 en comparación con los modelos anteriores. Sin embargo, vale la pena señalar que la única medida en la que el método no funciona mejor que los modelos anteriores es la sensibilidad. Esta diferencia muestra que, aunque el modelo es muy eficaz en aspectos de precisión y exactitud, todavía existe la necesidad de mejorar la capacidad del modelo para identificar correctamente todas las instancias positivas de la detección. Esto quiere decir, que el modelo presentando tiene una capacidad del 88% de detectar verdaderos positivos entre la suma de verdaderos positivos y falsos negativos y, por tanto, es indispensable minimizar los falsos negativos en el diagnóstico clínico.

Esto puede ser mejorado con un aumento de la base de datos, como lo muestra la tabla comparativa: hay una relación entre la cantidad de imágenes utilizadas y valor de la sensibilidad entre los autores del estado del arte. Zhang et al, por ejemplo, utilizan más del doble de imágenes (1500) en el entrenamiento y obtienen una sensibilidad del 93.50%. Por otro lado, los autores que utilizaron mayor cantidad de imágenes como Varcin y colaboradores en el año 2021 (2746) obtienen una sensibilidad del 98%. Esto indica que un punto de mejora inmediata para aumentar la sensibilidad del modelo, debe ser el aumento de la cantidad imágenes con casos en la base de datos para el reentrenamiento del modelo.

En el caso de Varcin y colaboradores del año 2019, podemos observar alta sensibilidad debido a que los autores se enfocan en la tarea de visión de computadoras que realiza una clasificación y no una detección específica de puntos de interés. Por tanto, se obtiene una sensibilidad alta para la detección de una columna normal o con espondilolistesis, pero no se calcula el grado de espondilolistesis, que conlleva mayor complejidad.

En el trabajo realizado por Liao y colaboradores en 2015, muestran resultados altamente precisos y eficientes en comparación a las mediciones manuales de expertos clínicos para la detección de vértebras y medición de espondilolistesis ya que utilizaron un método de etiquetado automático de la columna que combina la apariencia de la imagen y la información geométrica,

donde obtuvieron una exactitud total de 94.4% en detección de vértebras y medición de espondilolistesis, pero se basaron en imágenes de tomografía computarizada específicamente, aprovechando la ventaja que estas presentan en cuanto a la claridad para la detección de vértebras. Los métodos que se enfocan en imágenes de rayos X enfrentan mayores desafíos por la superposición de cuerpos vertebrales y mayor variabilidad en la calidad de las imágenes.

En la investigación realizada por Nguyen y colaboradores en el año 2021, se llevó a cabo una comparativa de los resultados de la clasificación de Meyerding realizada por un médico experto y por un algoritmo de clasificación. Para la detección de vértebras, se utilizaron redes neuronales convolucionales (CNN), específicamente dos modelos: uno para la detección de las esquinas de las vértebras y otro para la detección de las esquinas desde una perspectiva de izquierda a derecha, mejorando así la precisión en la ubicación de los puntos clave. Se compararon los resultados de la clasificación de Meyerding obtenidos por el algoritmo propuesto con los realizados por un médico experto, mostrando una variación mínima entre el diagnóstico del experto y el algoritmo, lo cual indica la alta precisión del método automático. Además, los autores llevaron a cabo la medición de los ángulos de segmentación vertebral utilizando el algoritmo. Para validar estos resultados, realizaron una validación con expertos, separando los resultados en función del ángulo de desviación con respecto a las mediciones manuales realizadas por los especialistas.

Partiendo del modelo de detección de vértebras, el algoritmo propuesto es capaz de realizar mediciones precisas que representan el porcentaje de traslación vertebral y así categorizar el grado de espondilolistesis según la clasificación de Meyerding. En cuanto a la medición del porcentaje de traslación vertebral, los resultados mostrados en la gráfica de dispersión en la Ilustración 12 y el gráfico de Bland-Altman (Ilustración 13) muestran que las mediciones automáticas son precisas y consistentes en comparación con las mediciones manuales. Un R^2 de 0.97 indica que el 97% de la variabilidad en las mediciones manuales puede ser explicada por las mediciones automáticas, lo cual es un indicador de una fuerte relación lineal. Se obtiene un Error Absoluto Medio de 1.38% y Error Cuadrático Medio de 2.61% estas métricas sugieren una alta concordancia entre las mediciones manuales y automáticas, con un R^2 cercano a 1 indicando que el modelo automático explica bien la variabilidad de las mediciones manuales.

Producto de la comparativa entre las mediciones realizadas por el algoritmo y los diagnósticos emitidos por los especialistas demostrada en la Tabla 13 en la cual se obtiene un coeficiente Kappa de Fleiss de 0.628 para L5-S1 y de 0.686 para L4-L5 que representa concordancia sustancial, validamos la necesidad de la objetividad en las mediciones, ya que los diagnósticos entre los distintos especialistas mostraron variaciones elevadas, especialmente en casos en los cuales el porcentaje de traslación se encuentra en los límites de las condiciones para la clasificación, es decir, cercano a las marcas de 25, 50, 75 y 100%.

En una de las imágenes de prueba, el algoritmo da como resultado una traslación de 26.18% entre las vértebras L5 y S1, sin embargo, 2 de los 3 especialistas aseguran que no se presenta espondilolistesis en dicha imagen y el otro menciona ser un caso de grado 1. El error en la medición por parte del algoritmo surge del etiquetado incorrecto de la vértebra S1, ya que en la imagen original se observa que las vértebras siguen la curvatura natural, pero la esquina superior derecha de la S1 esta etiquetada de manera incorrecta mostrada en el *Anexo 26*, alterando los cálculos. Este es un error común que puede surgir cuando se trabaja segmentando o etiquetando cualquier tipo de imágenes, pero especialmente cuando se trabaja con imágenes de tipo médico (Panayides et al., 2020).

Aunque el algoritmo presente algunos fallos, estos pueden considerarse fallos puntuales. Una posible causa de estos errores y punto de mejorar para este proyecto es la calidad de las imágenes, donde el brillo excesivo puede dificultar la apreciación de las características principales de las vértebras, tales como sus delimitaciones y esquinas. Es posible también reducir estos errores mediante un procesamiento de imágenes adecuado que elimine el ruido y mejore el contraste, permitiendo una mejor visualización de los bordes y detalles anatómicos. Además, la anatomía de las vértebras puede influir significativamente en las mediciones. Patologías como la degeneración discal pueden provocar estrechamientos en los espacios intervertebrales. Estos estrechamientos pueden alterar la detección de las vértebras debido a su proximidad o incluso solapamiento, dificultando la identificación precisa mediante técnicas de segmentación automática

VII. CONCLUSIONES

El algoritmo para la detección de la espondilolistesis y su clasificación según el sistema de Meyerding presenta una gran aplicabilidad como herramienta de apoyo en la medición precisa del porcentaje de traslación entre una vértebra y otra. Esta innovación es muy útil para los radiólogos, ya que actualmente se usa un método de medición subjetiva sin herramientas precisas. La implementación de este algoritmo puede aportar un grado significativo de objetividad en los resultados, proporcionando un respaldo visual que facilita al radiólogo la comunicación del diagnóstico preliminar al neurocirujano. Esta medición precisa es crucial, ya que el porcentaje de traslación puede determinar si el paciente requiere una intervención quirúrgica.

Para llevar a cabo el reentrenamiento del modelo de aprendizaje profundo para el etiquetado de puntos de interés en estudios de rayos X, se clasificaron y preprocesaron las imágenes de espondilolistesis tanto de la base de datos pública como de la base interna, donde se amplió hasta en 700 imágenes la base de datos usada en el entrenamiento. Permitiendo mejorar la precisión del modelo en la detección de puntos clave en diversos conjuntos de datos.

Para identificar los puntos de interés con precisión y optimizar el rendimiento del modelo al detectar vértebras, se seleccionó y ajustó las épocas de entrenamiento como parámetro de refinamiento, realizando múltiples entrenamientos con variaciones en dicho parámetro que influyó en un mejoramiento de las 5 métricas de evaluación que determinaron las capacidades del modelo.

Se desarrolló e implementó con éxito un algoritmo que utiliza los puntos de interés para realizar la medición del porcentaje de traslación vertebral y la clasificación automática de Meyerding como herramienta de apoyo para los radiólogos en brindar un diagnóstico preliminar preciso basado en mediciones puntuales.

De la mano de 3 especialistas divididos en radiología y neurocirugía, el método de detección automática de espondilolistesis se validó en términos de su capacidad para clasificar según sistema de clasificación de Meyerding y su fiabilidad en la detección de espondilolistesis en imágenes de la columna vertebral.

VIII. RECOMENDACIONES

Con el propósito de proyectar los pasos necesarios para continuar avanzando en la investigación y desarrollo de métodos de apoyo para facilitar la labor de los especialistas y asegurar la coherencia en los diagnósticos emitidos, se plantean las siguientes recomendaciones:

Para mejorar la robustez y fiabilidad del modelo de detección de vértebras, se recomienda ampliar el conjunto de datos utilizado ya que en la actualidad una de las principales limitaciones para las investigaciones alineadas al aprendizaje profundo en imágenes médicas es la falta de bases de datos extensas. Evaluando la posibilidad de expandir las capacidades a imágenes no solo de rayos X, sino a estudios de tomografía computarizada y resonancia magnética, lo que ayudaría a reducir la variabilidad en los datos y mejorar la aplicación del modelo. También la utilización de distintas arquitecturas de redes neuronales y técnicas de aprendizaje profundo como el aprendizaje por transferencia, para mejorar la detección de puntos clave en las vértebras.

Se recomienda desarrollar una interfaz de usuario intuitiva y fácil de usar para permitir a los especialistas usar el método propuesto como herramienta eficiente, facilitando la visualización y análisis de los resultados de la detección de espondilolistesis. De esta forma, se propone SpineGrade (*Anexo 61 a Anexo 67*) diseñada para la detección y clasificación de la espondilolistesis, en ella se podrá observar el porcentaje de traslación exacto y bajo que clasificación recae utilizando el sistema de Meyerding. SpineGrade permitirá a los usuarios cargar una imagen o un conjunto de imágenes, posibilitando un análisis exhaustivo. Además, la aplicación se integrará con Google Drive, proporcionando un acceso seguro a las imágenes almacenadas en la nube. Esta conectividad en la nube asegura que los usuarios puedan gestionar y recuperar sus datos de imágenes fácilmente desde cualquier lugar. SpineGrade combina algoritmos de cálculo de distancia utilizando puntos clave y líneas de referencia en las imágenes con un almacenamiento en la nube conveniente, ofrece a los profesionales de la salud una herramienta eficiente para el diagnóstico y la clasificación precisa de la espondilolistesis. Así permitiría realizar estudios de validación clínica más amplios con mayor participación de especialistas de diferentes especialidades. Evaluando la capacidad de integración con sistemas de información hospitalarios para la adopción de la tecnología propuesta en la práctica clínica cotidiana.

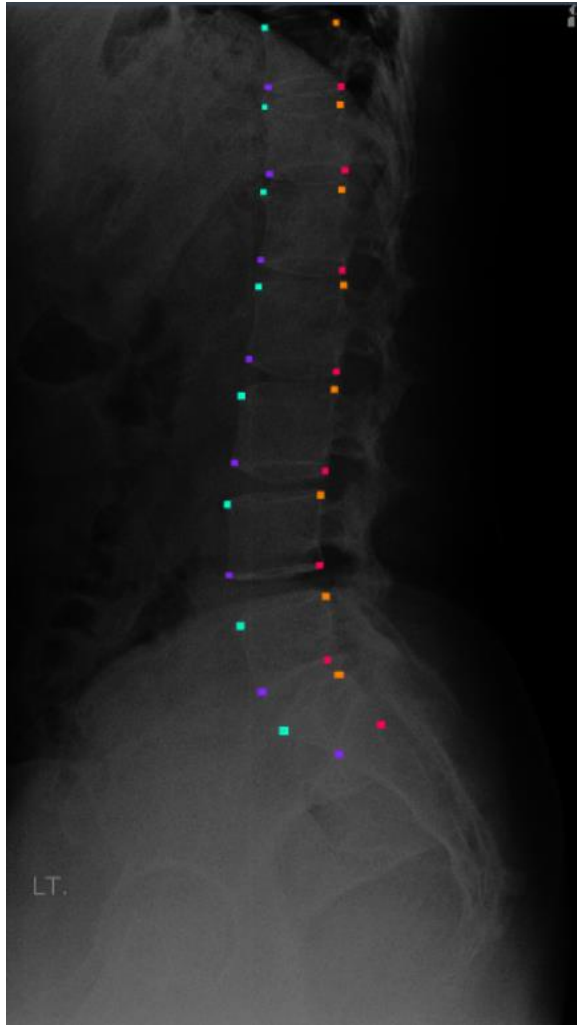
Además de la espondilolistesis, el enfoque y los métodos utilizados para llevar a cabo el proyecto pueden aplicarse a la investigación de otras condiciones médicas que se benefician de la detección automatizada en imágenes médicas, como ser la medición del ángulo de Cobb para la detección de escoliosis, la detección de fracturas vertebrales y otras patologías relacionadas con la columna vertebral.

Anexo 1 Anotaciones manuales con Keypoints en Roboflow



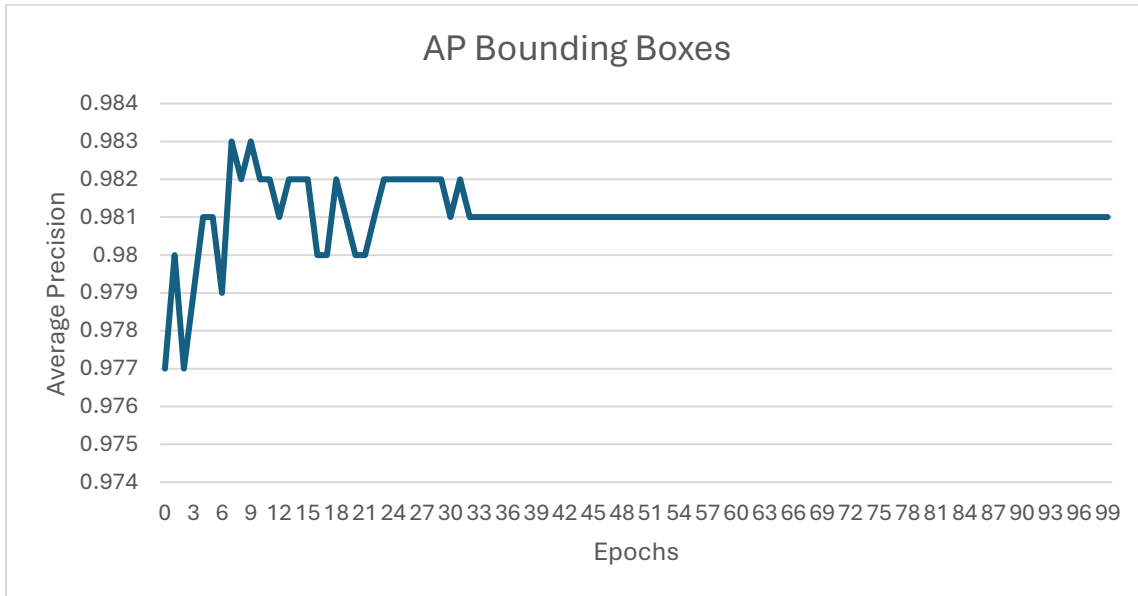
Fuente: Elaboración propia obtenida en Roboflow

Anexo 2 Anotaciones manuales con bounding boxes



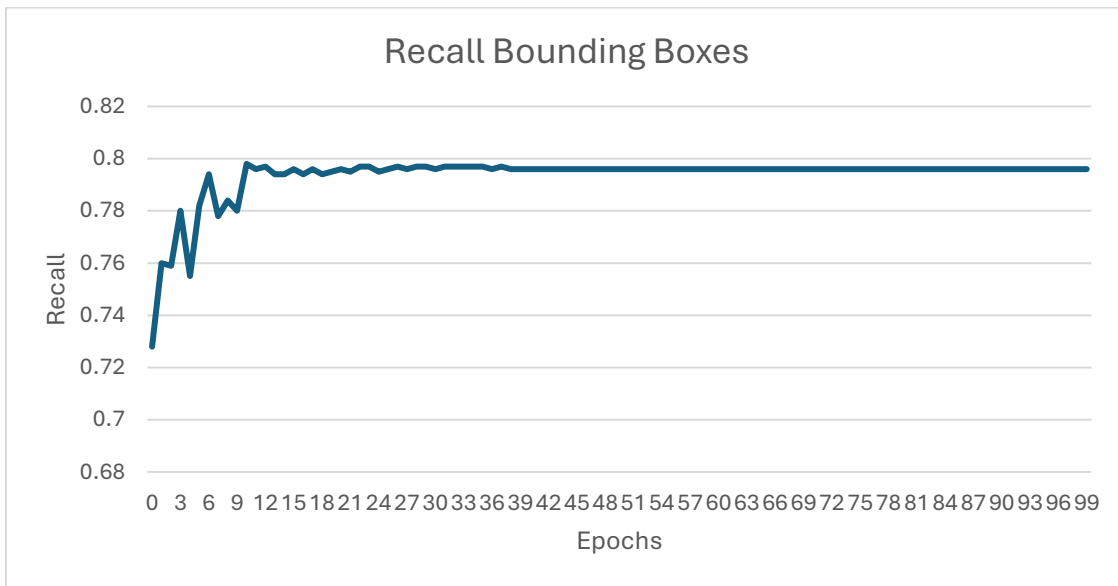
Fuente: Sugiyama & García (2023)

Anexo 3 Métrica de precisión promedio del modelo con 100 épocas en bounding boxes



