

Universidad del Valle de Guatemala

Data Science

Proyecto 2

Análisis exploratorio: Estimación de la Edad Ósea (EO)

Jennifer Sandoval	18962
Esteban del Valle	18221
Luis Quezada	18028

Situación problemática

La estimación de edad ósea con métodos convencionales puede llegar a generar errores debido a que las observaciones son realizadas por médicos, por lo que la edad ósea es estimada con un porcentaje de error correspondiente al factor humano, además, estos métodos pueden llegar a ser muy tardados. Utilizar un sistema de reconocimiento de imágenes automatizado permite eliminar este tipo de errores en las observaciones y predecir las edades óseas con mayor precisión. Para desarrollar un modelo de predicción que se encuentre basado en imágenes de rayos-x que pueda estimar la edad ósea de un paciente, es necesario que las imágenes proporcionadas al modelo estas deben tener una consistencia uniforme, una buena calidad, mismo tamaño, limpias de ruido, enfocadas en la estructura que se desea estudiar y resalte y contraste adecuado. Es por esto que es necesario llevar a cabo una limpieza y exploración de las imágenes y los datos asociados que puedan proporcionar información útil en la generación del modelo.

Problema científico

La edad ósea es una medida del grado de desarrollo esquelético en un niño. La maduración esquelética se ve afectada por las hormonas y por esta razón hay una correlación entre el desarrollo esquelético y el inicio de la pubertad, la cual se ve influenciada también por hormonas. La determinación de la edad ósea por medio de una radiología es utilizada generalmente como un indicador efectivo para diagnosticar enfermedades y así mismo determinar el mejor momento de tratamiento (Lepe, et. al, 2018).

Algunas enfermedades que pueden ser diagnosticadas son la deficiencia en la hormona de crecimiento, hipotiroidismo, pubertad precoz, trastornos de la glándula adrenal. También algunos trastornos de crecimiento genético como Síndrome de Turner y por último, la edad ósea puede ser útil en personas con problemas ortopédicos u ortodoncia ya que la edad ósea determinará cuál es el mejor momento en el cual se pueden utilizar herramientas y/o llevar a cabo cirugías asegurándose que la conformación de los huesos ya no cambiará (Hospital de San Diego, 2011).

Objetivos

Se quiere tomar los datos proporcionados y realizar un análisis exploratorio de los mismos. Se hace esto para aprender acerca de la forma y comportamiento de todos los datos proporcionados. Se pretende estudiar no solo los datos categóricos y cuantitativos, sino las imágenes dentro del dataset para obtener características que podrían resultar valiosas al momento de querer realizar un modelo o usar las imágenes para otros propósitos.

Descripción de los datos

Los datos proporcionados incluyen un test set y un training set con más de 12,000 fotografías de radiografías de manos. Adicionalmente, cada imagen tiene un id que se relaciona con algunas variables contenidas en un dataset proporcionado en formato csv que contienen las siguientes variables:

Variable	Clasificación	Tipo de dato
id	Categórica	int
boneage	Continua	int
sex	Categórica	int

Se realizó una modificación en la variable que indicaba el sexo de las personas en las muestras ya que se encontró que en los archivos correspondientes a la muestra de training y test tenían una inconsistencia en el formato de esta variable por lo que se decidió decidir la variable como 'sex' y asignar los valores 0 -> mujer y 1 -> hombre.

Análisis exploratorio

	id	sex	boneage
count	12611.000000	12611.000000	12611.000000
mean	8537.653001	0.541829	127.320752
std	4108.763993	0.498267	41.182021
min	1377.000000	0.000000	1.000000
25%	5074.500000	0.000000	96.000000
50%	8565.000000	1.000000	132.000000
75%	12091.500000	1.000000	156.000000
max	15610.000000	1.000000	228.000000

Figura 1. Estadística descriptiva para el conjunto de datos de entrenamiento

Es posible observar datos interesantes en la estadística descriptiva de los datos como por ejemplo, la edad ósea máxima es de 228 meses, 19 años. Esto nos indica que generalmente se estima la edad ósea de niños hasta aproximadamente la edad en la que se espera que hayan entrado a la pubertad, a partir de esta edad en adelante la edad ósea es calculada a partir de radiologías de esternón y escápula. También es posible observar que el promedio de edad ósea estimada es de 11 años y esto puede deberse a que esta edad se encuentra asociada al inicio de la pubertad y esto puede ser determinado por medio de la estimación ósea. La importancia del estudio de la edad ósea en estas edades se debe a que puede ser útil para diagnosticar desórdenes en el sistema endocrino y síndromes pediátricos..

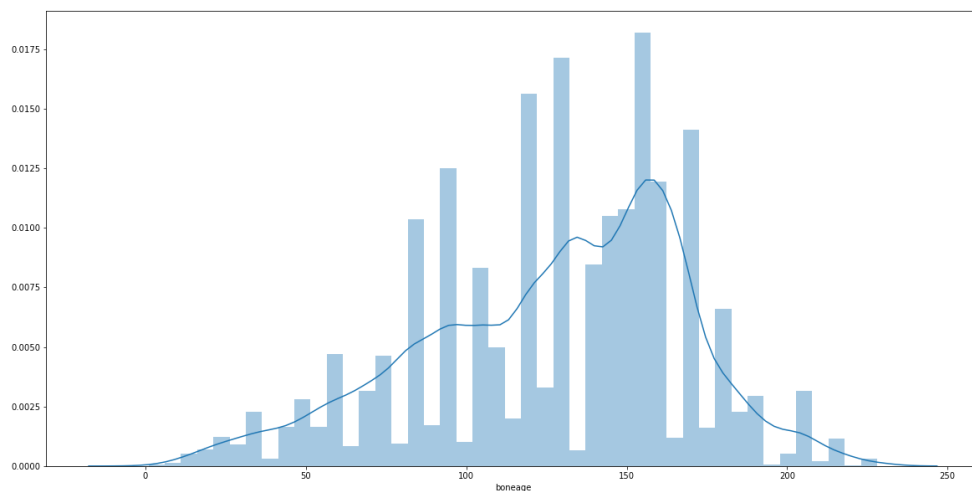


Figura 2. Histograma de distribución de las edades óseas

En la figura 2 es posible observar que la distribución de las estimaciones de edades óseas tiene un sesgo hacia la derecha, lo cual nos indica la importancia de la edad ósea en edades de

aproximadamente 12 años. La frecuencia de estimaciones en estas edades puede deberse a que es importante conocer el desarrollo de los niños en estas edades pues es cuando generalmente se da inicio a la pubertad y este tipo de estudios puede ayudar a determinar si existe algún retraso en el desarrollo óseo y si este tiene relación con el desarrollo hormonal.

