



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Geodésica,
Cartográfica y Topográfica

Desarrollo de un Modelo Aprendizaje Profundo para la
Clasificación de Malformaciones Craneales a Partir de
Mallado 3D

Trabajo Fin de Máster

Máster Universitario en Ingeniería Geomática y Geoinformación

AUTOR/A: Simarro González, Adrián

Tutor/a: Lerma García, José Luis

CURSO ACADÉMICO: 2023/2024

Agradecimientos

Este trabajo no hubiera sido posible sin el apoyo incondicional de mis padres y mi pareja, sin las enseñanzas de mis tutores y sin el respaldo de mis compañeros. Gracias al Prof. José Luis Lerma, Omar Christian Quispe y a GIFLE por brindarme la oportunidad y facilitarme los datos para poder realizar este proyecto.

Los autores agradecen la financiación recibida con cargo al Proyecto de I+D+i en Salud PI22/01416

Compromiso

El presente documento ha sido realizado completamente por el firmante; no ha sido entregado como otro trabajo académico previo y todo el material tomado de otras fuentes ha sido convenientemente entrecomillado y citado su origen en el texto, así como referenciado en la bibliografía

Resum

En aquest treball final de mester (TFM) es presenta un enfocament integral per a la classificació de malformacions cranials utilitzant núvols de punts 3D generades a partir d'escanejos fotogramètrics. S'empren tècniques avançades de processament de malles i anàlisis de concavitat per a transformar les dades i agrupar les estructures cranials en clusters. Inicialment, s'implementen algorismes de normalització i reconstrucció de superfícies per a netejar i preparar les malles. Posteriorment, s'utilitza l'algorisme de K-Means per a clusteritzar el núvol de punts en regions específiques basades en punts de llavor seleccionats. Les dades clusteritzades i processaments s'utilitzaran per a entrenar un model d'aprenentatge profund que classifica les malformacions cranials en categories com braquicefalia, dolicocefalia, normocefalia, plagiocefàlia, turricefalia i trigonocefalia. L'objectiu final és desenvolupar un model robust i precís que pugui assistir en el diagnòstic i classificació d'aquestes condicions en un entorn clínic. Per a això, es començarà l'estudi amb l'anàlisi de núvols sintètics controlats per a ensenyar al sistema.

Summary

In this Master's Thesis, a comprehensive approach to cranial malformation classification using 3D point clouds generated from photogrammetric scans is presented. Advanced mesh processing and concavity analysis techniques are employed to transform the data and group the cranial structures into clusters. Initially, surface normalization and reconstruction algorithms are implemented to clean and prepare the meshes. Subsequently, the K-Means algorithm is used to cluster the point cloud into specific regions based on selected seed points. The clustered and processed data will be used to train a deep learning model that classifies cranial malformations into categories such as brachycephaly, dolichocephaly, normocephaly, plagiocephaly, turriccephaly, and trigonocephaly. The ultimate goal is to develop a robust and accurate model that can assist in the diagnosis and classification of these conditions in a clinical setting. To this end, the study will begin with the analysis of controlled synthetic clouds to teach the system.

Resumen

En este Trabajo Final de Máster (TFM) se presenta un enfoque integral para la clasificación de malformaciones craneales utilizando nubes de puntos 3D generadas a partir de escaneos fotogramétricos. Se emplean técnicas avanzadas de procesamiento de mallas y análisis de concavidad para transformar los datos y agrupar las estructuras craneales en clústeres. Inicialmente, se implementan algoritmos de normalización y reconstrucción de superficies para limpiar y preparar las mallas. Posteriormente, se utiliza el algoritmo de K-Means para clusterizar la nube de puntos en regiones específicas basadas en los puntos semilla seleccionados. Los datos segmentados y procesados se utilizarán para entrenar un modelo de aprendizaje profundo que clasifica las malformaciones craneales en categorías como braquicefalia, dolicocefalia, normocefalia, plagiocefalia, turricefalia y trigonocefalia. El objetivo final es desarrollar un modelo robusto y preciso que pueda asistir en el diagnóstico y clasificación de estas condiciones en un entorno clínico. Para ello, se comenzará el estudio con el análisis de nubes sintéticas controladas para enseñar al sistema.

Índice de Figuras

- Figura 1: Deformidades craneales, sacado de “Neuropediatra.org” – Página 9**
- Figura 2: Imágenes 2D y 3D, sacado de (Obuchowicz et al., 2024) – Página 10**
- Figura 3: Nube de puntos 3D – Página 13**
- Figura 4: Carpeta con nomenclaturas – Página 13**
- Figura 5: Organigrama procesamiento datos – Página 15**
- Figura 6: Nube de puntos 3D – Página 16**
- Figura 7: Reconstrucción superficie – Página 17**
- Figura 8: Organigrama segmentación y análisis de concavidad – Página 18**
- Figura 9: Clusterización de 6 puntos semilla – Página 19**
- Figura 10: Normalización y ajuste de dimensiones – Página 20**
- Figura 11: Datos para el entrenamiento del modelo de aprendizaje – Página 22**
- Figura 12: Modelo primero – Página 23**
- Figura 13: Modelo segundo – Página 24**
- Figura 14: Combinaciones de hiperparámetros – Página 24**
- Figura 15: Mejores conjuntos de hiperparámetros – Página 25**
- Figura 16: Organigrama de validación y evaluación – Página 26**
- Figura 17: Modelo 90/05/05 – Página 31**
- Figura 18: Reporte clasificación 90/05/05 – Página 32**
- Figura 19: Modelo 70/15/15 – Página 33**
- Figura 20: Reporte clasificación 70/15/15 – Página 34**
- Figura 21: Modelo 60/20/20 – Página 35**
- Figura 22: Reporte clasificación 60/20/20 – Página 36**
- Figura 23: Modelo 80/10/10 – Página 37**
- Figura 24: Reporte clasificación 80/10/10 – Página 38**
- Figura 25: Modelo 50/25/25 – Página 39**
- Figura 26: Reporte clasificación 50/25/25 – Página 40**
- Figura 27: Reporte Tabla salarial del convenio de ingenierías y oficinas técnicas –
Página 40**
- Figura 28: Archivos código – Página 51**

Índice de Tablas

- Tabla 1: Dataset – Página 21**
- Tabla 2: Combinaciones de porcentajes – Página 27**
- Tabla 3: Porcentajes y precisiones – Página 29**
- Tabla 4: Definición métricas matriz de confusión – Página 29**
- Tabla 5: Clases codificadas – Página 31**
- Tabla 6: Matriz de confusión 90/05/05 – Página 48**
- Tabla 7: Matriz de confusión porcentaje 90/05/05 – Página 31**
- Tabla 8: Métricas mc 90/05/05 – Página 32**
- Tabla 9: Matriz de confusión 70/15/15 – Página 48**
- Tabla 10: Matriz de confusión porcentaje 70/15/15 – Página 33**
- Tabla 11: Métricas mc 70/15/15 – Página 34**
- Tabla 12: Matriz de confusión 60/20/20 – Página 49**
- Tabla 13: Matriz de confusión porcentaje 60/20/20 – Página 35**
- Tabla 14: Métricas mc 60/20/20 – Página 36**
- Tabla 15: Matriz de confusión 80/10/10 – Página 49**
- Tabla 16: Matriz de confusión porcentaje 80/10/10 – Página 37**
- Tabla 17: Métricas mc 80/10/10 – Página 38**
- Tabla 18: Matriz de confusión 50/25/25 – Página 50**
- Tabla 19: Matriz de confusión porcentaje 50/25/25 – Página 39**
- Tabla 20: Métricas mc 50/25/25 – Página 40**
- Tabla 21: Precisiones conjuntos de prueba – Página 41**
- Tabla 22: Costes salariales – Página 42**

Índice

1. Introducción	9
2. Objetivos	12
3. Datos	13
4. Metodología	15
5. Resultados	29
6. Presupuesto	42
7. Discusión	43
8. Conclusiones	45
9. Referencias	47
10. Anexos	49

1.Introducción

Contexto y motivación

Las malformaciones craneales representan un conjunto de deformidades del cráneo que no solo afectan a la apariencia física sino que también puede tener implicaciones graves en el desarrollo neurológico y cognitivo de los individuos. Entre las deformidades más comunes se encuentran la braquicefalia, dolicocefalia, normocefalia, plagiocefalia, turricefalia y trigonocefalia (Foster et al., 2020). Estas condiciones pueden surgir debido a diversas causas, incluyendo factores genéticos, anomalías congénitas y la craneosinostosis, que implica el cierre prematuro de las suturas craneales (Struck et al., 2019).

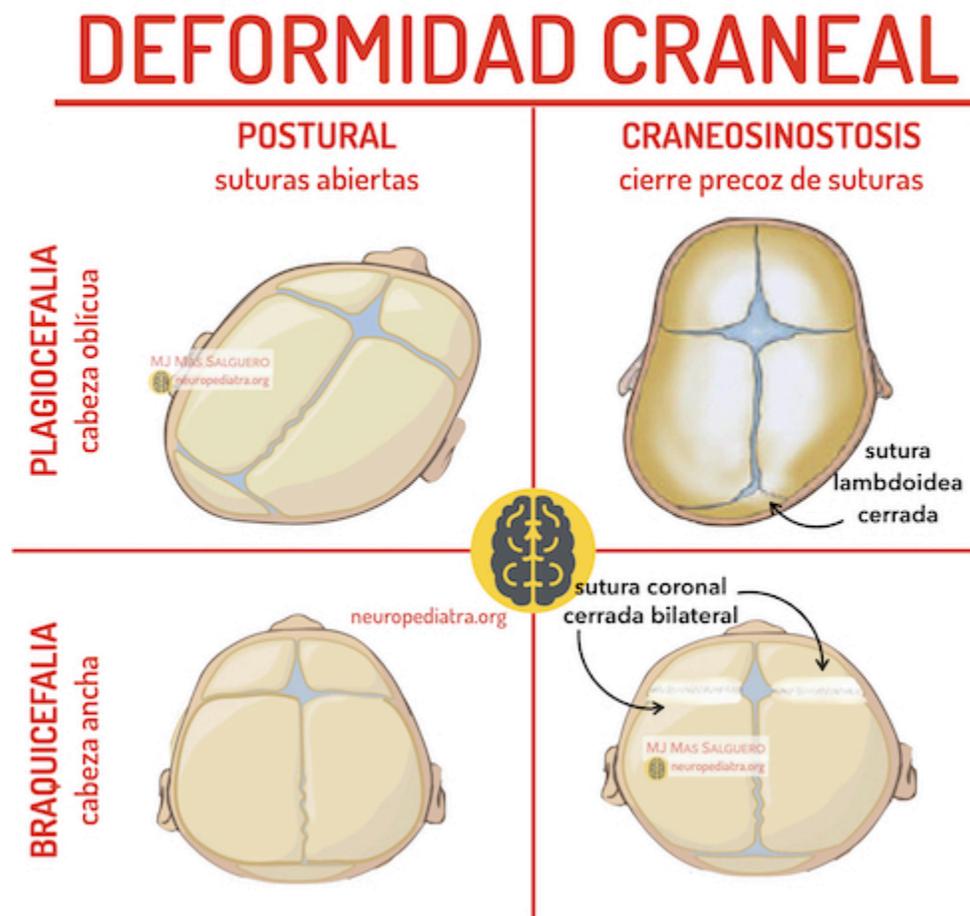


Figura 1 - Deformidades Craneales, sacada de (Mas, 2021)

La detección temprana y su clasificación precisa de estas deformaciones son cruciales para su intervención oportuna, ya que un diagnóstico certero puede guiar tratamientos que varían desde la terapia no quirúrgica hasta procedimientos quirúrgicos complejos (Quispe-Enriquez et al., 2023). Sin embargo, los métodos tradicionales de diagnóstico, que dependen de evaluación

clínica y técnicas de imagen como la tomografía computerizada (CT) y la resonancia magnética (RM), presentan limitaciones en términos de accesibilidad, coste y exposición a la radiación (Barbero-García et al., 2019).

Estado actual y antecedentes

En los últimos años, las técnicas de procesamiento de nubes de puntos 3D y el aprendizaje profundo han emergido como herramientas prometedoras para superar las limitaciones de los métodos tradicionales. La fotogrametría utiliza imágenes fotográficas para reconstruir modelos tridimensionales, ha demostrado ser una alternativa no invasiva y sin radiación, especialmente cuando se combina con dispositivos accesibles como los teléfonos inteligentes (Ey-Chmielewska et al., 2015). Estos avances han permitido la creación de un modelo 3D detallado del cráneo, que son fundamentales para un diagnóstico más preciso de las malformaciones craneales (Geisler et al., 2021).