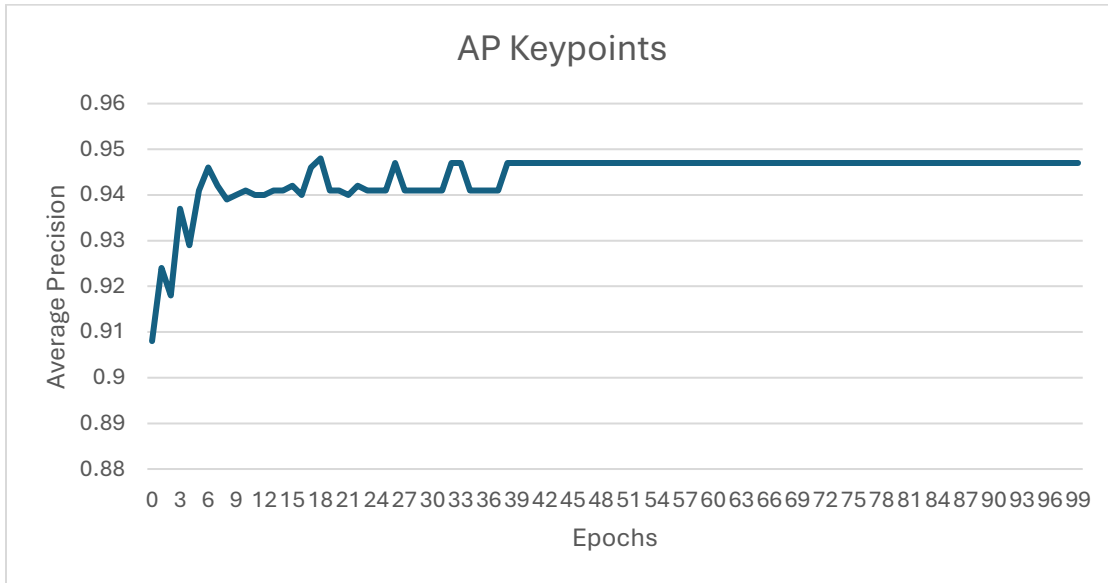
Fuente: Elaboración Propia

Anexo 4 Métrica de sensibilidad del modelo con 100 épocas en bounding boxes

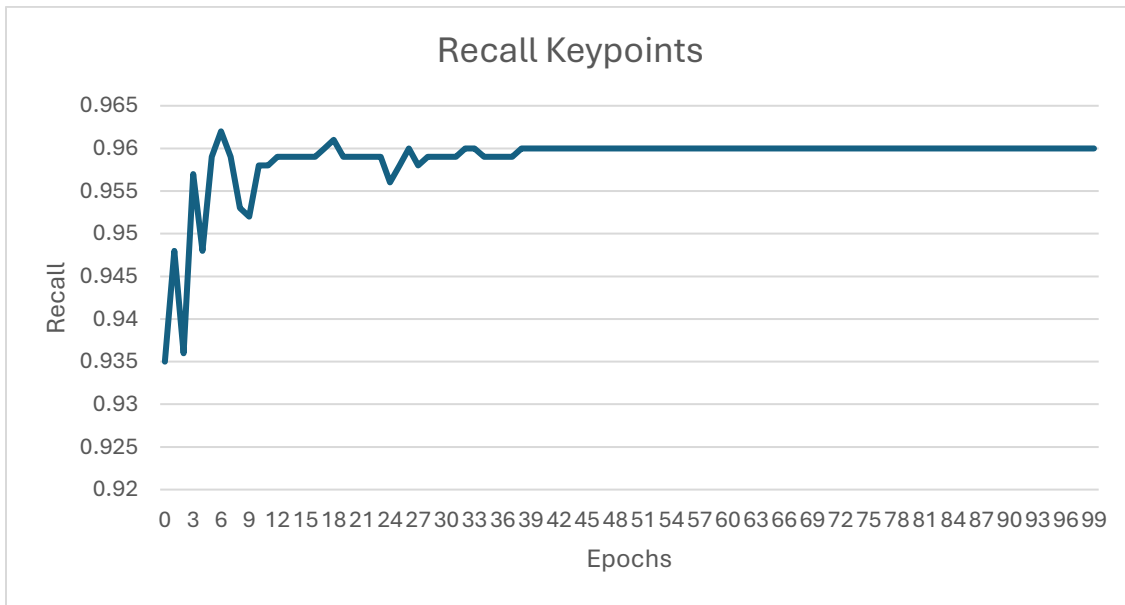


Fuente: Elaboración Propia

Anexo 5 Métrica de precisión promedio del modelo con 100 épocas en Keypoints

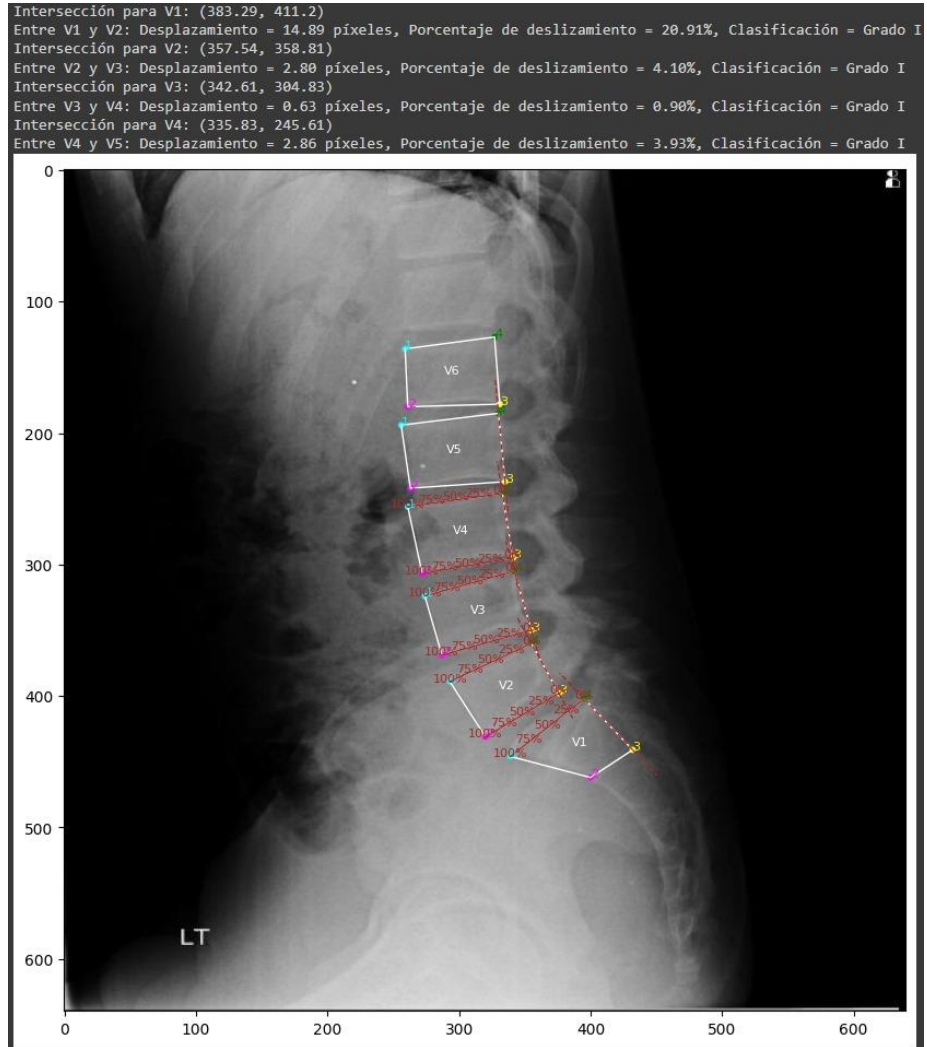


Anexo 6 Métrica de sensibilidad del modelo con 100 épocas en Keypoints



Fuente: Elaboración Propia

Anexo 7 Resultados del módulo de detección de espondilolistesis



Fuente: Elaboración Propia

Anexo 8 Resultados del módulo de detección de espondilolistesis

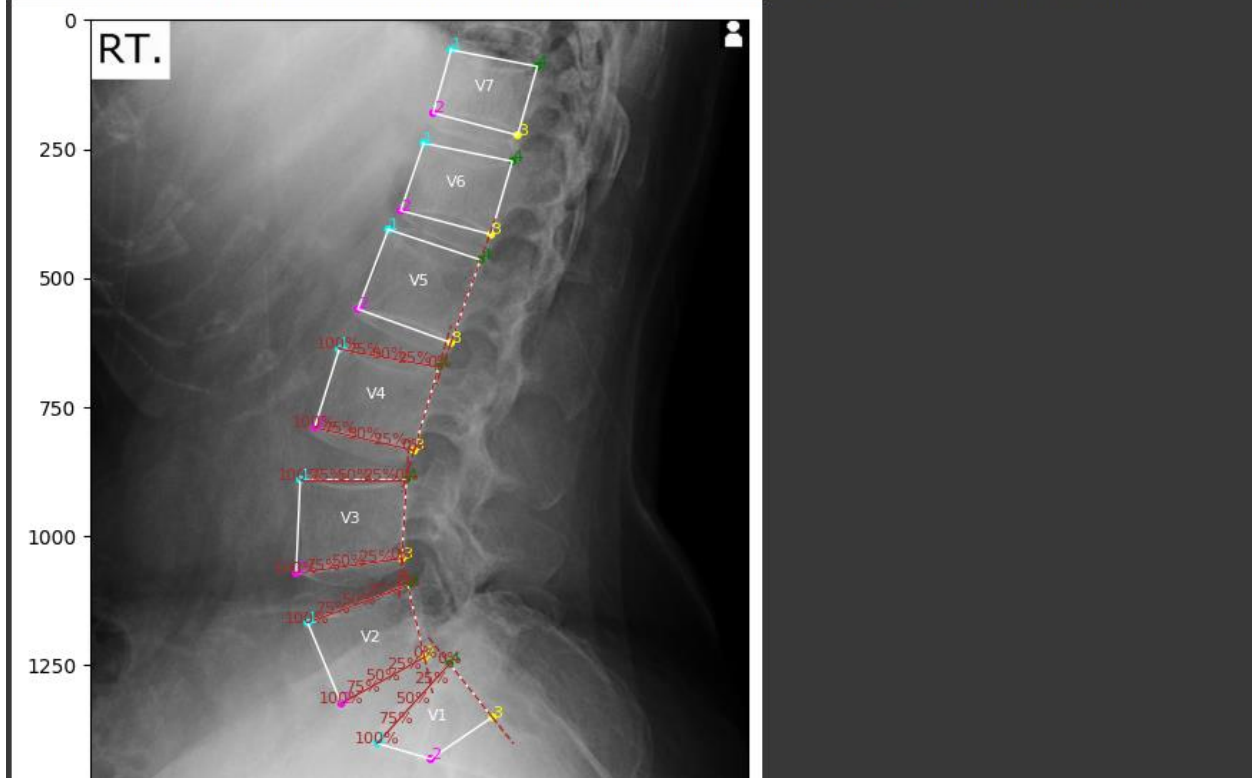
Intersección para V1: (1490.6, 1144.76)
Entre V1 y V2: Desplazamiento = 19.68 píxeles, Porcentaje de deslizamiento = 11.44%, Clasificación = Grado I
Intersección para V2: (1377.42, 1012.1)
Entre V2 y V3: Desplazamiento = 28.37 píxeles, Porcentaje de deslizamiento = 15.92%, Clasificación = Grado I
Intersección para V3: (1244.33, 845.4)
Entre V3 y V4: Desplazamiento = 14.85 píxeles, Porcentaje de deslizamiento = 7.24%, Clasificación = Normal
Intersección para V4: (1199.95, 663.0)
Entre V4 y V5: Desplazamiento = 3.34 píxeles, Porcentaje de deslizamiento = 1.56%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 9 Resultados del módulo de detección de espondilolistesis

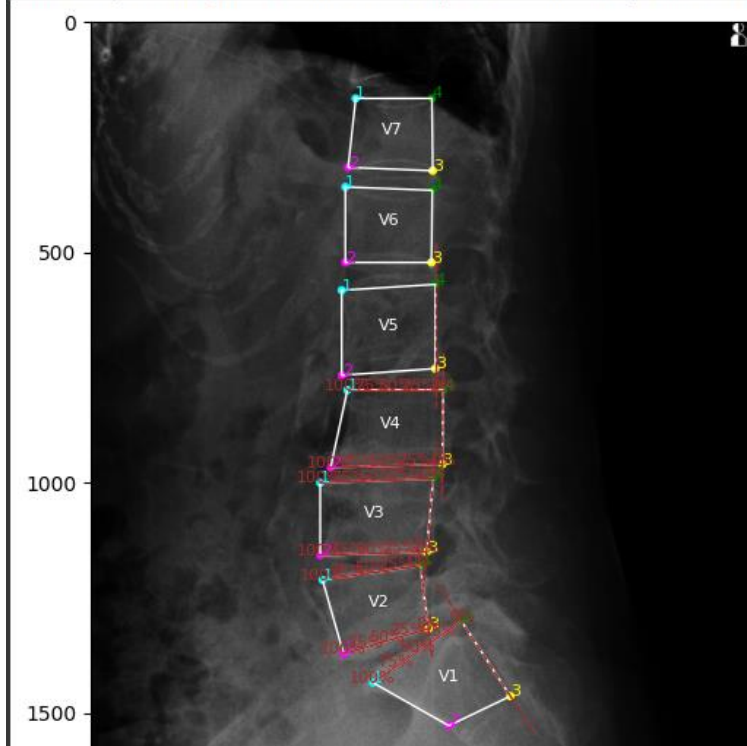
Intersección para V1: (659.63, 1287.31)
Entre V1 y V2: Desplazamiento = 54.54 píxeles, Porcentaje de deslizamiento = 26.18%, Clasificación = Grado II
Intersección para V2: (599.83, 1101.05)
Entre V2 y V3: Desplazamiento = 16.85 píxeles, Porcentaje de deslizamiento = 8.13%, Clasificación = Normal
Intersección para V3: (610.74, 890.14)
Entre V3 y V4: Desplazamiento = 2.33 píxeles, Porcentaje de deslizamiento = 1.13%, Clasificación = Normal
Intersección para V4: (681.56, 673.42)
Entre V4 y V5: Desplazamiento = 6.45 píxeles, Porcentaje de deslizamiento = 3.29%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 10 Resultados del módulo de detección de espondilolistesis

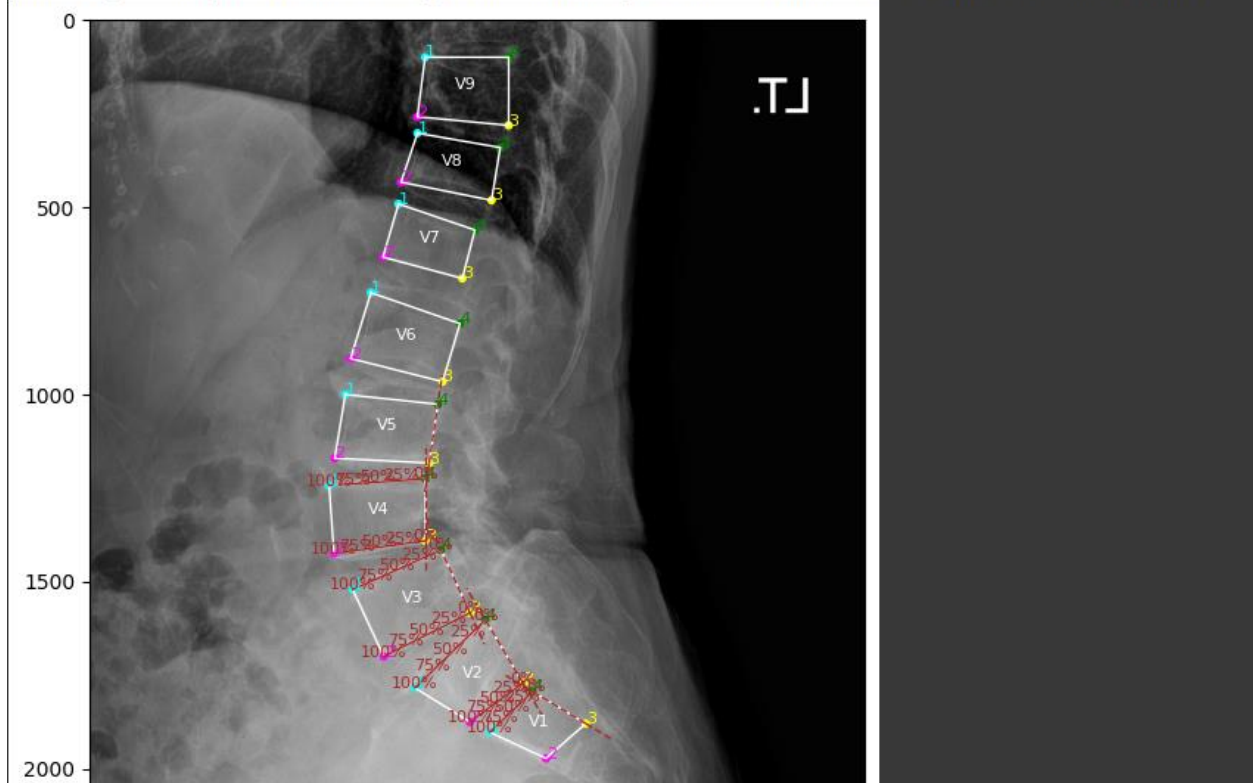
Intersección para V1: (738.53, 1347.22)
Entre V1 y V2: Desplazamiento = 82.02 píxeles, Porcentaje de deslizamiento = 34.98%, Clasificación = Grado II
Intersección para V2: (727.3, 1178.43)
Entre V2 y V3: Desplazamiento = 11.27 píxeles, Porcentaje de deslizamiento = 5.25%, Clasificación = Normal
Intersección para V3: (763.96, 990.32)
Entre V3 y V4: Desplazamiento = 19.41 píxeles, Porcentaje de deslizamiento = 7.89%, Clasificación = Normal
Intersección para V4: (748.96, 798.43)
Entre V4 y V5: Desplazamiento = 17.18 píxeles, Porcentaje de deslizamiento = 8.25%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 11 Resultados del módulo de detección de espondilolistesis

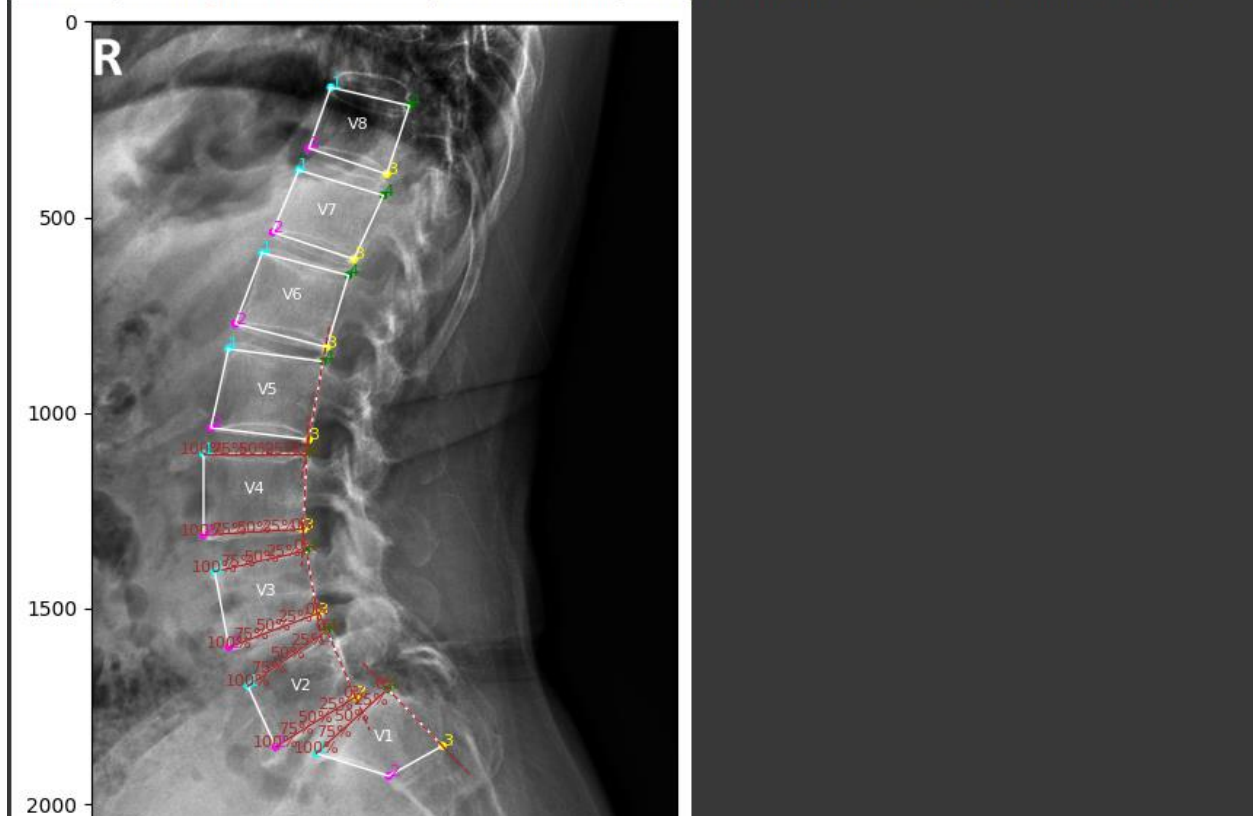
Intersección para V1: (1177.9, 1800.98)
Entre V1 y V2: Desplazamiento = 10.53 píxeles, Porcentaje de deslizamiento = 6.68%, Clasificación = Normal
Intersección para V2: (1036.96, 1623.73)
Entre V2 y V3: Desplazamiento = 31.08 píxeles, Porcentaje de deslizamiento = 11.87%, Clasificación = Grado I
Intersección para V3: (898.52, 1434.16)
Entre V3 y V4: Desplazamiento = 46.02 píxeles, Porcentaje de deslizamiento = 17.78%, Clasificación = Grado I
Intersección para V4: (902.85, 1224.69)
Entre V4 y V5: Desplazamiento = 4.35 píxeles, Porcentaje de deslizamiento = 1.69%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 12 Resultados del módulo de detección de espondilolistesis

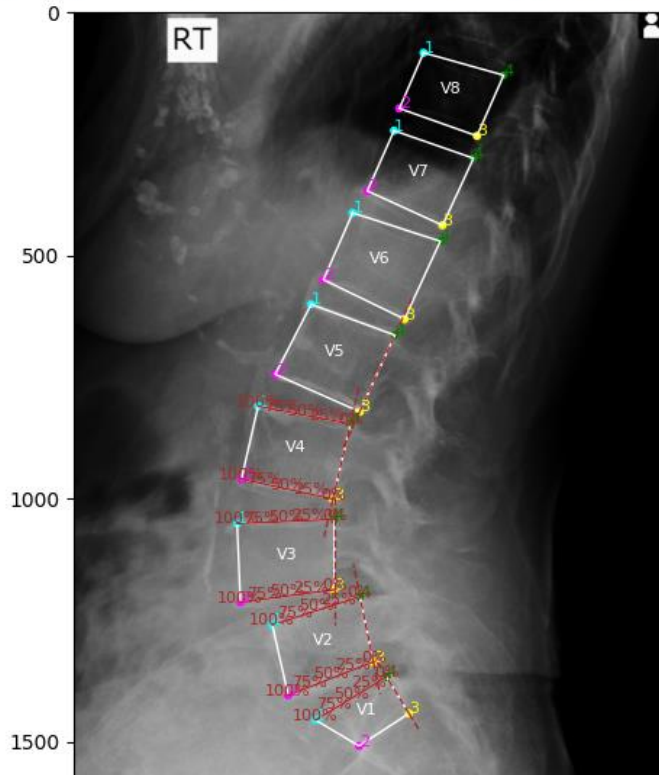
Intersección para V1: (692.48, 1765.46)
Entre V1 y V2: Desplazamiento = 88.72 píxeles, Porcentaje de deslizamiento = 36.03%, Clasificación = Grado II
Intersección para V2: (587.2, 1567.78)
Entre V2 y V3: Desplazamiento = 18.95 píxeles, Porcentaje de deslizamiento = 7.67%, Clasificación = Normal
Intersección para V3: (539.93, 1354.61)
Entre V3 y V4: Desplazamiento = 10.14 píxeles, Porcentaje de deslizamiento = 4.22%, Clasificación = Normal
Intersección para V4: (549.56, 1105.29)
Entre V4 y V5: Desplazamiento = 1.51 píxeles, Porcentaje de deslizamiento = 0.57%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 13 Resultados del módulo de detección de espondilolistesis

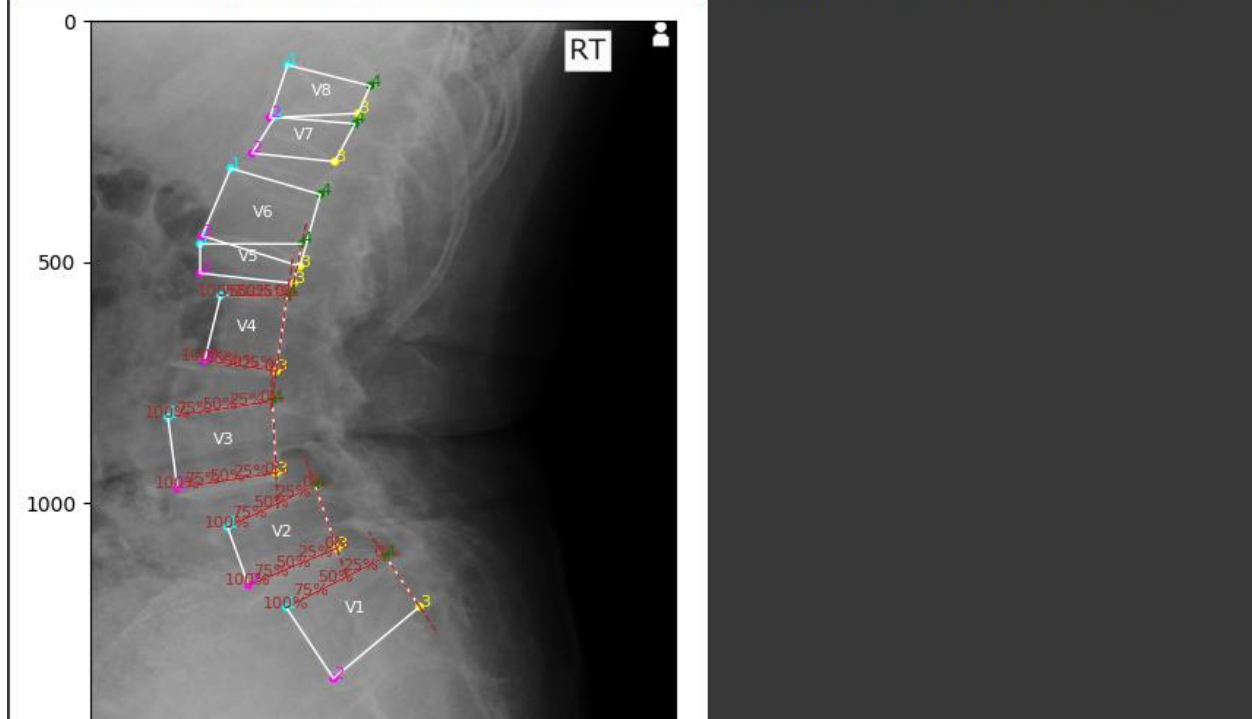
Intersección para V1: (626.02, 1377.94)
Entre V1 y V2: Desplazamiento = 21.66 píxeles, Porcentaje de deslizamiento = 12.67%, Clasificación = Grado I
Intersección para V2: (536.58, 1218.64)
Entre V2 y V3: Desplazamiento = 56.00 píxeles, Porcentaje de deslizamiento = 29.50%, Clasificación = Grado II
Intersección para V3: (524.28, 1043.14)
Entre V3 y V4: Desplazamiento = 12.31 píxeles, Porcentaje de deslizamiento = 6.15%, Clasificación = Normal
Intersección para V4: (577.28, 847.33)
Entre V4 y V5: Desplazamiento = 7.33 píxeles, Porcentaje de deslizamiento = 3.79%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 14 Resultados del módulo de detección de espondilolistesis

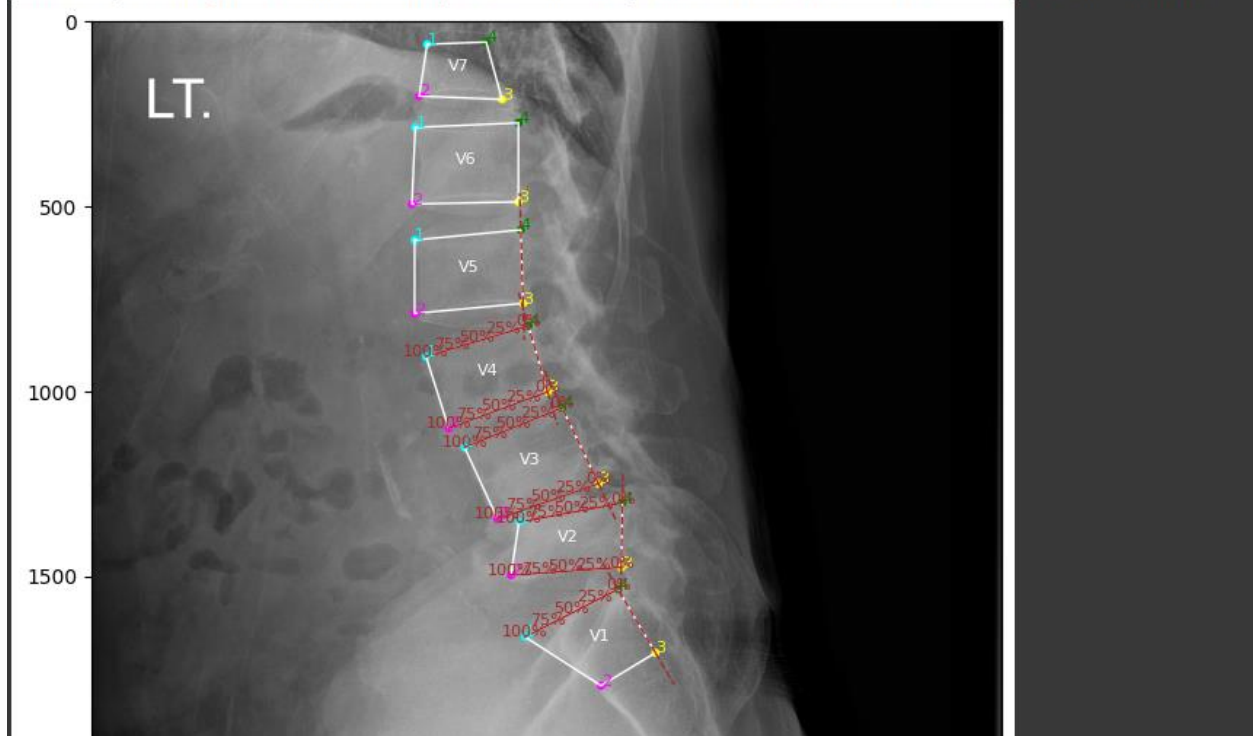
Intersección para V1: (533.16, 1152.7)
Entre V1 y V2: Desplazamiento = 90.75 píxeles, Porcentaje de deslizamiento = 39.07%, Clasificación = Grado II
Intersección para V2: (388.75, 1001.06)
Entre V2 y V3: Desplazamiento = 86.37 píxeles, Porcentaje de deslizamiento = 42.96%, Clasificación = Grado II
Intersección para V3: (377.17, 788.54)
Entre V3 y V4: Desplazamiento = 1.57 píxeles, Porcentaje de deslizamiento = 0.71%, Clasificación = Normal
Intersección para V4: (416.58, 570.77)
Entre V4 y V5: Desplazamiento = 7.49 píxeles, Porcentaje de deslizamiento = 5.41%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 15 Resultados del módulo de detección de espondilolistesis