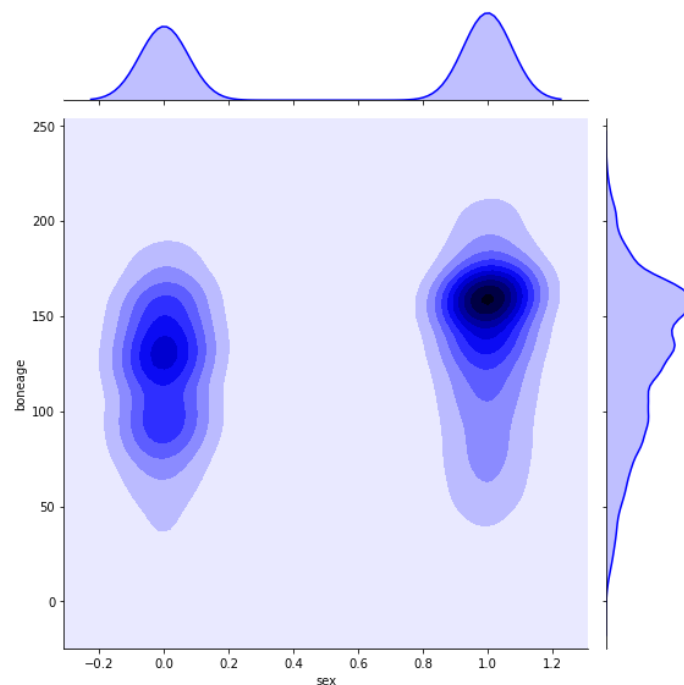


Figura 3. Relación entre las edades óseas y el sexo

Es posible observar que las estimaciones de edad ósea en hombres tienen una mayor frecuencia en edades avanzadas aproximadamente a los 14 años y esto se debe a que esta es la edad del inicio de la pubertad tardía en hombres, la cual varía en comparación de las mujeres que inicia a los 13 años aproximadamente. Para estas edades se esperaría la fusión de las falanges distales, fusión de los metacarpianos, fusión de las falanges proximales y fusión de las falanges medias. Este factor es importante ya que según estudios los adolescentes que entran a la pubertad más tarde de lo normal podrían enfrentar problemas en la salud ósea (Elhakeem , 2019).

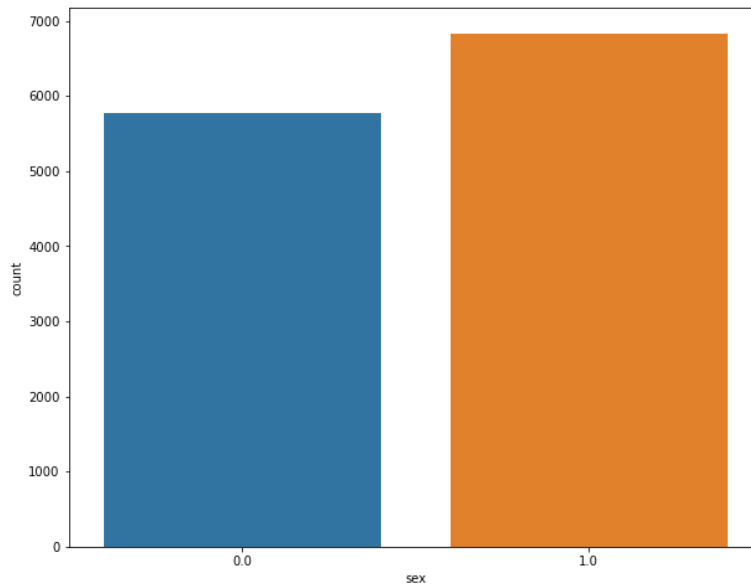


Figura 4. Frecuencias de sexos en el conjunto de datos

En esta gráfica se pueden observar la cantidad de personas que existen en la base de datos clasificado por sexo. Es posible observar que existe una cantidad mayor de hombres por alrededor 1,000. Esto puede afectar analisis en caso no se tome en cuenta el sexo ya que existe la probabilidad de que el sexo afecte la edad ósea a un punto significativo.

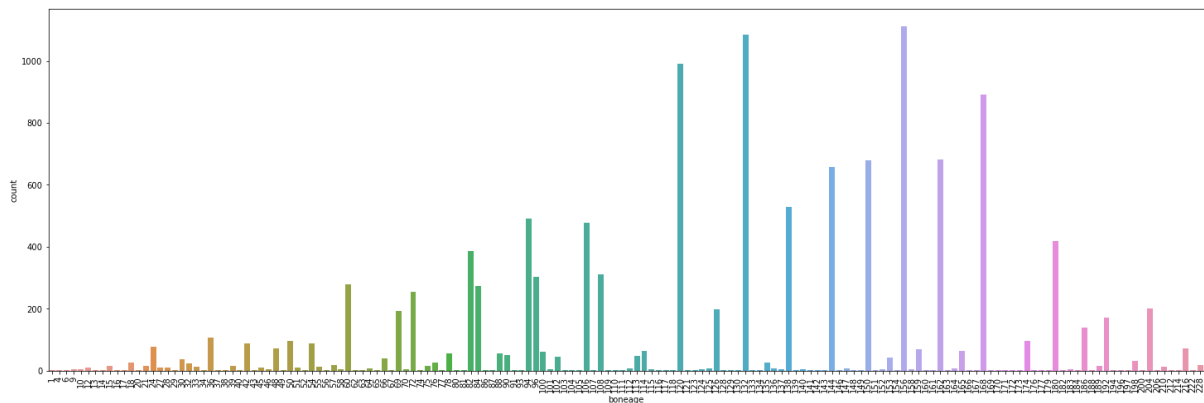


Figura 5. Frecuencias de edades óseas estimadas

En esta tabla se puede observar la frecuencia de todas las edades óseas presentes en el análisis. Es posible observar que la mayor parte de personas poseen una edad ósea media alta relativa al dataset, siendo 156 meses, 13 años, la edad de mayor frecuencia.

EDA Imágenes

```
*-----*
number of images          | 12611
dtype                     | uint8
channels                  | [1]
extensions                | ['png']

min height                | 640
max height                | 2989
mean height               | 1664.247878835937
median height             | 1624

min width                 | 521
max width                 | 3001
mean width                | 1319.3485052731742
median width              | 1277

mean height/width ratio   | 1.2614164280205484
median height/width ratio | 1.2717306186374315
recommended input size(by mean) | [1664 1320] (h x w, multiples of 8)
recommended input size(by mean) | [1664 1312] (h x w, multiples of 16)
recommended input size(by mean) | [1664 1312] (h x w, multiples of 32)

channel mean(0~1)         | 0.18266279810284672
channel std(0~1)          | 0.1951994388377542
*-----*
eda ended in 00 hours 07 minutes 25 seconds
```

Figura 6. Descripción de imágenes

De las imágenes se puede observar que el mínimo de altura es de 640 y esto podría representar una restricción al momento de hacer un modelo si es que se quiere que las imágenes sean uniformes. A la misma vez se observa un mínimo de ancho de 521 que también puede ser una restricción en la uniformidad de las imágenes. Sobre estas restricciones se podría tomar en cuenta el definir un tamaño que encaje dentro de estas restricciones o bien se podría definir un tamaño basándose en las medias de los valores de altura y ancho, dejando fuera del modelo las imágenes que no cumplan con el reajuste del tamaño. También es posible observar que la media del canal de pixeles es de 0.18, un valor muy cercano a 0 el cual representa el color negro en las imágenes de las radiologías.

