
Proyecto de Identificación Automatizada de Lesiones Dentales mediante Redes Neuronales Convolucionales

Introducción:

La presente investigación tiene como objetivo principal desarrollar un sistema avanzado de diagnóstico dental mediante el empleo de Redes Neuronales Convolucionales (CNN). Estas redes, especializadas en el procesamiento de imágenes, se adaptarán para la detección de patrones asociados con caries, lesiones periapicales y dientes impactados en radiografías dentales.

Objetivos Específicos:

1. Diseñar e implementar una arquitectura de CNN adaptada específicamente para el procesamiento de radiografías dentales.
2. Entrenar la CNN utilizando un conjunto de datos cuidadosamente preparado, ajustando parámetros e hiperparámetros para optimizar el rendimiento del modelo.
3. Evaluar el modelo utilizando un conjunto de datos de prueba independiente, midiendo su precisión y rendimiento en la detección de caries, lesiones periapicales y dientes impactados.



Contextualización Clínica:



La radiografía panorámica es una herramienta fundamental en odontología para el diagnóstico de diversas afecciones dentales. Al llegar a una consulta odontológica, la obtención de una radiografía panorámica es una práctica común y esencial. Esto permite una evaluación integral de la salud bucal, identificación temprana de problemas dentales, enfermedades periodontales y planificación precisa del tratamiento.

Fase Inicial del Proyecto: Recolección y Preparación del Conjunto de Datos:

Se recopilaron radiografías dentales panorámicas de tres instituciones diferentes, utilizando condiciones clínicas estándar. El conjunto de datos, denominado DENTEX 2023, abarca pacientes de 12 años en adelante, seleccionados aleatoriamente para garantizar la privacidad. Incluye datos de cuadrante parcialmente anotados, cuadrante-enumeración parcialmente anotados y cuadrante-enumeración-diagnóstico completamente anotados. Para simplificar y generalizar el modelo, se clasificaron las imágenes en tres categorías: caries, dientes impactados y lesiones periapicales.

Diseño de Modelos Específicos:

Con el fin de abordar cada categoría de lesión de manera individual y especializada, se crearon tres modelos distintos:

1. Modelo 1: Identifica radiografías con presencia o ausencia de caries.
2. Modelo 2: Detecta radiografías con dientes impactados.
3. Modelo 3: Enfocado en la identificación de lesiones periapicales.



Fase Inicial del Proyecto: Recolección y Preparación del Conjunto de Datos:
Se recopilaron radiografías dentales panorámicas de tres instituciones diferentes, utilizando condiciones clínicas estándar.

El conjunto de datos, denominado DENTEX 2023, abarca pacientes de 12 años en adelante, seleccionados aleatoriamente para garantizar la privacidad. Incluye datos de cuadrante parcialmente anotados, cuadrante-enumeración parcialmente anotados y cuadrante-enumeración-diagnóstico completamente anotados. Para simplificar y generalizar el modelo, se clasificaron lasArquitectura de la CNN:

La arquitectura del modelo sigue un enfoque de red convolucional:

- Tres capas convolucionales con diferentes tamaños de filtros (32, 64 y 128), activadas por ReLU para la detección de características.
- Capas de MaxPooling para reducir la dimensionalidad espacial y conservar las características más importantes.
- Capa de aplanamiento (Flatten) para transformar la salida bidimensional en un vector unidimensional.
- Dos capas densas totalmente conectadas para combinar características y una capa de Dropout para mitigar el sobreajuste.
- Capa de salida con función de activación sigmoide para la clasificación binaria.

Compilación y Entrenamiento del Modelo:

Se utilizó el optimizador Adam con una tasa de aprendizaje de 0.001 y la función de pérdida 'binary_crossentropy'.