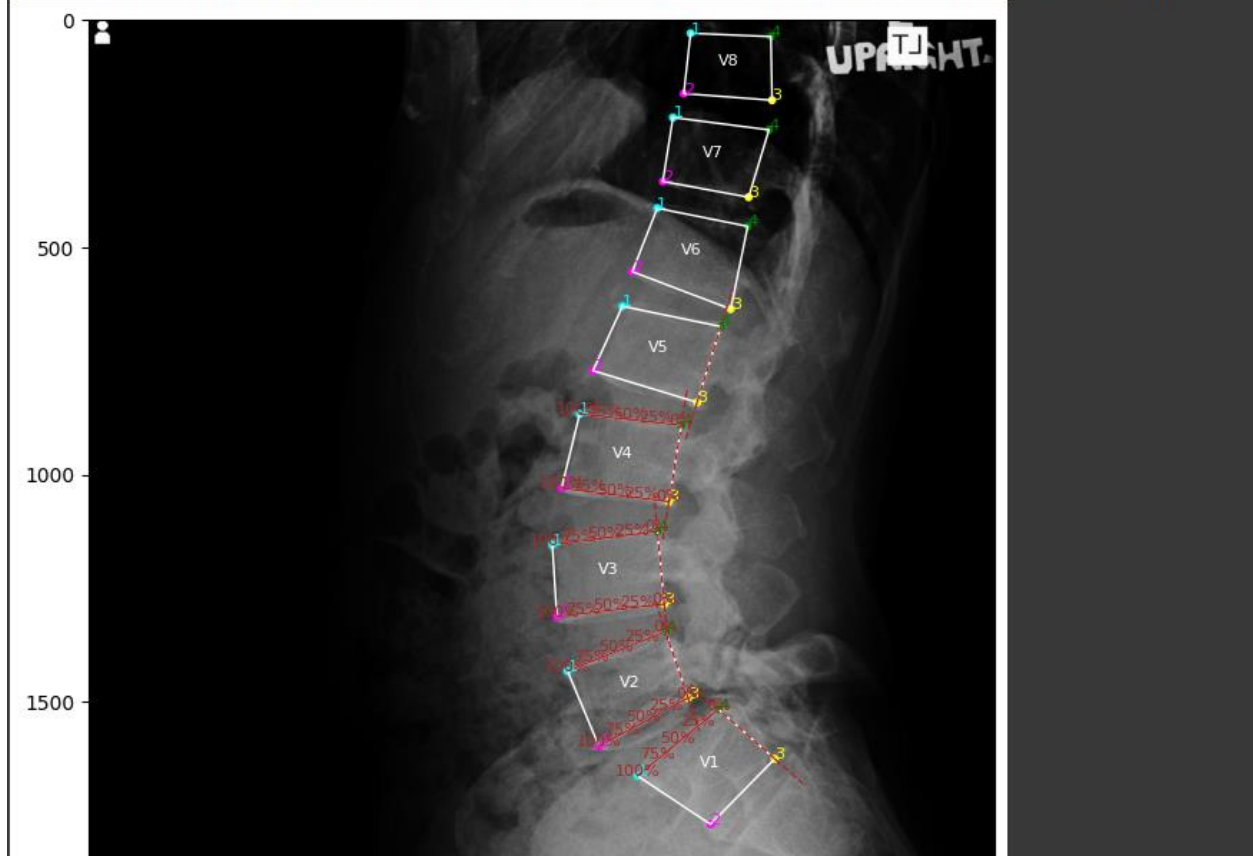
Intersección para V1: (1433.59, 1530.71)
Entre V1 y V2: Desplazamiento = 8.72 píxeles, Porcentaje de deslizamiento = 3.08%, Clasificación = Normal
Intersección para V2: (1401.23, 1308.14)
Entre V2 y V3: Desplazamiento = 37.04 píxeles, Porcentaje de deslizamiento = 13.00%, Clasificación = Grado I
Intersección para V3: (1250.82, 1053.63)
Entre V3 y V4: Desplazamiento = 25.76 píxeles, Porcentaje de deslizamiento = 9.03%, Clasificación = Normal
Intersección para V4: (1170.72, 826.32)
Entre V4 y V5: Desplazamiento = 12.84 píxeles, Porcentaje de deslizamiento = 4.44%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 16 Resultados del módulo de detección de espondilolistesis

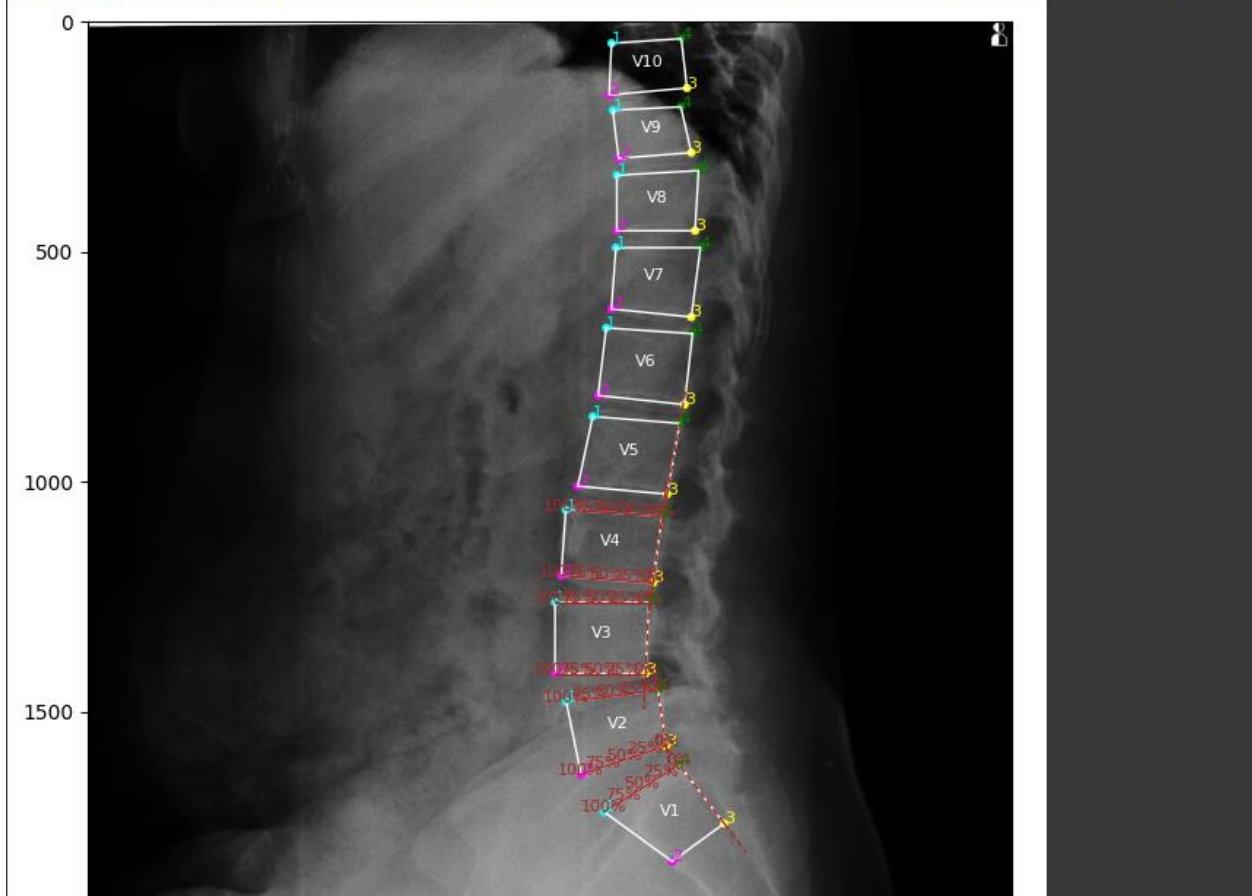
Intersección para V1: (1342.17, 1554.23)
Entre V1 y V2: Desplazamiento = 58.90 píxeles, Porcentaje de deslizamiento = 25.51%, Clasificación = Grado II
Intersección para V2: (1273.89, 1344.5)
Entre V2 y V3: Desplazamiento = 2.84 píxeles, Porcentaje de deslizamiento = 1.22%, Clasificación = Normal
Intersección para V3: (1268.23, 1122.97)
Entre V3 y V4: Desplazamiento = 14.93 píxeles, Porcentaje de deslizamiento = 6.39%, Clasificación = Normal
Intersección para V4: (1324.24, 893.49)
Entre V4 y V5: Desplazamiento = 18.38 píxeles, Porcentaje de deslizamiento = 8.15%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 17 Resultados del módulo de detección de espondilolistesis

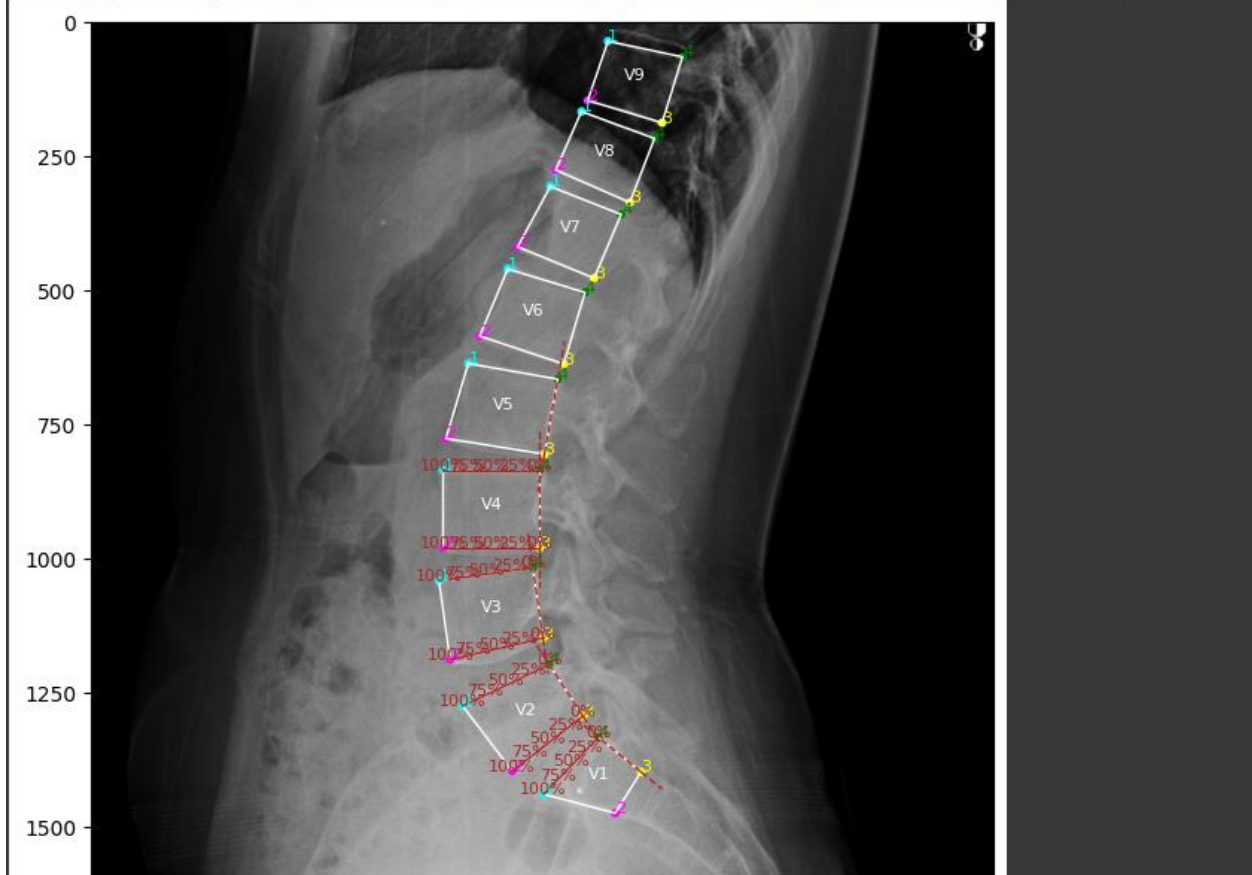
Intersección para V1: (1266.68, 1628.32)
Entre V1 y V2: Desplazamiento = 24.08 píxeles, Porcentaje de deslizamiento = 12.49%, Clasificación = Grado I
Intersección para V2: (1212.14, 1457.34)
Entre V2 y V3: Desplazamiento = 29.37 píxeles, Porcentaje de deslizamiento = 14.57%, Clasificación = Grado I
Intersección para V3: (1224.5, 1261.73)
Entre V3 y V4: Desplazamiento = 2.93 píxeles, Porcentaje de deslizamiento = 1.43%, Clasificación = Normal
Intersección para V4: (1252.63, 1074.02)
Entre V4 y V5: Desplazamiento = 2.38 píxeles, Porcentaje de deslizamiento = 1.13%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 18 Resultados del módulo de detección de espondilolistesis

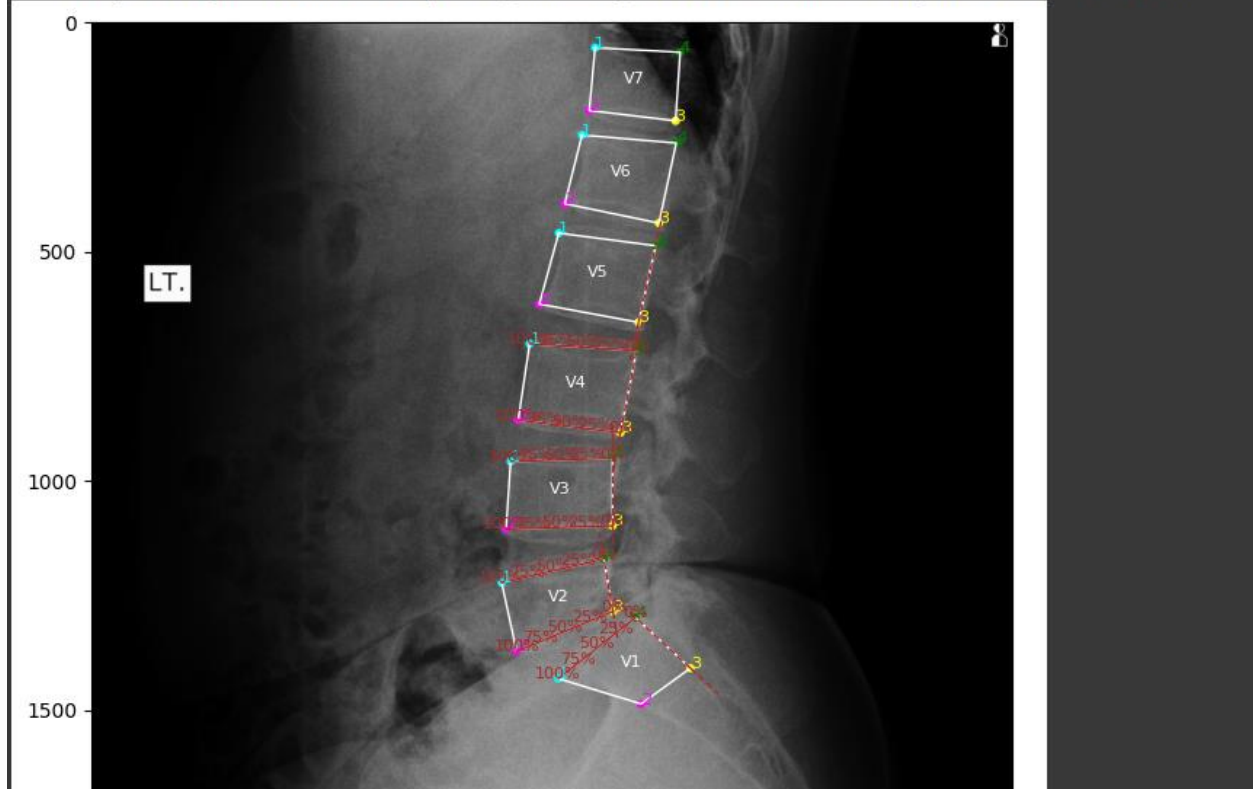
Intersección para V1: (945.24, 1334.71)
Entre V1 y V2: Desplazamiento = 2.74 píxeles, Porcentaje de deslizamiento = 1.86%, Clasificación = Normal
Intersección para V2: (850.7, 1198.96)
Entre V2 y V3: Desplazamiento = 5.06 píxeles, Porcentaje de deslizamiento = 2.83%, Clasificación = Normal
Intersección para V3: (837.31, 1014.07)
Entre V3 y V4: Desplazamiento = 12.77 píxeles, Porcentaje de deslizamiento = 7.22%, Clasificación = Normal
Intersección para V4: (840.86, 834.1)
Entre V4 y V5: Desplazamiento = 3.55 píxeles, Porcentaje de deslizamiento = 1.97%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 19 Resultados del módulo de detección de espondilolistesis

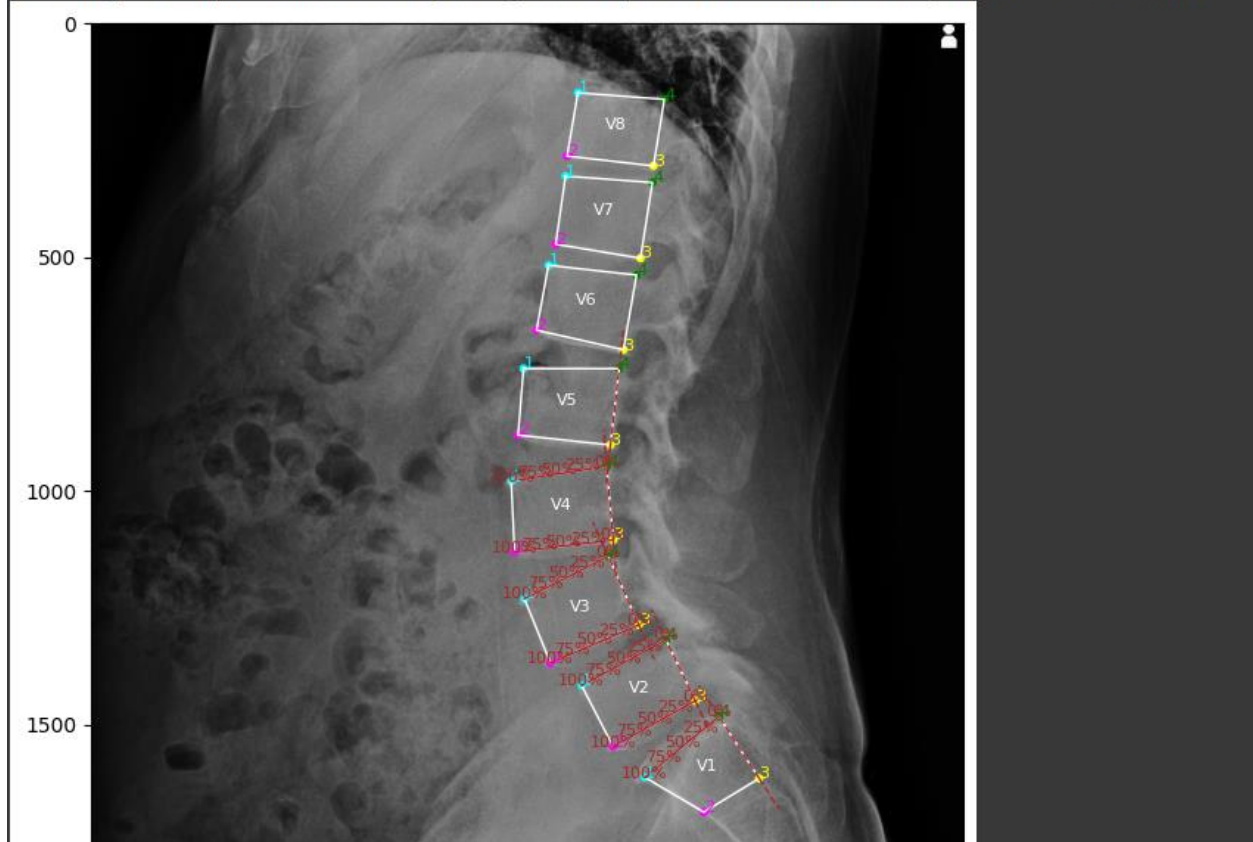
Intersección para V1: (1147.11, 1330.58)
Entre V1 y V2: Desplazamiento = 52.23 píxeles, Porcentaje de deslizamiento = 24.35%, Clasificación = Grado I
Intersección para V2: (1138.09, 1165.71)
Entre V2 y V3: Desplazamiento = 19.53 píxeles, Porcentaje de deslizamiento = 8.53%, Clasificación = Normal
Intersección para V3: (1144.66, 949.9)
Entre V3 y V4: Desplazamiento = 6.57 píxeles, Porcentaje de deslizamiento = 2.95%, Clasificación = Normal
Intersección para V4: (1181.48, 714.98)
Entre V4 y V5: Desplazamiento = 9.18 píxeles, Porcentaje de deslizamiento = 3.93%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 20 Resultados del módulo de detección de espondilolistesis

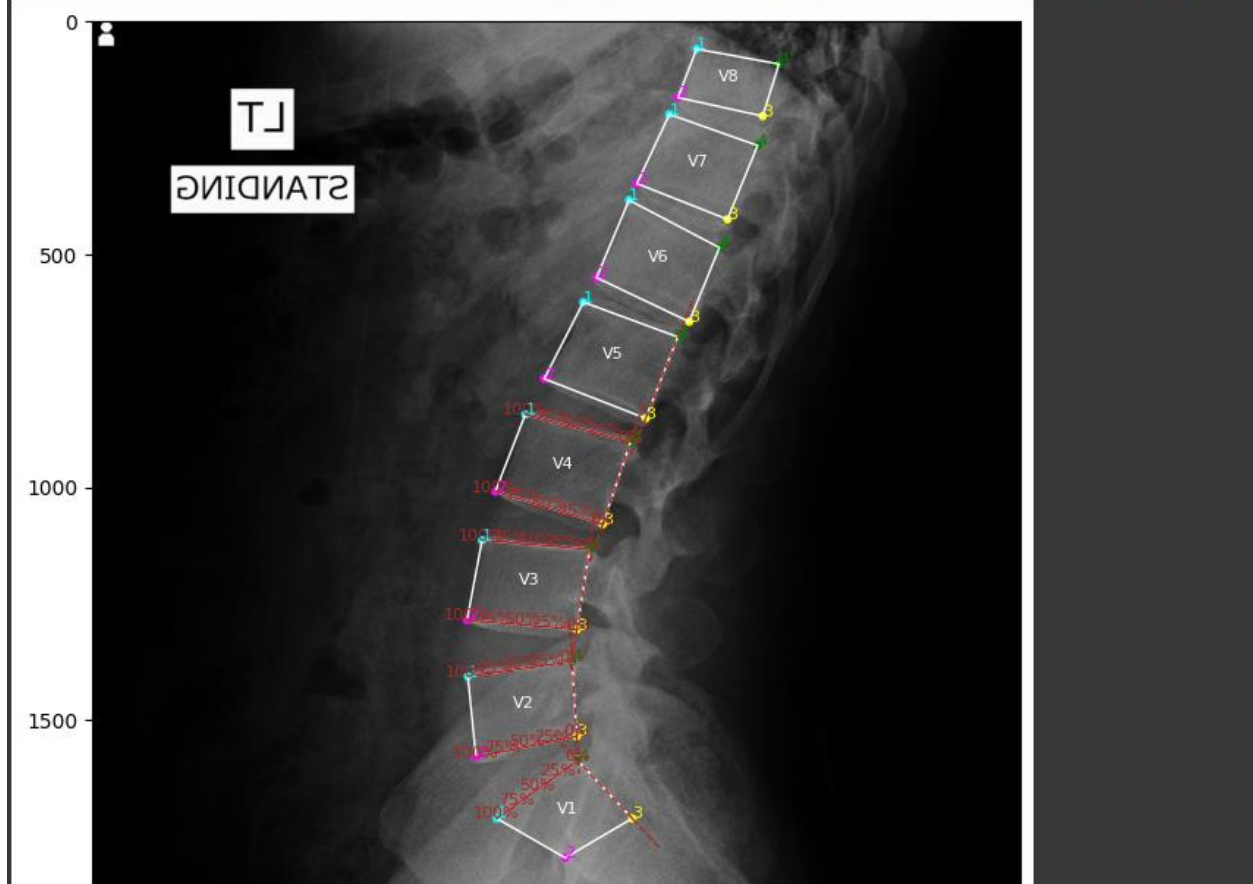
Intersección para V1: (1320.15, 1502.75)
Entre V1 y V2: Desplazamiento = 32.11 píxeles, Porcentaje de deslizamiento = 15.66%, Clasificación = Grado I
Intersección para V2: (1196.29, 1337.86)
Entre V2 y V3: Desplazamiento = 40.24 píxeles, Porcentaje de deslizamiento = 19.59%, Clasificación = Grado I
Intersección para V3: (1121.05, 1133.24)
Entre V3 y V4: Desplazamiento = 14.60 píxeles, Porcentaje de deslizamiento = 7.25%, Clasificación = Normal
Intersección para V4: (1107.51, 946.3)
Entre V4 y V5: Desplazamiento = 2.97 píxeles, Porcentaje de deslizamiento = 1.44%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 21 Resultados del módulo de detección de espondilolistesis

