



UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

FACULTAD DE INGENIERÍA
MAESTRÍA EN CIENCIAS EN INTELIGENCIA ARTIFICIAL

“Segmentación de aneurismas cerebrales a partir de nubes de puntos por medio de técnicas de aprendizaje profundo”

Tesis que como parte de los requisitos para obtener el grado
de la Maestría en Ciencias en Inteligencia Artificial

Presenta:
Luis Felipe Estrella Ibarra

Dirigido por:
Dr. Saúl Tovar Arriaga

Co-dirigido por:
M.C. Alejandro De León Cuevas

Querétaro, Qro. Mayo 2024

La presente obra está bajo la licencia:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



CC BY-NC-ND 4.0 DEED

Atribución-NoComercial-SinDerivadas 4.0 Internacional

Usted es libre de:

Compartir — copiar y redistribuir el material en cualquier medio o formato

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

Bajo los siguientes términos:



Atribución — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



NoComercial — Usted no puede hacer uso del material con [propósitos comerciales](#).



SinDerivadas — Si [remezcla, transforma o crea a partir](#) del material, no podrá distribuir el material modificado.

No hay restricciones adicionales — No puede aplicar términos legales ni [medidas tecnológicas](#) que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

Avisos:

No tiene que cumplir con la licencia para elementos del material en el dominio público o cuando su uso esté permitido por una [excepción o limitación](#) aplicable.

No se dan garantías. La licencia podría no darle todos los permisos que necesita para el uso que tenga previsto. Por ejemplo, otros derechos como [publicidad, privacidad, o derechos morales](#) pueden limitar la forma en que utilice el material.



UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

FACULTAD DE INGENIERÍA
MAESTRÍA EN CIENCIAS EN INTELIGENCIA ARTIFICIAL

Segmentación de aneurismas cerebrales a partir de nubes de puntos por medio de técnicas de aprendizaje profundo

Tesis que como parte de los requisitos para obtener el grado de la Maestría en
Ciencias en Inteligencia Artificial

Presenta:

Luis Felipe Estrella Ibarra

Dirigido por:

Dr. Saúl Tovar Arriaga

Co-dirigido por:

M.C. Alejandro De León Cuevas

Dr. Saúl Tovar Arriaga
Presidente

M.C. Alejandro De León Cuevas
Secretario

Dr. Jesús Carlos Pedraza Ortega
Vocal

Dr. Juan Manuel Ramos Arreguín
Suplente

Dr. Gendry Alfonso Francia
Suplente

INDICE

	Página
<u>Resumen</u>	<u>I</u>
<u>Abstract</u>	<u>II</u>
<u>Agradecimientos</u>	<u>III</u>
<u>Índice de cuadros</u>	<u>IV</u>
<u>Índice de figuras</u>	<u>V</u>
1. <u>INTRODUCCION</u>	<u>1</u>
2. <u>DESCRIPCION DEL PROBLEMA</u>	<u>5</u>
3. <u>JUSTIFICACION</u>	<u>6</u>
4. <u>HIPOTESIS</u>	<u>7</u>
5. <u>OBJETIVOS</u>	<u>7</u>
6. <u>ANTECEDENTES</u>	<u>8</u>
7. <u>FUNDAMENTACION TEORICA</u>	<u>15</u>
<u>7.1 Sistema cerebrovascular</u>	<u>15</u>
<u>7.2 Anomalías Cerebrovasculares</u>	<u>16</u>
<u>7.3 Aneurismas intracraneales</u>	<u>16</u>
<u>7.4 Aneurisma sacular</u>	<u>17</u>
<u>7.5 Aneurisma fusiforme</u>	<u>17</u>
<u>7.6 Procedimiento Clínico</u>	<u>18</u>
<u>7.7 Imágenes médicas</u>	<u>19</u>
<u>7.8 Visión por computadora</u>	<u>20</u>
<u>7.9 Preprocesamiento de imágenes médicas</u>	<u>21</u>
<u>7.10 Transformación de histograma y filtrado de ruido</u>	<u>21</u>

7.11	Reconstrucción de anatomía	24
7.12	Inteligencia artificial en el área médica	26
7.13	Inteligencia artificial en planeación preoperatoria	26
7.14	Inteligencia artificial como guía intraoperatoria	27
7.15	Machine learning	27
7.16	Redes neuronales artificiales	28
7.17	Deep learning	30
7.18	Transformers	31
7.19	Aprendizaje contrastivo	32
7.20	Función de pérdida	33
7.21	Segmentación	34
7.22	Segmentación 3D	34
7.23	Métricas de evaluación	35
8.	METODOLOGIA	38
8.1	Conjunto de Datos	38
8.2	Aprendizaje contrastivo	39
8.3	Definición de Borde	40
8.4	Particionamiento de los Puntos Vecinos al Borde	42
8.5	Nested Contrastive Boundary Learning (NCBL)	44
8.6	Nested Contrastive Boundary Learning Point Transformer	45
8.7	Entrenamiento	49
9.	RESULTADOS	50
10.	COCLUSIONES	52
11.	REFERENCIAS	53