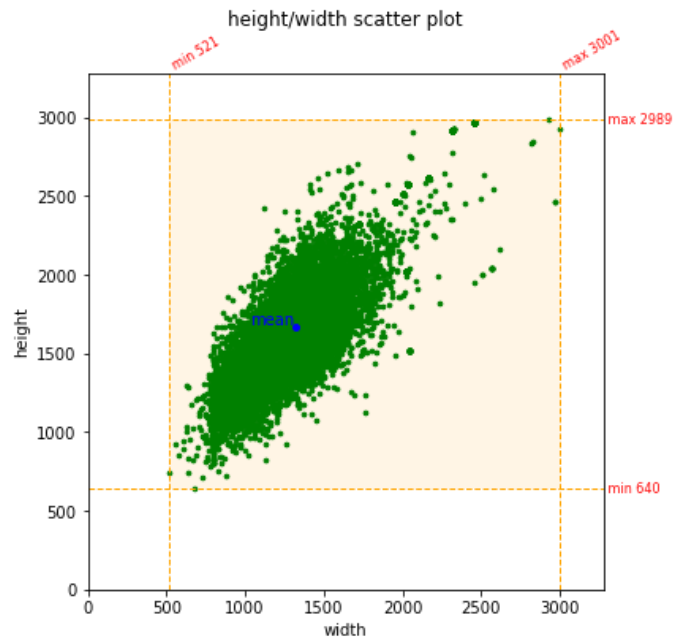


Figura 7. Distribución de ancho y altura de las imágenes de entrenamiento

En esta distribución es posible observar que la mayoría de imágenes se encuentran concentradas cercanas a la media de tamaños. Este es un buen indicador ya que se podría establecer un reajuste del tamaño en las imágenes y con esto incluir la mayoría de imágenes proporcionadas para que el modelo que se produzca sea más robusto en cuanto a la cantidad de imágenes con las que es entrenado.

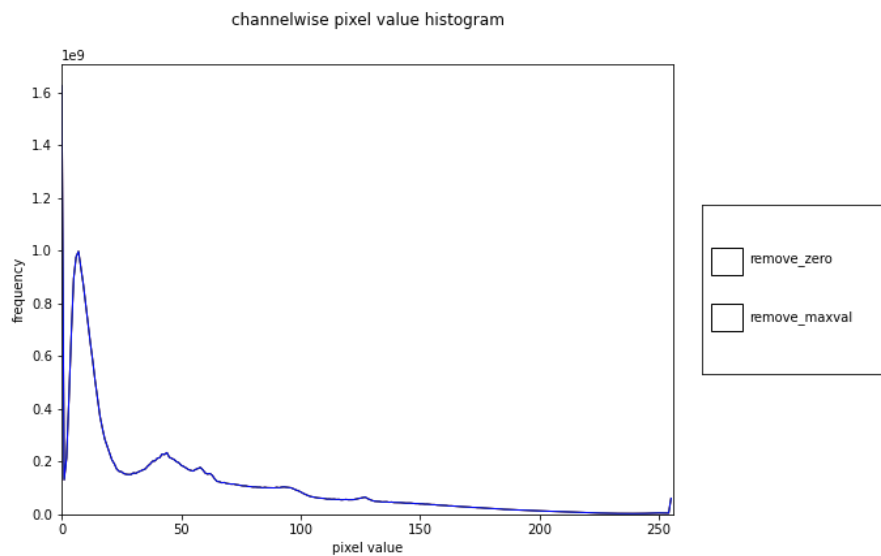


Figura 8. Histograma de píxeles por canal

Es posible observar que únicamente hay un canal en las imágenes y esto se debe a que las imágenes se encuentran en blanco y negro y a su vez es posible observar una mayor frecuencia de valores de píxeles cercanos a 0, esto también puede ser explicado porque el fondo de las radiografías tiene distintos tonos de negro y estos píxeles abarcan una gran mayoría en toda la imagen. Sin embargo, en este caso los píxeles de nuestro interés son los píxeles con un valor cercano a 1, representando el color blanco en el cual es posible observar la estructura de los huesos.

Hallazgos del EDA

Al realizar un análisis exploratorio profundo sobre los datos proporcionados, se pudieron encontrar varios factores importantes dentro de las 2 áreas claves de la base de datos: las imágenes, y los datos asociados correspondientes a la edad ósea.

En cuanto a los datos de edad ósea, se pudo observar que la gran mayoría de datos pertenecen a adolescentes desde el inicio hasta el final de la pubertad. También fue posible observar que la mayoría de datos pertenecen a hombres. Estos son datos importantes ya que se observa que al final de la pubertad es cuando la mano se termina de formar. También es importante considerar que la mayoría de los datos se encuentran alrededor de los 13 años y al no haber una distribución pareja entre todas las edades se podría ver afectado un modelo de predicción.

En cuanto a las imágenes incluidas en el dataset, inicialmente se presentó una preocupación relacionada al tamaño de la imagen ya que existía variación y no se presentaba un tamaño óptimo para hacer un posible modelo. Sin embargo, se logró identificar que la mayoría de imágenes se acercaban al tamaño promedio, lo cual es positivo ya que se podría ajustar el tamaño para crear uniformidad sin sacrificar contenido en el mayor de los casos. Por último, se tiene la ventaja de que todas las imágenes, por ser radiografías, ya están en blanco y negro. Esto facilita el preprocesamiento y hace más factible la creación de un modelo preciso.

Se encontró una relación de la edad ósea con el sexo de la persona, por lo general, se realizan estimaciones de edad ósea mayores en los hombres. Como se mencionó anteriormente, esto puede deberse a que la maduración esquelética y desarrollo de la pubertad en hombres es más tardada y es importante identificar el desarrollo óseo en estas edades pues en relación a su edad cronológica se esperaría observar los huesos completamente desarrollados, de lo contrario esto indicaría una

anomalía que debería estudiarse para determinar si hay una enfermedad causante del retraso en el desarrollo. Con respecto a las imágenes, fue posible analizar únicamente 1 canal debido a que estas ya están en blanco y negro. Se encontró que hay un pequeño cuadro identificador de las imágenes que podría causar ruido a la hora de crear un modelo predictivo.

Investigación de algoritmos de machine learning para reconocimiento de imágenes

SURF (Speeded up robust features)

Este algoritmo funciona como un detector y descriptor de características locales patentado. Se puede utilizar para reconocimiento de objetos, registro de imágenes, clasificación o reconstrucción 3D. Está parcialmente inspirado en el descriptor de transformación de características invariantes de escala (SIFT). Para detectar puntos de interés, SURF utiliza una aproximación entera del determinante del detector de manchas de Hesse, que se puede calcular con 3 operaciones enteras utilizando una imagen integral precalculada. Su descriptor de características se basa en la suma de la respuesta de la ondícula de Haar alrededor del punto de interés. Estos también se pueden calcular con la ayuda de la imagen integral (Bay, Tuytelaars y Gool, 2019).

Phase Stretch Transform-based algorithm

Este algoritmo de código abierto fue desarrollado por la Universidad de California (UCLA), por medio de operaciones matemáticas que permiten transformar las imágenes a un plano en 2D y a partir de esto el algoritmo es capaz de reconocer objetos, bordes y detectar otras características encontradas en las imágenes. Además, este algoritmo se caracteriza por tener la capacidad de identificar características en las imágenes que no son posibles de observar con el ojo humano, por ejemplo, las imágenes astronómicas que generalmente son editadas para aumentar el brillo de la imagen y poder observar las estrellas que a simple vista no se ven (Chin, 2016).

Regresión lineal

Un algoritmo de regresión lineal consiste en modelar una relación entre variables ajustando una ecuación lineal a los datos observados. Para este tipo de modelos se asume que existe una dependencia lineal de las variables dependientes a las variables independientes. Este modelo comúnmente es utilizado para la predicción de valores específicos y da resultados muy eficientes cuando en efecto las variables independientes y dependientes tienen una relación lineal entre sí (Gavrilova y Stryungis, 2020).