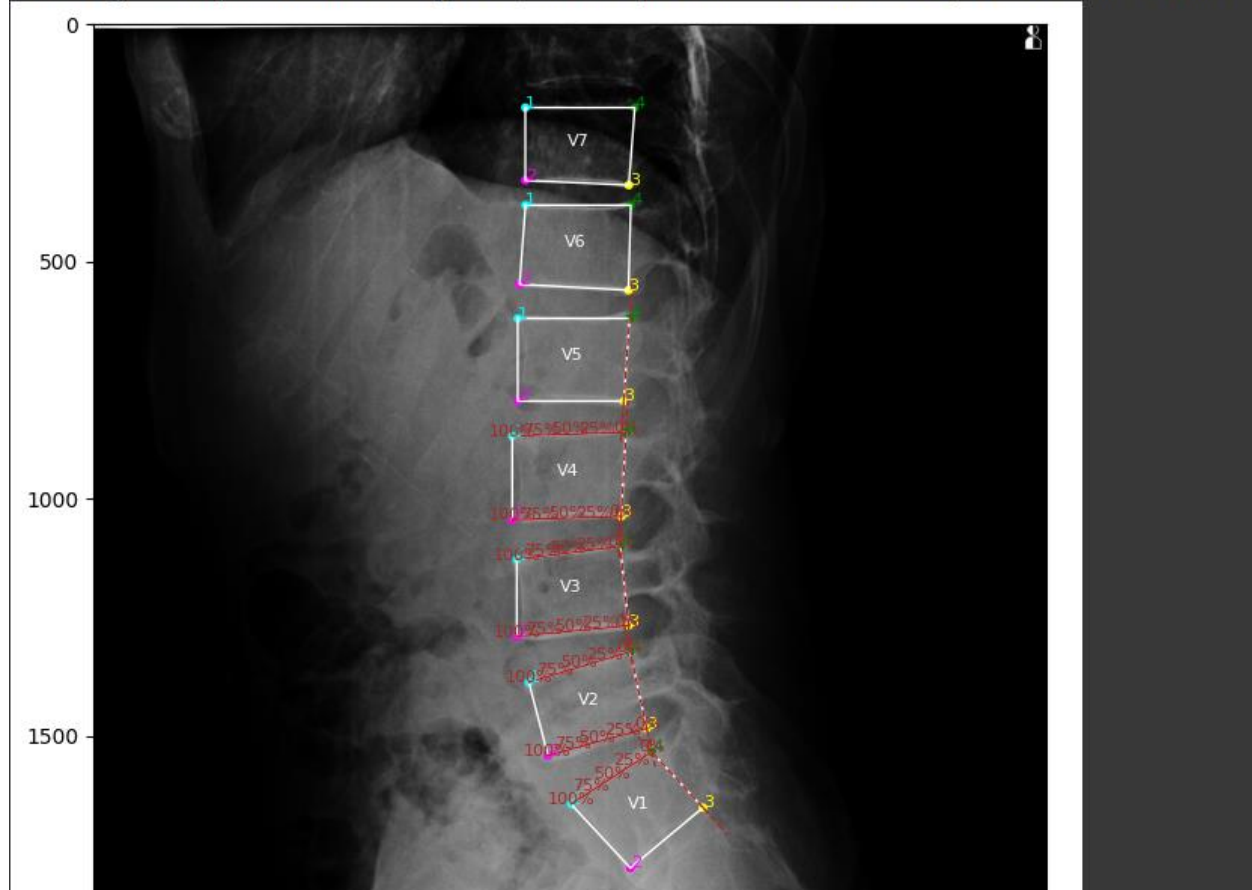
Intersección para V1: (1045.27, 1587.07)
Entre V1 y V2: Desplazamiento = 0.87 píxeles, Porcentaje de deslizamiento = 0.41%, Clasificación = Normal
Intersección para V2: (1031.52, 1374.49)
Entre V2 y V3: Desplazamiento = 0.61 píxeles, Porcentaje de deslizamiento = 0.27%, Clasificación = Normal
Intersección para V3: (1079.49, 1134.41)
Entre V3 y V4: Desplazamiento = 9.72 píxeles, Porcentaje de deslizamiento = 4.19%, Clasificación = Normal
Intersección para V4: (1168.83, 906.02)
Entre V4 y V5: Desplazamiento = 11.33 píxeles, Porcentaje de deslizamiento = 4.84%, Clasificación = Normal



Fuente: Elaboración Propia

Anexo 22 Resultados del módulo de detección de espondilolistesis

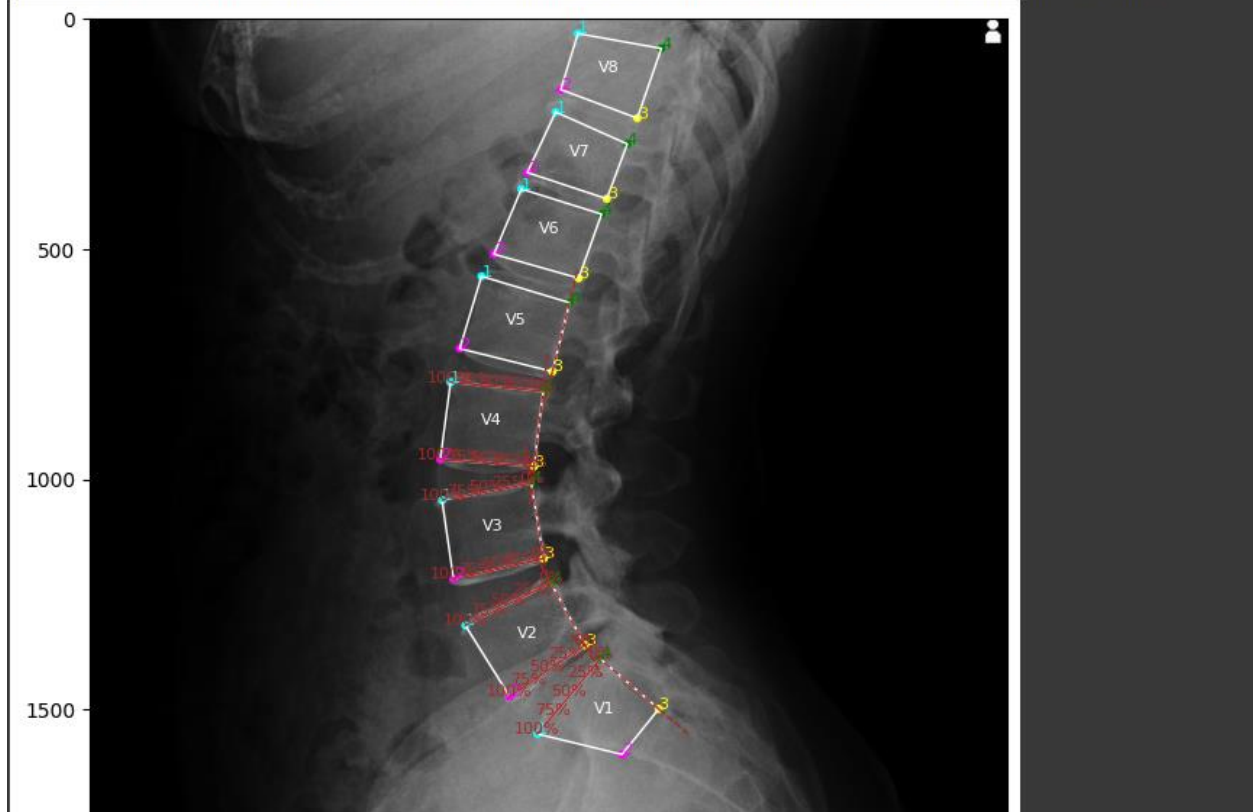
Intersección para V1: (1178.63, 1534.12)
Entre V1 y V2: Desplazamiento = 0.10 píxeles, Porcentaje de deslizamiento = 0.05%, Clasificación = Normal
Intersección para V2: (1135.04, 1321.74)
Entre V2 y V3: Desplazamiento = 2.53 píxeles, Porcentaje de deslizamiento = 1.13%, Clasificación = Normal
Intersección para V3: (1110.35, 1100.67)
Entre V3 y V4: Desplazamiento = 1.49 píxeles, Porcentaje de deslizamiento = 0.68%, Clasificación = Normal
Intersección para V4: (1114.45, 860.48)
Entre V4 y V5: Desplazamiento = 9.49 píxeles, Porcentaje de deslizamiento = 3.95%, Clasificación = Normal



Fuente: Elaboración Propia

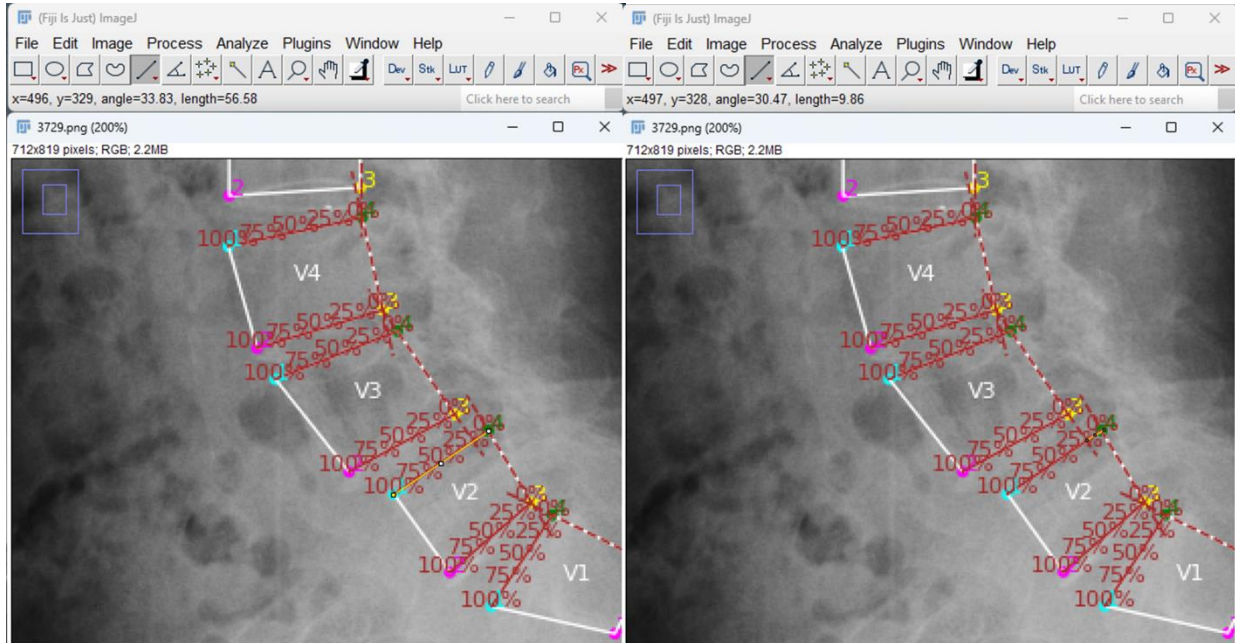
Anexo 23 Resultados del módulo de detección de espondilolistesis

Intersección para V1: (1101.03, 1402.47)
Entre V1 y V2: Desplazamiento = 15.56 píxeles, Porcentaje de deslizamiento = 7.31%, Clasificación = Normal
Intersección para V2: (997.37, 1228.71)
Entre V2 y V3: Desplazamiento = 6.68 píxeles, Porcentaje de deslizamiento = 3.22%, Clasificación = Normal
Intersección para V3: (962.38, 1008.18)
Entre V3 y V4: Desplazamiento = 1.80 píxeles, Porcentaje de deslizamiento = 0.91%, Clasificación = Normal
Intersección para V4: (995.86, 809.8)
Entre V4 y V5: Desplazamiento = 9.23 píxeles, Porcentaje de deslizamiento = 4.56%, Clasificación = Normal



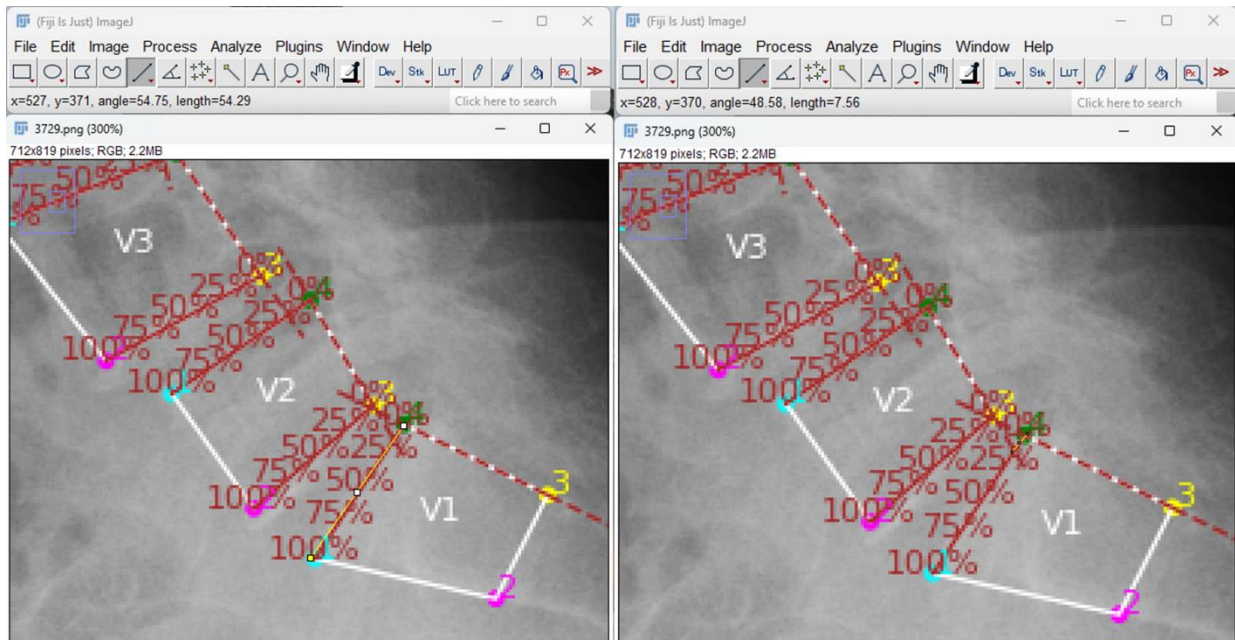
Fuente: Elaboración Propia

Anexo 24 Medición con ImageJ (3729 L4-L5)



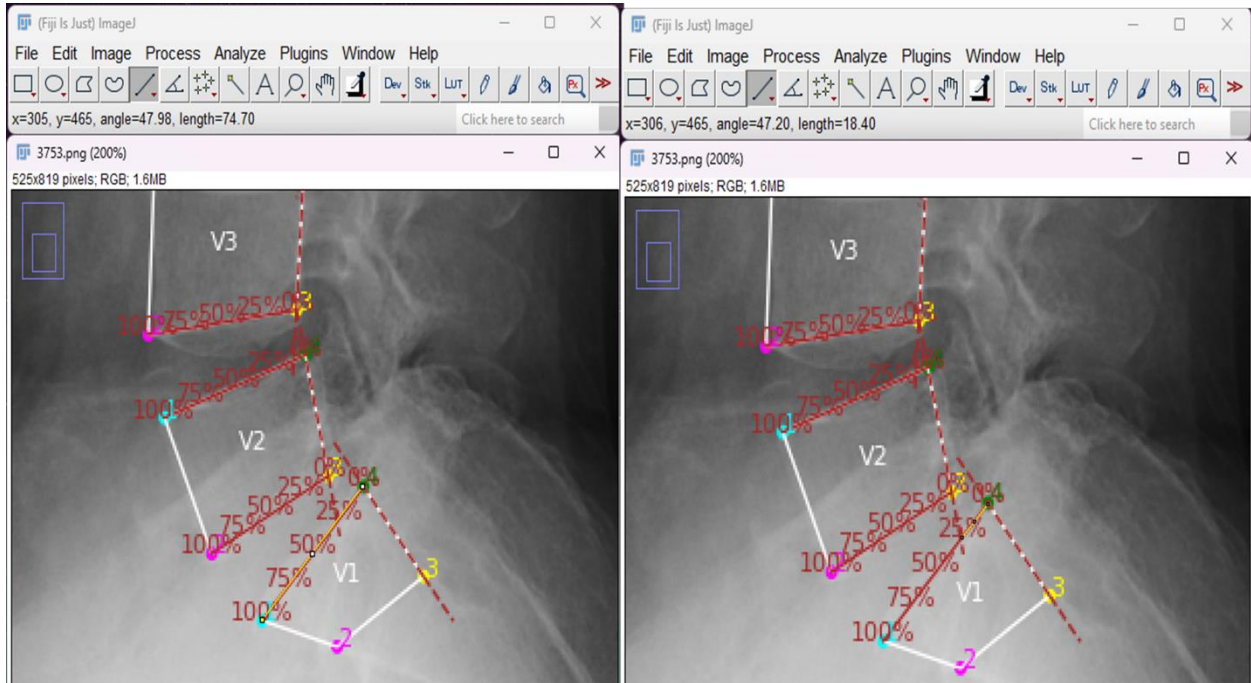
Fuente: Elaboración Propia

Anexo 25 Medición con ImageJ (3729 L5-S1)



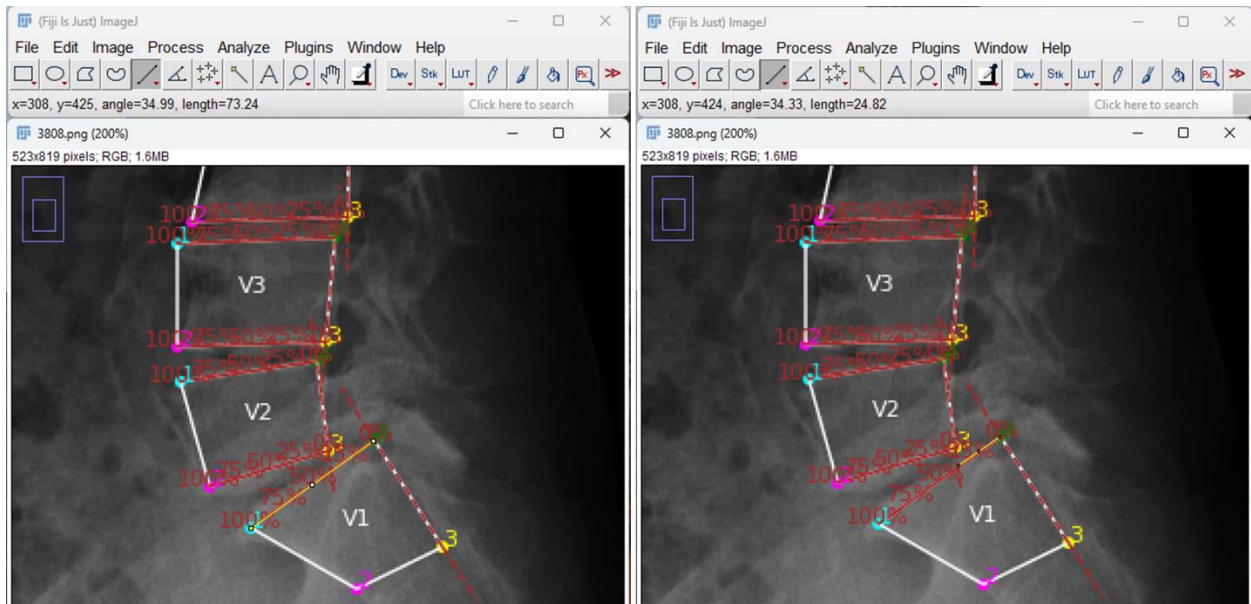
Fuente: Elaboración Propia

Anexo 26 Medición con ImageJ (3753 L5-S1)



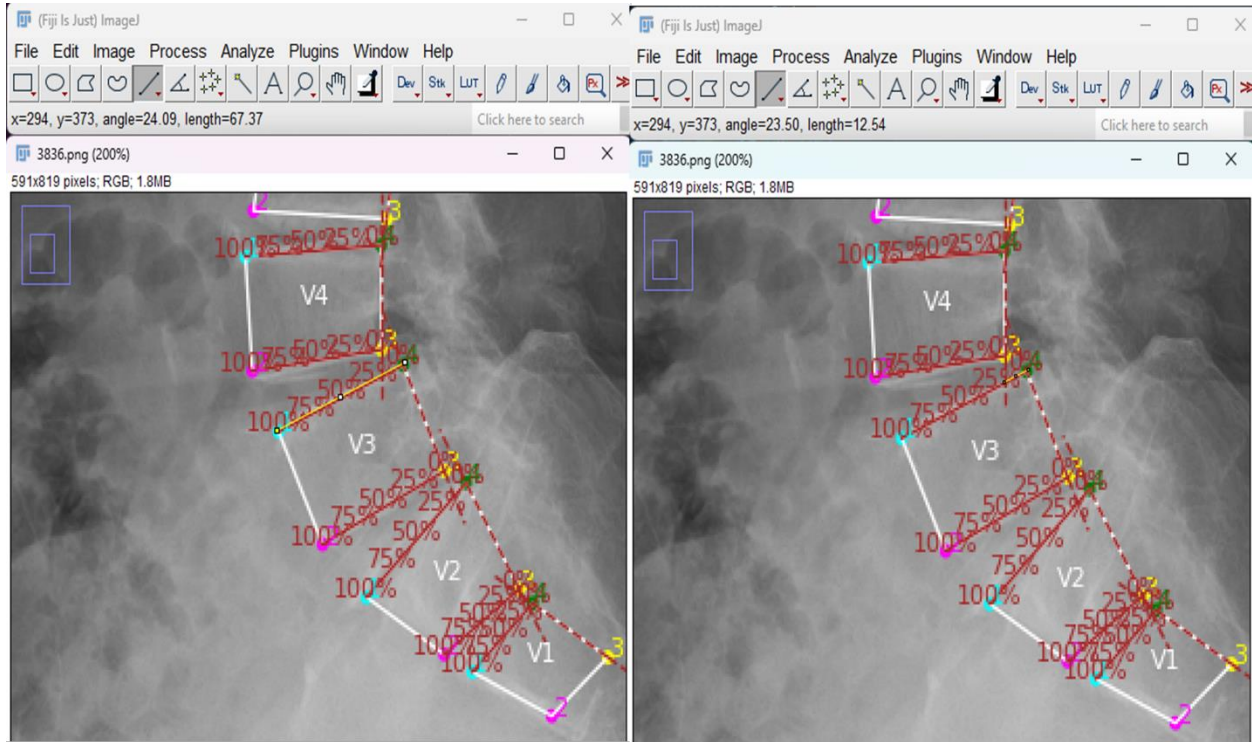
Fuente: Elaboración Propia

Anexo 27 Medición con ImageJ (3808 L5-S1)