Resumen

En el ámbito de la segmentación 3D, los modelos actuales basados en nubes de puntos muestran un rendimiento alentador, aunque enfrentan dificultades en la delineación precisa de clases alrededor de intersecciones, desafío inherente de los modelos de segmentación. La identificación precisa de límites entre clases dentro del espacio tridimensional es crucial, especialmente en aplicaciones médicas. Este estudio presenta el Punto Transformador de Aprendizaje de Fronteras Contrastivas Anidadas (NCBL-PT), un marco especializado diseñado para segmentaciones de nubes de puntos. Al aprovechar el poder del aprendizaje contrastivo, NCBL-PT mejora las representaciones de los puntos cercanos a la intersección de clases mediante la maximización de la similitud de características entre puntos que pertenecen a la misma clase. Este enfoque incorpora una distinción entre los puntos pertenecientes a esta frontera, discerniendo si constituyen el límite o no. Esta delineación permite al modelo obtener información de puntos que se extienden más allá de la proximidad inmediata, mitigando así la ambigüedad semántica dentro de características de la misma clase y fomentando relaciones de características enriquecidas. El objetivo es evitar que el modelo relacione características entre puntos de diferentes clases en la zona de intersección de clases, donde similitudes de características inducidas por la proximidad pueden llevar a asociaciones erróneas y a espacios semánticos difíciles de generalizar. NCBL-PT opera dentro de cada nube de puntos submuestreada en cada etapa de los bloques codificadores de la arquitectura del punto transformer. Se aplica autoatención dentro de vecindarios locales representados por subconjuntos $X(i) \subseteq X$, definiendo un conjunto de puntos dentro del vecindario local de x_i . El uso de $k=16$ vecinos más cercanos para la autoatención mantienen consistencia con el mismo vecindario definido en los cálculos de NCBL, asegurando una regularización de autoatención del contexto del vecindario local. En consecuencia, NCBL-PT mejora la segmentación de nubes de puntos 3D, reduciendo efectivamente errores en áreas de intersección de clases. La aplicación de este enfoque resultó en una mejora notable del 3.31% haciendo uso de la métrica de evaluación Intersección sobre Unión (IOU) para la segmentación de aneurismas en comparación con el modelo base punto transformer.

(Palabras clave: Aneurisma intracraneal Segmentación, Aprendizaje contrastivo)

Abstract

In the realm of 3D segmentation, current point-based models exhibit commendable performance, yet they encounter difficulties in precise class delineation around intersections—an inherent challenge in segmentation models. The accurate identification of boundaries in 3D space is vital, particularly in medical applications. It plays a pivotal role in influencing patient care, surgical planning, assisted surgery, and improved medical education and training facilitated by sophisticated simulation systems. This study introduces the Nested Contrastive Boundary Learning Point Transformer (NCBL-PT), a specialized framework tailored for 3D point cloud segmentations. By harnessing the power of contrastive learning, NCBL-PT enhances the representation of boundary points through the maximization of feature similarity among points belonging to the same class. The framework incorporates a border-aware distinction for points within the same class, discerning whether they constitute the boundary or not. This delineation empowers the model to glean insights from points extending beyond immediate proximity, thereby mitigating semantic ambiguity within same class features and fostering enriched feature relationships. The goal is to prevent the model from confusing feature relationships among points of different classes in the ambiguous class intersection zone, where proximity-induced feature similarities may lead to erroneous associations. NCBL-PT operates within each subsampled point cloud at each stage of the encoder blocks of the point transformer architecture. Self-attention is applied within local neighborhoods represented by subsets $X(i) \subseteq X$, defining a set of points within the local neighborhood of x_i . Using $k=16$ nearest neighbors for self-attention maintains consistency with the same neighborhood defined in NCBL calculations, ensuring a self-attention regularization of the local neighborhood context. Consequently, the NCBL-PT improves 3D point cloud segmentation, effectively reducing errors in areas of class intersection. The application of this framework resulted in a notable improvement of +3.31% in Intersection over Union (IOU) for aneurysm segmentation compared to the base point transformer model.

(Key words: Intracranial Aneurysm, Point cloud segmentation, Contrastive learning)

AGRADECIMIENTOS

Agradezco profundamente a mi familia, a mi papá y a mi mamá, por darme la oportunidad y el privilegio de perseguir lo que amo; todo logro mío es verdaderamente logro de ellos. A la doctora Paulina Estrella, gran fuente de inspiración, mentora y la mejor hermana que podría tener. A mis amigos, quienes han estado a mi lado en cada paso del camino. A mi mejor amiga, Fernanda Tovar, cuyo apoyo incondicional y consejo me han dado el valor para tomar decisiones que han marcado mi crecimiento académico y personal de formas inexplicables. Agradezco al maestro Alejandro de León Cuevas por su guía a lo largo de estos años, su apoyo y paciencia, así como al doctor Saúl Tovar Arriaga, excelente coordinador, profesor y guía en cada paso del proceso. Sin su guía, no hubiera sido capaz de concluir. También extiendo mi agradecimiento por el apoyo brindado por Luis Aguilar, Alejandro Ávalos, Jair García de LAVIS, y Erick Pasaye, así como a Cesar Arturo Domínguez Frausto de LANIREM en INB UNAM Juriquilla. También a CONAHCYT por proporcionar la beca necesaria para llevar a cabo esta investigación.

INDICE DE CUADROS

Cuadro		Página
2.1	Trabajos antecedentes sobre detección y segmentación de aneurismas por medio de algoritmos de aprendizaje profundo.	12
3.1	Clasificación de valores predictivos.	36
7.1	Comparativa del desempeño del método propuesto contra el estado del arte y el modelo base bajo el esquema de 2048 puntos.	51

INDICE DE FIGURAS

Figura		Página
3.1	Esquema de las principales ramas de las arterias vertebrales y basilar, el círculo de Willis y las arterias que contribuyen a la formación de este.	16
3.2	(a) aneurisma sacular, (b) aneurisma fusiforme.	17
3.3	Clips implantables intracraneales sobre aneurismas.	18
3.4	Ejemplos de terapia endovascular para tratamiento de Aneurisma intracraneal.	19
3.5	Transformación de histograma de imágenes de 16 bits.	22
3.6	Kernel y visualización de concepto de vecindad por medio de un filtro.	23

3.7	Ejemplo de filtrado de ruido en imagen de 8 bits.	24
3.8	Ejemplo de reconstrucción anatómica de la estructura cerebro vascular a partir del filtrado de imágenes médicas tipo TOF-MRA.	25
3.9	Estructura básica de una red neuronal.	29
3.10	Aprendizaje contrastivo para la maximización de similitud de clases.	32
3.11	Segmentación bidimensional y tridimensional de un aneurisma intracraneal.	35
3.12	valores predictivos en la segmentación de imágenes.	37
6.1	Definición de Borde.	41
6.2	partición de vecindad del borde.	43
6.3	Nested Contrastive Boundary Learning Point Transformer (NCBL-PT).	45