Redes neuronales convolucionales

Una red neuronal convolucional (CNN) es una clase de redes neuronales de deep learning. Son utilizadas comúnmente para analizar imágenes visuales y con frecuencia trabajan detrás de escena en la clasificación de imágenes. Un CNN cuenta con los siguientes elementos:

- Capas Convolucionales
- ReLU layers
- Pooling layers
- Una capa completamente conectada

Una CNN agrupa características aprendidas con datos de entrada y utiliza capas convolucionales 2D. Esto significa que este tipo de red es ideal para procesar imágenes 2D. En comparación con otros algoritmos de clasificación de imágenes. Esto significa que pueden aprender los filtros que tienen que ser diseñados desde cero con otros algoritmos. Las CNN tienen una variedad de aplicaciones, desde reconocimiento de imágenes y video, clasificación de imágenes y sistemas de recomendación hasta procesamiento de lenguaje natural y análisis de imágenes médicas (Silva y Freire, 2019).

Selección de algoritmos

Para este conjunto de imágenes y datos se decidió elegir específicamente trabajar con algoritmos de **redes neuronales convolucionales** con diferentes topologías para poder obtener la topología que sea más adecuada para este conjunto de datos. La razón por la cual decidimos utilizar este tipo de algoritmo fue porque consideramos que es muy importante la flexibilidad que este algoritmo nos permite al poder variar las funciones que pueden ser aplicadas a las imágenes para obtener una mejor predicción y a la misma vez esta flexibilidad nos asegura que no ocurrirá un 'overfitting'.

Además, se ha demostrado en varias investigaciones y proyectos que las redes neuronales convolucionales son ideales para el reconocimiento de imágenes de carácter médico ya que estas soportan grandes volúmenes de datos y en comparación de otros algoritmos son muy eficientes (Yadav y Jadhav, 2019). Otro factor importante es que generalmente los algoritmos de reconocimiento de imágenes son utilizados para la 'clasificación' de objetos y en este proyecto el objetivo es llevar a cabo una predicción y con las CNN es posible obtener como valor de salida una

predicción. También es importante mencionar que contamos con un conocimiento básico de las redes neuronales por lo que consideramos que es más ventajoso utilizar algoritmos conocidos para poder utilizar todos los recursos que las CNN nos proporcionan en beneficio de cumplir con los objetivos del proyecto.

Modelos desarrollados

Inicialmente se decidió poner a prueba modelos que tuvieran la estructura de un algoritmo de CNN pero variando las topologías entre cada modelo para poder analizar qué topologías podrían ser las más adecuadas, cuáles necesitan un mayor desarrollo y cuáles deben ser descartadas. Para futuros avances del proyecto se llevará a cabo una comparación de los resultados obtenidos con los algoritmos de CNN y un modelo de regresión lineal. Para el entrenamiento y validación de cada modelo se utilizaron 5000 imágenes. A continuación se detallan las estructuras de los modelos de CNN desarrollados, cada una de las siguientes estructuras se conservó en un archivo con los pesos correspondientes para poder hacer uso de estos modelos sin tener que entrenarlos nuevamente.

Modelo 1

- Neuronas de entrada 128x128x1
- Convolutional Layer (32)
- Convolutional Layer (64)
- Max Pooling
- Dropout y flatten
- Dense de 200 neuronas
- Dense de 200 neuronas
- Dropout final
- Dense (1 neurona final)

Modelo 2

- Neuronas de entrada 128x128x1
- Convolutional Layer (32)
- Convolutional Layer (64)
- Max Pooling
- Dropout y flatten
- Dense de 200 neuronas
- Dense de 400 neuronas
- Dense de 50 neuronas
- Dense de 100 neuronas
- Dropout final
- Dense (1 neurona final)

Modelo 3

- Neuronas de entrada 128x128x1
- Convolutional Layer (32)
- Convolutional Layer (64)
- Max Pooling
- Dropout y flatten
- Dense de 500 neuronas
- Dense de 100 neuronas
- Dense de 100 neuronas
- Dense de 50 neuronas
- Dense de 100 neuronas
- Dropout final
- Dense (1 neurona final)

Modelo 4

- Neuronas de entrada 128x128x1
- Convolutional Layer (32)
- Convolutional Layer (64)
- Max Pooling
- Dropout y flatten

- Dense de 512 neuronas con regularizador L2
- Dense de 1024 neuronas con regularizador L2
- Dense de 512 neuronas con regularizador L2
- Dropout final
- Dense (1 neurona final)

Modelo 5

- Neuronas de entrada 128x128x1
- Convolutional Layer (32)
- Convolutional Layer (64)
- Max Pooling
- Dropout y flatten
- Dense de 200 neuronas
- Dense de 400 neuronas
- Dense de 200 neuronas
- Dense de 100 neuronas
- Dropout final
- Dense (1 neurona final)

Modelo 6

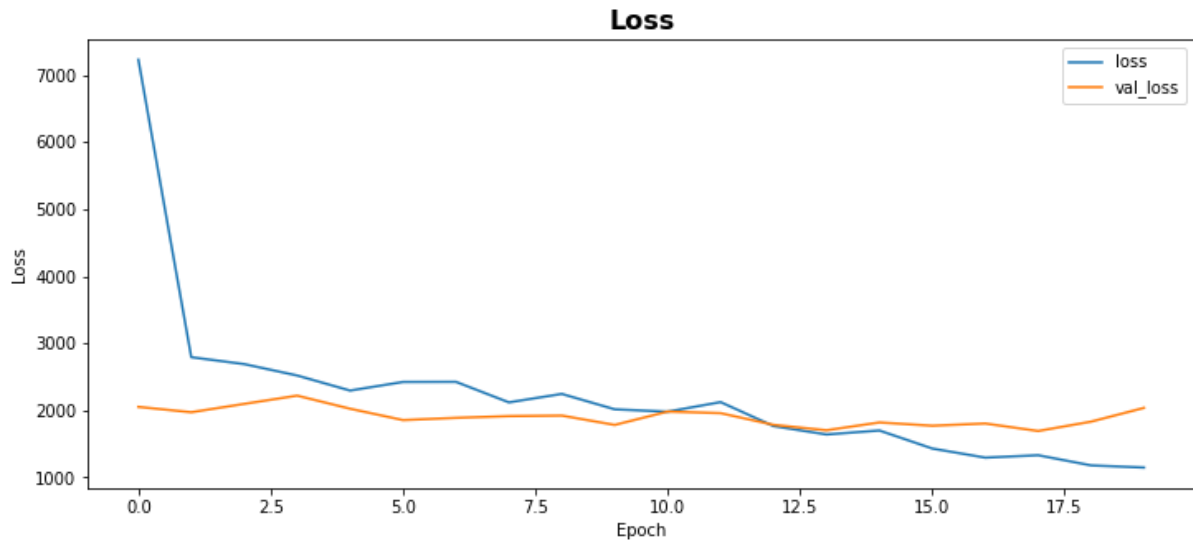
- Neuronas de entrada 128x128x1
- Convolutional Layer (16)
- Dropout (0.3)
- Convolutional Layer (32)
- Dropout (0.25)
- Convolutional Layer (32)
- Convolutional Layer (64)
- Convolutional Layer (128)
- Dropout (0.3)
- Flatten
- Dense de 300 neuronas
- Dense de 100 neuronas
- Dense (1 neurona final)