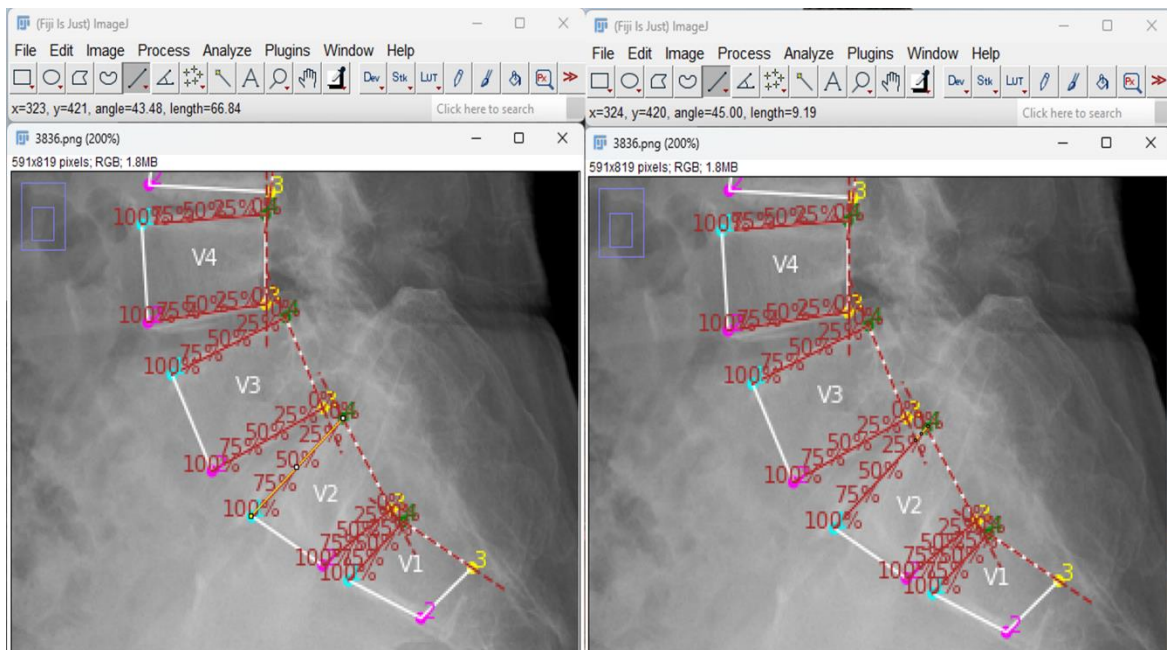
Fuente: Elaboración Propia

Anexo 28 Medición con ImageJ (3836 L3-L4)



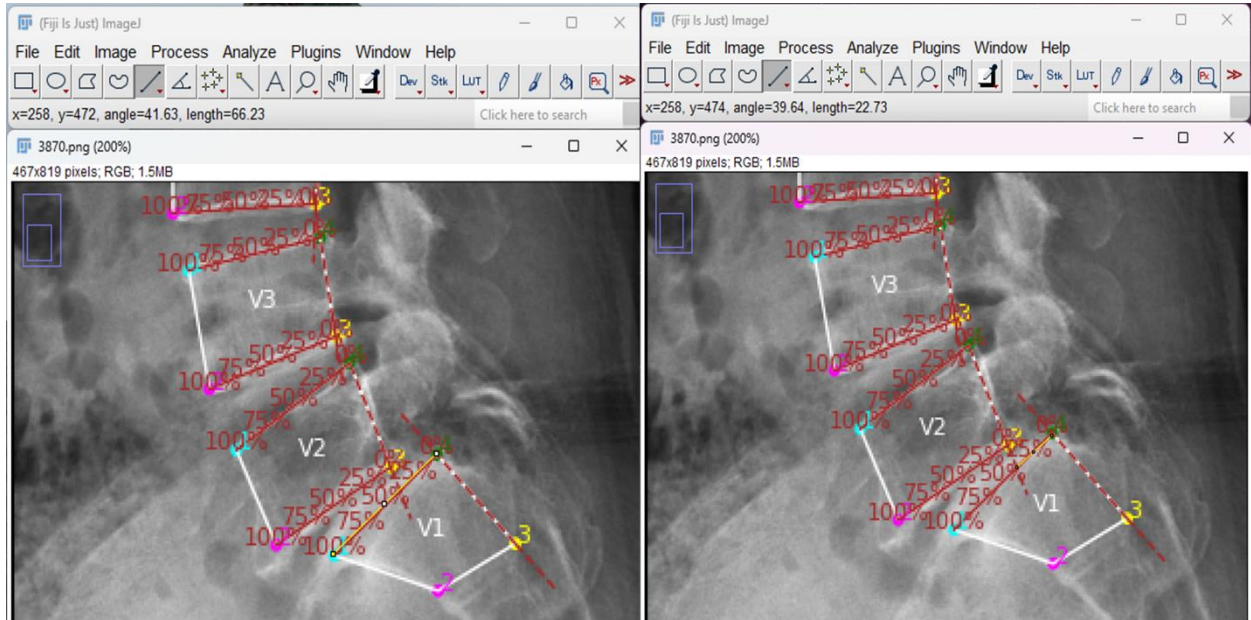
Fuente: Elaboración Propia

Anexo 29 Medición con ImageJ (3836 L4-L5)



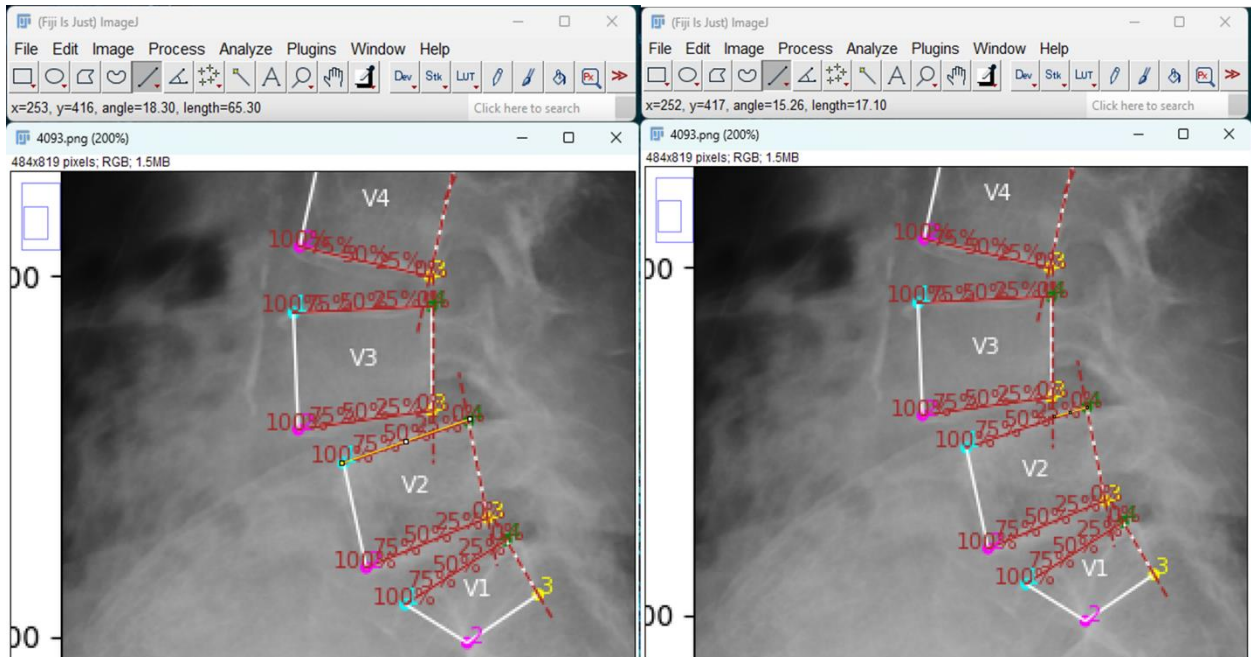
Fuente: Elaboración Propia

Anexo 30 Medición con ImageJ (3870 L5-S1)



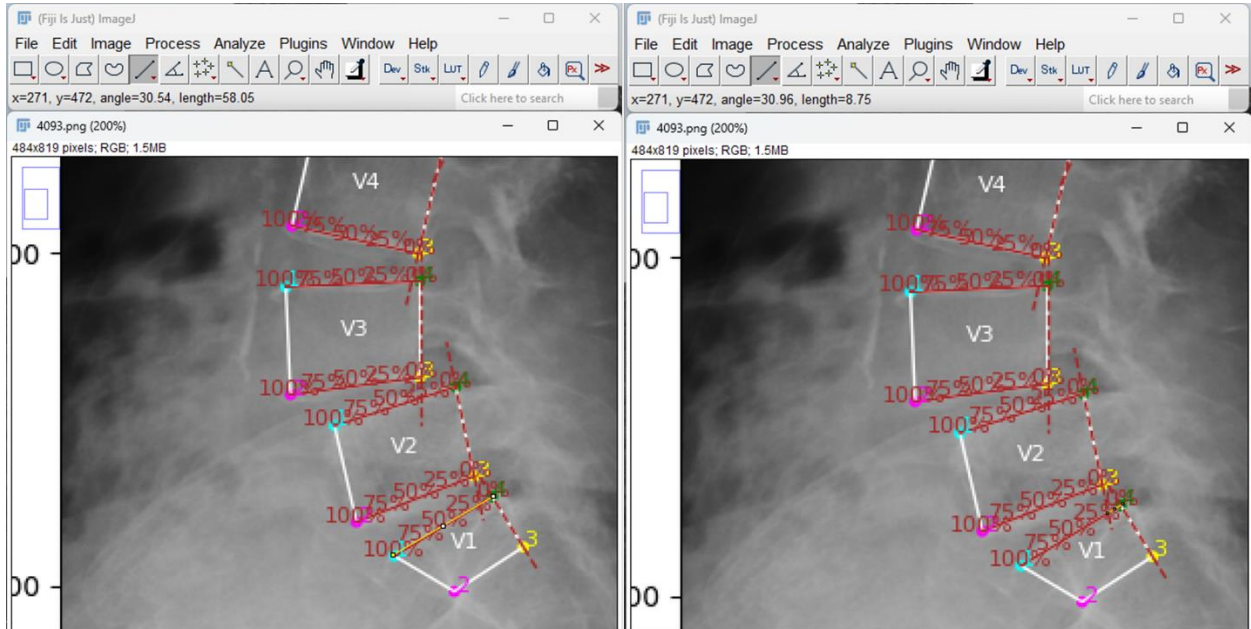
Fuente: Elaboración Propia

Anexo 31 Medición con ImageJ (4093 L4-L5)



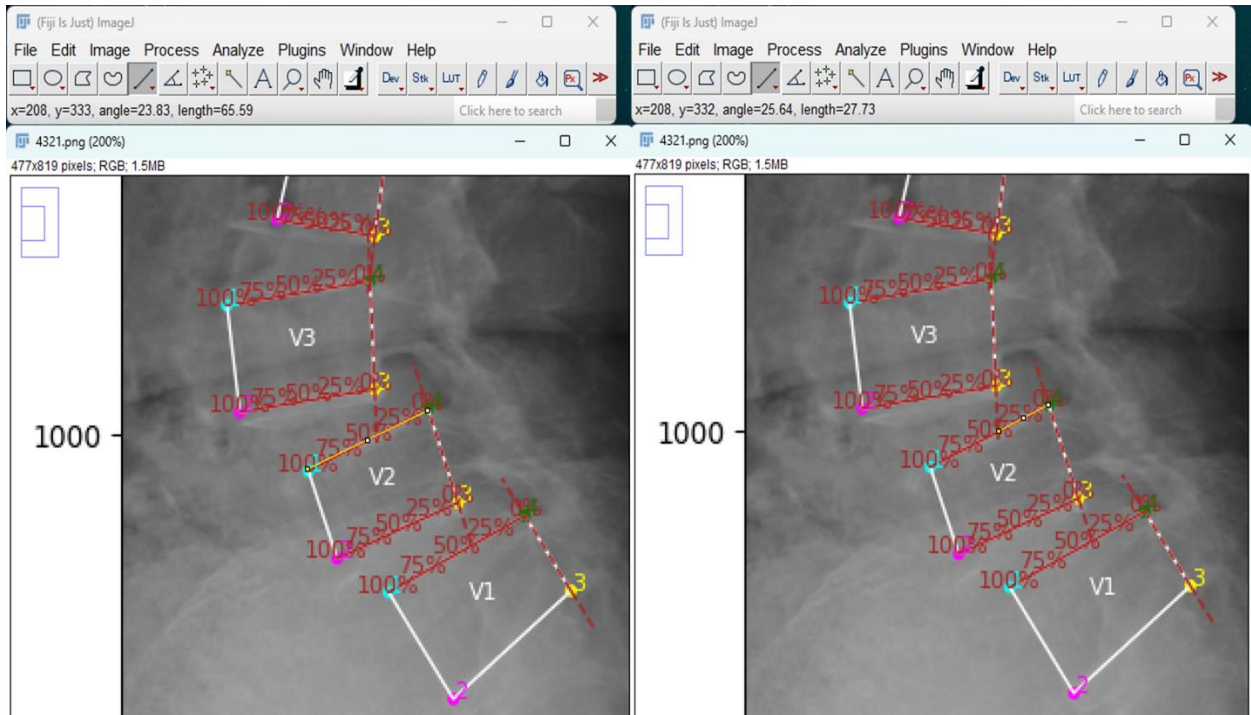
Fuente: Elaboración Propia

Anexo 32 Medición con ImageJ (4093 L5-S1)



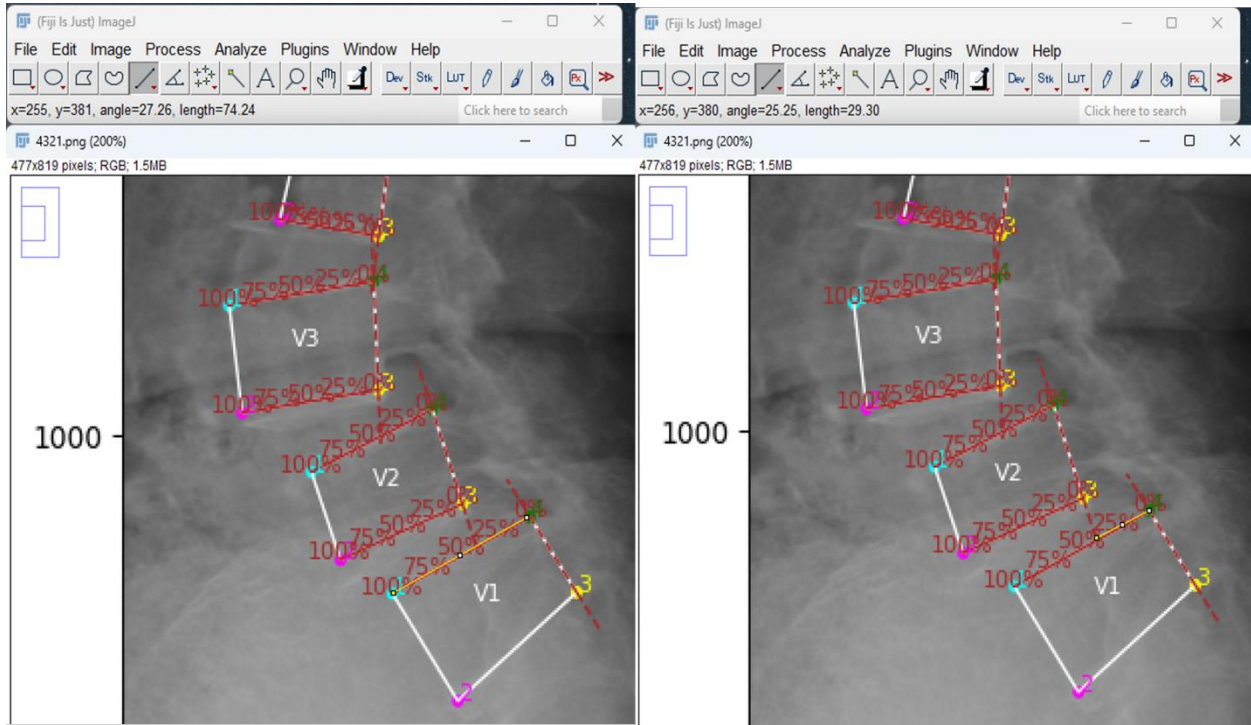
Fuente: Elaboración Propia

Anexo 33 Medición con ImageJ (4321 L4-L5)



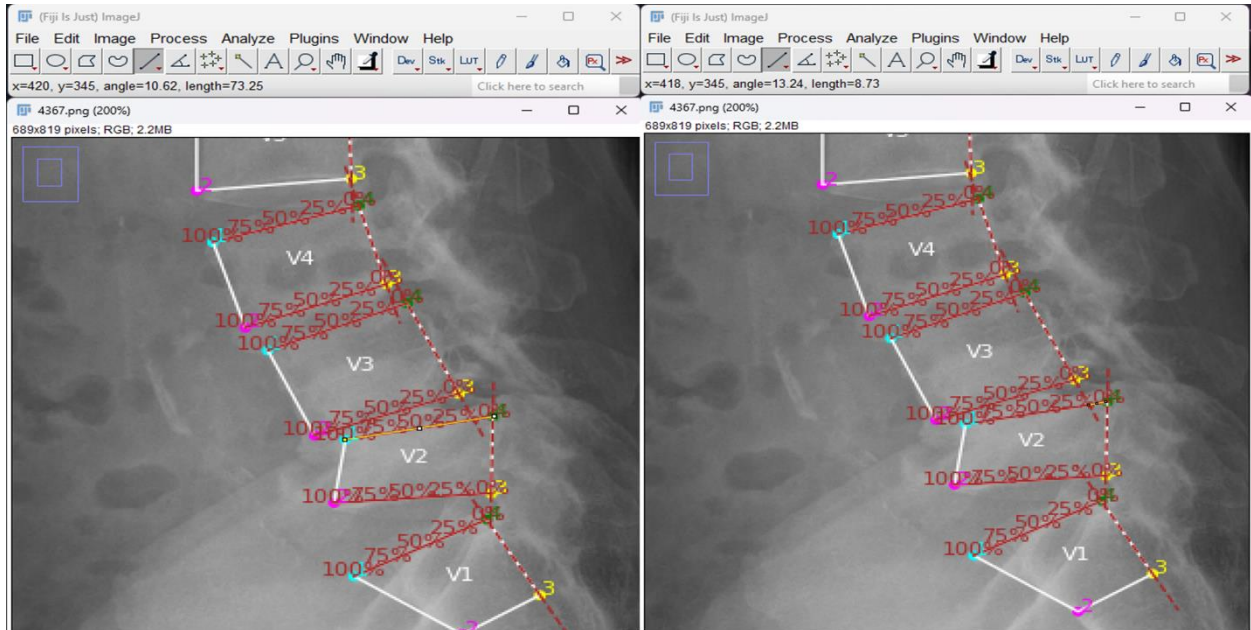
Fuente: Elaboración Propia

Anexo 34 Medición con ImageJ (4321 L5-S1)



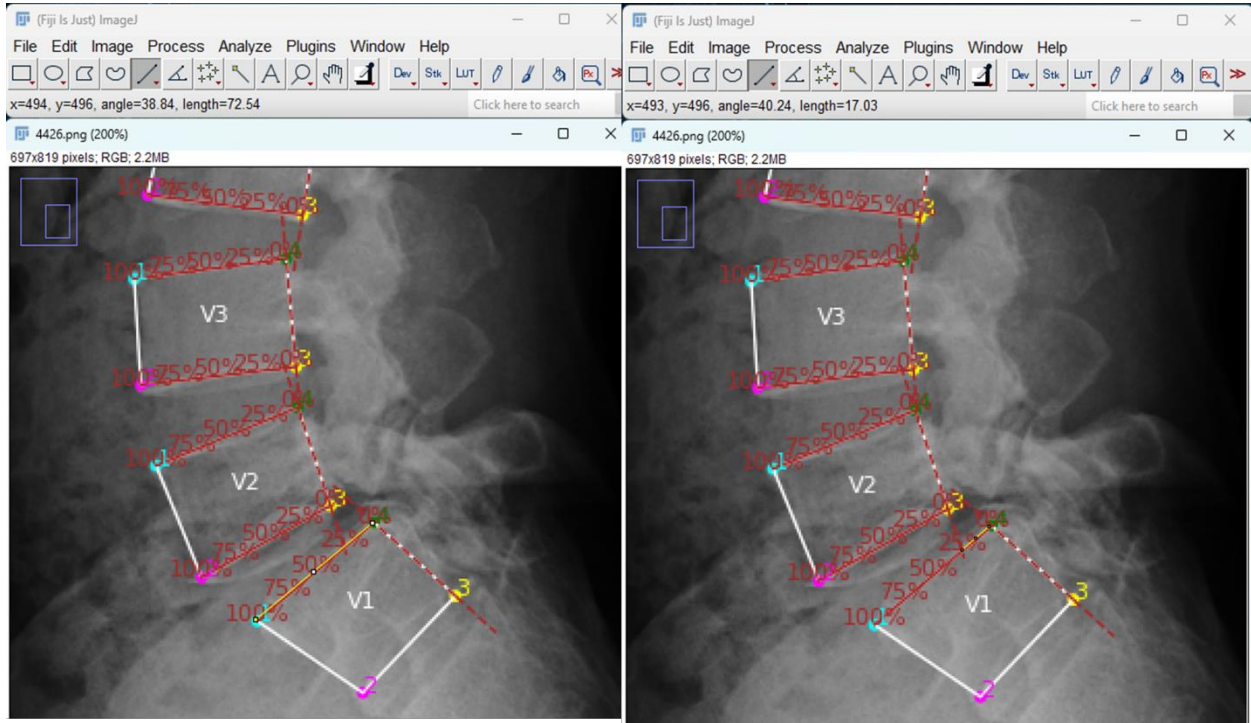
Fuente: Elaboración Propia

Anexo 35 Medición con ImageJ (4367 L4-L5)



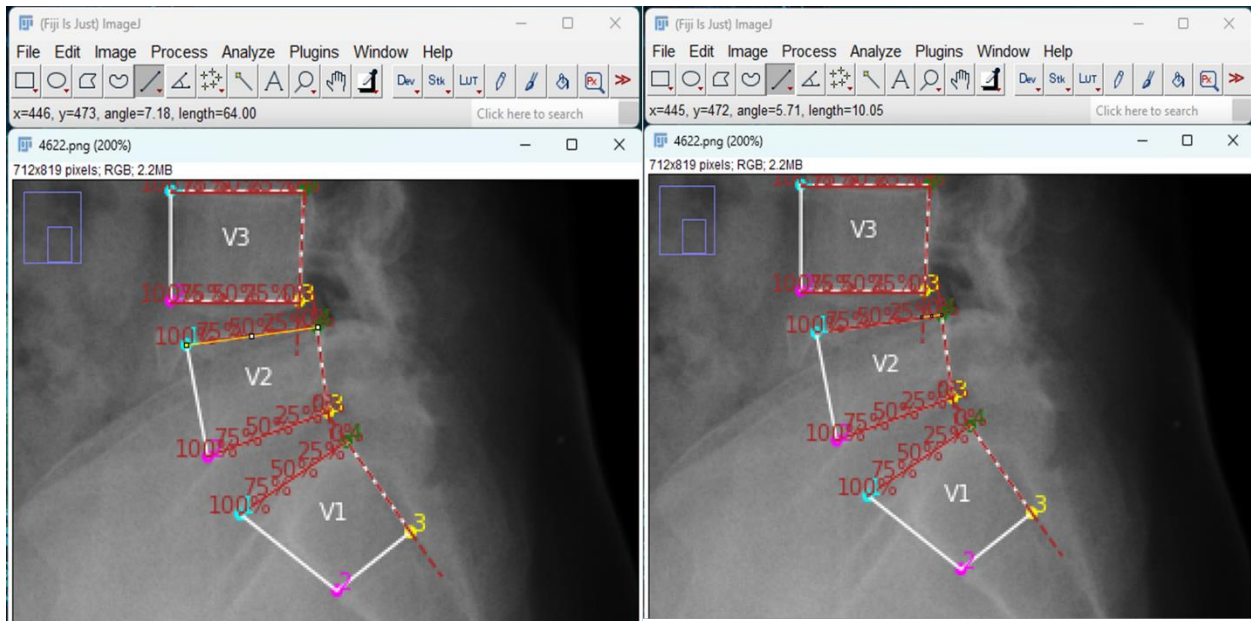
Fuente: Elaboración Propia

Anexo 36 Medición con ImageJ (4426 L5-S1)