6.4	Propagación de etiquetas de clase a través de la reducción de la dimensionalidad del espacio del codificador.	46
6.5	Propagación de etiquetas de clase “borde” a través de la reducción de la dimensionalidad del espacio del codificador.	47
6.6	Propagación de las etiquetas de clase a través de Farthest Point Sampling (FPS) y KNN.	47
6.7	Nested Contrastive Boundary Learning Point Transformer (NCBL-PT) Diagrama de flujo que muestra lo que sucede en cada bloque para llegar a los cálculos del NCBL.	48
7.1	Explorando los resultados de la segmentación: análisis comparativo del etiquetado manual contra el modelo base y el método propuesto.	51

I. INTRODUCCION

Los aneurismas intracraneales, caracterizados como anomalías cerebrovasculares, se presentan como una condición médica que, en su estado previo a la ruptura, frecuentemente permanece oculta debido a su naturaleza asintomática. Este aspecto sigiloso de los aneurismas contrasta dramáticamente con el peligro inminente que representan una vez que se produce una ruptura, desencadenando una hemorragia subaracnoidea (HSA), un evento a menudo catastrófico. Estos aneurismas generalmente se descubren de forma incidental durante evaluaciones médicas para otras condiciones, permaneciendo latentes hasta que ocurre la ruptura [1].

La gestión de los riesgos asociados con la ruptura de aneurismas intracraneales es un desafío complejo que involucra múltiples factores, incluyendo el tamaño y la ubicación del aneurisma, así como características específicas del paciente, como la edad y la historia de hipertensión. Una monitorización y evaluación mediante estudios de imagen es esencial para detectar y determinar la necesidad y el momento óptimo para una intervención [1], [2].

Afortunadamente, el campo de la visión por computadora ha experimentado un cambio revolucionario en la última década, impulsado por avances significativos en el desarrollo de la inteligencia artificial y el aprendizaje profundo. Este progreso comenzó en 1998 con el desarrollo de LeNet-5 [3], una de las primeras redes neuronales convolucionales (CNN). Desde entonces, las CNN se han convertido en una herramienta esencial en el procesamiento de imágenes cuando de algoritmos de inteligencia artificial se trata, estableciéndose como el método predominante en las dos primeras décadas del siglo XXI, y manteniéndose hasta la fecha como el más popular, gracias a su habilidad en el mapeo y generalización de características dentro de las imágenes.

La evolución continua de estos algoritmos de aprendizaje profundo ha estimulado la investigación y el desarrollo en el ámbito de la inteligencia artificial aplicada a imágenes, progresando de manera constante y mejorando en términos de precisión y eficiencia.

Un avance significativo en este campo, aunque no era evidente en su momento, tuvo lugar en 2017 con la introducción de los modelos Transformer [4], los cuales incorporan conceptos de atención. Este enfoque revolucionario en el procesamiento de información se ha convertido en un pilar fundamental para la mejora de tareas relacionadas con la

inteligencia artificial. Esta técnica permite asignar ponderaciones variables a diferentes partes de la entrada, lo que mejora significativamente la capacidad del sistema para captar y comprender información relevante de forma contextual.

En términos simples, la atención en este contexto se asemeja al enfoque selectivo que aplicamos al leer un texto. Así como dirigimos nuestra atención a palabras clave para comprender mejor el significado de una oración, los modelos Transformer utilizan la atención para destacar componentes cruciales de los datos.

Aunque inicialmente estos modelos se concibieron con la finalidad de procesar el lenguaje natural, su capacidad de focalización selectiva ha demostrado ser de gran interés en el continuo desarrollo y la incesante curiosidad científica. Este enfoque no solo se restringe a la función originalmente prevista, sino que, de manera sorprendente pero comprobada, estos algoritmos exhiben una capacidad extraordinaria en campos adicionales, como la visión por computadora y muchos otros [5].

Esta versatilidad demuestra la impresionante eficacia y adaptabilidad de los modelos de inteligencia artificial, abriendo un amplio abanico de posibilidades y aplicaciones en diversas áreas del conocimiento y la industria.

El mecanismo de atención de los Transformers permite manejar las dependencias contextuales de manera más flexible y eficiente, un aspecto crucial en el análisis de imágenes.

La convergencia de algoritmos avanzados de inteligencia artificial, que incorporan conceptos de atención, junto con técnicas convencionales como las redes neuronales convolucionales (CNN) y otros desarrollos en algoritmos, ha generado nuevas posibilidades y enfoques en el campo de la visión por computadora [6]. Esta fusión está impulsando innovaciones en el procesamiento y análisis de imágenes en diversos formatos, superando retos considerados difíciles anteriormente. Este avance tecnológico ha sustancialmente ampliado las capacidades y aplicaciones de los modelos utilizados en la detección y segmentación de anomalías en imágenes.

La detección en este contexto implica la identificación precisa de la presencia de una anomalía, mientras que la segmentación se enfoca en delinear y clasificar con exactitud la región afectada. Ambos procesos son fundamentales para obtener información detallada sobre la extensión y ubicación exacta de las anomalías. Esta información es crucial, ya que proporciona a los profesionales médicos herramientas valiosas que no

solo aceleran el diagnóstico y tratamiento, sino que también permiten una comprensión más profunda y precisa de la patología subyacente [7], [8].

La aplicación de estos avances tecnológicos en el campo médico representa un salto cualitativo en la manera en que se manejan y se interpretan los datos visuales. Las técnicas de visión por computadora, enriquecidas por el aprendizaje profundo, no solo mejoran la eficiencia de los diagnósticos médicos, sino que también abren la puerta a nuevos métodos de análisis y comprensión de complejas condiciones médicas. Gracias a estas herramientas avanzadas, los profesionales de la salud pueden obtener una comprensión más completa y detallada de diversas enfermedades, especialmente a través de métodos de segmentación. Esto conduce a un avance notable en la calidad del cuidado y tratamiento que reciben los pacientes, optimizando así los resultados y la eficacia de las intervenciones médicas.