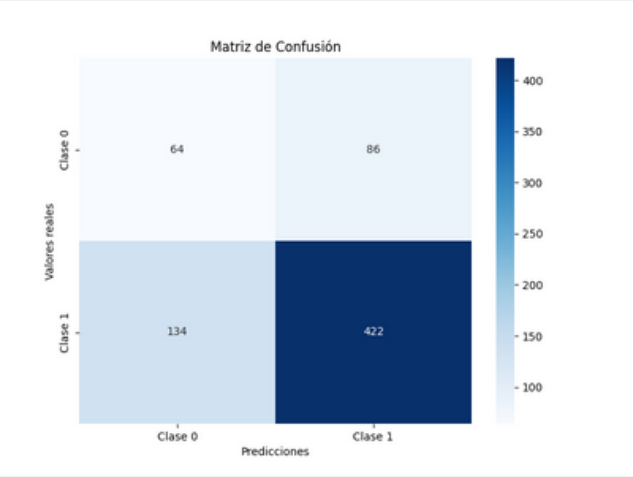
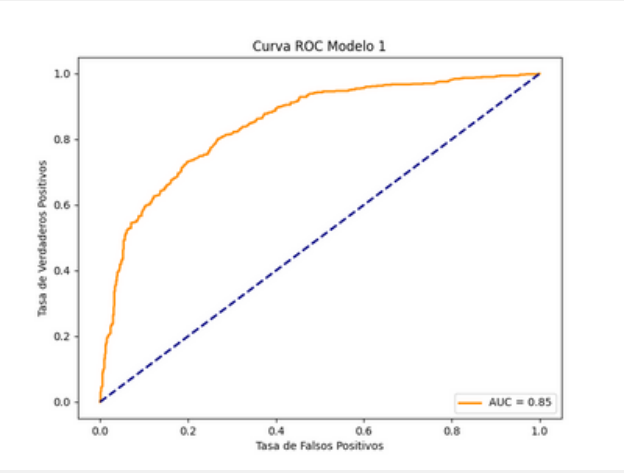
El entrenamiento incluyó EarlyStopping para evitar sobreentrenamiento y se ajustaron hiperparámetros como class_weight para abordar el desbalance de clases.

Modelo 1			
Cat	Etiqueta	Rx	Class_weight
1	Caries	2767	1:2.0
0	Sin Caries	762	0:4.0

Modelo 3			
Cat	Etiqueta	Rx	Class_weight
1	Sin Lesión Periapical	3371	1:2.0
0	Con Lesión Periapical	158	0:2.8

Modelo 2			
Cat	Etiqueta	Rx	Class_weight
1	No impactado	2925	1:2.0
0	Impactado	604	0:6.0