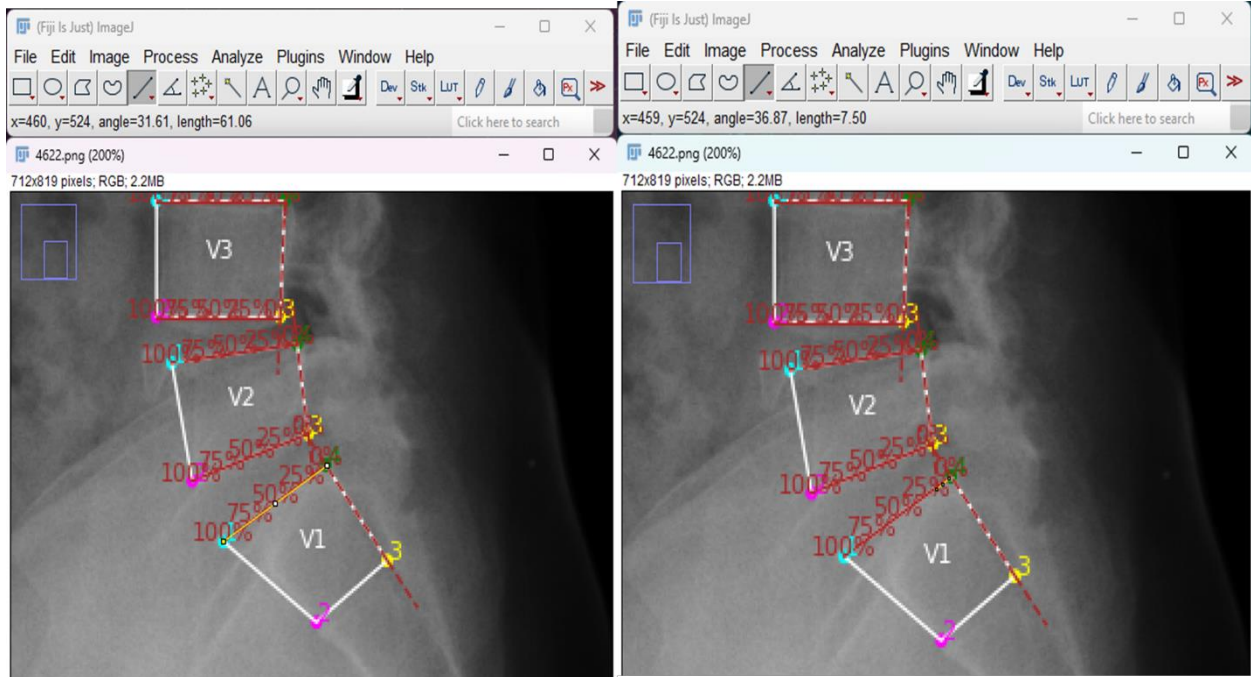
Fuente: Elaboración Propia

Anexo 37 Medición con ImageJ (4622 L4-L5)



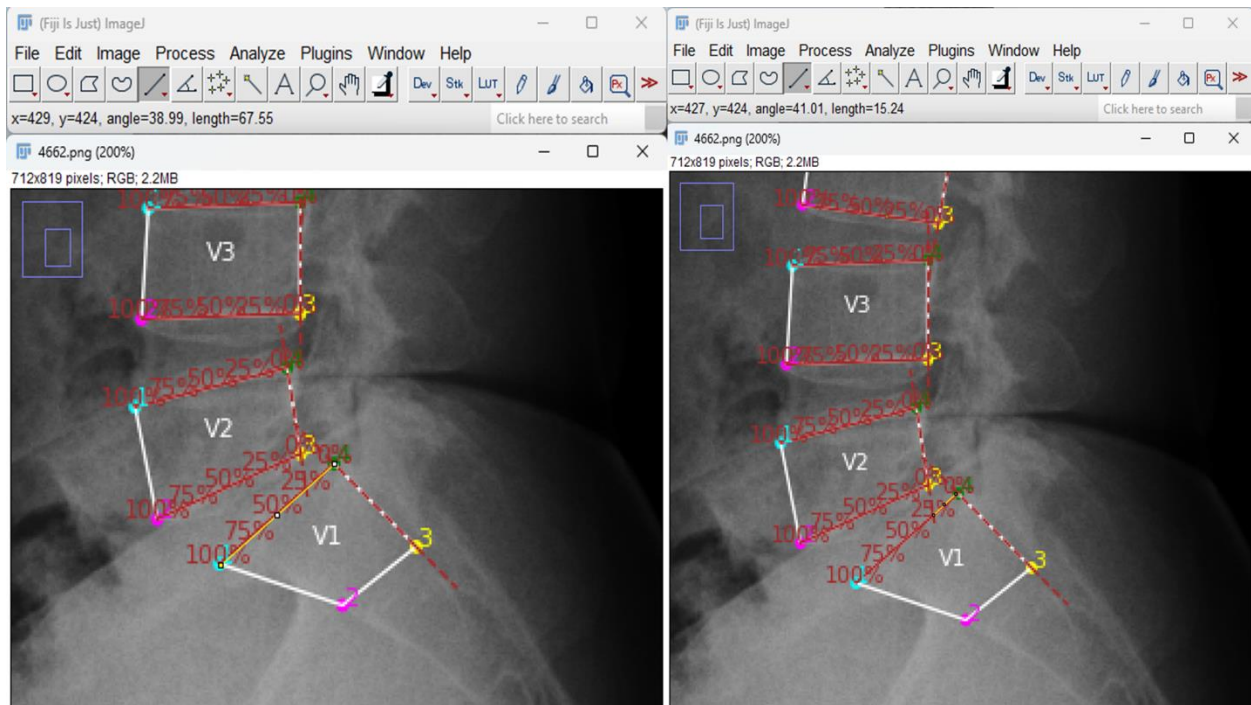
Fuente: Elaboración Propia

Anexo 38 Medición con ImageJ (4622 L5-S1)



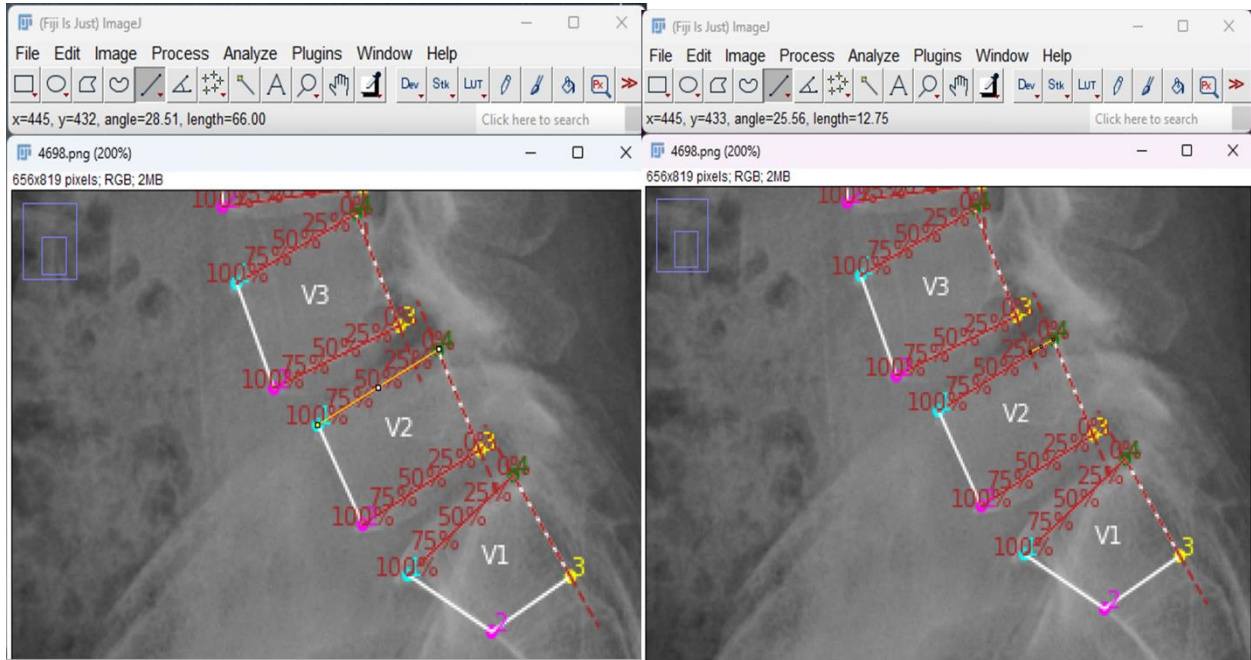
Fuente: Elaboración Propia

Anexo 39 Medición con ImageJ (4662 L5-S1)



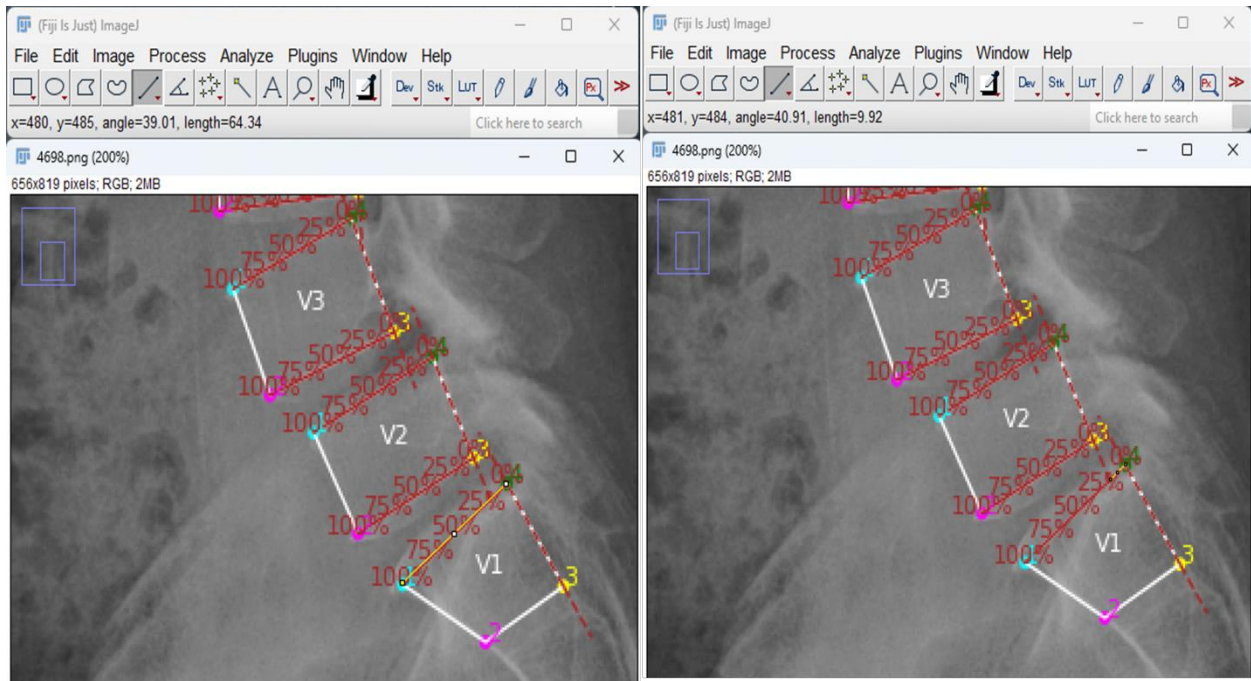
Fuente: Elaboración Propia

Anexo 40 Medición con ImageJ (4698 L4-L5)



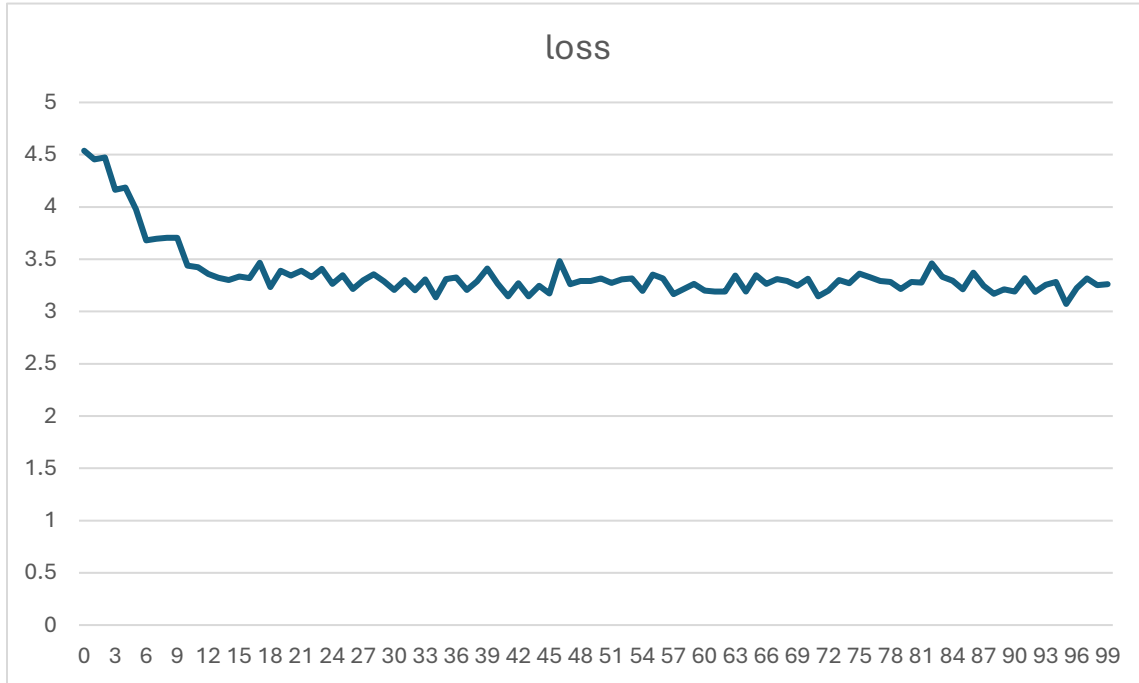
Fuente: Elaboración Propia

Anexo 41 Medición con ImageJ (4698 L5-S1)



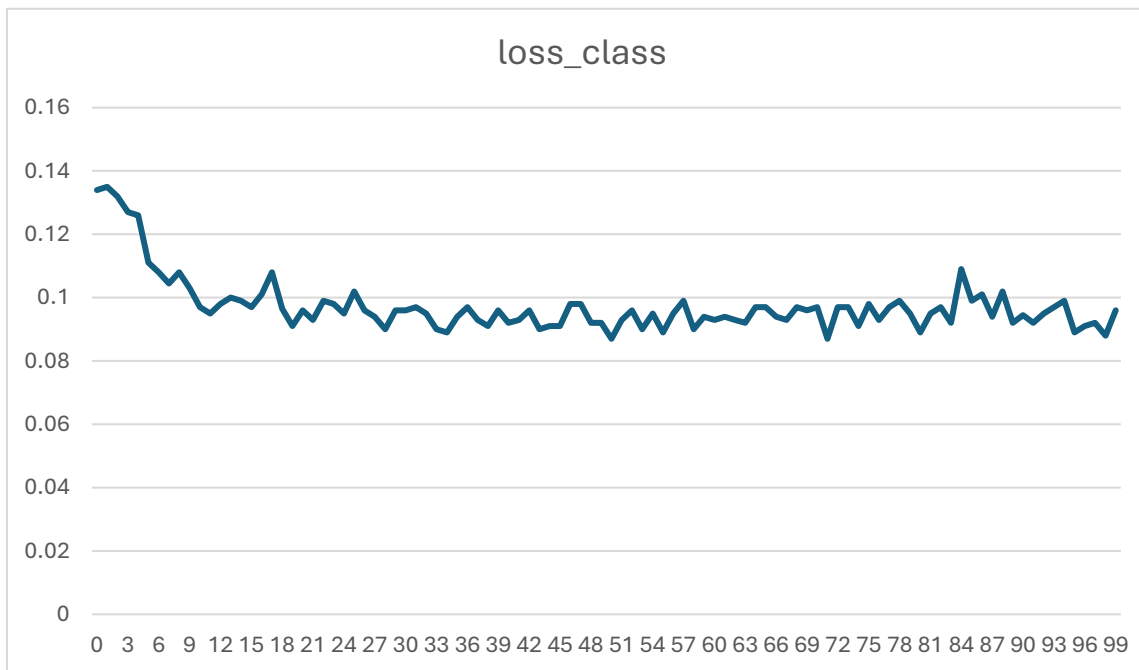
Fuente: Elaboración propia

Anexo 42 Métrica de pérdida del modelo con 100 épocas



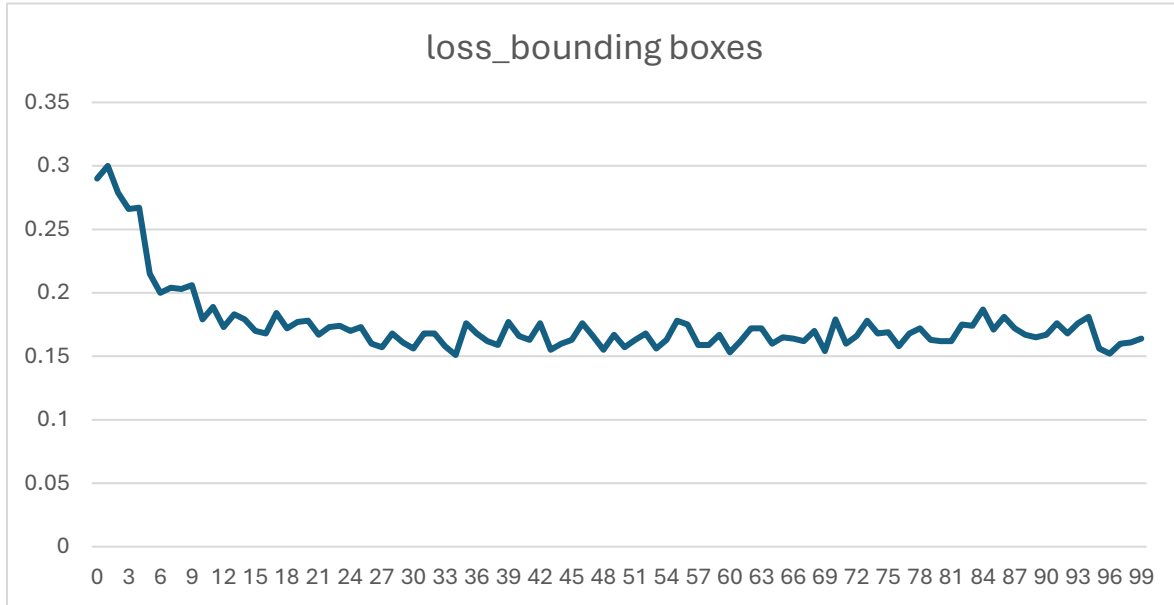
Fuente: Elaboración Propia

Anexo 43 Métrica de pérdida de clasificación del modelo con 100 épocas



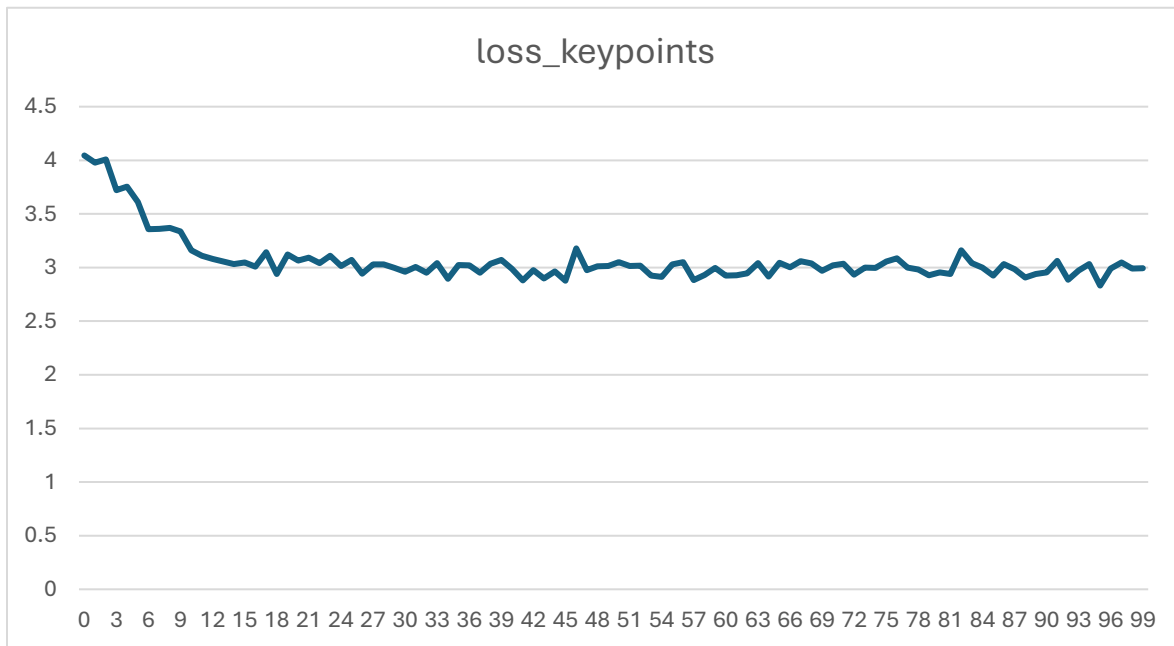
Fuente: Elaboración propia

Anexo 44 Métrica de pérdida de bounding boxes del modelo con 100 épocas



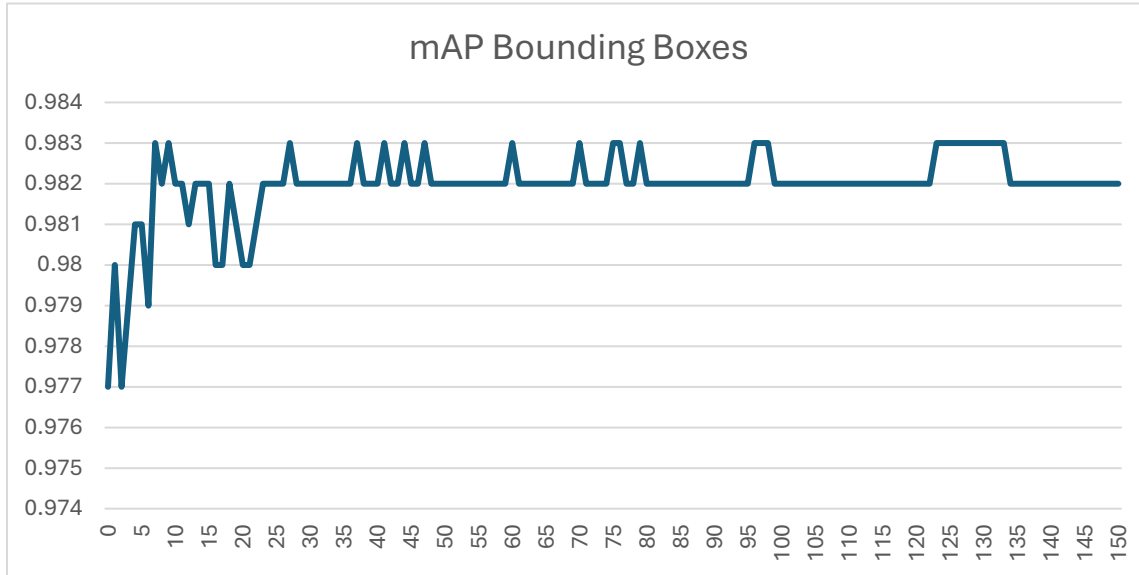
Fuente: Elaboración propia

Anexo 45 Métrica de pérdida de keypoints del modelo con 100 épocas



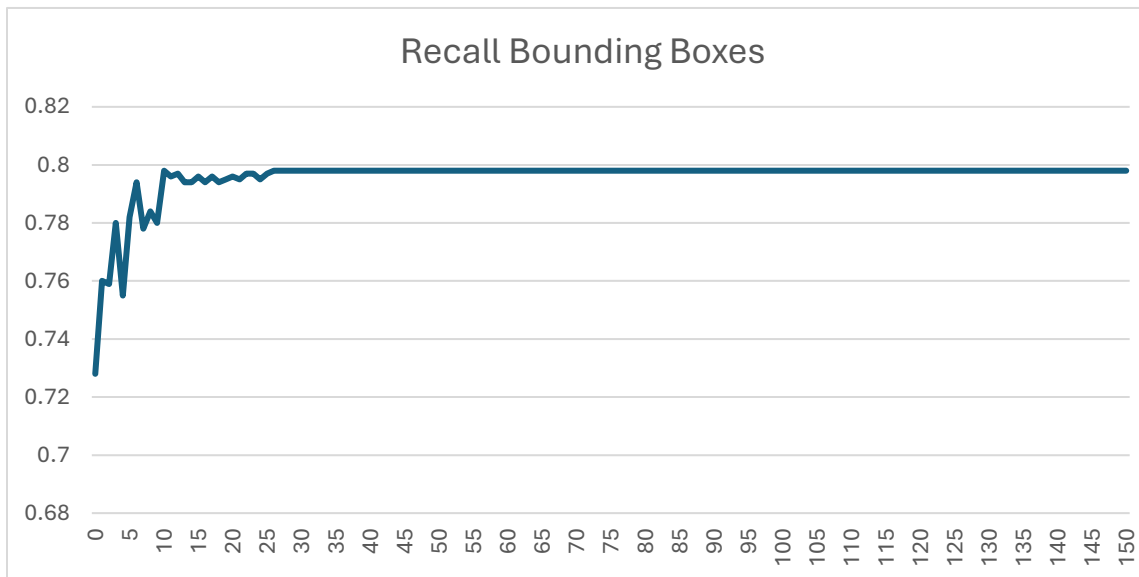
Fuente: Elaboración Propia

Anexo 46 Métrica de precisión promedio del modelo con 150 épocas en bounding boxes