Modelo 7

- Neuronas de entrada 128x128x1
- Convolutional Layer (16)
- Dropout (0.3)
- Convolutional Layer (32)
- Dropout (0.3)
- Convolutional Layer (32)
- Convolutional Layer (64)
- Convolutional Layer (128)
- Dropout (0.3)
- Flatten
- Dense de 200 neuronas
- Dense de 100 neuronas
- Dense de 10 neuronas
- Dense (1 neurona final)

Eficiencia de los modelos

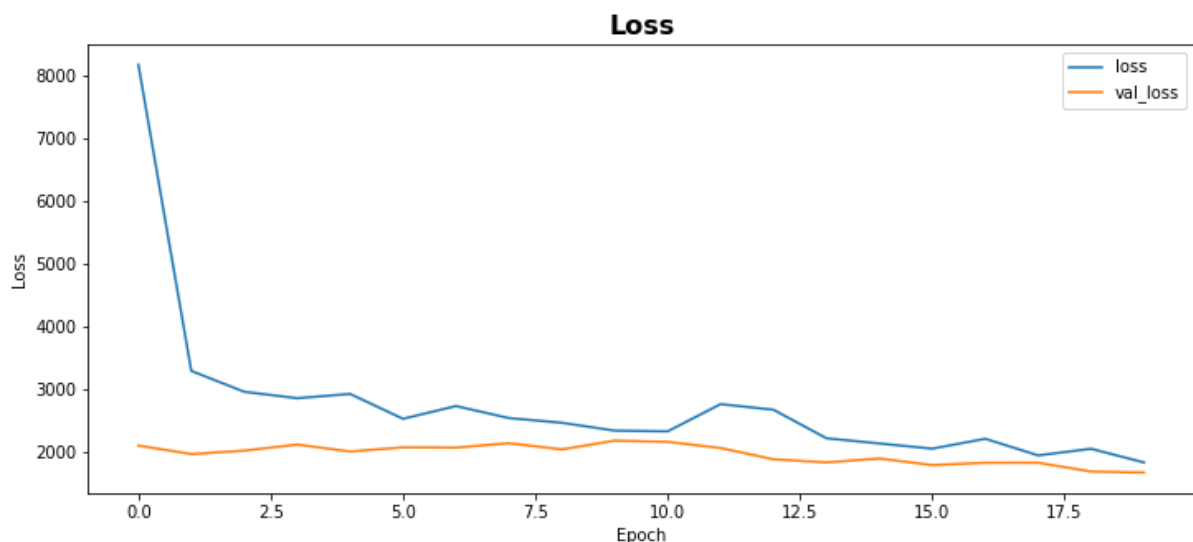
Modelo 1



Gráfica 9. Error de predicción en el Modelo 1

Para el modelo 1, es posible observar que alrededor del Epoch 2.5, empieza a disminuir el error de predicción del modelo y se mantiene así hasta el final del proceso. No obstante, luego de esto, se empieza a generar una mayor separación entre ambas líneas. Esto significa que el modelo presenta cierta eficiencia pero cerca del final del modelo presenta una falta de fidelidad leve pero lo suficientemente notable. El modelo converge en el epoch 12.5 y luego se divide. Por esta razón, se podría asumir que en este modelo existe un poco de overfitting.

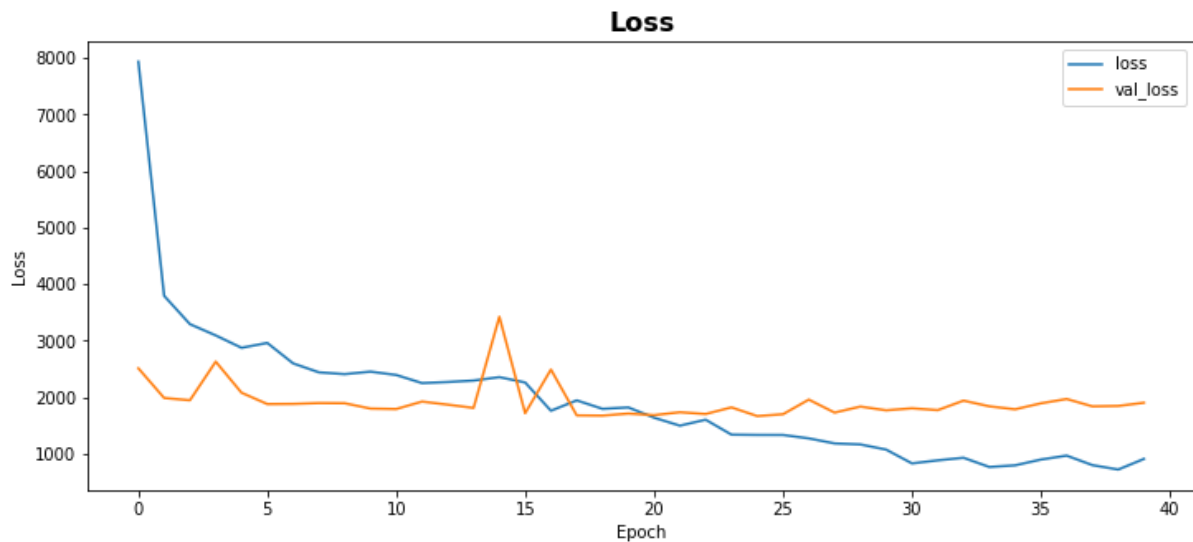
Modelo 2



Gráfica 10. Error de predicción en el Modelo 2

En el modelo 2, se puede observar que el error de predicción baja constantemente al inicio del modelo hasta casi converger completamente en el final del entrenamiento. Esto significa que el modelo debería de lograr predecir con bastante certeza edad ósea. Cabe mencionar que alrededor del epoch 11, se observa un leve pico en la línea de predicción. Este se remedia rápidamente en el epoch 12.5 y el modelo continuo en su patrón de convergencia.