Estudios recientes han explorado el uso de redes neuronales convolucionales (CNN) para la clasificación automática de imágenes médicas, incluyendo aquellas obtenidas a partir de datos 2D y 3D (Obuchowicz et al., 2024). Por ejemplo, un estudio realizado por el “*Departamento de Cirugía Oral y Maxilofacial de la Universidad de Heidelberg*”, se emplearon CNNs para clasificar deformidades craneofaciales con una alta precisión, utilizando imágenes 2D derivadas de fotogrametría (Anas et al., 2019)

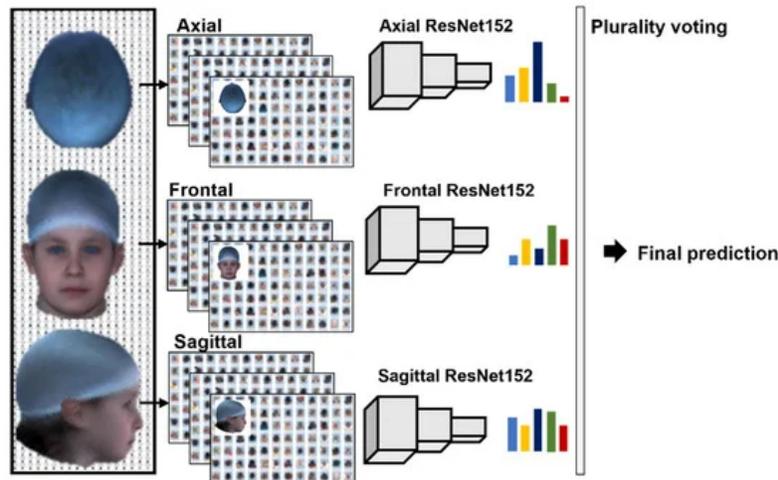


Figura 2 - Imágenes 2D y 3D, sacado de [9]

Esta tendencia se ha visto reforzada por investigaciones que implementan técnicas de fotogrametría y escáner 3D en dispositivos móviles, subrayando su viabilidad para el diagnóstico clínico en contextos diversos y no invasivos (Schaufelberger et al., 2022). La combinación de estas tecnologías no solo mejora la precisión del diagnóstico, sino que también democratiza el acceso a herramientas avanzadas de análisis morfológico (Nord et al., 2015).

Desafíos y avances recientes

A pesar de los avances, existen desafíos significativos en la implementación de estas tecnologías en entornos clínicos. Uno de los principales retos es la variabilidad en la calidad de los datos obtenidos a partir de diferentes dispositivos y condiciones de captura. La resolución, la iluminación y la geometría del objeto pueden afectar significativamente la precisión del modelo 3D resultante (Nishimoto et al., 2023). Además, el procesamiento y análisis de grandes volúmenes de datos 3D requiere recursos computacionales considerables, lo que puede ser un obstáculo en instalaciones con recursos limitados (Patias, 2002).

Sin embargo, estudios recientes han mostrado avances prometedores en la mitigación de estos desafíos. Por ejemplo, la implementación de algoritmos de procesamiento de mallas y análisis de concavidad ha mejorado la calidad de las nubes de puntos, permitiendo una segmentación más precisa de las estructuras craneales (Rasool & Bhat, 2023), (De Buhan & Nardoni, 2018). Asimismo, el uso de técnicas de clusterización, como K-Means, ha facilitado la agrupación de datos en regiones específicas del cráneo, lo que es crucial para el entrenamiento de modelos de aprendizaje profundo (Ng et al., 2006), (Shrivastava et al., 2014).

Además, la fotogrametría de corto alcance ha demostrado ser particularmente efectiva en la adquisición de datos precisos para el análisis de deformidades craneales. Investigadores como Barbero-García y Baselga han demostrado que aplicaciones basadas en smartphones pueden capturar datos 3D de alta calidad en neonatos, ofreciendo una alternativa accesible y precisa para la monitorización del crecimiento craneal (Barbero-García et al., 2020), (Barbero-García et al., 2017).

Importancia del estudio

El presente estudio tiene el potencial de transformar el diagnóstico y tratamiento de las malformaciones craneales al desarrollar un sistema automatizado y preciso para su clasificación utilizando datos tridimensionales obtenidos a partir de escaneos fotogramétricos (Barbero-García et al., 2020).

Al integrar técnicas avanzadas de procesamiento de mallas, análisis de concavidad y aprendizaje profundo, este enfoque promete no solo mejorar la precisión y consistencia del diagnóstico, sino también reducir la carga de trabajo de los profesionales de la salud, permitiéndoles enfocar sus esfuerzos en intervenciones más complejas y efectivas en un entorno clínico.

Además, este trabajo contribuye al campo del aprendizaje profundo aplicado a la medicina, demostrando cómo las técnicas avanzadas de procesamiento de nubes de puntos pueden integrarse eficazmente en la práctica clínica diaria. La capacidad de detectar y analizar anomalías craneales con mayor precisión y consistencia que los métodos tradicionales subraya la relevancia de este enfoque en la mejora de los resultados clínicos para los pacientes (Barbero-García et al., 2017), (Yang & Su, 2021).

2. Objetivos

El objetivo principal del proyecto es el procesado de los datos en bruto para introducirlos en un modelo de red neuronal y desarrollar un modelo de aprendizaje profundo robusto y preciso. Este modelo se utilizará para la clasificación de malformaciones craneales a partir de nubes de puntos 3D generadas a partir de soluciones fotogramétricas, facilitando el diagnóstico y la clasificación de estas afecciones en un entorno clínico.

Un objetivo complementario es contribuir al campo del aprendizaje profundo en medicina demostrando cómo las técnicas avanzadas de aprendizaje profundo e inteligencia artificial pueden integrarse en la práctica clínica diaria.

3. Datos

Los datos de partida son una serie de nubes de puntos 3D generados a partir de escaneos fotogramétricos usando un teléfono inteligente, y variantes (datos sintéticos) ligeramente manipuladas que permiten facilitar el aprendizaje de la red. Dichos datos fueron proporcionados por el grupo GIFLE. Estos escaneos producen datos en formato .ply que contiene las coordenadas de cada punto en el espacio tridimensional. Este formato es comúnmente utilizado en reconstrucción 3D y se presta bien al análisis y procesamientos adicionales. Véase la Figura 3.

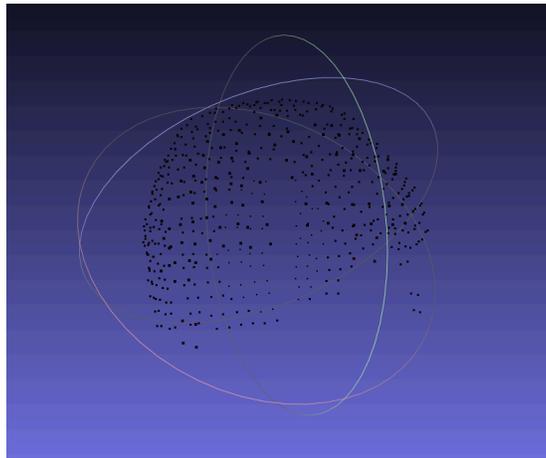


Figura 3 - Nube de puntos 3D

Los datos se agrupan en función del factor de escala y del ruido aleatorio:

Factor de escala: [0.988,0.989, 0.990, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999, 1, 1.001, 1.002, 1.003, 1.004, 1.005, 1.006, 1.007, 1.008, 1.009, 1.010, 1.011,1.012]

Cantidad de ruido aleatorio:

- Caso a: -1mm a 1 mm
- Caso b: -2mm a 2mm
- Caso c: -3 mm a 3 mm

Estos datos se organizan en carpetas que siguen una nomenclatura específica, como se muestra en la Figura 4. La nomenclatura de las carpetas incluye el tipo de datos ("5_datos_sinteticos_escalado_ruido"), el nivel de ruido aplicado (por ejemplo, "-1_1mm", "-2_2mm", "-3_3mm"), y el número de casos en cada carpeta ("1800casos").

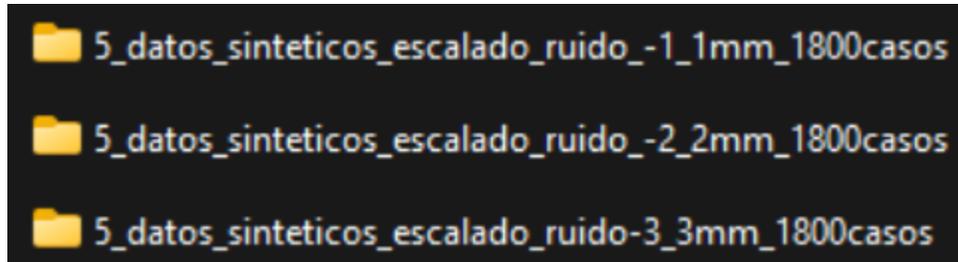


Figura 4 - Carpeta con nomenclaturas

Los tipos de malformaciones craneales (labels) son:

- Brachycephaly (Etiqueta 0 en el modelo)
- Dolichocephaly (Etiqueta 1 en el modelo)
- Normocephaly (Etiqueta 2 en el modelo)
- Plagiocephaly (Etiqueta 3 en el modelo)
- Towering (Etiqueta 4 en el modelo)
- Trigonocephaly (Etiqueta 5 en el modelo)

La nomenclatura de los ficheros es la siguiente:

Ejemplo: 2079_XY_0.998_ruido

Descripción del fichero

- 2079: código del caso
- XY,XYZ, Z : Indica los ejes en los que fueron aplicados los factores de escala
- 0.998: Factor de escalado aplicado

4. Metodología

Librerías usadas

Las librerías usadas para procesar datos y realizar modelado son las siguientes: **os** gestiona archivos y directorios, mientras que **pandas** y **numpy** manejan datos tabulares y operaciones numéricas. **pymeshlab** y **trimesh** se encargan del procesamiento y manipulación de mallas 3D. **sklearn** se utiliza para la normalización, clustering (KMeans), y preprocesamiento de datos, y **keras** junto con **tensorflow** sirven para construir y entrenar modelos de deep learning. Además, **joblib** optimiza el procesamiento paralelo y la serialización de objetos, **re** y **ast** procesan cadenas de texto, e **itertools** realiza combinaciones y cálculos en arrays. Finalmente, **openpyxl** se usa para trabajar con archivos Excel.

Preprocesamiento de Datos

El proceso de preprocesamiento de datos en este estudio es un paso crucial para asegurar la integridad y calidad de las nubes de puntos 3D utilizadas en el análisis. Como se ilustra en la Figura 5, este proceso se divide en dos fases principales: el clonado de la estructura de carpetas y el procesamiento de los archivos. Fichero usado, mirar Figura 28: 00_simarro_ply.py.

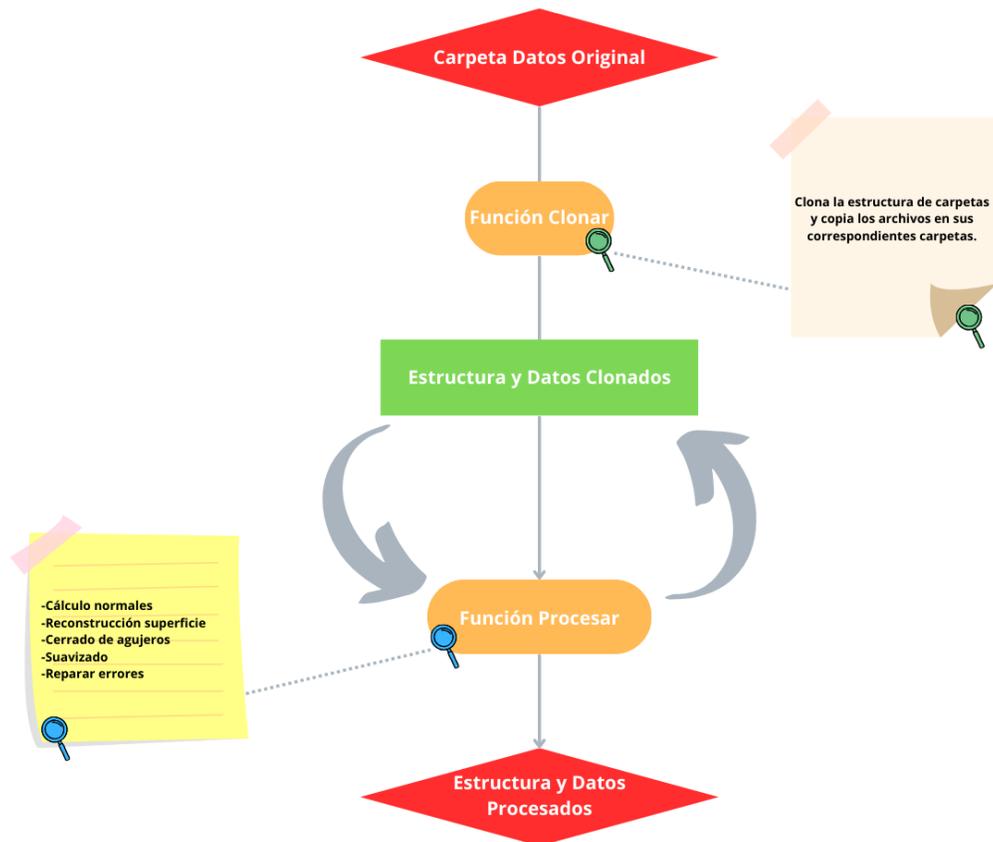


Figura 5 - Organigrama procesamiento datos

Clonado de Estructura de Carpetas y Procesamiento de Archivos

El primer paso en el procesamiento de los datos es clonar la estructura de carpetas y procesar los archivos .ply, como se observa en la Figura 6. Este proceso asegura que la disposición de los datos originales se mantenga intacta, mientras se procesan las nubes de puntos para obtener un mallado libre de errores. Todo este procesamiento se ejecuta con la librería pymeshlab.