La variabilidad en los formatos utilizados en la segmentación es crucial para entender las distintas aplicaciones y sus capacidades. Tradicionalmente, se han utilizado imágenes bidimensionales procesadas, pero con el avance tecnológico, han surgido representaciones más complejas y detalladas, como las tridimensionales, a través de métodos como la voxelización o las nubes de puntos [9]. Estos formatos ofrecen ventajas y desventajas únicas, siendo el formato tridimensional particularmente beneficioso para una comprensión más profunda y detallada de las estructuras, lo que posibilita el estudio de relaciones espaciales en mayor profundidad. Este enfoque ha abierto nuevas posibilidades en la segmentación de datos, ofreciendo perspectivas únicas y valiosas en diversas áreas de investigación y aplicaciones prácticas.

Estos avances tecnológicos han sido cruciales en el desarrollo de herramientas para la detección y segmentación precisa de anomalías específicas, como los aneurismas intracraneales [10 - 18]. La heterogeneidad de estas anomalías, manifestándose en una amplia gama de formas y tamaños, subraya la importancia de una comprensión profunda de las implicaciones geométricas inherentes a estas variaciones. La capacidad de los modelos de inteligencia artificial para generalizar y adaptarse a esta diversidad es de vital importancia. A pesar de los avances en la sofisticación y precisión de los sistemas y algoritmos diseñados para estas tareas, sigue siendo un desafío desarrollar métodos que manejen efectivamente la complejidad de los datos. En este contexto, surge la necesidad de explorar enfoques innovadores que permitan no solo una mejor gestión de los datos

complejos, sino también que fomenten una asociación y generalización adecuada de las formas.

Frecuentemente, esta complejidad conlleva a la implementación de modelos con modularizaciones específicas, diseñadas para abordar la segmentación desde diversas perspectivas, cada una con su propio nivel de complejidad [18]. Este enfoque puede transformar la tarea de segmentación en un proceso de aprendizaje intrincado, lo que puede representar un obstáculo para su aplicabilidad y portabilidad en diferentes áreas o patologías. Por lo tanto, es crucial que los investigadores y desarrolladores en este campo continúen buscando soluciones que no solo sean eficaces en la tarea específica de segmentación de anomalías, sino que también sean flexibles y adaptables a una variedad de contextos clínicos y diagnósticos.

Este desafío nos lleva a explorar las relaciones espaciales intrínsecas a las anomalías, iniciando con preguntas fundamentales sobre cómo diferenciamos partes de un mismo sistema. A nivel humano, esta diferenciación puede parecer intuitiva, pero entender cómo nuestra mente realiza estas asociaciones es un reto mayor. La percepción humana es excepcionalmente buena en identificar patrones y formas, incluso en condiciones de variabilidad y ambigüedad. Sin embargo, este nivel de adaptabilidad y comprensión intuitiva presenta un desafío significativo para los sistemas de IA, que deben ser entrenados para reconocer y segmentar una gama diversa de estructuras anatómicas en imágenes médicas [8].

La segmentación de objetos puede ser relativamente sencilla cuando se trata de geometrías simples. Por ejemplo, diferenciar un círculo de un cuadrado es una tarea directa. Sin embargo, en el caso de formas más complejas, como una estructura orgánica con múltiples curvaturas y ramificaciones, la tarea de segmentación se torna exponencialmente más complicada. Esto se debe a que estas formas no solo varían en términos de su estructura básica, sino también en su orientación, escala y relación con el entorno circundante. En el contexto médico, esto se manifiesta en la diversidad de formas que pueden asumir los aneurismas intracraneales, cada uno con sus propias peculiaridades geométricas y espaciales.

Estos desafíos son inherentes en los modelos de segmentación, especialmente al tratar de delinear con precisión las clases en las intersecciones. Los algoritmos de inteligencia artificial, especialmente aquellos basados en aprendizaje profundo, deben ser capaces de discernir estas sutiles diferencias y variaciones, algo que en el ámbito humano se realiza

con una mezcla de percepción visual y experiencia cognitiva. Los avances en el área de inteligencia artificial están abriendo caminos para el desarrollo de soluciones que puedan generalizar en sistemas complejos, adaptándose a la variabilidad inherente de las estructuras anatómicas.

Estos avances ofrecen un amplio potencial, no solo en técnicas quirúrgicas asistidas por computadora, sino también en el ámbito educativo y formativo [19], [20]. La realidad aumentada, en particular, se está convirtiendo en una herramienta valiosa para los profesionales médicos, proporcionando visualizaciones en tiempo real que mejoran la precisión y eficacia de los procedimientos quirúrgicos. Además, los sistemas de simulación avanzados están revolucionando la formación médica, ofreciendo a los futuros cirujanos un entorno seguro y controlado para perfeccionar sus habilidades y comprensión de las estructuras anatómicas complejas.

Este estudio se centra en mejorar la segmentación de aneurismas intracraneales en modelos 3D, con el objetivo de avanzar en la comprensión y aplicación de este campo crítico.

II. DESCRIPCION DEL PROBLEMA

Al enfrentar la complejidad geométrica y espacial inherente de las imágenes, los modelos de segmentación se enfrentan a un desafío importante. El error de asociación por proximidad entre estructuras de clases distintas y la dificultad para generalizar. Este problema se vuelve aún más prominente al considerar la variabilidad intrínseca de las estructuras anatómicas y sus potenciales afecciones. Es por ello que la investigación naturalmente tiende a recurrir a la modularización adicional extensa que busca abordar estos problemas, sin percatarse de que esto podría crear nuevos desafíos.

Las soluciones complejas y específicas resultan difíciles de aplicar en diferentes afecciones, áreas e incluso modelos, lo cual vuelve difícil la portabilidad y aplicación. Por esta razón, el desarrollo de métodos ligeros y funcionales se presentan como un área de oportunidad en crecimiento.

III. JUSTIFICACION

Es crucial subrayar la impredecibilidad en la tasa de crecimiento de estos aneurismas, donde incluso aquellos de tamaño pequeño no están exentos del riesgo de ruptura. Esta ruptura transforma la HSA en una condición de alta mortalidad, con tasas que fluctúan alarmantemente entre el 25% y el 50%. Además, un impactante 50% de los sobrevivientes de HSA enfrentan discapacidades permanentes, dejando solo un tercio de los pacientes con un desenlace favorable [2], [21].