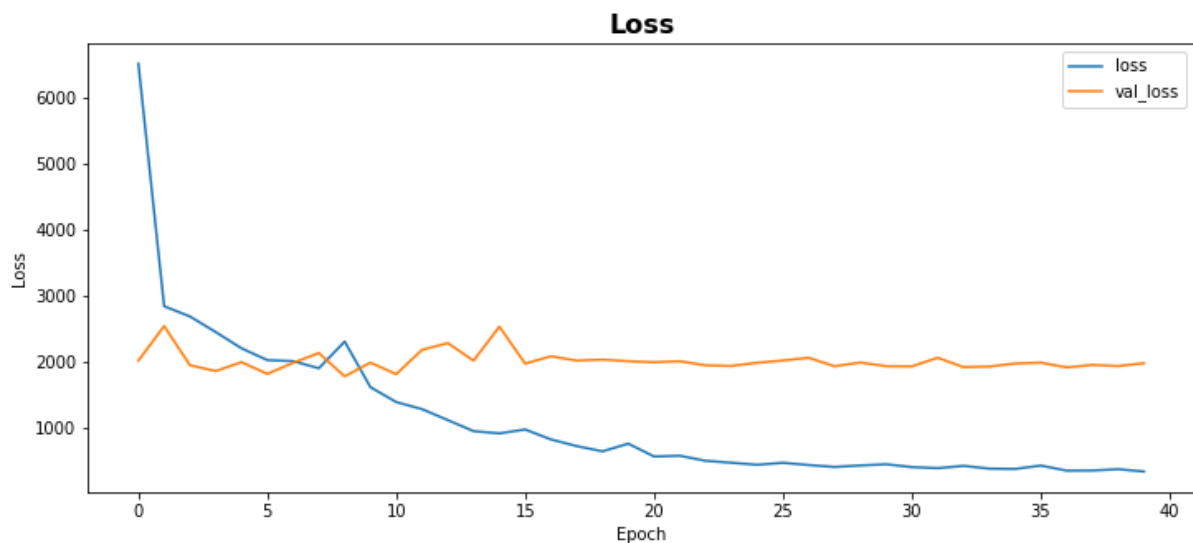
Modelo 3



Gráfica 11. Error de predicción en el Modelo 3

En el modelo 3, similarmente a los 2 modelos pasados, el error de predicción disminuye constantemente hasta el final del entrenamiento. No obstante, las líneas no convergen al final. Se observa una convergencia en el epoch 20 y, a partir de ese punto, las líneas empiezan a separarse. Esto puede significar que al entrenar el modelo, hubo bastante overfitting ya que hay divergencia muy temprano. Esto se verá reflejado en las predicciones que no serán tan exactas debido a este overfitting. Cabe mencionar que alrededor del epoch 15 se observan 2 picos que parecen converger. Sin embargo, si se hubiera cortado en ese punto el entrenamiento, se hubiera dado una divergencia con pendientes mucho más pronunciadas y el modelo hubiera tenido una mucho peor eficiencia.

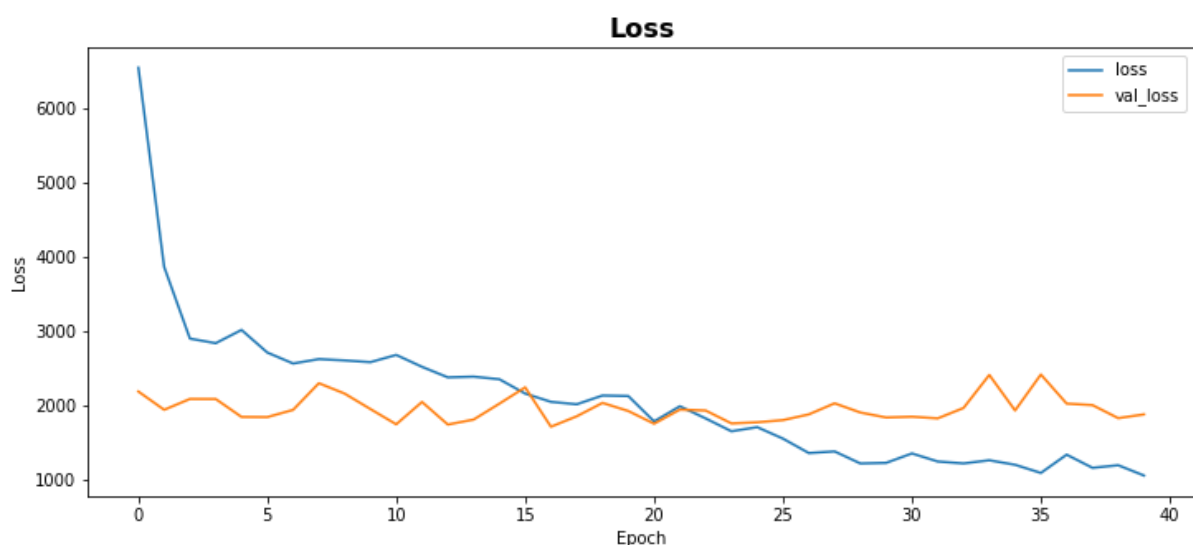
Modelo 4



Gráfica 12. Error de predicción en el Modelo 4

El modelo 4 también decrece constantemente en el error de predicción. A simple vista, es posible observar que este modelo podría ser el menos certero de los 4. Esto es ya que se observa que la pérdida de predicción y la de validación divergen de manera excesiva de manera muy temprana en el modelo. Se puede observar una convergencia temprana alrededor del epoch 5. Luego de esto, las líneas se separan drásticamente. Esto podría significar una cantidad significativa de overfitting y se vería reflejado en el uso de este modelo con resultados poco confiables y muchas veces equivocados.

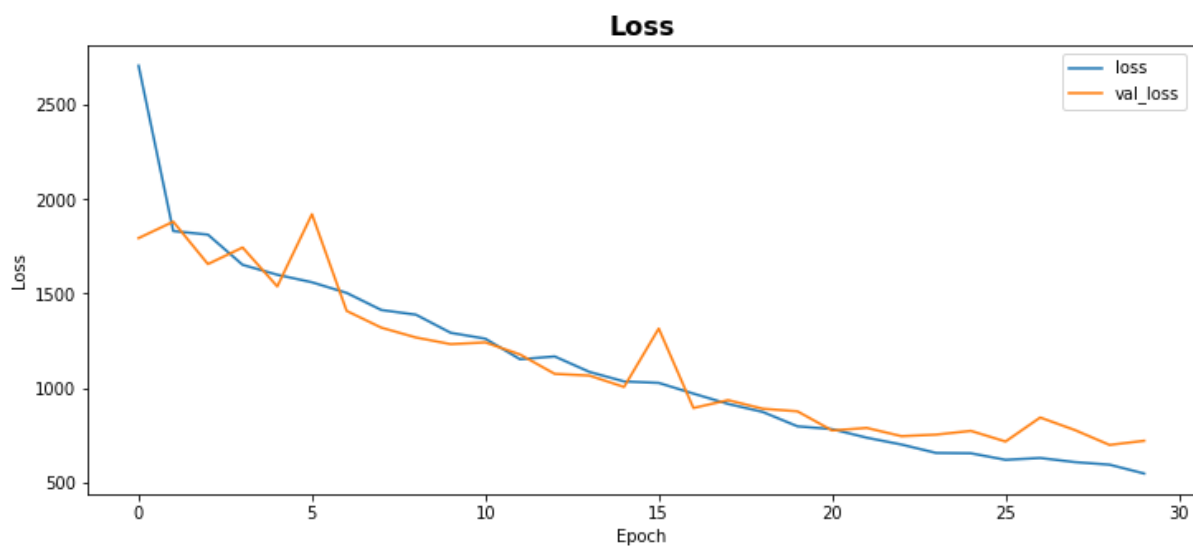
Modelo 5



Gráfica 13. Error de predicción en el Modelo 5

En el modelo 5, se puede observar que el error de predicción decrece constantemente, sin embargo, ambas líneas se mantienen separadas en la mayoría de epoch a excepción del epoch 20 en el cual parecen cruzarse y finalmente ambas divergen, manteniéndose el valor de pérdida aparentemente constante.