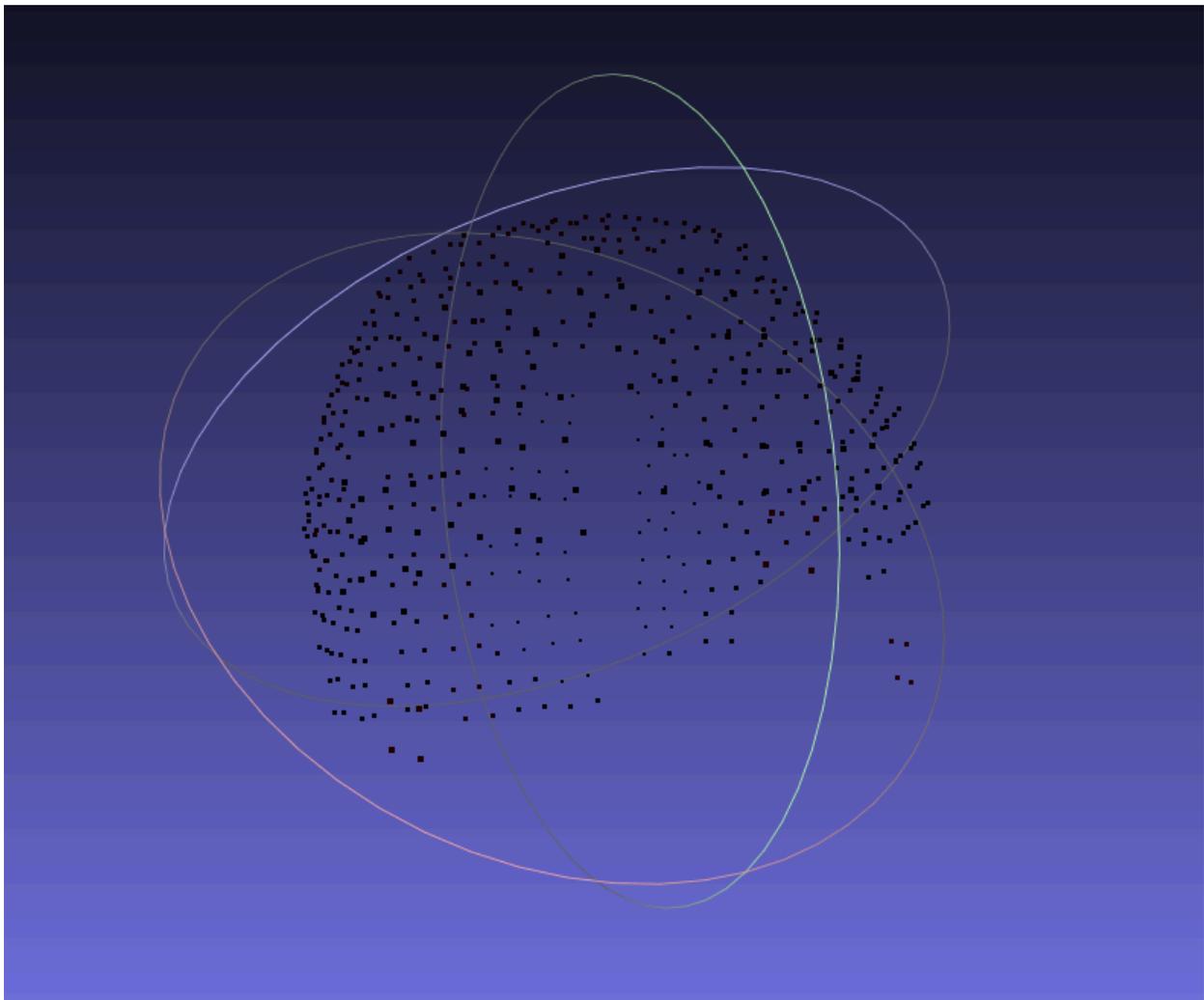


Figura 6 - Nube de puntos 3D

Normalización y reconstrucción de las superficies

Una vez copiada la estructura de los archivos originales se procedió a la normalización de las normales de los vértices, la reconstrucción de superficies utilizando el algoritmo de "Ball Pivoting" como se ve en la Figura 7. Este método permite generar superficies lisas y continuas

a partir de las nubes de puntos. Posteriormente, se cerraron los agujeros, se suavizó el resultado y se comprobó que no hubiesen errores en el mallado.

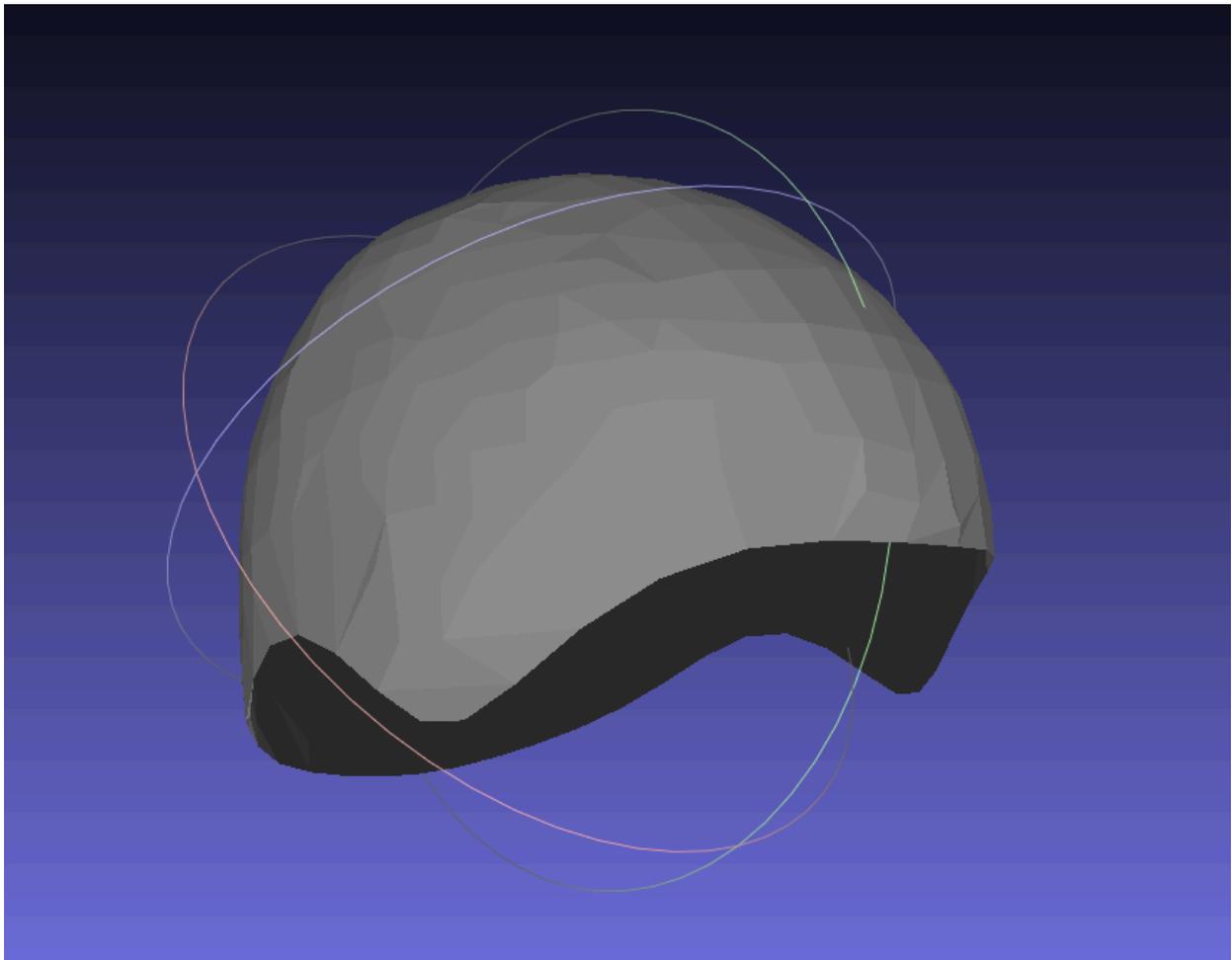


Figura 7 - Reconstrucción superficie

Segmentación y Análisis de Concavidad

El proceso de segmentación y análisis de concavidad es un componente fundamental en el estudio de las malformaciones craneales. Como se detalla en la Figura 8, este proceso se lleva a cabo utilizando técnicas de clusterización y análisis geométrico para identificar y caracterizar áreas específicas del cráneo. Fichero usado, mirar Figura 28: 01_simarro_segmentaciones_6.py.

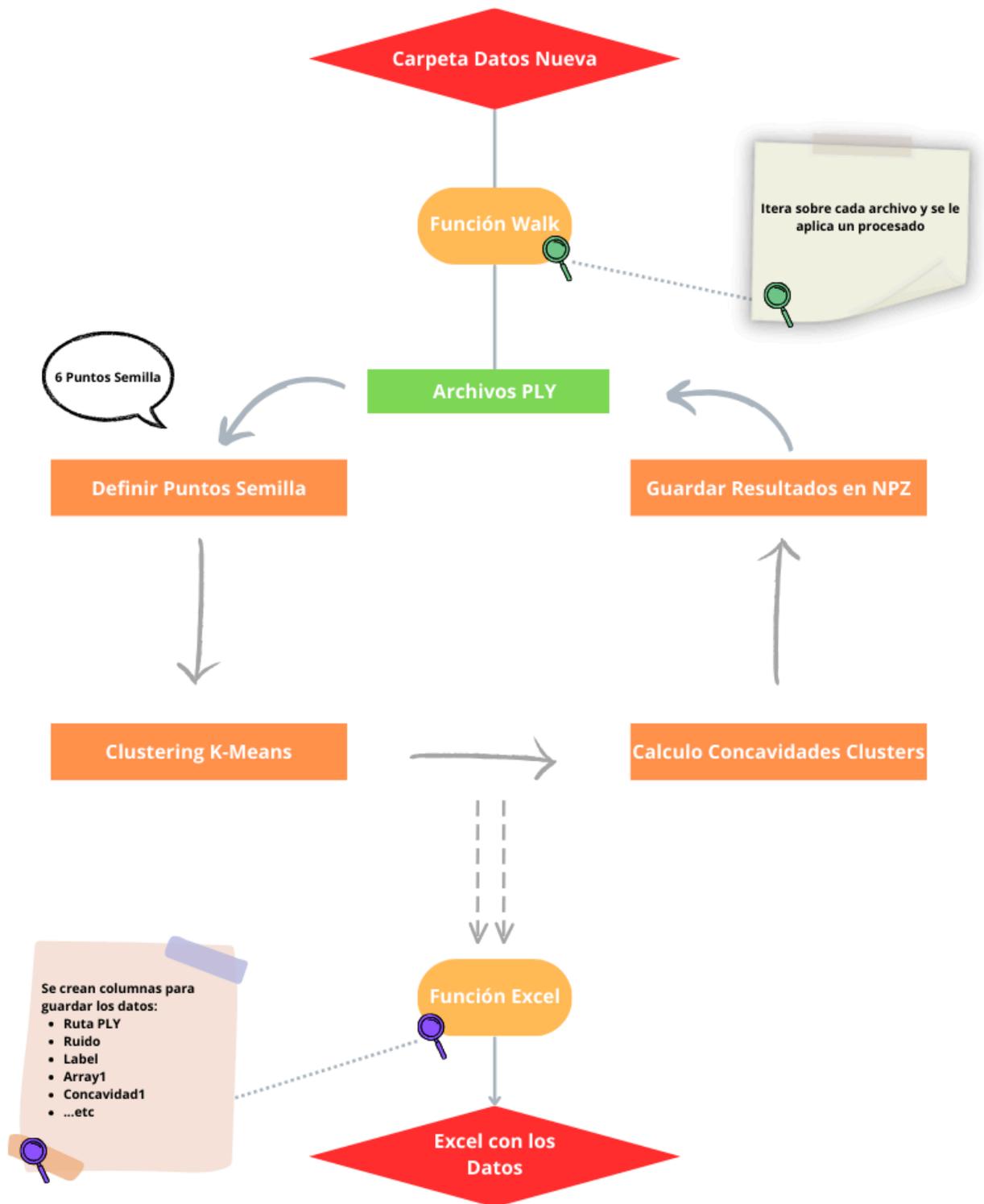


Figura 8 - Organigrama segmentación y análisis de concavidad

Clusterización basado en K-Means

Las nubes de puntos se agruparon mediante el algoritmo K-Means basado en la selección manual de puntos semilla. Estos puntos semilla se eligen para representar áreas específicas del cráneo, garantizando una acumulación uniforme.

Se probaron diferentes configuraciones para los puntos semillas, 4 puntos variando la división del cráneo en segmentos frontales y posteriores. La configuración final de 6 puntos semilla dividido en pares, como se muestra en la Figura 9, resultó ser la más efectiva para la clusterización y análisis de malformaciones craneales:

- Frontal Izquierdo y Frontal Derecho
- Centro Izquierdo y Centro Derecho
- Posterior Izquierdo y Posterior Derecho

El proceso de experimentación y la elección de los clústeres en este trabajo se fundamentaron en la necesidad de reducir el riesgo de sobreajuste del modelo. Si se hubiesen utilizado los datos como una sola pieza sin aplicar la técnica de clúster, no habríamos podido realizar un análisis más controlado y específico de las regiones craneales. Esto, a su vez, ha permitido desarrollar un modelo más robusto, entrenado en diversas condiciones, lo que ha resultado en una mejor clasificación.

Con cada área representada se procede a calcular su concavidad, obteniendo así la información que pueda ayudar al modelo a distinguir entre los tipos de malformaciones craneales. Finalmente se guarda toda la información en un fichero Excel como se plantea en la Figura 8 de cada modelo y división clúster calculada.