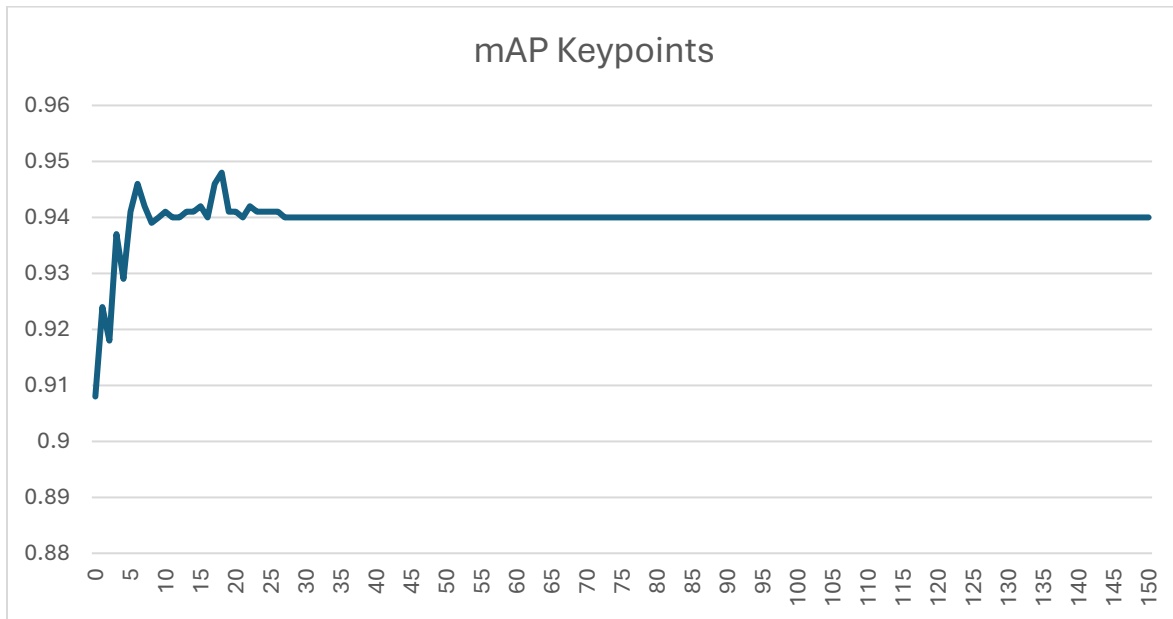
Fuente: Elaboración propia

Anexo 47 Métrica de sensibilidad del modelo con 150 épocas en bounding boxes



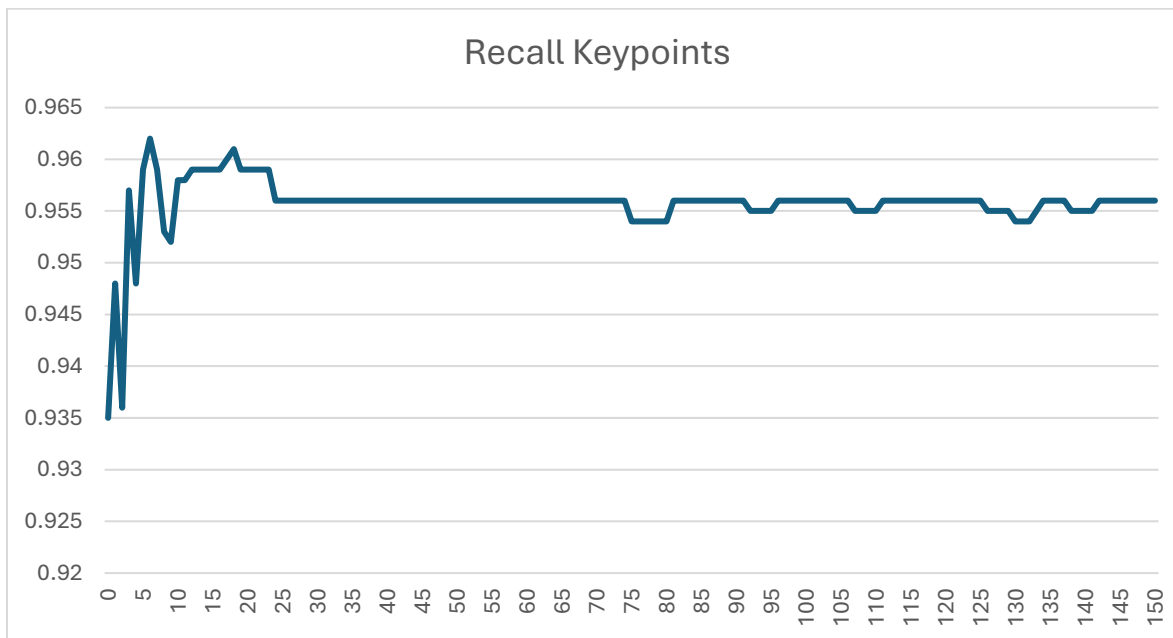
Fuente: Elaboración propia

Anexo 48 Métrica de precisión promedio del modelo con 150 épocas en keypoints



Fuente: Elaboración Propia

Anexo 49 Métrica de sensibilidad del modelo con 150 épocas en Keypoints



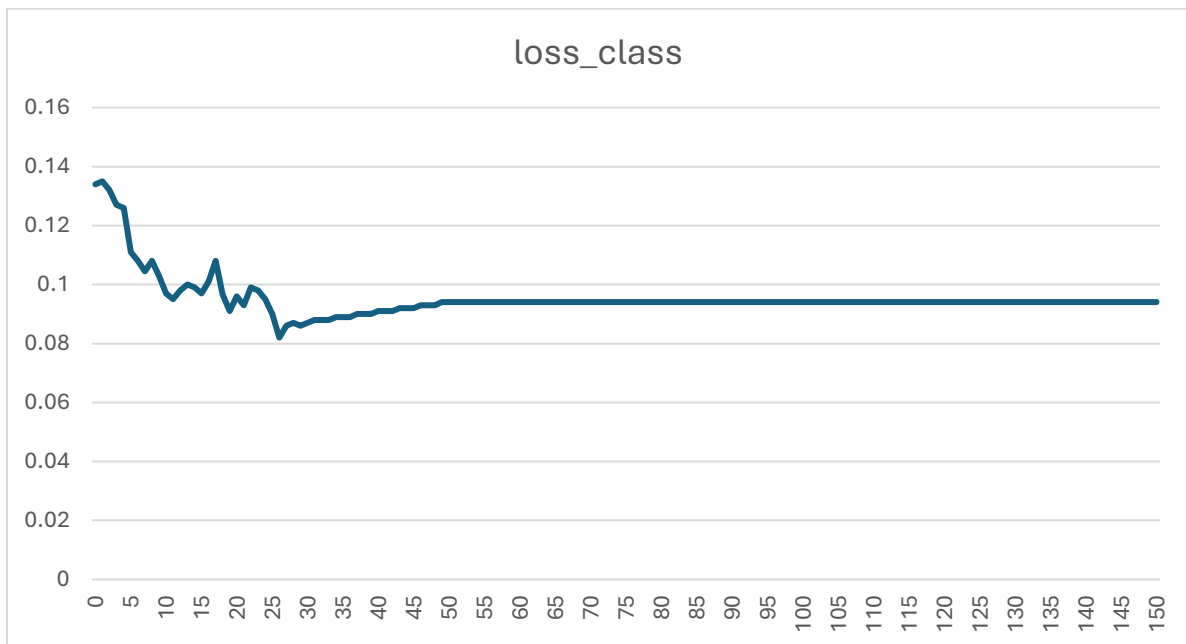
Fuente Elaboración propia

Anexo 50 Métrica de pérdida del modelo con 150 épocas



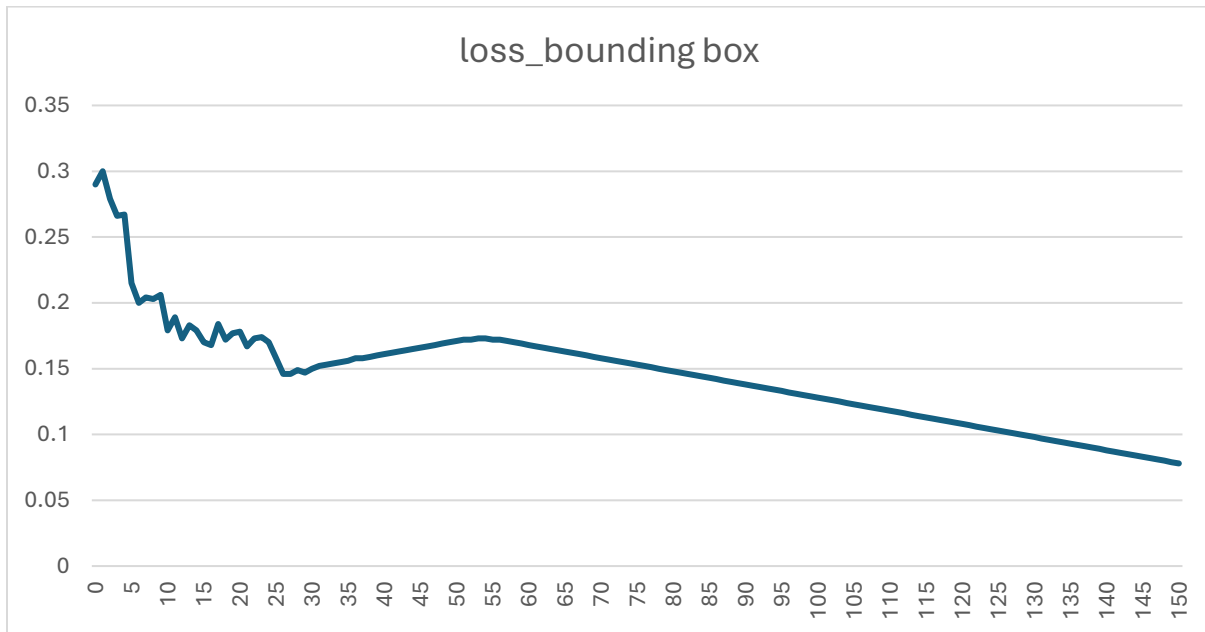
Fuente: Elaboración Propia

Anexo 51 Métrica de pérdida de clasificación del modelo con 150 épocas



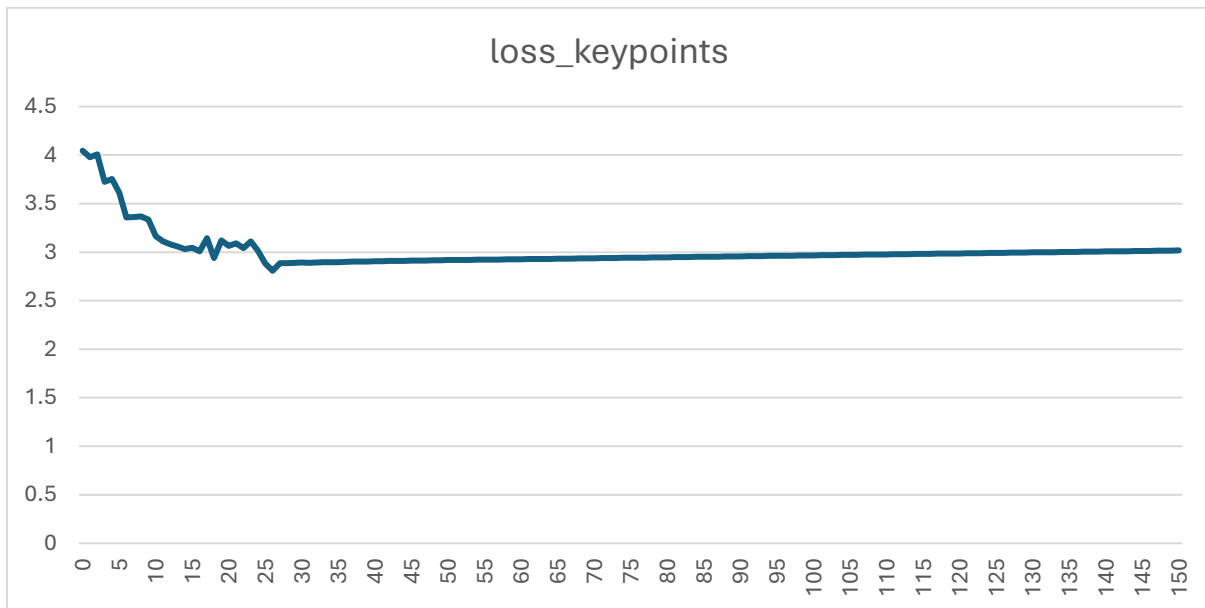
Fuente: Elaboración propia

Anexo 52 Métrica de pérdida de bounding boxes del modelo con 150 épocas



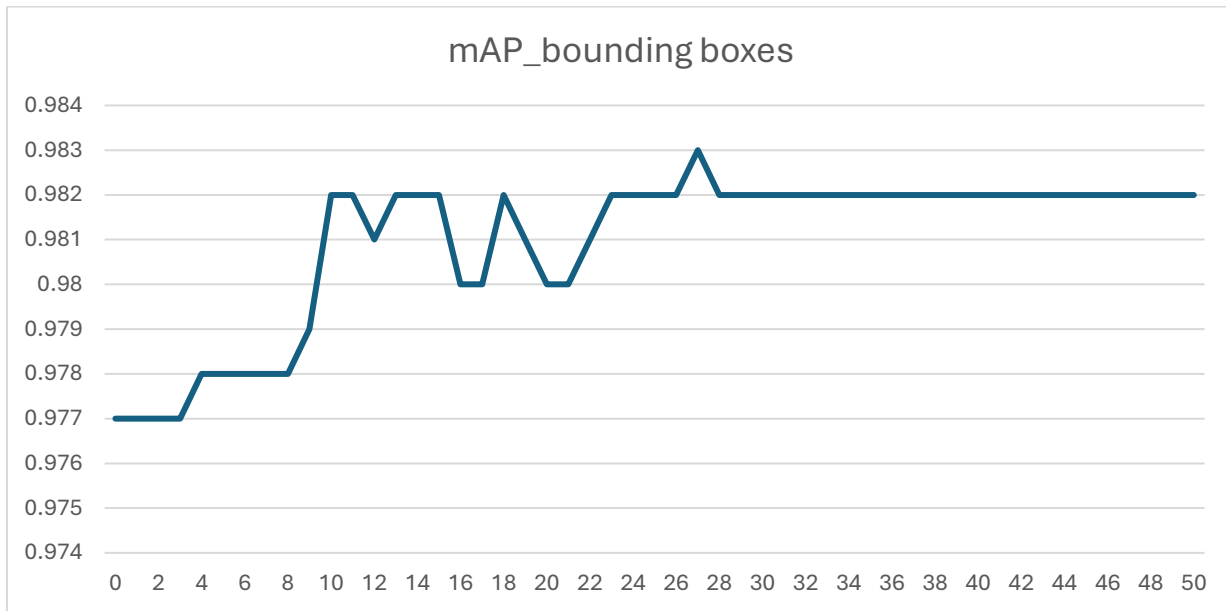
Fuente: Elaboración propia

Anexo 53 Métrica de pérdida de keypoints del modelo con 150 épocas



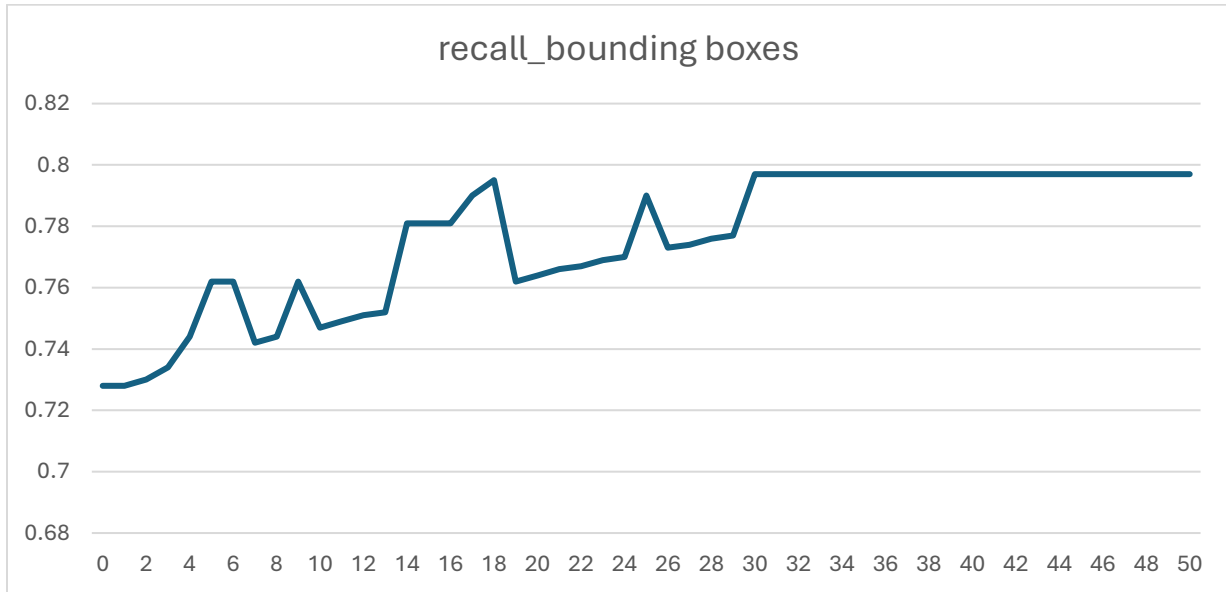
Fuente: Elaboración propia

Anexo 54 Métrica de precisión promedio del modelo con 50 épocas en bounding boxes