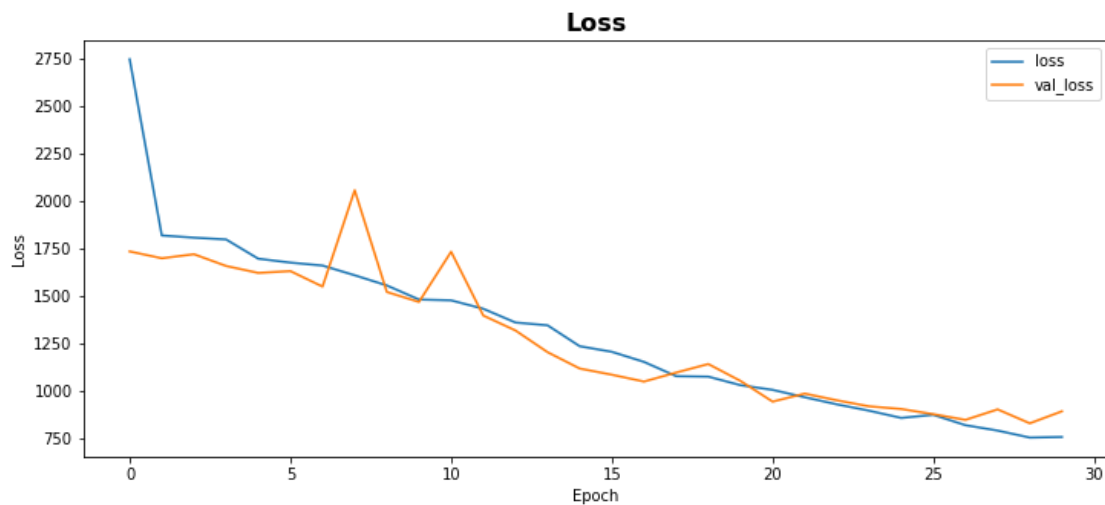
Modelo 6



Gráfica 14. Error de predicción en el Modelo 6

En el modelo 6 se realizó una variación en las capas, colocando más capas convolucionales, seguida de un dropout. Además se utilizaron 12,000 imágenes para realizar el entrenamiento de este modelo. Se obtuvo un MSE de 720 y se puede observar que se logró una disminución constante del error conforme se dió el entrenamiento del modelo. Por el epoch 25 se puede observar que se estabiliza y es por eso que se decidió hacer únicamente 30 epochs, ya que de ese punto en adelante la mejora es mínima. Se concluye que no hay necesidad de usar regularizadores porque no hay overfitting observable en la gráfica.

Modelo 7



Gráfica 15. Error de predicción en el Modelo 7

En el modelo 7 se realizó una variación en las capas del modelo 6 en busca de un MSE menor, sin embargo no se logró ya que se obtuvo un MSE de 893. Se puede observar que se logró una disminución constante del error conforme se dio el entrenamiento del modelo. Por el epoch 25 se puede observar que se estabiliza y es posible concluir que no hay necesidad de usar regularizadores porque no hay overfitting observable en la gráfica.

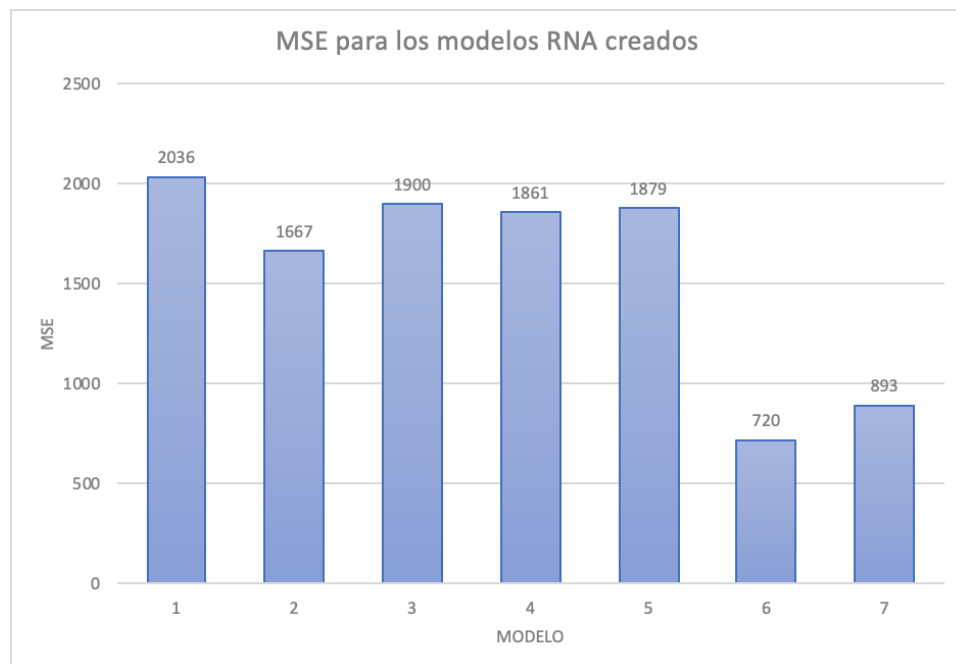


Gráfico 16. Comparación de MSE de los modelos de RNA desarrollados

Visualización de Predicciones con el Modelo 6

Nota: Estos datos pertenecen al set de datos de test, por lo que son imágenes con las que el modelo no fue entrenado. Se muestran las predicciones para 300 imágenes a forma de ejemplo.