Dada la gravedad de estos aneurismas, es imperativo destacar que representan una amenaza dual: una alta mortalidad y una carga significativa de discapacidad para los sobrevivientes. Las estadísticas muestran que un 12% de los pacientes con ruptura de aneurisma fallecen antes de recibir atención médica, un 25% no sobrevive las primeras 24 horas, y entre el 40% y el 60% no sobrevive más allá de los primeros 30 días. Por lo tanto, la detección temprana de aneurismas intracraneales es un aspecto crítico para mejorar las tasas de supervivencia y mitigar las consecuencias discapacitantes de esta condición [1].

IV. HIPOTESIS

El uso de técnicas basadas en aprendizaje contrastivo en conjunto a un modelo de Aprendizaje profundo permitirá la segmentación de aneurismas intracraneales en imágenes, representadas como nubes de puntos.

V. OBJETIVOS

Objetivo General

Implementar técnicas basadas en aprendizaje contrastivo en conjunto a un modelo de Aprendizaje profundo para la segmentación de aneurismas intracraneales en imágenes, representadas como nubes de puntos.

Objetivos específicos

- a) Estudiar el estado del arte para entender las últimas conceptualizaciones que se han implementado para segmentar la afección estudiada.
- b) Elegir el conjunto de datos con el que se realizará el estudio.
- c) Seleccionar un modelo de segmentación basado en técnicas de aprendizaje profundo.
- d) Adaptar la arquitectura del modelo para el uso de técnicas basadas en técnicas de aprendizaje contrastivo para la segmentación de aneurismas intracraneales.
- e) Evaluar el modelo utilizando métricas cuantitativas específicas para la segmentación de aneurismas y la precisión de las máscaras de segmentación correspondientes.

VI. ANTECEDENTES

Como se discutió en la sección anterior, la investigación y el desarrollo en el ámbito de la visión por computadora juegan un papel que podría ser de importancia crítica, especialmente en el contexto clínico. Esta relevancia no es casual, dado que identificar con precisión estructuras anatómicas críticas es clave para la implementación efectiva de la asistencia médica. En particular, la tarea de detectar y segmentar estructuras complejas, como los aneurismas en las vasculaturas cerebrales, representa un notable desafío. En este contexto, el alto rendimiento del modelo es esencial para brindar apoyo efectivo y preciso, en lugar de limitarse a una noción puramente teórica. Es en este marco donde la investigación adquiere relevancia, centrándose en el progreso de la inteligencia artificial aplicada a la visión por computadora. El objetivo es desarrollar y avanzar en técnicas de segmentación 3D que faciliten la delineación de estructuras complejas. Afortunadamente, en los últimos años, el campo de la segmentación de modelos 3D mediante técnicas de aprendizaje profundo, especialmente en el contexto de aneurismas, ha experimentado un resurgimiento de interés después de un período de relativo abandono.

En el ámbito de la segmentación, al igual que en la adopción de tecnologías y algoritmos novedosos, se observa una fase de adaptación y adopción progresiva. Aunque estén disponibles nuevos desarrollos, su implementación y exploración no son inmediatas. Esto es evidente en el caso de los algoritmos de segmentación, que, a pesar de estar ya desarrollados, inicialmente se enfocaban en la detección para su uso en la identificación de aneurismas. Esta tendencia se hace visible al analizar una línea de tiempo de las investigaciones, comenzando con el desarrollo de T. Nakao et al [17]. En su estudio de 2017, crearon una red neuronal convolucional para detectar aneurismas utilizando imágenes bidimensionales derivadas de imágenes tipo "3D TOF-MRA". Aplicaron métodos como el crecimiento de regiones y la proyección de intensidad máxima (MIP) desde varios ángulos, buscando lograr una detección bidimensional con información recopilada desde una perspectiva tridimensional de la estructura vascular.

Esta tendencia inicial hacia la detección refleja cómo, en aquel momento, las investigaciones no se centraban sustancialmente en la segmentación, y esto se hará más evidente conforme avancemos en el tiempo y en las investigaciones.

Sin embargo, la introducción en 2015 de modelos como la Fully Convolutional Network (FCN) por J. Long et al. [22] y U-Net por O. Ronneberger et al. [23], marcó un cambio gradual en el paradigma de la investigación, orientándose también hacia la segmentación.

Pero la adopción de estas nuevas metodologías no fue instantánea, como se ejemplifica claramente con el antecedente científico de las investigaciones que mencionaremos.

Un ejemplo de esto es el estudio de W. Hou et al. [24] realizado en 2020, tres años después del trabajo de T. Nakao et al. [17]. Para ese momento, la segmentación mediante aprendizaje profundo ya se había consolidado incluso con la introducción de modelos específicos en el ámbito médico, como DeepMedic en 2017 por K. Kamnitsas et al. [25], e incluso el desarrollo de un modelo especializado para la segmentación 3D de aneurismas intracraneales en 2019, denominado HeadXNet por A. Park et al. [16] con el cual obtuvieron un puntaje F1 de .80 en promedio. No obstante, el estudio de W. Hou y colaboradores se centró, como muchas otras investigaciones, en perfeccionar la detección, introduciendo un modelo modificado del método propuesto por T. Nakao y otros, que empleaba el método de Acumulación de Píxeles Proyectados (PAP). Este enfoque mejoró el procesamiento de datos al generar vectores de píxeles acumulados de las imágenes resultantes del método MIP, ampliando así la investigación previa de 2017 con un enfoque renovado en la detección.

Este desarrollo evolutivo evidencia cómo incluso, en el campo del aprendizaje profundo aplicado a aneurismas intracraneales, la integración de nuevos modelos y técnicas, como la segmentación, ha sido un proceso gradual y constante, marcado por la adaptación y evolución de las investigaciones, incluso cuando las tecnologías emergentes ya estaban accesibles.