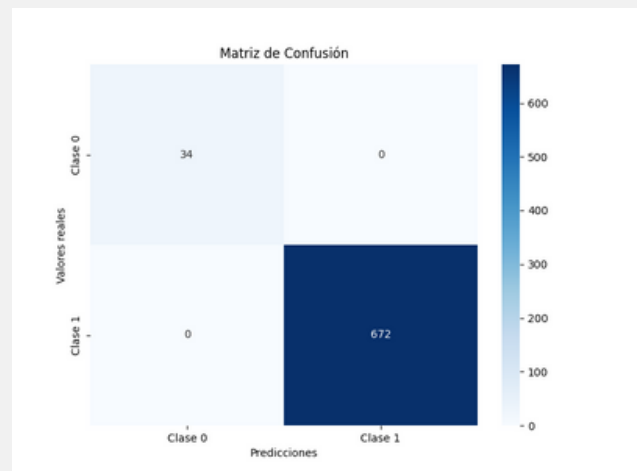
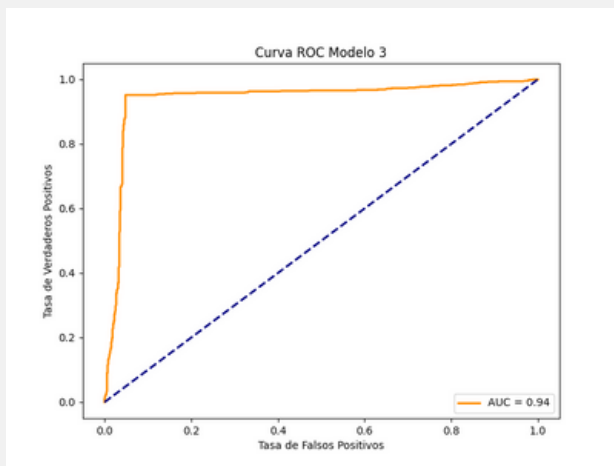
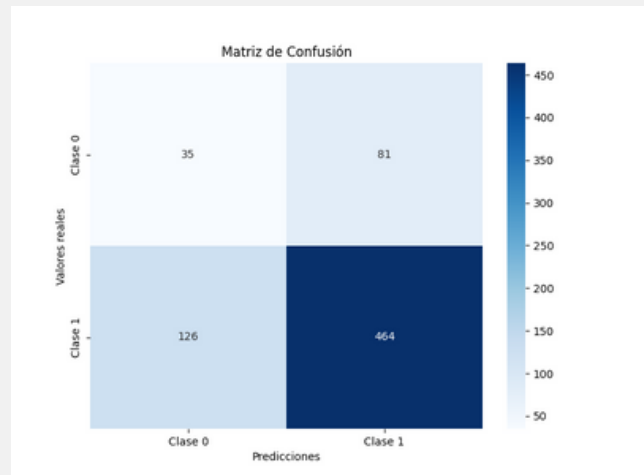
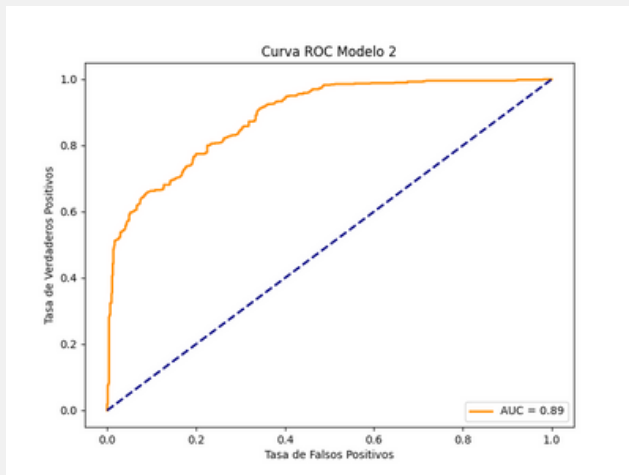
Finalmente creo la función `pred_model` que se ha diseñado con el propósito de realizar predicciones utilizando tres modelos de redes neuronales convolucionales entrenados para la identificación automatizada de lesiones dentales. Al recibir como entrada un conjunto de datos de prueba y un número específico de imagen, la función ejecuta los modelos respectivos y presenta las probabilidades asociadas a la presencia de caries, una pieza dental impactada y una pieza dental con lesión periapical en la imagen seleccionada.



Si la curva ROC se acerca al rincón superior izquierdo del gráfico, indica un mejor rendimiento del modelo, ya que tiene una alta Tasa de Verdaderos Positivos y una baja Tasa de Falsos Positivos.

Un AUC cercano a 1 sugiere que el modelo puede distinguir efectivamente entre clases positivas y negativas.

Si la curva ROC está cerca de la línea diagonal, el modelo no es mejor que una predicción aleatoria.



Resultados y Evaluación:

- Modelo 1: Pérdida en conjunto de prueba: 0.4954, Precisión: 75.64%.
- Modelo 2: Pérdida en conjunto de prueba: 0.4149, Precisión: 77.76%.
- Modelo 3: Pérdida en conjunto de prueba: 0.2207, Precisión: 95.18%.

Se empleó la Curva ROC para evaluar el rendimiento, considerando la Tasa de Falsos Positivos (FPR) y la Tasa de Verdaderos Positivos (TPR).

Conclusiones y Recomendaciones:

La implementación de modelos CNN mostró un rendimiento satisfactorio, destacando especialmente en la detección de lesiones periapicales. Se sugiere explorar ajustes en la arquitectura de los modelos y abordar el desbalance de clases para mejorar la generalización. Se destaca la importancia de ajustar umbrales de decisión y explorar técnicas de clasificación multiclase para mejorar el rendimiento global. La expansión del conjunto de datos con imágenes de diversas instituciones es recomendada para fortalecer la robustez del modelo en entornos clínicos variados.

Fuente de datos : <https://www.kaggle.com/datasets/truthisneverlinear/dentex-challenge-2023>