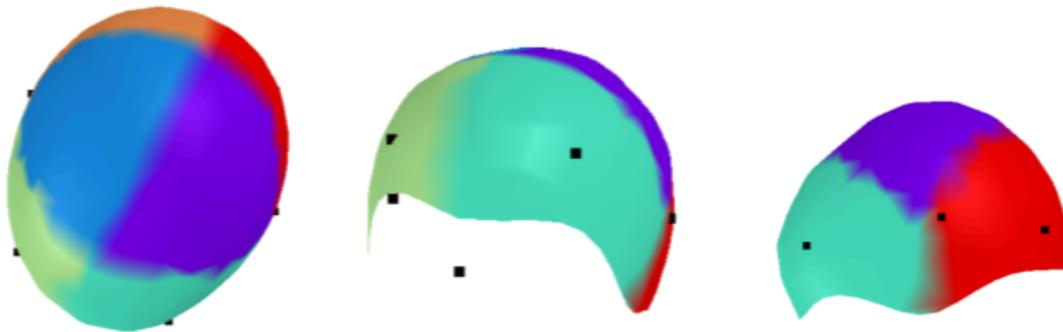


Figura 9 - Clusterización de 6 puntos semilla

Normalización y ajuste de dimensiones

Una vez agrupados los datos, se comprobó que no hubiese duplicaciones entre arrays y en el mismo, se normalizaron y ajustaron las dimensiones de los arrays para garantizar que todos los grupos tuvieran la misma longitud y pudieran utilizarse por igual en el modelo de aprendizaje

profundo, tal como se muestra en la Figura 10. Se opta por rellenar con ceros aquellos arrays con dimensiones menores para evitar datos perdidos, mantener la coherencia de las medidas y no perder de vista la importancia de los datos. Finalmente se realizó una diferencia entre todas las posibles combinaciones de arrays para introducir al modelo una variable que fuera el resultado de la relación espacial entre arrays, enriquecer el modelo y mejorar la generalización a nuevos datos para no depender de valores absolutos únicamente. Fichero usado, mirar Figura 28: 02_simarro_newExcel_Zero.py.

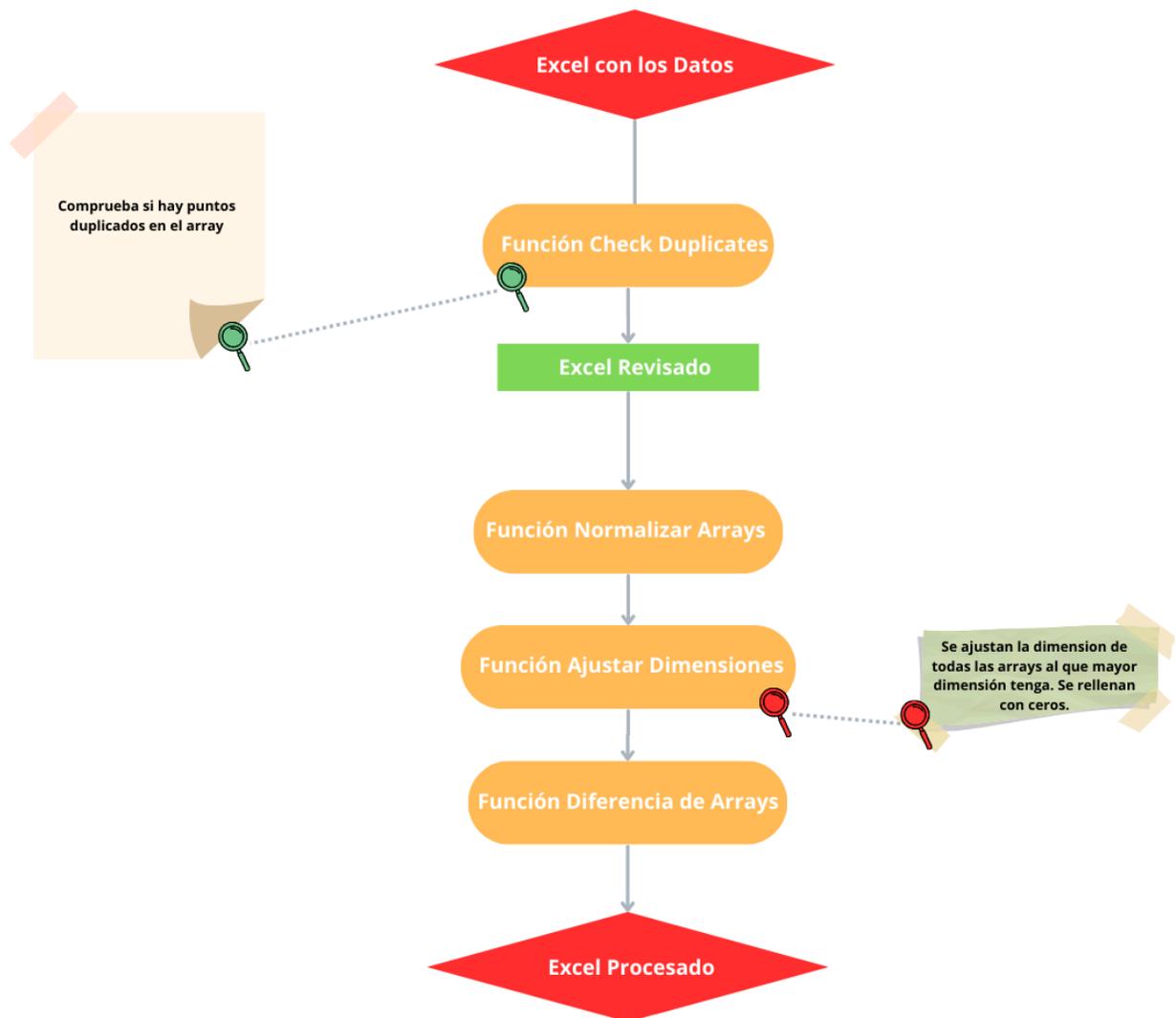


Figura 10 - Normalización y ajuste de dimensiones

Desarrollo y entrenamiento del modelo de aprendizaje profundo

Datos para el entrenamiento del modelo de aprendizaje

Los datos recopilados para su posterior introducción al modelo de aprendizaje profundo son los siguientes:

Columna	Descripción
Ruta PLY	La ruta del archivo PLY original.
Ruido	Información sobre el nivel de ruido presente en la nube de puntos.
Label	Etiqueta asociada a la malformación craneal específica.
Arrays	Segmentos específicos de la nube de puntos después de aplicar el algoritmo de segmentación basado en K-Means.
Concavidades	Valores de concavidad calculados para cada uno de los segmentos de la nube de puntos.
Arrays Normalizados	Segmentos de la nube de puntos después de ser normalizados.
Diferencia Arrays	Diferencias entre pares específicos de arrays normalizados.

Tabla 1 (a) - Dataset

Ruta PLY	Ruido	Label	Array 1	Array 2	Array 3	Array 4	Array 5	Array 6	Concavidad 1	Concavidad 2	Concavidad 3	Concavidad 4	Concavidad 5	Concavidad 6	Array 1 Normalizado	Array 2 Normalizado	Array 3 Normalizado	Array 4 Normalizado	Array 5 Normalizado	Array 6 Normalizado	Diferencia Array 1-2	Diferencia Array 1-3	Diferencia Array 1-4	Diferencia Array 1-5	Diferencia Array 1-6													
new_data_5_datos_s_brachyee[[2.30241[[9.33074[[4.71241[[1.212138[[40.7547[[4.01024	0.568522404	0.548539316	0.55247662	0.624191815	0.623360418	0.617683713	[[-0.01555	[[0.03856	[[0.14272	[[0.08898	[[-0.20402	[[-0.08945	[[0.34674	[[0.232174	[[-0.11457	[[-0.65801	[[0.17805	[[0.07389	new_data_5_datos_s_brachyee[[-2.28457[[22.7050[[50.0969[[7.20739[[59.8990[[42.9350	0.573814794	0.617831399	0.515511996	0.502429666	0.56535395	0.611532424	[[-0.01210	[[0.09572	[[0.13366	[[0.02060	[[-0.18506	[[-0.08808	[[0.318736	[[0.221751	[[-0.09698	[[0.074668	[[0.108891	[[0.07538	
new_data_5_datos_s_brachyee[[31.5584[[21.5617[[43.6480[[1.28431[[-67.6636[[-38.0904	0.548592362	0.623774914	0.627899045	0.582326638	0.574253781	0.59112026	[[0.144096	[[0.121475	[[0.113205	[[0.054402	[[-0.10061	[[0.287301	[[0.213816	[[-0.07348	[[0.067077	[[0.155015	[[0.24471	new_data_5_datos_s_brachyee[[1.53004[[-2.03895[[42.7703[[22.1394[[-64.0990	[[-11.9506	0.649167352	0.48831603	0.575739337	0.548949055	0.637368856	0.58021808	[[0.084811	[[-0.01194	[[0.148064	[[0.058206	[[-0.13699	[[-0.04111	[[0.285056	[[0.189175	[[-0.09587	[[-0.07015	[[0.09932	[[0.12592	
new_data_5_datos_s_brachyee[[36.1881[[-9.05041[[44.0000[[1.85564[[-67.0131[[-40.8222	0.610482426	0.648290618	0.595194467	0.564317072	0.58024125	0.626080709	[[0.183323	[[-0.05151	[[0.117675	[[0.008002	[[-0.16803	[[-0.10362	[[0.285716	[[0.221304	[[-0.06441	[[-0.05952	[[0.111631	[[0.28694	new_data_5_datos_s_brachyee[[4.04433[[5.44314[[32.3142[[36.2046[[-60.0650	[[-38.5873	0.644121987	0.552884722	0.59991664	0.633099749	0.597357819	0.628334874	[[0.023806	[[0.026046	[[0.143581	[[0.087756	[[-0.17318	[[-0.09069	[[0.316771	[[0.234275	[[-0.08249	[[-0.06170	[[0.178442	[[0.11449
new_data_5_datos_s_brachyee[[-1.48319[[6.27151[[45.0991[[9.96979[[-9.35791[[-38.5766	0.466496888	0.638828835	0.509523823	0.574831398	0.554104377	0.572611289	[[-0.09846	[[0.029605	[[0.15254	[[0.027315	[[-0.02925	[[-0.08855	[[0.181802	[[0.238104	[[0.056305	[[0.002285	[[0.112876	[[0.07707	new_data_5_datos_s_brachyee[[11.2832[[6.73716[[44.9271[[12.7456[[45.56270[[-46.3710	0.61498445	0.568671519	0.543396273	0.54821775	0.542759362	0.593592059	[[0.049593	[[0.039976	[[0.181134	[[0.023931	[[-0.16242	[[-0.11899	[[0.343555	[[0.299227	[[-0.04433	[[0.007043	[[0.151024	[[0.16749	
new_data_5_datos_s_brachyee[[-4.30820[[20.5039[[44.5548[[7.03151[[-63.8865[[-38.7792	0.54790917	0.550947687	0.540838937	0.570218586	0.588342865	0.599697478	[[-0.00262	[[0.094242	[[0.149772	[[0.019424	[[-0.17068	[[-0.09138	[[0.320451	[[0.241151	[[-0.07929	[[0.074811	[[0.110806	[[0.08875	new_data_5_datos_s_brachyee[[6.13758[[4.47492[[47.7614[[7.04805[[-59.3709	[[-39.6501	0.604303799	0.643742235	0.600923714	0.552080759	0.596999539	0.642313688	[[0.042105	[[0.021451	[[0.133176	[[0.022715	[[-0.17628	[[-0.08470	[[0.309454	[[0.217877	[[-0.09157	[[-0.00126	[[0.107422	[[0.12681
new_data_5_datos_s_brachyee[[26.1845[[-2.36016[[4.17576[[1.223339[[-65.6296[[-3.96783	0.559283118	0.614214458	0.591385973	0.628972488	0.56948881	0.625352407	[[0.146926	[[0.009276	[[0.107306	[[0.113234	[[-0.16151	[[-0.10091	[[0.268816	[[0.208213	[[-0.06060	[[-0.10396	[[0.214147	[[0.24783	new_data_5_datos_s_brachyee[[25.7174[[-3.7678[[38.0481[[12.7955[[-65.2495	[[-41.9692	0.582765498	0.643631382	0.610905422	0.562599641	0.595012751	0.617700993	[[0.143339	[[-0.01849	[[0.094286	[[0.064386	[[-0.16782	[[-0.10323	[[0.262116	[[0.197522	[[-0.06459	[[-0.08287	[[0.167616	[[0.24657
new_data_5_datos_s_brachyee[[36.60779[[77.8179[[34.7258[[14.6138[[67.6280[[34.7425	0.581815166	0.626450016	0.611588248	0.54670453	0.601888247	0.542173277	[[0.137883	[[0.091742	[[0.069056	[[0.165248	[[0.189718	[[0.257237	[[0.188075	[[0.166810	[[0.057885	[[0.157688	[[0.17319	new_data_5_datos_s_brachyee[[36.60779[[77.8179[[34.7258[[14.6138[[67.6280[[34.7425	0.581815166	0.626450016	0.611588248	0.54670453	0.601888247	0.542173277	[[0.137883	[[0.091742	[[0.069056	[[0.165248	[[0.189718	[[0.257237	[[0.188075	[[0.166810	[[0.057885	[[0.157688	[[0.17319			

Tabla 1(b) - Excel

En la Figura 11 se muestra el flujo de trabajo efectuado en la preparación y el procesamiento de datos antes de su introducción en un modelo de aprendizaje profundo. Fichero usado, mirar Figura 28: 03_modelo_tfm+.py.