Para abordar más propuestas en el ámbito de la inteligencia artificial, y al mismo tiempo ilustrar lo ya mencionado, podemos proceder gradualmente. En este contexto, I. Rahmany et al, en su estudio de 2018 [15], exploraron el campo del aprendizaje profundo utilizando la CNN desarrollada por Google, conocida como Inception-v3. Este modelo se entrenó usando la base de datos de acceso libre "ImageNet", y aplicando la técnica de "transfer learning", se potenció la capacidad del modelo para detectar aneurismas. Como resultado, se creó un modelo capaz de detectar dichas anomalías vasculares en imágenes médicas digitales de angiografía por sustracción (2D-DSA), logrando una notable sensibilidad del 100%. A pesar de ser impresionante, es crucial recordar que la sensibilidad no toma en cuenta los falsos positivos y se limita a indicar el porcentaje de casos verdaderos positivos identificados. Uno de los aspectos más sobresalientes de este estudio es el limitado conjunto de datos empleado para el desarrollo del modelo, que se basó en apenas 30 casos de pacientes con la afección.

Otra propuesta destacada es el estudio de J. N. Stember et al. [11], realizado en 2018, tres años después del lanzamiento del reconocido modelo de código abierto U-net anteriormente mencionado. Este desarrollo marco un punto significativo en la evolución de la tecnología de segmentación. U-net, con su estructura de codificador-decodificador, se especializa en la segmentación semántica, creando máscaras alrededor de objetos identificados. Este estudio en particular se centró en imágenes bidimensionales derivadas de 3D TOF-MRA, empleando el método de Proyección de Máxima Intensidad (MIP) y técnicas de aumento de datos como rotaciones y redimensionamientos. El proceso de entrenamiento implicó iteraciones con imágenes de aneurismas, utilizando pesos iniciados al azar, para predecir grupos de píxeles y sus probabilidades de corresponder a un aneurisma. Se generó una malla de puntos alrededor de la máscara pronosticada y se midió la distancia máxima entre pares de puntos, lo que sirvió como indicador del tamaño lineal máximo del aneurisma. Posteriormente, se compararon estas predicciones con imágenes manualmente etiquetadas, logrando una sensibilidad del 87% en detección, una divergencia de 2.02 mm o 30% en la dimensión lineal, y una discrepancia en la predicción del área de la máscara de 8.1 m² o 27%, en comparación con las anotaciones de un radiólogo.

Resulta crucial volver destacar que, a partir de la introducción de modelos de segmentación como U-net, se intensificó la investigación en su aplicación a patologías específicas como la que se estará abordando. Sin embargo, en estas etapas iniciales, aún se percibía una carencia en la evaluación exhaustiva y rigurosa de estos modelos. A pesar de que las técnicas utilizadas en aquel momento representaban un avance, todavía existía un amplio margen para mejorar en cuanto a métricas cuantitativas de evaluación, en especial en lo que respecta a la precisión y exactitud de las máscaras generadas para delinear las patologías.

T. Sichtermann et al. (2019) [12] se concentraron en el impacto del preprocesamiento de imágenes 3D TOF-MRA en la identificación de aneurismas. Desarrollaron cuatro modelos de preprocesamiento y los aplicaron utilizando la 3D-CNN de código abierto 'The Deep Medic', entrenada para la localización y segmentación de aneurismas. Este estudio reveló variaciones en la sensibilidad dependiendo del modelo y tamaño del aneurisma, aunque sin cambios significativos en la ubicación.

D. Ueda et al. (2019) [10] propusieron un método de detección adaptado al amplio volumen de imágenes TOF MRA, provenientes de distintas instituciones y bajo diversas

condiciones. El enfoque era crear un algoritmo aplicable a diferentes sistemas y condiciones de adquisición de imágenes de resonancia magnética. Para ello, reentrenaron una arquitectura de detección de objetos existente, 'ResNet-18', que ofrece predicciones probabilísticas para la detección de aneurismas.

R. Shahzad et al. (2020) [14] introdujeron un método para la detección automática de aneurismas intracraneales asociados a hemorragias subaracnoideas, utilizando angiografías por tomografía computarizada (CTA) y la 3D-CNN 'The Deep Medic'. Implementaron tres modelos basados en diferentes estrategias de entrenamiento, además de un modelo adicional que combinaba los resultados de los primeros tres, logrando sensibilidades superiores al 70%.

F. Claux, M. Baudouin, C. Bogey y A. Rouchaud (2022) [26] desarrollaron un método que emplea una 3D-CNN semántica basada en la arquitectura 'U-net'. Este sistema consta de dos redes neuronales convolucionales con un esquema codificador-decodificador, una para la segmentación de vasos sanguíneos y otra para la localización y segmentación tridimensional de aneurismas. Se consideraron tres clases para la detección semántica: vasos sanguíneos, aneurismas y fondo, resultando en un modelo que combina preprocesamiento y procesamiento en la detección de aneurismas con una sensibilidad del 78%.

El avance en el campo de la inteligencia artificial se distingue por su rica diversidad de métodos y tecnologías innovadoras. Esta evolución, en línea con los antecedentes ya mencionados, se manifiesta en la notable variedad de formatos de entrada utilizados. Un ejemplo destacado es la segmentación de nubes de puntos, un enfoque que se aparta del formato de imagen convencional y representa una nueva dirección en el procesamiento y análisis de datos.

La introducción de métodos avanzados como PointNet [27] y PointNet++ [28] marcó un cambio significativo en esta dirección. PointNet, conocido por su procesamiento directo de nubes de puntos y reconocimiento de características globales, preparó el escenario para PointNet++, que mejoró aún más la técnica al centrarse en estructuras locales a través de una red jerárquica. Esta era de innovación se vio enriquecida aún más por PointCNN [29], que introdujo una operación X-Conv única para el procesamiento estructurado de nubes de puntos, y RandLA-Net [30], aclamado por su eficiencia en el manejo de nubes de puntos a gran escala, datos en la nube a través de muestreo aleatorio y agregación de características locales.

Un ejemplo de esta tendencia es el método Dual Branch, desarrollado por Shao et al. (2022) [13]. Este enfoque alcanzó una eficacia media, medida como Intersección sobre Unión (IoU), del 67.05%, distinguiendo claramente entre dos categorías clave: aneurismas y vasos sanguíneos. Específicamente en la segmentación de vasos, el método demostró una eficiencia excepcional, logrando un IoU del 82.65%. Sin embargo, la segmentación de aneurismas presentó mayores desafíos, con un IoU de 51.45% en esta categoría específica.