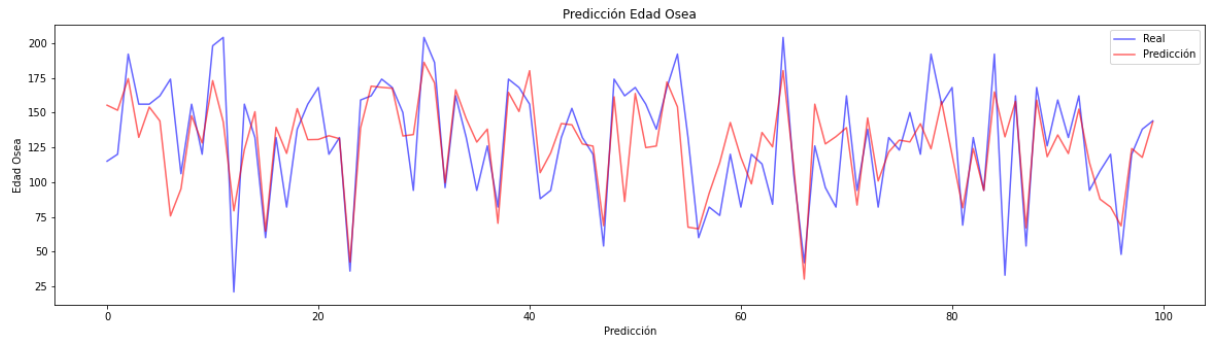


Gráfico 17. Predicción de Edad Ósea Imágenes 1-100

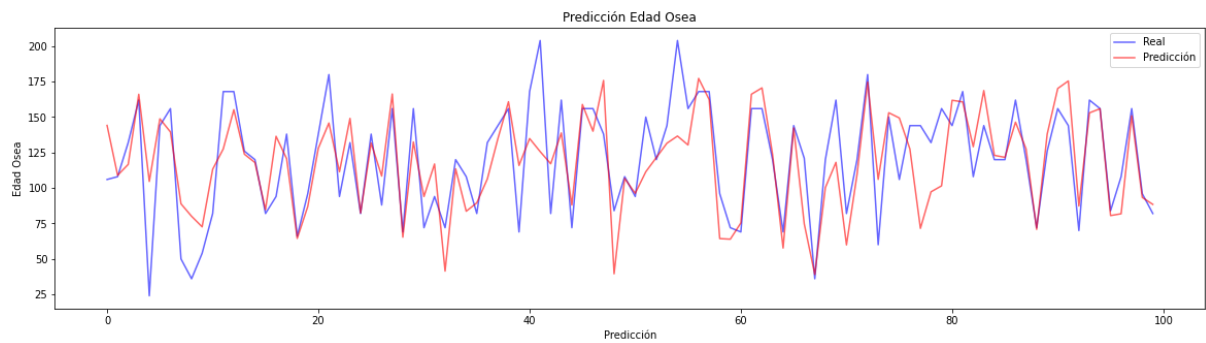


Gráfico 18. Predicción de Edad Ósea Imágenes 101-200

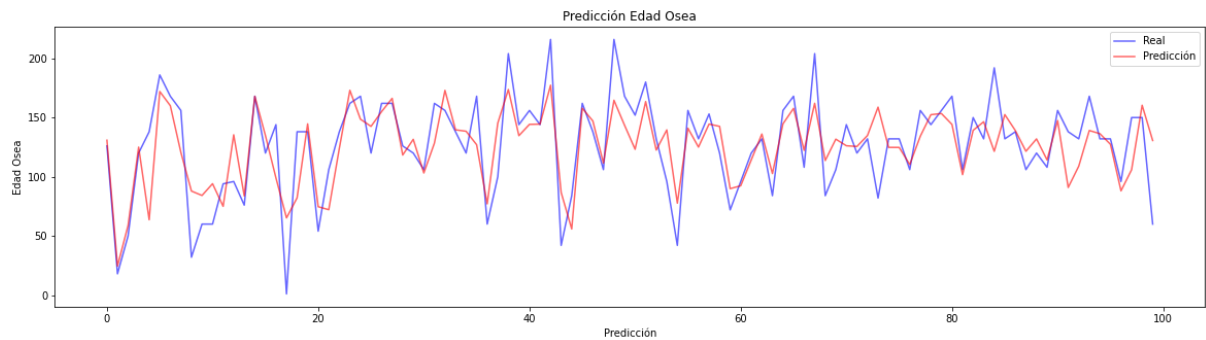


Gráfico 19. Predicción de Edad Ósea Imágenes 201-300

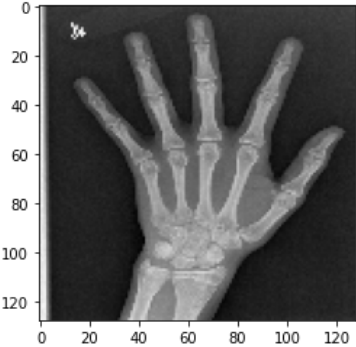
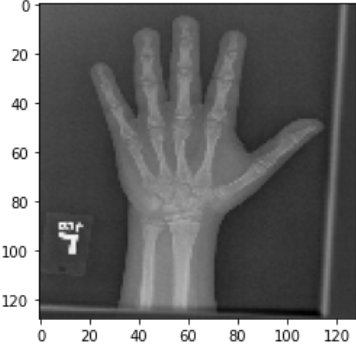
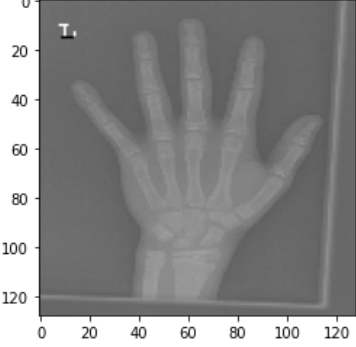
Imagen	Edad Real (años)	Edad Predicha (años)
	15	15.7
	11.5	11.9
	12.5	12.7

Tabla 1. Ejemplos visuales de predicciones hechas por el Modelo 6.

Conclusiones

Finalmente para evaluar la eficiencia de los modelos desarrollados de RNA, para cada modelo se calculó el error cuadrático medio (MSE) lo que nos permite determinar la eficiencia de predicción de cada modelo. Para los modelos 1, 2, 3, 4, 5, 6 y 7 se obtuvo un MSE de 2036, 1667, 1900, 1861, 720 y 893 respectivamente. A partir de esto fue posible observar que el modelo 6 fue el que presentó un valor menor de MSE lo cual nos indica que fue el modelo más eficiente de los modelos desarrollados. Se encontró una mejora significativa a la hora de realizar modelos con más capas convolucionales que capas densas. Además, para los últimos 2 modelos se logró utilizar 12,000 imágenes de entreno por lo que esto puede ser un factor determinante para reducir el error de predicción en los modelos.

A lo largo de estas semanas de trabajo lo que se realizó fue una gran cantidad de corridas, variando la topología y composición de cada modelo. Los 7 modelos que se analizan en este documento son los mejores casos de cada variación.

Como se puede observar en las gráficas 9, 10 y 11, las predicciones del modelo 6 se acercan bastante a las edades óseas verdaderas en una gran mayoría de los casos. Estas gráficas fueron realizadas con imágenes que el modelo no ha visto antes, por lo que es un buen indicador del funcionamiento del modelo. Claramente el modelo puede ser mejorado para reducir el MSE lo más posible, hacer que este alcanzara 0 sería lo ideal.

En conclusión, consideramos que el modelo 6 fue el que presentó mayor exactitud de predicción y este modelo se podría emplear en la predicción de edad ósea de una persona. A medida de facilitar su uso, se guardaron sus pesos y conformación en un archivo formato .h5 para su uso futuro.

Referencias

Bay, Tuytelaars y Gool. (2019). SURF: Speeded Up Robust Features. ETH Zurich. Extraído de: <https://people.ee.ethz.ch/~surf/>

Chin, M. (2016). UCLA researchers release open source code for a powerful image detection algorithm. UCLA Newsroom. Extraído de: <https://newsroom.ucla.edu/releases/ucla-researchers-release-open-source-code-for-powerful-image-detection-algorithm>

Elhakeem. (2019). Teens who hit puberty later could face bone health issues later in life, studies suggest. Extraído de: <https://theconversation.com/teens-who-hit-puberty-later-could-face-bone-health-issues-later-in-life-studies-suggest-121470>

Gavrilova y Stryungis. (2020). **Machine Learning Algorithm Classification for Beginners**. Serokell. Extraído de: <https://serokell.io/blog/machine-learning-algorithm-classification-overview>
Hospital de San Diego. (2011). Radiografía: edad ósea. Extraído de: <https://www.rchsd.org/health-articles/radiografa-estudio-de-la-edad-sea/>

Pose Lepe, Georgette, Villacrés, Fabián, Silva Fuente-Alba, Claudio, & Guilloff, Stefan. (2018). Correlación en la determinación de la edad ósea radiológica mediante el método de Greulich y Pyle versus la evaluación automatizada utilizando el software BoneXpert. *Revista chilena de pediatría*, 89(5), 606-611. <https://dx.doi.org/10.4067/S0370-41062018005000705>

S. Yang, P. Luo, C. C. Loy, and X. Tang, "From Facial Parts Responses to Face Detection: A Deep Learning Approach", in IEEE International Conference on Computer Vision (ICCV), 2015.
Silva y Freire. (2019). Introducción a las redes neuronales convolucionales. Medium. Extraído de: <https://medium.com/@bootcampai/redes-neuronales-convolucionales-5e0ce960caf8>

Yadav, S.S., Jadhav, S.M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. *J Big Data* 6, 113. <https://doi.org/10.1186/s40537-019-0276-2>