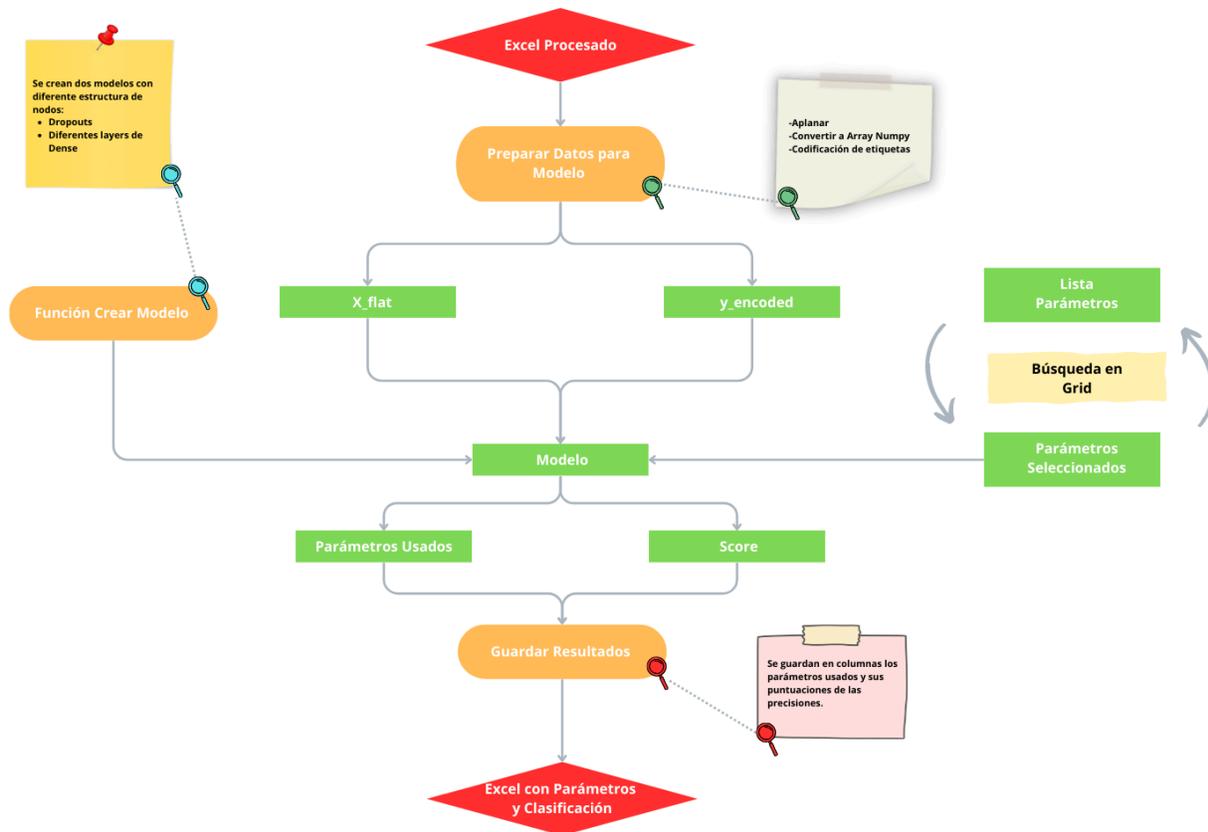


Figura 11 - Datos para el entrenamiento del modelo de aprendizaje

Desarrollo de los modelos de la red neuronal

Los datos se separaron en las variables que explicarán el modelo (x_flat) y las labels ($y_encoded$), las predicciones de lo que es el dato correcto. A pesar de que los datos ya estaban preparados, se tuvo que aplanar las listas de los datos de entrada y la codificación de las etiquetas para que el modelo las aceptara.

La codificación de las etiquetas es darle un valor numérico en vez de cadena de texto para una multclasificación.

Se desarrollaron varios modelos de red neuronal utilizando diferentes parámetros para identificar la mejor arquitectura y los mejores hiperparámetros posibles. El tipo de red neuronal es la red neuronal secuencial multicapa, como un Perceptrón Multicapa (MLP), consiste en capas ordenadas de forma secuencial: una capa de entrada, una o más capas ocultas y una capa de salida. La capa de entrada recibe los datos, las ocultas procesan la información con funciones de activación no lineales y la capa de salida genera la predicción final.

Modelo 1:

El primer modelo es una red neuronal secuencial implementada con Keras como se muestra en la Figura 12. Comienza con una capa de entrada que acepta un vector de características de tamaño `X_flat.shape(1)`. Sigue con una primera capa densa de 64 neuronas utilizando la función de activación ReLU, con restricción de norma máxima (MaxNorm) y regularización L2. Si la tasa de dropout es mayor a cero, se añade una capa de Dropout para prevenir el sobreajuste. A continuación, se incluye una segunda capa densa con 32 neuronas, también con activación ReLU, y las mismas técnicas de regularización. De nuevo, se agrega una capa de Dropout si se requiere. Finalmente, la capa de salida tiene tantas neuronas como clases de salida, utilizando la función de activación softmax para la clasificación multiclase. El modelo se compila con un optimizador específico, la pérdida categórica de entropía cruzada, y la métrica de precisión para evaluar el rendimiento.

```

model = Sequential()
model.add(Input(shape=(X_flat.shape[1],)))
model.add(Dense(64, activation='relu',
                kernel_constraint=MaxNorm(weight_constraint),
                kernel_regularizer=l2(l2_reg)))
if dropout_rate > 0.0:
    model.add(Dropout(dropout_rate))
model.add(Dense(32, activation='relu',
                kernel_constraint=MaxNorm(weight_constraint),
                kernel_regularizer=l2(l2_reg)))
if dropout_rate > 0.0:
    model.add(Dropout(dropout_rate))
model.add(Dense(len(label_encoder.classes_), activation='softmax'))
model.compile(optimizer=optimizer, loss='categorical_crossentropy', metrics=
['accuracy'])

```

Figura 12 - Modelo primero

Modelo 2:

El segundo modelo de deep learning mostrado en la Figura 13 es una red neuronal secuencial, también implementada con Keras, pero con una arquitectura ligeramente más compleja en comparación con el modelo anterior. Este modelo comienza con una capa de entrada que toma un vector de características de tamaño `X_flat.shape(1)`. A continuación, se agrega una primera capa densa con 128 neuronas, utilizando la función de activación ReLU, con restricciones de norma máxima (MaxNorm) y regularización L2. Le sigue una segunda capa densa con 64 neuronas, nuevamente con activación ReLU y las mismas técnicas de regularización. Después, si la tasa de dropout es mayor que cero, se añade una capa de Dropout para mitigar el sobreajuste.

El modelo continúa con una tercera capa densa de 32 neuronas con activación ReLU, manteniendo las restricciones y la regularización anteriores. Otra capa de Dropout se agrega si se requiere. Finalmente, la capa de salida utiliza la función de activación softmax con tantas neuronas como clases, proporcionando una salida adecuada para problemas de clasificación

multiclase. El modelo se compila utilizando un optimizador específico, la función de pérdida de entropía cruzada categórica (`categorical_crossentropy`), y la métrica de precisión para evaluar su desempeño. Este modelo, al tener más capas y neuronas, está diseñado para capturar características más complejas en los datos de entrada, como se muestra en la Figura 13..

```
model = Sequential()
model.add(Input(shape=(X_flat.shape[1],)))
model.add(Dense(128, activation='relu',
kernel_constraint=MaxNorm(weight_constraint), kernel_regularizer=l2(12_reg)))
model.add(Dense(64, activation='relu',
kernel_constraint=MaxNorm(weight_constraint), kernel_regularizer=l2(12_reg)))
if dropout_rate > 0.0:
    model.add(Dropout(dropout_rate))
model.add(Dense(32, activation='relu',
kernel_constraint=MaxNorm(weight_constraint), kernel_regularizer=l2(12_reg)))
if dropout_rate > 0.0:
    model.add(Dropout(dropout_rate))
model.add(Dense(len(label_encoder.classes_), activation='softmax'))
model.compile(optimizer=optimizer, loss='categorical_crossentropy', metrics=
['accuracy'])
```

Figura 13 - Modelo segundo

Búsqueda de hiperparámetros

Se realizó una búsqueda en grid para encontrar los mejores hiperparámetros, valores configurados antes del entrenamiento del modelo que controla el proceso de aprendizaje, y se guardaron los resultados en un archivo Excel para su análisis.

Una búsqueda en grid consiste en la prueba de todas las combinaciones posibles de la lista aportada, Figura 14, lo que añadiendo un solo parámetro puede elevar enormemente las veces que tendría que evaluar con esos parámetros.

```
param_grid = {
    'model__optimizer': [RMSprop(learning_rate=0.001), Nadam(learning_rate=0.001)],
    'model__dropout_rate': [0.0],
    'model__weight_constraint': [2],
    'model__l2_reg': [0.0, 0.001],
    'batch_size': [64],
    'epochs': [50, 100, 150]
}
```

Figura 14 - Combinaciones de hiperparámetros

Se seleccionaron los parámetros que daban un mejor resultado en el modelo como se ve en la Figura 15. En esta figura se representan los valores óptimos que maximizan el rendimiento del modelo. Estos hiperparámetros incluyen el número de neuronas, las tasas de aprendizaje y los métodos de regularización, entre otros. La importancia de estos parámetros radica en su influencia directa sobre la precisión y la capacidad del modelo para generalizar. En concreto, la figura ayuda a visualizar el impacto de cada configuración, permitiendo seleccionar la que ofrece el mejor compromiso entre la capacidad de ajuste y la evitación del sobreajuste

<u>mean_test_score</u>	<u>std_test_score</u>	<u>rank_test_score</u>	<u>param_model_optimizer_name</u>
0,87705930722582	0,00887619676714482		1 <u>Nadam</u>
0,874470826052695	0,00958656398261482		2 <u>Nadam</u>
0,872432761160791	0,00690447790704092		3 <u>Nadam</u>
0,824107308048104	0,0346266340010237		4 <u>Nadam</u>
0,808178983794155	0,112224269982309		5 <u>RMSprop</u>
0,803555178675438	0,0910012498958975		6 <u>RMSprop</u>
0,797631137150101	0,0349764041033087		7 <u>Nadam</u>
0,79301367046973	0,0592615896076203		8 <u>RMSprop</u>
0,792453489567273	0,046732304563684		9 <u>RMSprop</u>
0,786704868605886	0,0434026968267678		10 <u>Nadam</u>

Figura 15 - Mejores conjuntos de hiperparámetros

Validación y evaluación

El diagrama en la Figura 16 muestra el proceso de división de datos, entrenamiento, validación, y evaluación del modelo utilizando diferentes semillas y porcentajes de datos , véase Tabla 2, para garantizar un rendimiento robusto y evitar el sobreajuste. Fichero usado, mirar Figura 28: 04_05_asimarrodev_evpyb.

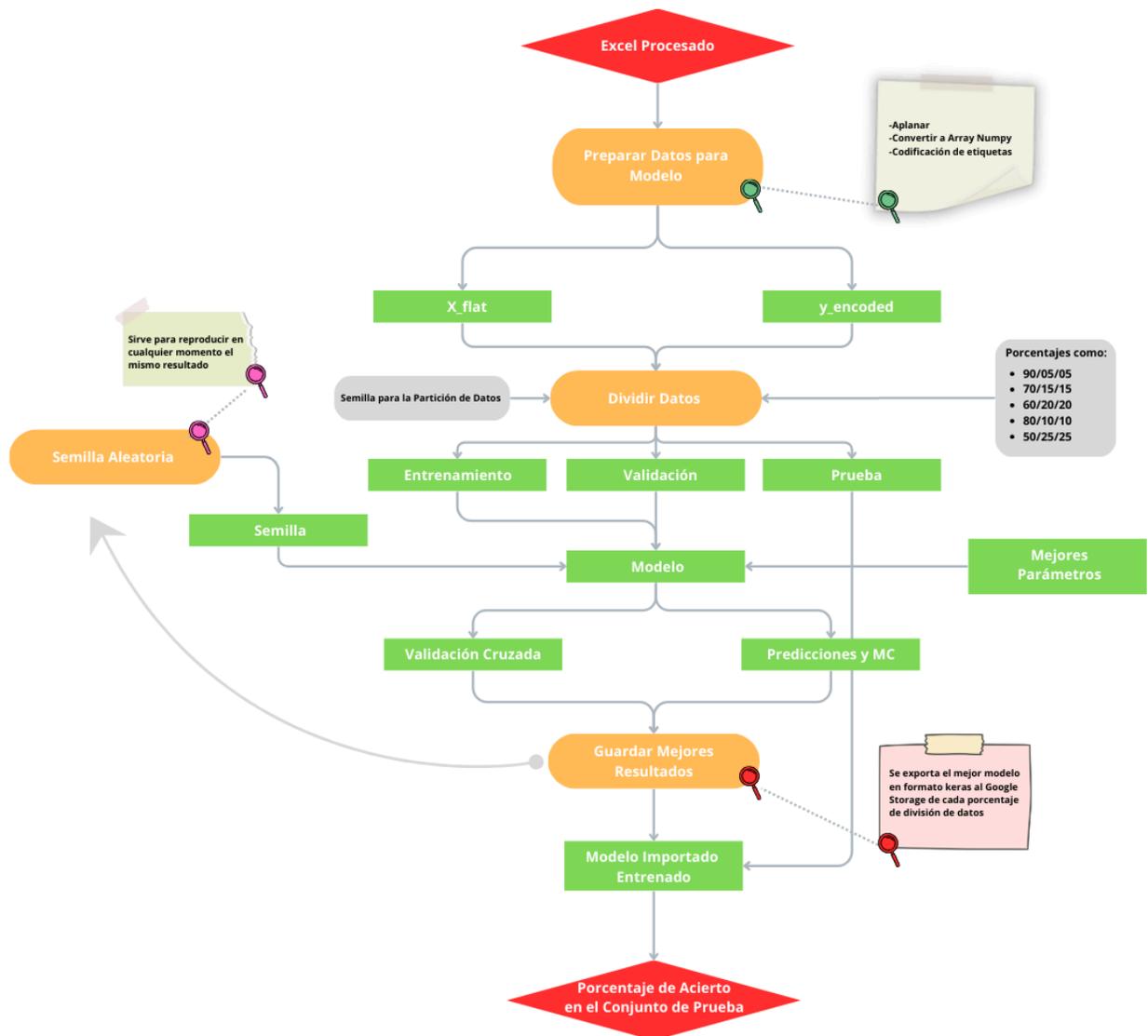


Figura 16 - Organigrama de validación y evaluación

Una vez identificados la estructura y los hiperparámetros que mejor definían el modelo, los datos se dividieron en tres conjuntos: entrenamiento, validación y prueba. Para garantizar la validez del modelo, evitar el problema del sobreajuste y así evaluar con precisión y exactitud su rendimiento.