Siguiendo estos avances, Yifan Liu et al. [18] presentaron una propuesta aún más reciente. Emplearon una combinación de dos modelos de transformadores de puntos y un enfoque modular más extenso para refinar la segmentación en la intersección de clases. Este enfoque innovador logró un IoU promedio de 91.52%, marcando un progreso significativo en la segmentación tanto de vasos sanguíneos, con un notable 95.46% de IoU, como de aneurismas, con un impresionante 87.58% de IoU.

A continuación, en el cuadro 2.1, se ofrecerá un resumen de todas las investigaciones referenciadas anteriormente, centradas en la detección y segmentación de aneurismas intracraneales. Esta tabla brindará una perspectiva clara y concisa de cada estudio, abarcando desde la metodología empleada hasta los resultados obtenidos. Además, proporcionará una visión general del desarrollo de estos algoritmos a lo largo del tiempo, ilustrando cómo ha evolucionado el enfoque en este campo y aplicación específica.

Cuadro 2.1. Trabajos antecedentes sobre detección y segmentación de aneurismas por medio de algoritmos de aprendizaje profundo.

Titulo	Algoritmos	Arquitectura	Tipo de Imágenes	Volumen de imágenes	Procedencia de la información	resultados
T. Nakao et al. 2017	<ul style="list-style-type: none"> • Deep learning: CNN • maximum intensity projection (MIP) algorithm • método region-growing 	Arquitectura propuesta para detección	3D TOF-MRA con MIP	<ul style="list-style-type: none"> • 300 para entrenamiento • 50 para validación • 100 para evaluación 	Hospital de la universidad de tokyo-departamento de radiología.	Sensibilidad del 70%

I.Rahmany et al. 2018	<ul style="list-style-type: none"> • Deep learning: CNN • Transfer learning 	Inception-v3 (uso libre) para detección	2D-DSA	<ul style="list-style-type: none"> • 20 para entrenamiento • 10 para evaluación 	Clínica” SOUKRA”, Tunisia	Sensibilidad del 100%
J. N. Stember et al. 2018	<ul style="list-style-type: none"> • Deep learning: CNN • maximum intensity projection (MIP) 	U-Net (uso libre) para segmentación 2D	TOF-MRA	<ul style="list-style-type: none"> • 500 para entrenamiento • 86 para evaluación 	Institución médica no especificada	Sensibilidad del 86%
T.Sichterman et al. 2019	•Deep learning: 3D-CNN	The DeepMedic (uso libre) para segmentación 3D	3D TOF-MRA	<ul style="list-style-type: none"> • 58 para entrenamiento • 10 para validación • 17 para evaluación 	Universidad RWTH Aachen-facultad de medicina y otras instituciones médicas externas	Sensibilidad del 90% Coeficiente Dice: 0.53
D. Ueda et al. 2019	• Deep learning: CNN	ResNET-18 (uso libre) para detección	TOF-MRA	<ul style="list-style-type: none"> • 683 para entrenamiento • 588 para evaluación • 	4 instituciones médicas diferentes	Sensibilidad del 91%
A. Park et al. 2019	• Deep learning: 3D-CNN	Arquitectura propuesta de nombre HeadXNet para segmentación 3D	CTA	<ul style="list-style-type: none"> • 611 para entrenamiento • 92 para validaciones • 115 para evaluación 	Escuela de medicina, departamento de radiología de la Universidad de Stanford	Sensibilidad del 89%
R. Shahzad et al. 2020	•Deep learning: 3D-CNN	The DeepMedic (uso libre) Para segmentación 3D	CTA	<ul style="list-style-type: none"> • 68 para entrenamiento • 185 para evaluación 	Hospital de la universidad de colonia-Alemania	Sensibilidad de hasta el 76% Coeficiente Dice:0.80
W. Hou et al. 2020	•Deep learning: 1D-CNN	Arquitectura propuesta basada en la de T. Nakao et al. 2017 Para detección	3D TOF-MRA	<ul style="list-style-type: none"> • 126 para entrenar • 18 para validación • 36 para evaluación 	Union Hospital of Tongji Medical College y Huazhong University of Science and Technology	Sensibilidad del 95%

Di Shao <i>et al.</i> 2022	•Deep learning: Unsupervised Dual-branch Learning	Encoder-MLP	Nube de puntos etiquetada	<ul style="list-style-type: none"> • 215 segmentos de aneurismas. • 	Deakin University - Australia	Mean IOU: 67.05% IOU-A: 51.45% IOU-V: 82.65%.
F. Claux <i>et al.</i> 2022	•Deep learning: 3D-CNN	U-Net (uso libre) para segmentación 3D	3D TOF-MRA	<ul style="list-style-type: none"> • 24 para entrenar • 25 para evaluación 	Hospital de la universidad de Limoges-departamento de radiología	Sensibilidad del 78%
Yifan Liu <i>et al.</i> 2023	•Deep learning: Transformer	Doble Point-cloud Transformer-modularizado	Nube de puntos etiquetada	<ul style="list-style-type: none"> • 215 segmentos de aneurismas. • 	Department of Electrical Engineering, City University of Hong Kong, Kowloon, Hong Kong, China	Mean IOU: 91.52% IOU-A: 87.58% IOU-V: 95.46%.

Este conjunto de investigaciones subraya la oportunidad de desarrollar marcos de trabajo más eficientes y ligeros en la segmentación de aneurismas y otras afecciones. La segmentación, lejos de limitarse a aplicaciones médicas específicas, presenta un vasto potencial en diversos campos.

Tal como hemos visto en la segmentación de aneurismas, se abre un amplio abanico de posibilidades en términos de desarrollo de algoritmos enfocados en la ligereza, eficiencia y estandarización en la evaluación de la segmentación.

Esta evolución refleja una curva de adaptación progresiva de los algoritmos de segmentación, un fenómeno que se ha desarrollado paralelamente a la creciente sofisticación y diversificación de las técnicas de inteligencia artificial en la visión por computadora. Este progreso, marcado por la continua integración de modelos y técnicas novedosas, destaca la relevancia de estos avances no solo en el ámbito clínico, donde la confiabilidad en resultado de la segmentación podría ser crítica, sino también en la expansión de estas tecnologías hacia aplicaciones más amplias y variadas.