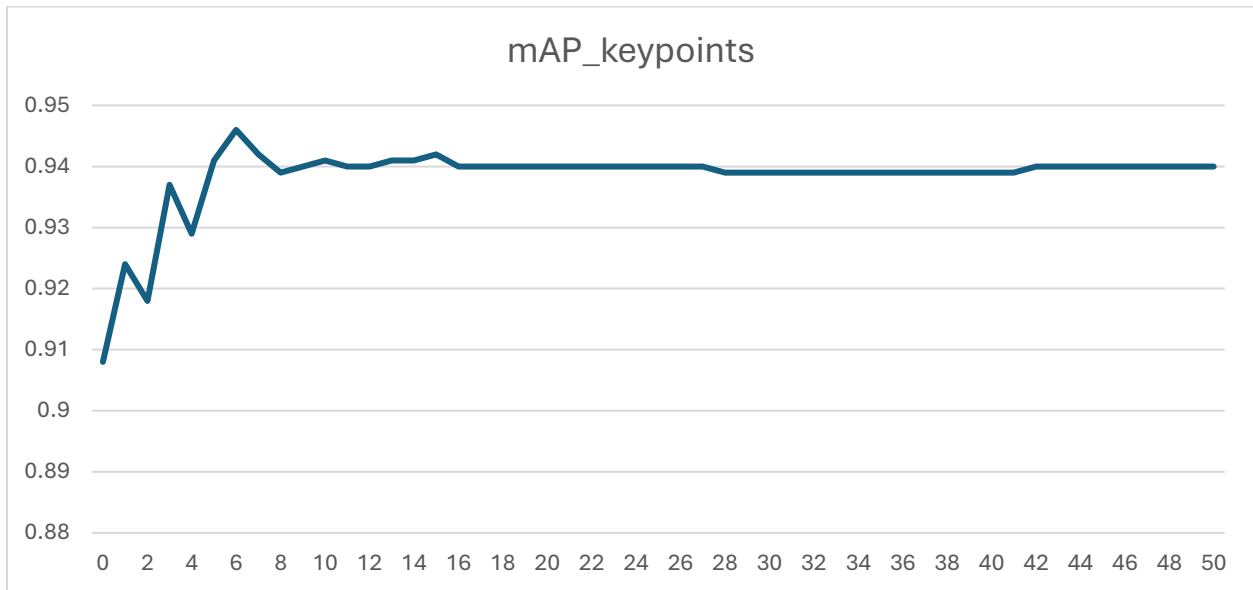
Fuente: Elaboración propia

Anexo 55 Métrica de sensibilidad del modelo con 50 épocas en bounding boxes



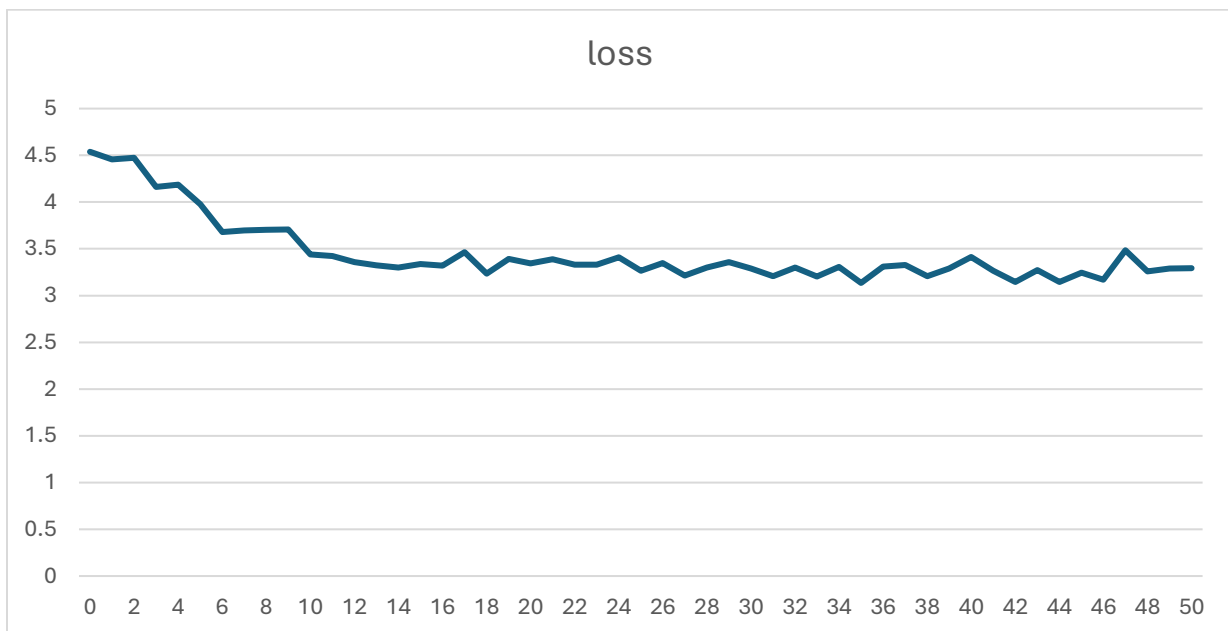
Fuente: Elaboración propia

Anexo 56 Métrica de precisión promedio del modelo con 50 épocas en keypoints



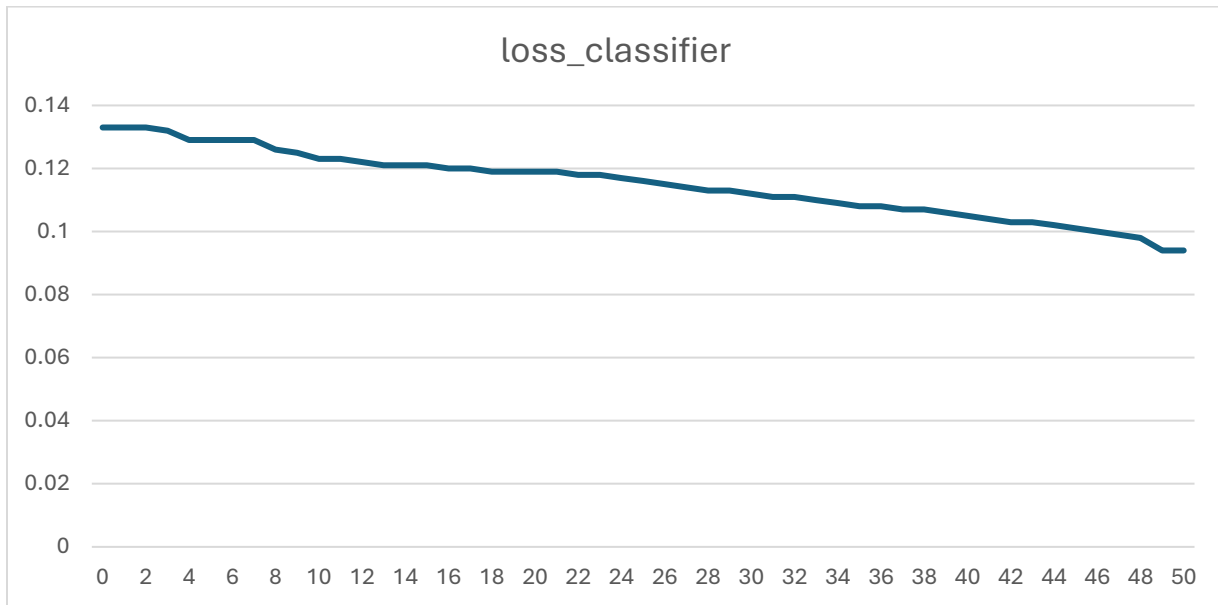
Fuente: Elaboración propia

Anexo 57 Métrica de pérdida del modelo con 50 épocas



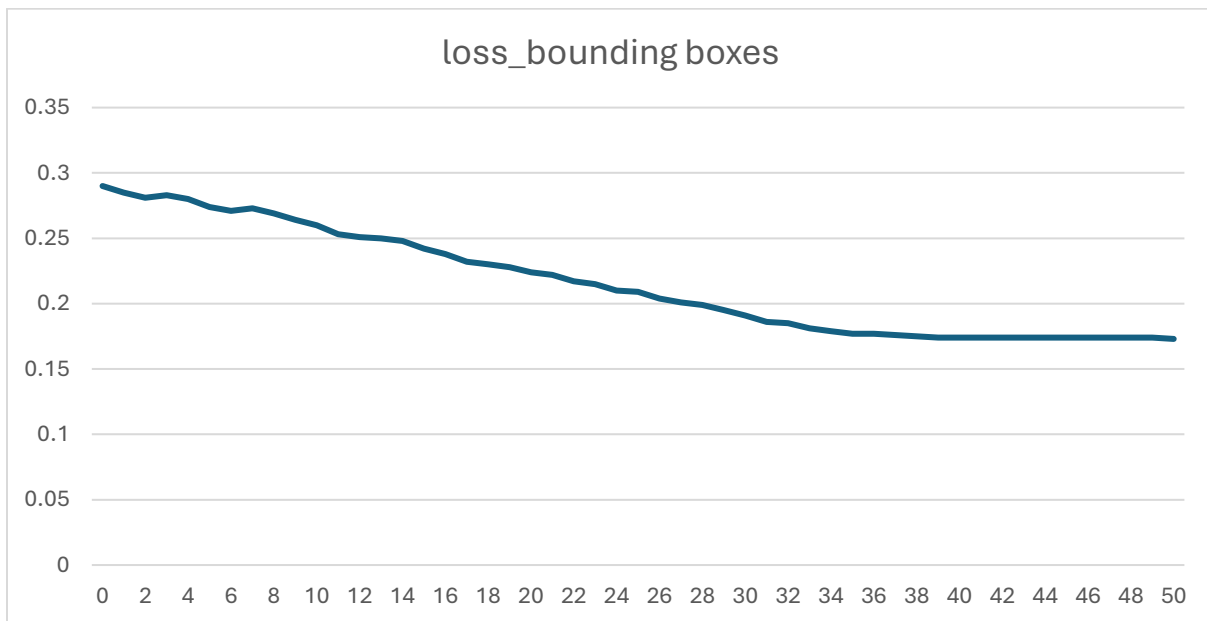
Fuente: Elaboración propia

Anexo 58 Métrica de pérdida de clasificación del modelo con 50 épocas



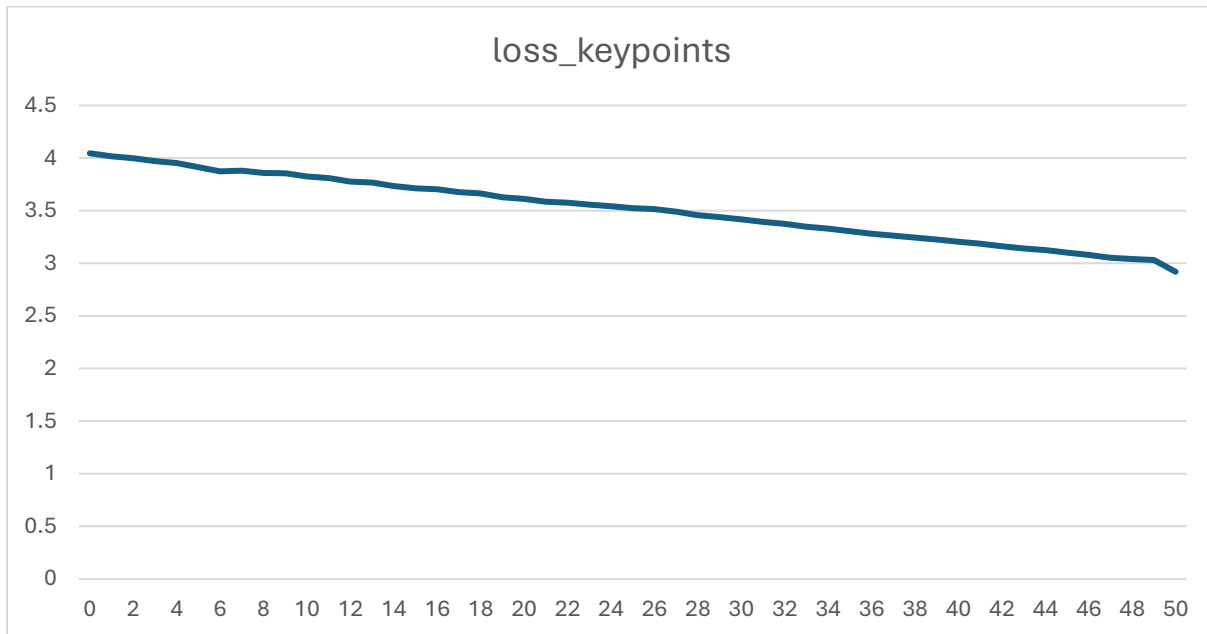
Fuente: Elaboración propia

Anexo 59 Métrica de pérdida de bounding boxes del modelo con 50 épocas



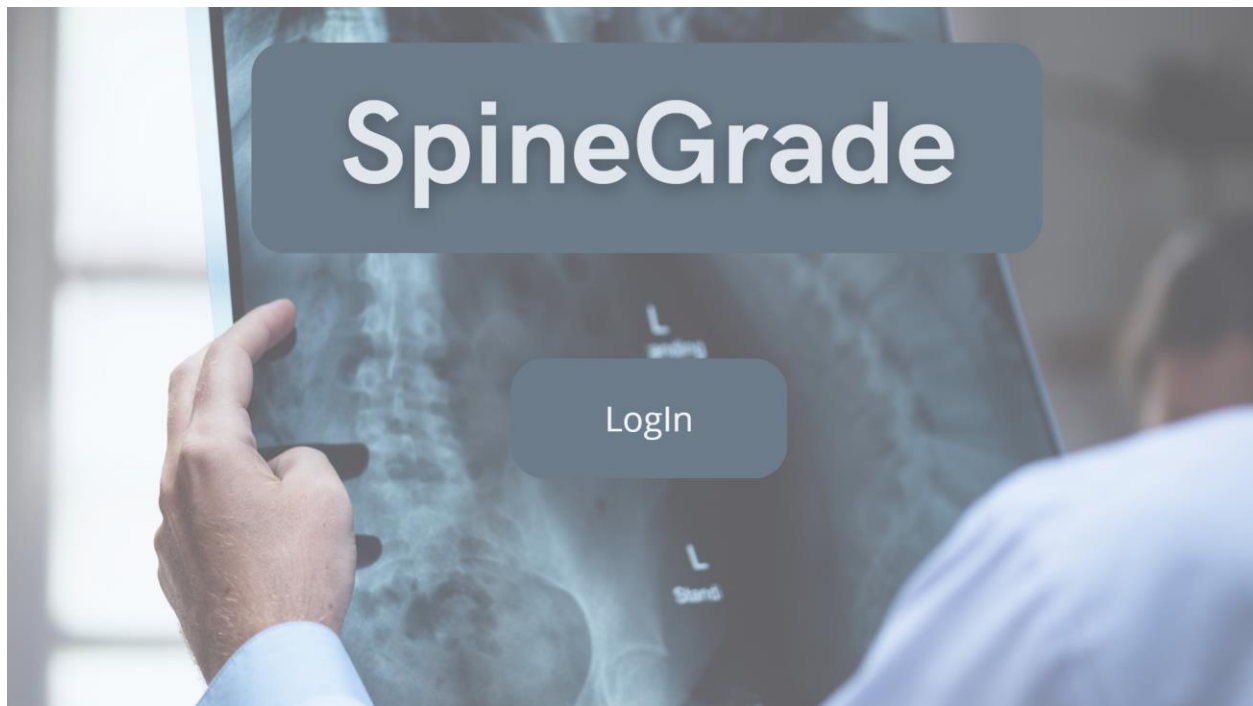
Fuente: Elaboración propia

Anexo 60 Métrica de pérdida de keypoints del modelo con 50 épocas



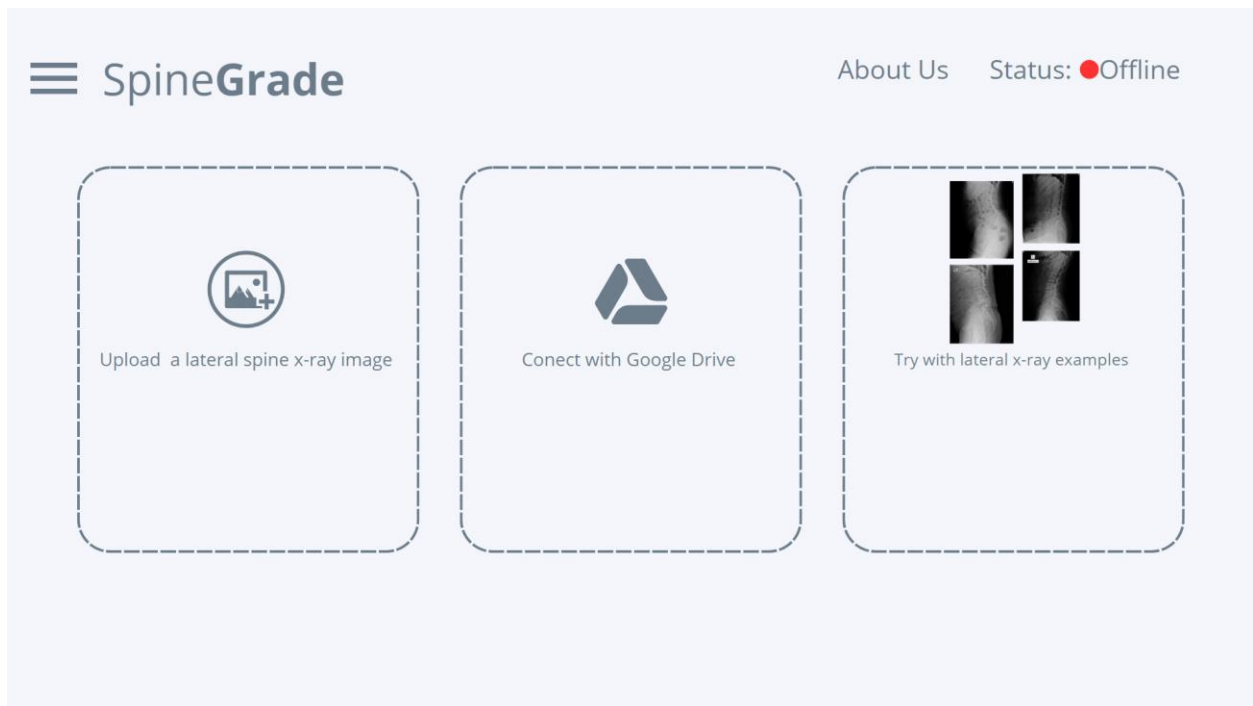
Fuente: Elaboración propia

Anexo 61 Mockup de Aplicación



Fuente: Elaboración propia

Anexo 62 Mockup de Aplicación



Fuente: Elaboración propia

Anexo 63 Mockup de Aplicación



Fuente: Elaboración propia

Anexo 64 Mockup de Aplicación



Fuente: Elaboración propia

Anexo 65 Mockup de Aplicación



Fuente: Elaboración propia

Anexo 66 Mockup de Aplicación

SpineGrade

About Us Status: ●Offline

Patient: Age:



Intersection for V1: (1490.6, 1144.76)
Between V1 and V2: Displacement = 19.68 pixels, Percentage of translation = 11.44%, Classification = Grade I
Intersection for V2: (1377.42, 1012.1)
Between V2 and V3: Displacement = 28.37 pixels, Percentage of translation = 15.92%, Classification = Grade I
Intersection for V3: (1244.33, 845.4)
Between V3 and V4: Displacement = 14.85 pixels, Percentage of translation = 7.24%, Classification = Normal
Intersection for V4: (1199.95, 663.0)
Between V4 and V5: Displacement = 3.34 pixels, Percentage of translation = 1.56%, Classification = Normal

Expand Image Save Image

Fuente: Elaboración propia

Anexo 67 Mockup de Aplicación


SpineGrade

About Us Status: ●Offline


About

SpineGrade

At SpineGrade, we offer a state-of-the-art app for precise measurement of spondylolisthesis translation percentages. Our app uses advanced vertebra detection models to annotate key points and calculate the horizontal displacement of vertebrae. By measuring the displacement relative to the adjacent vertebrae, we determine the degree of spondylolisthesis according to the Meyerding classification. SpineGrade provides healthcare professionals with an accurate and efficient tool for diagnosing and managing spondylolisthesis.



Authors of SpineGrade:



Diego Espinal & Carlos Mejia

Fuente: Elaboración propia

X. REFERENCIAS

- Balaniuk, R., Isupova, O., & Reece, S. (2020). Mining and tailings dam detection in satellite imagery using deep learning. *Sensors*, 20(23), 6936. <https://doi.org/10.3390/s20236936>
- 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>
- Albakkosh, I. (2024). Using Fleiss' Kappa Coefficient to Measure the Intra and Inter-Rater Reliability of Three AI Software Programs in the Assessment of EFL Learners' Story Writing. *International Journal of Educational Sciences and Arts*, 3(1), 69-96. <https://doi.org/10.59992/IJESA.2023.v3n1p4>
- Balzer, I., Mühlemann, M., Jokeit, M., Rawal, I. S., Snedeker, J. G., Farshad, M., & Widmer, J. (2022). A deep learning pipeline for automatized assessment of spinal MRI. *Computer Methods and Programs in Biomedicine Update*, 2, 100081. <https://doi.org/10.1016/j.cmpbup.2022.100081>
- Cai, Y., Leung, S., Warrington, J., Pandey, S., Shmuilovich, O., & Li, S. (2017). *Direct spondylolisthesis identification and measurement in MR/CT using detectors trained by articulated parameterized spine model* (M. A. Styner & E. D. Angelini, Eds.; p. 1013319). <https://doi.org/10.1117/12.2254072>

- Dlshad Ahmed, K., & Hawezi, R. (2023). Detection of bone fracture based on machine learning techniques. *Measurement: Sensors*, 27, 100723. <https://doi.org/10.1016/j.measen.2023.100723>
- 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. <https://doi.org/10.1371/journal.pone.0267851>
- Galbusera, F., Niemeyer, F., Wilke, H.-J., Bassani, T., Casaroli, G., Anania, C., Costa, F., Brayda-Bruno, M., & Sconfienza, L. M. (2019). Fully automated radiological analysis of spinal disorders and deformities: A deep learning approach. *European Spine Journal*, 28(5), 951-960. <https://doi.org/10.1007/s00586-019-05944-z>
- Gallagher, B., Moatz, B., & Tortolani, P. J. (2020). Classifications in Spondylolisthesis. *Seminars in Spine Surgery*, 32(3), 100802. <https://doi.org/10.1016/j.semss.2020.100802>
- Gordillo, J. J. T., & Rodríguez, V. H. P. (2009). CÁLCULO DE LA FIABILIDAD Y CONCORDANCIA ENTRE CODIFICADORES DE UN SISTEMA DE CATEGORÍAS PARA EL ESTUDIO DEL FORO ONLINE EN E-LEARNING. 27.
- Gorelik, N., & Gyftopoulos, S. (2021). Applications of Artificial Intelligence in Musculoskeletal Imaging: From the Request to the Report. *Canadian Association of Radiologists Journal*, 72(1), 45-59. <https://doi.org/10.1177/0846537120947148>
- Hoque, S., Arafat, Md. Y., Xu, S., Maiti, A., & Wei, Y. (2021). A Comprehensive Review on 3D Object Detection and 6D Pose Estimation With Deep Learning. *IEEE Access*, 9, 143746-143770. <https://doi.org/10.1109/ACCESS.2021.3114399>

- Horng, M.-H., Kuok, C.-P., Fu, M.-J., Lin, C.-J., & Sun, Y.-N. (2019). Cobb Angle Measurement of Spine from X-Ray Images Using Convolutional Neural Network. *Computational and Mathematical Methods in Medicine*, 2019, 1-18. <https://doi.org/10.1155/2019/6357171>
- Jaehan Koh, Chaudhary, V., & Dhillon, G. (2012). Automated boundary extraction of the spinal canal in mri based on dynamic programming. *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6559-6562. <https://doi.org/10.1109/EMBC.2012.6347497>
- Jamaludin, A., Kadir, T., & Zisserman, A. (2017). SpineNet: Automated classification and evidence visualization in spinal MRIs. *Medical Image Analysis*, 41, 63-73. <https://doi.org/10.1016/j.media.2017.07.002>
- Kim, K. C., Yun, H. S., Kim, S., & Seo, J. K. (2020). Automation of Spine Curve Assessment in Frontal Radiographs Using Deep Learning of Vertebral-Tilt Vector. *IEEE Access*, 8, 84618-84630. <https://doi.org/10.1109/ACCESS.2020.2992081>
- Koller, H., Ames, C., Mehdiian, H., Bartels, R., Ferch, R., Deriven, V., Toyone, H., Shaffrey, C., Smith, J., Hitzl, W., Schröder, J., & Robinson, Y. (2019). Characteristics of deformity surgery in patients with severe and rigid cervical kyphosis (CK): Results of the CSRS-Europe multi-centre study project. *European Spine Journal*, 28(2), 324-344. <https://doi.org/10.1007/s00586-018-5835-2>
- Koslosky, E., & Gendelberg, D. (2020). Classification in Brief: The Meyerding Classification System of Spondylolisthesis. *Clinical Orthopaedics & Related Research*, 478(5), 1125-1130. <https://doi.org/10.1097/CORR.0000000000001153>

- Liao, S., Zhan, Y., Dong, Z., Yan, R., Gong, L., Zhou, X. S., Salganicoff, M., & Fei, J. (2016). Automatic Lumbar Spondylolisthesis Measurement in CT Images. *IEEE Transactions on Medical Imaging*, 35(7), 1658-1669. <https://doi.org/10.1109/TMI.2016.2523452>
- Mohile, N. V., Kuczumarski, A. S., Lee, D., Warburton, C., Rakoczy, K., & Butler, A. J. (2022). Spondylolysis and Isthmic Spondylolisthesis: A Guide to Diagnosis and Management. *The Journal of the American Board of Family Medicine*, 35(6), 1204-1216. <https://doi.org/10.3122/jabfm.2022.220130R1>
- Moon, G., Kim, S., Kim, W., Kim, Y., Jeong, Y., & Choi, H.-S. (2022). Computer Aided Facial Bone Fracture Diagnosis (CA-FBFD) System Based on Object Detection Model. *IEEE Access*, 10, 79061-79070. <https://doi.org/10.1109/ACCESS.2022.3192389>
- 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. <https://doi.org/10.1016/j.bspc.2020.102371>
- Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsiftaris, S. A., Young, A., Foran, D., Do, N., Golemati, S., Kurc, T., Huang, K., Nikita, K. S., Veasey, B. P., Zervakis, M., Saltz, J. H., & Pattichis, C. S. (2020). AI in Medical Imaging Informatics: Current Challenges and Future Directions. *IEEE Journal of Biomedical and Health Informatics*, 24(7), 1837-1857. <https://doi.org/10.1109/JBHI.2020.2991043>

- Qi, Y., Zhao, J., Shi, Y., Zuo, G., Zhang, H., Long, Y., Wang, F., & Wang, W. (2020). Ground Truth Annotated Femoral X-Ray Image Dataset and Object Detection Based Method for Fracture Types Classification. *IEEE Access*, 8, 189436-189444. <https://doi.org/10.1109/ACCESS.2020.3029039>
- Rajpurkar, P., Irvin, J., Bagul, A., Ding, D., Duan, T., Mehta, H., Yang, B., Zhu, K., Laird, D., Ball, R. L., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2018). *MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs* (arXiv:1712.06957). arXiv. <http://arxiv.org/abs/1712.06957>
- Raraz-Vidal, J., & Raraz-Vidal, O. (2022). Aplicaciones de la inteligencia artificial en la medicina. *Revista Peruana de Investigación en Salud*, 6(3), 131-133. <https://doi.org/10.35839/repis.6.3.1559>
- Schindelin, J., Rueden, C. T., Hiner, M. C., & Eliceiri, K. W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular Reproduction and Development*, 82(7-8), 518-529. <https://doi.org/10.1002/mrd.22489>
- Singh, G., Anand, D., Cho, W., Joshi, G. P., & Son, K. C. (2022). Hybrid Deep Learning Approach for Automatic Detection in Musculoskeletal Radiographs. *Biology*, 11(5), 665. <https://doi.org/10.3390/biology11050665>
- Sun, Y., Xing, Y., Zhao, Z., Meng, X., Xu, G., & Hai, Y. (2022). Comparison of manual versus automated measurement of Cobb angle in idiopathic scoliosis based on a deep learning keypoint detection technology. *European Spine Journal*, 31(8), 1969-1978. <https://doi.org/10.1007/s00586-021-07025-6>

- Trinh, G. M., Shao, H.-C., Hsieh, K. L.-C., Lee, C.-Y., Liu, H.-W., Lai, C.-W., Chou, S.-Y., Tsai, P.-I., Chen, K.-J., Chang, F.-C., Wu, M.-H., & Huang, T.-J. (2022). Detection of Lumbar Spondylolisthesis from X-ray Images Using Deep Learning Network. *Journal of Clinical Medicine*, 11(18), 5450. <https://doi.org/10.3390/jcm11185450>
- Tu, Y., Wang, N., Tong, F., & Chen, H. (2019). Automatic measurement algorithm of scoliosis Cobb angle based on deep learning. *Journal of Physics: Conference Series*, 1187(4), 042100. <https://doi.org/10.1088/1742-6596/1187/4/042100>
- Vanti, C., Ferrari, S., Guccione, A. A., & Pillastrini, P. (2021). Lumbar spondylolisthesis: STATE of the art on assessment and conservative treatment. *Archives of Physiotherapy*, 11(1), 19. <https://doi.org/10.1186/s40945-021-00113-2>
- Varcin, F., Erbay, H., Cetin, E., Cetin, I., & Kultur, T. (2019). Diagnosis of Lumbar Spondylolisthesis via Convolutional Neural Networks. *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 1-4. <https://doi.org/10.1109/IDAP.2019.8875988>
- Varçin, F., Erbay, H., Çetin, E., Çetin, İ., & Kültür, T. (2021). End-To-End Computerized Diagnosis of Spondylolisthesis Using Only Lumbar X-rays. *Journal of Digital Imaging*, 34(1), 85-95. <https://doi.org/10.1007/s10278-020-00402-5>
- Wiltse, L. L. (1981). *CLASSIFICATION, TERMINOLOGY AND MEASUREMENTS IN SPONDYLOLISTHESIS* +.
- 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. <https://doi.org/10.3389/fpubh.2023.1044525>

- Yang, R., & Yu, Y. (2021). Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis. *Frontiers in Oncology*, *11*, 638182. <https://doi.org/10.3389/fonc.2021.638182>
- Zhang, J., Lin, H., Wang, H., Xue, M., Fang, Y., Liu, S., Huo, T., Zhou, H., Yang, J., Xie, Y., Xie, M., Cheng, L., Lu, L., Liu, P., & Ye, Z. (2023). Deep learning system assisted detection and localization of lumbar spondylolisthesis. *Frontiers in Bioengineering and Biotechnology*, *11*, 1194009. <https://doi.org/10.3389/fbioe.2023.1194009>
- Zhao, S., Wu, X., Chen, B., & Li, S. (2019). Automatic spondylolisthesis grading from MRIs across modalities using faster adversarial recognition network. *Medical Image Analysis*, *58*, 101533. <https://doi.org/10.1016/j.media.2019.101533>
- Werth, A., Oliver, K., West, C., & Lewandowski, H. (2022). Engagement in collaboration and teamwork using google colaboratory. <https://doi.org/10.1119/perc.2022.pr.werth>