Este modelo se utilizó repetidamente para encontrar el mejor ajuste de pesos, ya que, aunque el modelo sea el mismo, la distribución de pesos varía en cada iteración lo que conduce a diferentes resultados. En consecuencia, se desarrolló un código que ejecuta un bucle, guarda los resultados y la semilla correspondiente, permitiendo reproducir los mismos resultados en futuras ejecuciones. Para garantizar la obtención del mejor modelo posible, se realizaron cinco

iteraciones para cada combinación de porcentajes (%), lo que dio lugar a un total de cinco modelos.

Entrenamiento	Validacion	Prueba
90	05	05
70	15	15
60	20	20
80	10	10
50	25	25

Tabla 2 - Combinaciones de porcentajes

Para la construcción del modelo se utilizó la semilla 42. Sin embargo, para la evaluación del modelo se utilizaron semillas diferentes. Esta decisión se tomó para evitar cualquier posible sesgo que pudiera introducirse debido a la misma distribución de los datos utilizada durante la fase de entrenamiento. El conjunto de prueba compuesto por datos que el modelo nunca ha visto durante el entrenamiento, se utilizó para la evaluación final del rendimiento del modelo. Al emplear diferentes semillas para la evaluación, se garantiza que el modelo se prueba en múltiples escenarios y no se sobreajusta a un único conjunto de datos de prueba.

Los modelos que se usaron no estaban entrenados, es decir, solo tenían la arquitectura y los hiperparámetros seleccionados anteriormente. El proceso de entrenamiento se realizó utilizando los conjuntos de datos de entrenamiento y validación, mientras que el conjunto de prueba se reservó para la evaluación del modelo.

Por último, se cargó el modelo importado y se empleó el conjunto de prueba para la evaluación, garantizando así una medida precisa de su rendimiento en datos completamente nuevos. Este enfoque metodológico garantiza una evaluación rigurosa y objetiva.

Implementación y evaluación en entorno clínico

Implementación en Google Cloud

Debido a la alta demanda computacional y el requerimiento de tener el ordenador encendido, se utilizó Google Cloud para entrenar y evaluar los modelos, además de almacenar los archivos generados, todo esto en la nube. Lo que permitió aprovechar recursos adicionales y reducir los tiempos de procesamiento. Se usó Google Enterprise para mantener el código online y Google Storage para almacenar la información.

Evaluación clínica

Una buena propuesta para el futuro o continuación del estudio sería implementar el modelo en un entorno clínico para validar su utilidad práctica. Esta etapa incluye recoger feedback de profesionales de la salud y realizar ajustes necesarios para mejorar la precisión y la facilidad de uso del sistema.

5. Resultados

Los resultados obtenidos son 5 redes neuronales secuenciales multicapa (MLP), que se han probado con diferente porcentaje de datos (entrenamiento, validación y prueba). Se realizó un bucle de 50 iteraciones por cada porcentaje de manera que se pudiera determinar la semilla que daba los mejores resultados. En total se obtuvieron 250 resultados, de los que se guardan aquellos con la mayor precisión (porcentaje). Además, se ha generado un informe para cada uno de ellos donde nos indica la precisión media, la semilla para su reproducibilidad, la desviación estándar, la matriz de confusión y el informe de la clasificación.

Porcentajes	90/05/05	70/15/15	60/20/20	80/10/10	50/25/25
Precisión Media Modelos	90.14%	86.82	87.9	89.18	89.4

Tabla 3 - Porcentajes y precisiones

La matriz de confusión nos permite evaluar el rendimiento del modelo en la clasificación, donde muestra la relación entre las predicciones por el modelo y los valores reales de la clase para entender la distribución de los errores del modelo. Está compuesto por True Positives (TP), True Negatives (TN), False Positives (FP) y False Negatives (FN). A partir de esta matriz podemos obtener algunos aspectos claves para comprender mejor o poder llegar a diferenciar mejor modelos que a primera vista podrían parecer similares:

Matriz de Confusión		
Métrica	Descripción	Fórmula
Exactitud (Accuracy)	Es la proporción de predicciones correctas, Valores superiores a 0.90 son considerados buenos.	Exactitud = $(TP + TN) / (TP + TN + FP + FN)$
Precisión (Precision)	Indica cuántas de las predicciones positivas son realmente positivas. Valores altos mejores.	Precisión = $TP / (TP + FP)$
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	Indica cuántos de los casos positivos reales fueron capturados por el modelo. Los valores altos son mejores.	Sensibilidad = $TP / (TP + FN)$
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	Indica cuántos de los casos negativos reales fueron capturados por el modelo. Los valores altos son mejores.	Especificidad = $TN / (TN + FP)$
Tasa de Falsos Positivos (FPR)	Indica cuántos de los casos negativos reales fueron incorrectamente clasificados como positivos. Los valores bajos son mejores.	FPR = $FP / (FP + TN)$
Tasa de Falsos Negativos (FNR)	Indica cuántos de los casos positivos reales fueron incorrectamente clasificados como negativos. Los valores bajos son mejores.	FNR = $FN / (FN + TP)$
Puntaje F1 (F1 Score)	Es la media armónica de la precisión y la sensibilidad. Los valores altos son mejores.	F1 Score = $2 * (Precisión * Sensibilidad) / (Precisión + Sensibilidad)$

Tabla 4 - Definición métricas matriz de confusión

El reporte de clasificación proporciona una evaluación cuantitativa del rendimiento del modelo con métricas para cada clase.

Las clases están repartidas del 0 al 5, es decir, 6 clases que equivalen a las siguientes etiquetas:

Clase	Etiqueta
0	brachycephaly
1	dolichocephaly
2	normocephaly
3	plagiocephaly
4	towering
5	trigonocephaly

Tabla 5 - Clases codificadas

Modelo 90/05/05 - 90,14%

```

Mejor semilla: 789706
Resultados de la mejor validación cruzada: [0.8703703880310059, 0.9413580298423767, 0.8209876418113708, 0.9351851940155029, 0.9393004179000854]
Precisión media: 0.9014403343200683
Desviación estándar: 0.04817364672653798
Matriz de confusión:
[[145.2  1.4  0.6  1.2  6.8  7.6]
 [  2. 163.8  0.6  3.2  5.4  4.4]
 [  0.6  0.4 69.2  0.2  7.8  0.6]
 [  1.2  3.  0. 130.6  3.2  8.2]
 [  3.  3.6  4.  4. 194.2  1.4]
 [  5.8  4.2  0.6  3.4  7.4 173.2]]
Reporte de clasificación:
      precision    recall  f1-score   support

 0         0.95     0.94     0.94       163
 1         0.97     0.93     0.95       179
 2         0.94     0.92     0.93        79
 3         0.99     0.92     0.95       146
 4         0.91     0.96     0.94       211
 5         0.90     0.95     0.92       194

 accuracy          0.94
 macro avg         0.94     0.94     0.94       972
 weighted avg     0.94     0.94     0.94       972
  
```

Figura 17 - Modelo 90/05/05

Resultados validación cruzada:

- Precisión media: 0,9014
- Desviación estándar: 0,0482

Matriz de confusión porcentaje (%):

	0	1	2	3	4	5
0	89.19	0.86	0.37	0.74	4.18	4.67
1	1.11	90.80	0.33	1.77	2.99	2.44
2	0.74	0.50	85.64	0.25	9.65	0.74
3	0.80	2.01	0.00	87.53	2.14	5.50
4	1.40	1.68	1.87	1.87	90.66	0.65
5	2.91	2.10	0.30	1.70	3.71	86.77

Tabla 7 - Matriz de confusión porcentaje 90/05/05

En este modelo se presentan dificultades en la discriminación de algunas clases. Se observa que un 4,67% de los elementos de la clase 0 fueron erróneamente clasificados como la clase 5. Asimismo, un 9,65% de los elementos de la clase 2 se confundieron con la clase 4. La clase 3 muestra confusión significativa con la clase 5 con un 5,50% de los elementos de la clase 3. Recíprocamente, la clase 5 también presenta una equivocación del 3,71% de sus elementos con la clase 4 y un 2,91% con la clase 0.

Métricas adicionales MC (Adaptación Figura 17 y cálculo nuevas métricas):

Matriz de Confusión	
Métrica	Valor
Exactitud (Accuracy)	0.940
Precisión (Precision)	0.940
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	0.940
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	0.982
Tasa de Falsos Positivos (FPR)	0.018
Tasa de Falsos Negativos (FNR)	0.060
Puntaje F1 (F1 Score)	0.940

Tabla 8 - Métricas mc 90/05/05

Reporte de clasificación:

```
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.95       0.94       0.94       163
   1       0.97       0.93       0.95       179
   2       0.94       0.92       0.93        79
   3       0.99       0.92       0.95       146
   4       0.91       0.96       0.94       211
   5       0.90       0.95       0.92       194

 accuracy          0.94       0.94       0.94       972
 macro avg         0.94       0.94       0.94       972
 weighted avg      0.94       0.94       0.94       972
```

Figura 18 - Reporte clasificación 90/05/05

La precisión y recall generales son cercanos o superiores a 0.90 para todas las clases, la clase 3 muestra la mayor precisión del 0.99 y la exactitud general es del 94%.

70/15/15 - 86,82%

```
Mejor semilla: 928062
Resultados de la mejor validación cruzada: [0.8095238208770752, 0.8108465671539307, 0.9563491940498352, 0.8558201193809509, 0.9086092710494995]
Precisión media: 0.8682297945022583
Desviación estándar: 0.057076714796248824
Matriz de confusión:
[[114.4  3.  0.6  1.2  6.8  4.4]
 [ 3.6 120.4  0.8  2.6  4.4  5.8]
 [ 1.6  0.6  52.8  0.  5.8  1.8]
 [ 1.4  3.8  0.6  93.8  3.4  6.4]
 [ 6.  3.  4.4  1.4 147.8  3.4]
 [ 5.6  5.8  1.8  3.6  6. 127. ]]
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.93       0.89       0.91       131
   1       0.91       0.94       0.92       137
   2       0.86       0.92       0.89        62
   3       0.94       0.88       0.91       110
   4       0.90       0.92       0.91       166
   5       0.89       0.91       0.90       149

 accuracy          0.91       0.91       0.91       755
 macro avg         0.91       0.91       0.91       755
 weighted avg      0.91       0.91       0.91       755
```

Figura 19 - Modelo 70/15/15

Resultados validación cruzada:

- Precisión media: 0,8682
- Desviación estándar: 0,0571

Matriz de confusión porcentaje (%):

	0	1	2	3	4	5
0	87.73	2.30	0.46	0.92	5.21	3.37
1	2.60	86.87	0.58	1.88	3.17	4.18
2	2.48	0.93	81.73	0.00	8.98	2.79
3	1.25	3.38	0.53	83.45	3.02	5.69
4	3.53	1.76	2.59	0.82	86.94	2.00
5	3.62	3.75	1.16	2.33	3.88	82.04

Tabla 10 - Matriz de confusión porcentaje 70/15/15

En este modelo se presentan dificultades en la discriminación de algunas clases. Se observa que un 5,21% de los elementos de la clase 0 fueron erróneamente clasificados como la clase 4. Asimismo, la clase 4 muestra confusiones notables del 3,53% con la clase 0.

Métricas adicionales MC (Adaptación Figura 19 y cálculo nuevas métricas)

Matriz de Confusión	
Métrica	Valor
Exactitud (Accuracy)	0.910
Precisión (Precision)	0.910
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	0.910
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	0.974
Tasa de Falsos Positivos (FPR)	0.026
Tasa de Falsos Negativos (FNR)	0.090
Puntaje F1 (F1 Score)	0.910

Tabla 11 - Métricas mc 70/15/15

Reporte de clasificación:

```
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.93       0.89       0.91       131
   1       0.91       0.94       0.92       137
   2       0.86       0.92       0.89        62
   3       0.94       0.88       0.91       110
   4       0.90       0.92       0.91       166
   5       0.89       0.91       0.90       149

 accuracy          0.91       0.91       0.91       755
 macro avg         0.91       0.91       0.91       755
 weighted avg      0.91       0.91       0.91       755
```

Figura 20 - Reporte clasificación 70/15/15

La precisión y recall son buenos, pero ligeramente inferiores al modelo con 90,14% de precisión. La exactitud general es del 91%.

60/20/20 - 87,9%

```
Mejor semilla: 976486
Resultados de la mejor validación cruzada: [0.8070987462997437, 0.8441358208656311, 0.9027777910232544, 0.9228395223617554, 0.9182098507881165]
Precisión media: 0.8790123462677002
Desviación estándar: 0.04562893812901879
Matriz de confusión:
[[ 97.6  1.  0.8  0.4  4.2  6.6]
 [  3.4 100.2  1.2  3.  3.4  4.6]
 [  1.6  1.  42.8  0.  5.  2.4]
 [  2.  3.  1.2  87.8  2.4  4. ]
 [  2.6  2.8  3.  0.6 125.2  3.4]
 [  3.8  2.6  1.4  2.4  4.6 116. ]]
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.91       0.93       0.92       111
   1       0.88       0.94       0.91       116
   2       0.94       0.87       0.90        53
   3       0.95       0.93       0.94       100
   4       0.93       0.91       0.92       137
   5       0.92       0.91       0.92       131

 accuracy          0.92       0.92       0.92       648
 macro avg         0.92       0.91       0.92       648
 weighted avg      0.92       0.92       0.92       648
```

Figura 21 - Modelo 60/20/20

Resultados validación cruzada:

- Precisión media: 0,8790
- Desviación estándar: 0,0456

Matriz de confusión porcentaje (%):

	0	1	2	3	4	5
0	88.25	0.90	0.72	0.36	3.80	5.97
1	2.91	85.79	1.03	2.57	2.91	3.94
2	2.92	1.82	78.10	0.00	9.12	4.38
3	1.93	2.90	1.16	84.91	2.32	3.87
4	1.84	1.98	2.12	0.42	88.42	2.40
5	2.80	1.91	1.03	1.77	3.39	85.42

Tabla 13 - Matriz de confusión porcentaje 60/20/20

En este modelo se presentan dificultades en la discriminación de algunas clases. Se observa que un 3,80% de los elementos de la clase 0 fueron erróneamente clasificados como la clase 4, y un 5,97% se confundieron con la clase 5. Asimismo, la clase 1 muestra confusiones del 3,94% con la clase 5. Por otro lado, la clase 2 se confunde en un 9,12% con la clase 4.

Métricas adicionales MC (Adaptación Figura 21 y cálculo nuevas métricas)

Matriz de Confusión	
Métrica	Valor
Exactitud (Accuracy)	0.920
Precisión (Precision)	0.920
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	0.920
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	0.978
Tasa de Falsos Positivos (FPR)	0.022
Tasa de Falsos Negativos (FNR)	0.080
Puntaje F1 (F1 Score)	0.920

Tabla 14 - Métricas mc 60/20/20

Reporte de clasificación:

```
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.91      0.93      0.92     111
   1       0.88      0.94      0.91     116
   2       0.94      0.87      0.90      53
   3       0.95      0.93      0.94     100
   4       0.93      0.91      0.92     137
   5       0.92      0.91      0.92     131

 accuracy          0.92     648
 macro avg         0.92     648
 weighted avg     0.92     648
```

Figura 22 - Reporte clasificación 60/20/20

La precisión y recall son consistentes y buenos en las clases 0, 3 y 5. La exactitud general es del 92%.

80/10/10 - 89,18%

```
Mejor semilla: 13216
Resultados de la mejor validación cruzada: [0.8472222089767456, 0.9224537014961243, 0.9953703880310059, 0.8611111044883728, 0.8333333134651184]
Precisión media: 0.8918981432914734
Desviación estándar: 0.06002294182457724
Matriz de confusión:
[[117.6  2.8  2.6  1.8  3.4 13.4]
 [ 1.2 143.8  0.8  4.8  2.2  6.4]
 [ 0.4  1.4  63.8  0.2  4.2  1. ]
 [ 0.8  3.  0.6 119.2  1.  4.8]
 [ 2.2  3.6  4.8  2.4 172.4  4.2]
 [ 3.2  4.2  1.6  5.4  5. 153.8]]
Reporte de clasificación:
      precision    recall  f1-score   support

   0       0.88      0.82      0.85     141
   1       0.82      0.88      0.85     159
   2       0.79      0.75      0.77      71
   3       0.84      0.83      0.84     130
   4       0.84      0.85      0.85     190
   5       0.80      0.82      0.81     173

 accuracy          0.83     864
 macro avg         0.83     864
 weighted avg     0.83     864
```

Figura 23 - Modelo 80/10/10

Resultados validación cruzada:

- Precisión media: 0,8919
- Desviación estándar: 0,06

Matriz de confusión porcentaje (%):

	0	1	2	3	4	5
0	83.05	1.98	1.84	1.27	2.40	9.46
1	0.75	89.76	0.50	3.00	1.37	4.00
2	0.55	1.92	87.40	0.27	5.75	1.37
3	0.60	2.27	0.45	90.03	0.76	3.63
4	1.14	1.86	2.48	1.24	89.05	2.17
5	1.80	2.36	0.90	3.03	2.81	86.31

Tabla 16 - Matriz de confusión porcentaje 80/10/10

En este modelo se presentan dificultades en la discriminación de algunas clases. Se observa que un 9,46% de los elementos de la clase 0 fueron erróneamente clasificados como la clase 5. Asimismo, la clase 1 muestra confusiones del 3,00% con la clase 3 y del 4% con la clase 5. Por otro lado, la clase 5 se confunde en un 3,03% con la clase 3 y en un 2,81% con la clase 4.

Métricas adicionales MC (Adaptación Figura 23 y cálculo nuevas métricas)

Matriz de Confusión	
Métrica	Valor
Exactitud (Accuracy)	0.830
Precisión (Precision)	0.830
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	0.820
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	0.963
Tasa de Falsos Positivos (FPR)	0.037
Tasa de Falsos Negativos (FNR)	0.180
Puntaje F1 (F1 Score)	0.830

Tabla 17 - Métricas mc 80/10/10

Reporte de clasificación:

```
Reporte de clasificación:
      precision    recall  f1-score   support

0         0.88        0.82        0.85        141
1         0.82        0.88        0.85        159
2         0.79        0.75        0.77         71
3         0.84        0.83        0.84        130
4         0.84        0.85        0.85        190
5         0.80        0.82        0.81        173

accuracy          0.83        864
macro avg         0.83        0.82        0.83        864
weighted avg     0.83        0.83        0.83        864
```

Figura 24 - Reporte clasificación 80/10/10

La precisión y recall son altos en general, pero la clase 0 tiene una precisión notablemente mas baja que el resto (0.88). La exactitud general es del 83%.

50/25/25 - 89,4%

```
Mejor semilla: 795917
Resultados de la mejor validación cruzada: [0.8129629492759705, 0.9759259223937988, 0.9018518328666687, 0.8981481194496155, 0.8814814686775208]
Precisión media: 0.8940740585327148
Desviación estándar: 0.05198923874726118
Matriz de confusión:
[[ 75.4  2.4  0.6  0.6  2.6  4. ]
 [  0.8 91.6  0.4  1.4  1.8  3.6]
 [  2.6  1.  35.8  0.6  3.8  0.4]
 [  1.4  3.  0.  76.8  1.8  2.2]
 [  2.8  2.  2.6  0.4 106.8  1.4]
 [  2.  4.6  0.8  2.6  3.  96.4]]
Reporte de clasificación:
      precision    recall  f1-score   support

0         0.84        0.87        0.86         85
1         0.88        0.88        0.88        100
2         1.00        0.75        0.86         44
3         0.89        0.89        0.89         85
4         0.87        0.94        0.90        116
5         0.88        0.87        0.88        110

accuracy          0.88        540
macro avg         0.89        0.87        0.88        540
weighted avg     0.88        0.88        0.88        540
```

Figura 25 - Modelo 50/25/25

Resultados validación cruzada:

- Precisión media: 0,8941
- Desviación estándar: 0,052

Matriz de confusión porcentaje (%):

	0	1	2	3	4	5
0	88.08	2.80	0.70	0.70	3.04	4.67
1	0.80	91.05	0.40	1.39	1.79	3.58
2	5.63	2.16	77.49	1.30	8.23	0.87
3	1.59	3.40	0.00	87.07	2.04	2.49
4	2.33	1.67	2.17	0.33	89.00	1.17
5	1.75	4.02	0.70	2.27	2.62	84.27

Tabla 19 - Matriz de confusión porcentaje 50/25/25

En este modelo se presentan dificultades en la discriminación de algunas clases. Se observa que un 3,58% de los elementos de la clase 1 fueron erróneamente clasificados como la clase 5. Por otro lado, la clase 5 se confunde en un 4,02% con la clase 1.

Métricas adicionales MC (Adaptación Figura 25 y cálculo nuevas métricas)

Matriz de Confusión	
Métrica	Valor
Exactitud (Accuracy)	0.880
Precisión (Precision)	0.880
Sensibilidad (Recall) o Tasa de Verdaderos Positivos (TPR)	0.870
Especificidad (Specificity) o Tasa de Verdaderos Negativos (TNR)	0.968
Tasa de Falsos Positivos (FPR)	0.032
Tasa de Falsos Negativos (FNR)	0.130
Puntaje F1 (F1 Score)	0.880

Tabla 20 - Métricas mc 50/25/25

Reporte de clasificación:

```
Reporte de clasificación:
      precision    recall  f1-score   support

0         0.84        0.87        0.86         85
1         0.88        0.88        0.88        100
2         1.00        0.75        0.86         44
3         0.89        0.89        0.89         85
4         0.87        0.94        0.90        116
5         0.88        0.87        0.88        110

accuracy                0.88        540
macro avg              0.89        0.87        0.88        540
weighted avg          0.88        0.88        0.88        540
```

Figura 26 - Reporte clasificación 50/25/25

La precisión y recall son buenos, especialmente en la clase 2 pero con un recall menor. La exactitud del modelo es del 88%.

Precisión del conjunto de prueba

La precisión del conjunto de pruebas se realiza para evaluar el rendimiento del modelo con datos que no fueron utilizados durante el entrenamiento. Para garantizar una evaluación realista y evitar sobre ajustes (cabecera columna) se ha utilizado diferentes semillas (cabecera fila) para la partición de los datos en lugar de usar la misma semilla que se empleó para construir los modelos.

	24	83	45	43
90/05/05	98.08	97.93	97.11	97.41
70/15/15	94.74	92.97	93.41	94.6
60/20/20	92.38	91.34	91.41	91.49
80/10/10	94.97	94.37	93.93	94.74
50/25/25	90.6	89.42	89.05	89.05

Tabla 21 - Precisiones conjuntos de prueba

6. Presupuesto

Se ha usado para el cálculo del presupuesto las horas dedicadas, siguiendo la tabla de la Figura 27, y la suscripción a la plataforma en la nube Google Cloud.

Niveles		Tabla salarial según art. 33		Plus convenio anual según art. 38 convenio	Total anual
		Mes x 14	Anual		
1	LICENCIADOS Y TITULADOS 2.º Y 3.º CICLO UNIVERSITARIO Y ANALISTA	1.712,42	23.973,88	2.349,69	26.323,57
2	DIPLOMADOS Y TITULADOS 1.º CICLO UNIVERSITARIO. JEFE SUPERIOR	1.291,04	18.074,56	2.349,69	20.424,25
3	TÉCNICO DE CÁLCULO O DISEÑO, JEFE DE 1.º Y PROGRAMADOR DE ORDENADOR	1.244,93	17.429,02	2.349,69	19.778,71
4	DELINEANTE-PROYECTISTA, JEFE DE 2.º Y PROGRAMADOR DE MAQ. AUXILIARES	1.141,36	15.979,04	2.349,69	18.328,73
5	DELINEANTE, TÉCNICO DE 1.º, OFICIAL 1.º ADMTVO. Y OPERADOR DE ORDENADOR	1.019,82	14.277,48	2.349,69	16.627,17
6	DIBUJANTE, TÉCNICO DE 2.º, OFICIAL 2.º ADMTVO., PERFORISTA, GRABADOR Y CONSERJE	878,63	12.300,82	2.349,69	14.650,51
7	TELEFONISTA-RECEPCIONISTA, OFICIAL 1.º OFICIOS VARIOS Y VIGILANTE	849,16	11.888,24	2.349,69	14.237,93
8	AUXILIAR TÉCNICO, AUX. ADMTVO., TELEFONISTA, ORDENANZA, PERSONAL DE LIMPIEZA Y OFICIAL 2.º OFICIOS VARIOS	790,36	11.065,04	2.349,69	13.414,73
9	AYUDANTE OFICIOS VARIOS	757,29	10.602,06	2.349,69	12.951,75

Figura 27 - Tabla salarial del convenio de ingenierías y oficinas técnicas

Sueldo base anual	Sueldo bruto	Coste seguridad social	Coste total anual de empresa	Coste por hora del ingeniero en geomática y topografía
18074,56€	20424,25€	8169,70€	28593,95€	16,25€

Tabla 22 - Costes salariales

El proyecto ha durado un total de 550 horas en total siendo el coste 8937,5€, al que le sumamos la suscripción a Google Cloud y Google Colab pro incrementa a 8957,5€.

7. Discusión

Este trabajo ha demostrado la eficacia del uso del algoritmo K-means para la clusterización de nubes de puntos 3D en la clasificación de malformaciones craneales. La elección de K-means fue adecuada debido a la estructura relativamente esférica de las superficies craneales. A diferencia de otros métodos como DBSCAN, que son más útiles para clústeres irregulares y densidades variables, K-means ofrece una clusterización precisa y coherente (Ding & He, 2004).

El análisis de concavidad permite obtener variables robustas que pueden resaltar características geométricas de las distintas malformaciones craneales que otras métricas podrían pasar por alto (Guo et al., 2020). Este enfoque es particularmente útil en la clasificación de formas geométricas complejas donde la concavidad puede indicar irregularidades. A diferencia de métricas más simples el análisis de concavidad proporciona una medida más detallada de la geometría de la superficie, lo que puede mejorar la precisión y la capacidad de los modelos de aprendizaje automático al distinguir entre diferentes tipos de malformaciones (Guo et al., 2020).

El análisis de los porcentajes de división de datos para entrenamiento, validación y prueba, se utilizaron cinco porcentajes diferentes (), como se observa en la Tabla 2. Cada configuración afecta a los resultados en cuanto a la precisión media y desviación estándar media, demostrando como la distribución de los datos en entrenamiento, validación y prueba afectan a la consistencia y rendimiento general del modelo.

El modelo 90/05/05 ha resultado ser el más eficaz en este estudio, mostrando la mayor precisión media (90.14%) y un recall elevado, además de una menor tasa de falsos positivos y negativos. La baja desviación estándar (0.0482) indica una consistencia sólida en los resultados, como se observa en la Tabla 7. Este modelo ha logrado la máxima precisión para cada semilla en la división de conjunto de datos, lo que lo convierte en el modelo más adecuado para la minimización de los errores y la correcta clasificación de los pacientes en la mayoría de los casos. Consiguiendo una precisión máxima en el conjunto de prueba del 98,08%, véase Tabla 21.

Finalmente, el modelo de aprendizaje profundo desarrollado mostró una precisión elevada superior al 90% en la mayoría de las configuraciones probadas, Tabla 21. Este es un resultado competitivo en comparación con estudios previos que utilizaron redes neuronales convolucionales (CNN) en imágenes 2D, donde las precisiones típicas varían entre el 85% y el 90% (Qi et al., 2017). Las redes neuronales, como las implementadas en este estudio, tienen la capacidad de trabajar directamente con datos tridimensionales como nubes de puntos, lo que les permite capturar relaciones espaciales más complejas que las CNN 2D, limitadas a datos en dos dimensiones. Al procesar información 3D directamente, estos modelos pueden explotar características geométricas adicionales, como la profundidad y la forma completa de un objeto, lo que mejora la precisión y la generalización del modelo. En tareas donde la estructura tridimensional es crítica, como la clasificación de malformaciones craneales, las redes

neuronales 3D suelen superar a las CNN 2D, que podrían perder información relevante al proyectar los datos en un espacio 2D (Qi et al., 2017).

Para validar completamente la aplicabilidad del modelo en un entorno clínico, sería necesario realizar pruebas adicionales con datos más diversos. El análisis futuro también podría explorar cómo diferentes métodos de análisis geométrico, más allá de la concavidad, podrían influir en la precisión y robustez de la clasificación.

8. Conclusiones

De acuerdo a los objetivos planeados, se ha conseguido desarrollar un modelo robusto y preciso de aprendizaje profundo para la clasificación de malformaciones craneales a partir de nubes de puntos 3D generadas a partir de soluciones fotogramétricas con teléfono móvil. El proceso implicó el preprocesamiento de datos, normalización y la utilización de técnicas avanzadas como el algoritmo K-Means para la clusterización, lo que permitió un análisis detallado de la concavidad de las áreas del cráneo.

La preparación de los datos incluyó la clonación de la estructura de carpetas, el procesamiento y mallado de las nubes de puntos, la normalización de las superficies y la segmentación basada en concavidad. Posteriormente se prepararon para la introducción al modelo.

Los resultados, como se muestra en la Tabla 21 y la Figura 17, mostraron que el modelo con una partición de datos del 90/05/05 presentó la mayor precisión media del 90.14%, la mayor exactitud y una baja desviación estándar, lo cual indica una alta consistencia en los resultados. Se realizó una clasificación con el conjunto de prueba obteniendo más de 97% de precisión en todos los conjuntos de partición de datos, como se detalla en la Tabla 22. Además, la implementación en Google Cloud permitió una mejor gestión de recursos computacionales, acelerando el proceso de entrenamiento y evaluación.

Observando la Tabla 7 podemos concluir que las clases que tienden a confundirse más entre sí son aquellas cuyas características geométricas son más similares. Por ejemplo, la clase 2 (Normocephaly) muestra una confusión con la clase 4 (Towering) en un 9,65% de los casos, y la clase 0 (Brachycephaly) se confunde con la clase 5 (Trigonocephaly) en un 4,67%. Estos errores sugieren que ciertas malformaciones con características geométricas parecidas podrían beneficiarse de un análisis más fino para mejorar aún más la precisión del modelo.

Sin embargo, es fundamental destacar que, a pesar de estas confusiones en algunos casos específicos todas las clases superan el 85% de precisión en la clasificación. Esta alta tasa de acierto demuestra la eficacia del modelo en manejar una tarea tan compleja como la clasificación de malformaciones craneales, especialmente considerando que algunas de estas condiciones presentan diferencias geométricas sutiles.

Además, algunas clases presentan una altísima precisión y especificidad, como la clase 3 (Plagiocephaly), que alcanza un 87,53% de acierto y una baja tasa de falsos positivos y negativos. Estos resultados validan la robustez y fiabilidad del modelo.

En resumen, el modelo ha demostrado ser altamente eficaz, logrando tasas de acierto que superan el 85% en todas las clases, y mostrando un rendimiento especialmente sólido en la clasificación de malformaciones más distintivas. Esto subraya su potencial para mejorar significativamente el diagnóstico y tratamiento de malformaciones craneales en la práctica médica.

A la luz de los resultados presentados en las Tablas 21 y Figura 17, se concluye que el modelo desarrollado no solo es capaz de clasificar correctamente las malformaciones craneales en la mayoría de los casos, sino que también puede integrarse eficazmente en la práctica clínica diaria, proporcionando diagnósticos rápidos y precisos. Este avance tiene el potencial de transformar el tratamiento de las malformaciones craneales, mejorando los resultados clínicos para los pacientes y contribuyendo significativamente al campo del aprendizaje profundo aplicado a la medicina.

Propuestas de mejoras

Uno de los aspectos que se podría mejorar sería un modelo con una fuente de datos mayor y más diverso incluyendo diferentes etnias, edades y tipos de malformaciones craneales. Esto permitiría al modelo generalizar y reducir posibles sesgos. Además, la validación del modelo en entornos clínicos reales ayudaría a recoger feedback para identificar áreas de mejora y ajustar el sistema para su uso. Por último, optimizar y consolidar los diferentes códigos desarrollados en este proyecto en un solo sistema o software fácil e intuitivo con el que los profesionales de la salud pudiesen usar.

9. Referencias

- [1] Foster, J., Ahluwalia, R., Sherburn, M., Kelly, K., Sellyn, G. E., Kiely, C., ... & Bonfield, C. M. (2020). Pediatric cranial deformations: Demographic associations. *Journal of Neurosurgery: Pediatrics*, 26(4), 415-420.
- [2] Struck, R., Cordoni, S., Aliotta, S., Pérez-Pachón, L., & Gröning, F. (2019). Application of photogrammetry in biomedical science. *Biomedical Visualisation: Volume 1*, 121-130.
- [3] Mas, M. J. (2021). Plagiocefalia, deformidad de la cabeza. *Neuronas en crecimiento*
- [4] Quispe-Enriquez, O. C., Valero-Lanzuela, J. J., & Lerma, J. L. (2023). Craniofacial 3D Morphometric Analysis with Smartphone-Based Photogrammetry. *Sensors*, 24(1), 230.
- [5] Barbero-García, I., Lerma, J. L., Miranda, P., & Marqués-Mateu, Á. (2019). Smartphone-based photogrammetric 3D modelling assessment by comparison with radiological medical imaging for cranial deformation analysis. *Measurement*, 131, 372-379.
- [6] Ey-Chmielewska, H., Chruściel-Nogalska, M., & Frączak, B. (2015). Photogrammetry and its potential application in medical science on the basis of selected literature. *Adv Clin Exp Med*, 24(4), 737-41.
- [7] Geisler, E. L., Agarwal, S., Hallac, R. R., Daescu, O., & Kane, A. A. (2021). A role for artificial intelligence in the classification of craniofacial anomalies. *Journal of Craniofacial Surgery*, 32(3), 967-969.
- [8] Obuchowicz, R., Strzelecki, M., & Piórkowski, A. (2024). Clinical Applications of Artificial Intelligence in Medical Imaging and Image Processing—A Review. *Cancers*, 16(10), 1870.
- [9] Anas, I. Y., Bamgbose, B. O., & Nuhu, S. (2019). A comparison between 2D and 3D methods of quantifying facial morphology. *Heliyon*, 5(6).
- [10] Schaufelberger, M., Ktihle, R. P., Kaiser, C., Wachter, A., Weichel, F., Hagen, N., ... & Nahml, W. (2022, July). Cnn-based classification of craniosynostosis using 2d distance maps. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 446-449). IEEE.
- [11] Nord, F., Ferjencik, R., Seifert, B., Lanzer, M., Gander, T., Matthews, F., ... & Lübbers, H. T. (2015). The 3dMD photogrammetric photo system in cranio-maxillofacial surgery: Validation of interexaminer variations and perceptions. *Journal of Cranio-Maxillofacial Surgery*, 43(9), 1798-1803.
- [12] Nishimoto, S., Saito, T., Ishise, H., Fujiwara, T., Kawai, K., & Kakibuchi, M. (2023). Three-Dimensional Craniofacial Landmark Detection in Series of CT Slices Using Multi-Phased Regression Networks. *Diagnostics*, 13(11), 1930.
- [13] Patias, P. (2002). Medical imaging challenges photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(5-6), 295-310.
- [14] Rasool, N., & Bhat, J. I. (2023). Unveiling the complexity of medical imaging through deep learning approaches. *Chaos Theory and Applications*, 5(4), 267-280.
- [15] De Buhan, M., & Nardoni, C. (2018). A facial reconstruction method based on new mesh deformation techniques. *Forensic sciences research*, 3(3), 256-273.
- [16] Ng, H. P., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2006, March). Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE southwest symposium on image analysis and interpretation* (pp. 61-65). IEEE.
- [17] Shrivastava, K., Gupta, N., & Sharma, N. (2014). Medical image segmentation using modified k means clustering. *International Journal of Computer Applications*, 103(16).
- [18] Barbero-García, I., Lerma, J. L., & Mora-Navarro, G. (2020). Fully automatic smartphone-based photogrammetric 3D modelling of infant's heads for cranial deformation analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166, 268-277.

- [19] Barbero-García, I., Lerma, J. L., Marqués-Mateu, Á., & Miranda, P. (2017). Low-cost smartphone-based photogrammetry for the analysis of cranial deformation in infants. *World neurosurgery*, *102*, 545-554.
- [20] Yang, W. F., & Su, Y. X. (2021). Artificial intelligence-enabled automatic segmentation of skull CT facilitates computer-assisted craniomaxillofacial surgery. *Oral Oncology*, *118*, 105360.
- [21] Ding, C., & He, X. (2004, July). K-means clustering via principal component analysis. In *Proceedings of the twenty-first international conference on Machine learning* (p. 29).
- [22] Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2020). Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, *43*(12), 4338-4364.
- [23] Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 652-660).

10. Anexos

Tablas:

Matriz de confusión (Adaptación Figura 17):

	0	1	2	3	4	5
0	145.2	1.4	0.6	1.2	6.8	7.6
1	2	163.8	0.6	3.2	5.4	4.4
2	0.6	0.4	69.2	0.2	7.8	0.6
3	1.2	3	0	130.6	3.2	8.2
4	3	3.6	4	4	194.2	1.4
5	5.8	4.2	0.6	3.4	7.4	173.2

Tabla 6 - Matriz de confusión 90/05/05

Matriz de confusión (Adaptación Figura 19):

	0	1	2	3	4	5
0	114.4	3	0.6	1.2	6.8	4.4
1	3.6	120.4	0.8	2.6	4.4	5.8
2	1.6	0.6	52.8	0	5.8	1.8
3	1.4	3.8	0.6	93.8	3.4	6.4
4	6	3	4.4	1.4	147.8	3.4
5	5.6	5.8	1.8	3.6	6	127

Tabla 9 - Matriz de confusión 70/15/15

Matriz de confusión (Adaptación Figura 21):

	0	1	2	3	4	5
0	97.6	1	0.8	0.4	4.2	6.6
1	3.4	100.2	1.2	3	3.4	4.6
2	1.6	1	42.8	0	5	2.4
3	2	3	1.2	87.8	2.4	4
4	2.6	2.8	3	0.6	125.2	3.4

5	3.8	2.6	1.4	2.4	4.6	116
---	-----	-----	-----	-----	-----	-----

Tabla 12 - Matriz de confusión 60/20/20

Matriz de confusión (Adaptación Figura 23):

	0	1	2	3	4	5
0	117.6	2.8	2.6	1.8	3.4	13.4
1	1.2	143.8	0.8	4.8	2.2	6.4
2	0.4	1.4	63.8	0.2	4.2	1
3	0.8	3	0.6	119.2	1	4.8
4	2.2	3.6	4.8	2.4	172.4	4.2
5	3.2	4.2	1.6	5.4	5	153.8

Tabla 15 - Matriz de confusión 80/10/10

Matriz de confusión (Adaptación Figura 25):

	0	1	2	3	4	5
0	75.4	2.4	0.6	0.6	2.6	4
1	0.8	91.6	0.4	1.4	1.8	3.6
2	2.6	1	35.8	0.6	3.8	0.4
3	1.4	3	0	76.8	1.8	2.2
4	2.8	2	2.6	0.4	106.8	1.4
5	2	4.6	0.8	2.6	3	96.4

Tabla 18 - Matriz de confusión 50/25/25

Códigos:

 00_simarro_ply.py

 01_simarro_segmentaciones_6.py

 02_simarro_newExcel_Zero.py

 03_modelo_tfm+.py

 04_05_asimarrodev_evp.ipynb

Figura 28 - Archivos código