VII. FUNDAMENTACION TEORICA

Tras revisar los antecedentes de investigación en este campo, se hace evidente la importancia de profundizar en ciertos conceptos fundamentales que subyacen al estudio en cuestión. La comprensión detallada de estos conceptos no solo es esencial para apreciar el progreso y las limitaciones de los trabajos previos, sino también para establecer un marco teórico sólido para la investigación. A continuación, se desglosará y explicarán estos conceptos claves con el fin de proporcionar una base teórica robusta y coherente que guíe el análisis y posteriores resultados de la investigación.

7.1 Sistema cerebrovascular

En el ámbito médico, el sistema de irrigación sanguínea del cerebro se clasifica en dos componentes principales: la circulación anterior, proveniente de las arterias carótidas internas, y la circulación posterior, originada en el sistema vertebrobasilar. Estas arterias, incluyendo las carótidas internas y la arteria basilar, se subdividen en ramificaciones de diámetro progresivamente menor. Estas ramas se extienden a lo largo de la superficie cerebral, recorren el espacio subaracnoideo e infiltran el tejido encefálico, desempeñando un papel crucial en su nutrición. [31].

Dentro de este sistema, las partes de las arterias cerebrales -anteriores, medias y posteriores- están interconectadas por las comunicantes anteriores y posteriores, formando una estructura conocida como el círculo o polígono de Willis. Este círculo actúa como un conducto anastomótico esencial, interconectando los sistemas de la carótida interna y el vertebrobasilar. El flujo sanguíneo en el cerebro depende del gradiente de presión arterial y de la resistencia vascular cerebral. Esta última está influenciada por factores como la viscosidad de la sangre y el calibre de los vasos cerebrales [31].

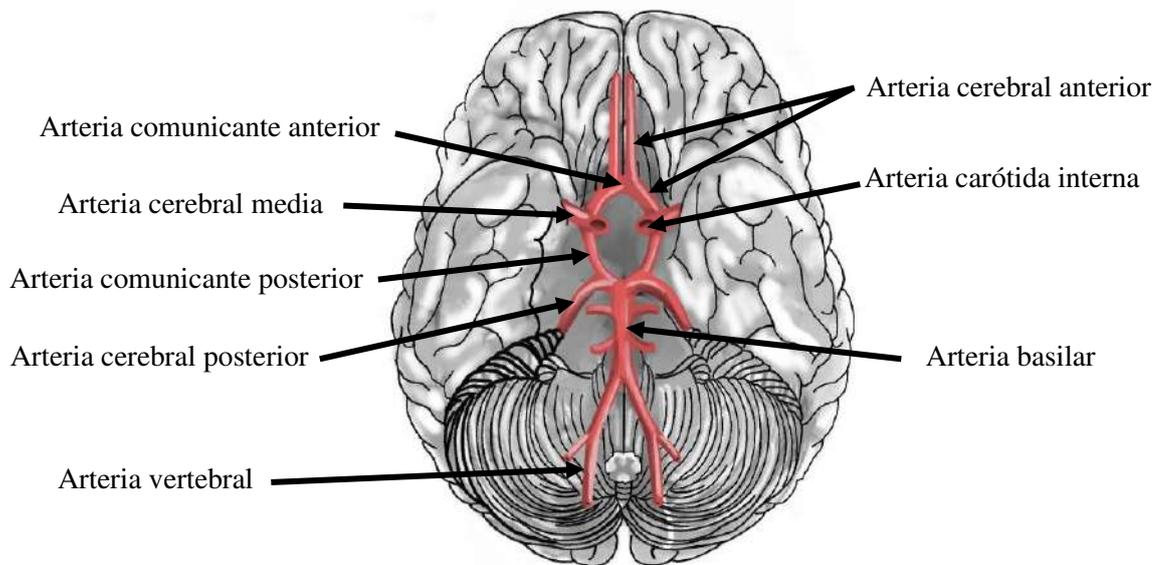


Figura 3.1. Esquema de las principales ramas de las arterias vertebrales y basilar, el círculo de Willis y las arterias que contribuyen a la formación de este. Imagen propia.

7.2 Anomalías Cerebrovasculares

Las anomalías cerebrovasculares, comúnmente conocidas como accidentes cerebrovasculares, representan la principal causa de daño cerebral. Entre estas, los infartos cerebrales son los más comunes, generalmente causados por la oclusión de vasos cerebrales debido a trombosis o embolia. Otro factor significativo son las hemorragias cerebrales, que ocurren mayormente a causa de la ruptura de una malformación congénita en los vasos sanguíneos del cerebro, conocida como aneurisma. Estas hemorragias pueden resultar en daños graves debido a la liberación de sangre en el tejido cerebral o en los espacios circundantes [31].

7.3 Aneurismas intracraneales

Un aneurisma intracraneal se define como una dilatación focal que se forma en las estructuras vasculares del cerebro, resultando en la debilitación de las paredes de los vasos sanguíneos. Esta condición puede ejercer presión sobre los nervios o tejidos adyacentes, y en casos graves, puede llevar a la ruptura del aneurisma, causando una hemorragia subaracnoidea al derramar sangre en los tejidos circundantes [32].

Los aneurismas intracraneales pueden presentarse en cualquier etapa de la vida, aunque son más comunes en adultos que en niños, y se observan con mayor frecuencia en mujeres que en hombres. En cuanto a su clasificación por tamaño, se categorizan de la siguiente manera: pequeños, de 5 mm o menos; medianos, de más de 5 mm pero no más de 15 mm; grandes, de más de 15 mm hasta menos de 25 mm; y gigantes, de 25 mm o más [33], [34].

7.4 Aneurisma sacular

Los aneurismas saculares, que se forman comúnmente en las bifurcaciones arteriales (15), constituyen la mayoría de los casos de aneurismas intracraneales, abarcando entre el 68% y el 98%. Estos aneurismas son frecuentemente responsables de hemorragias subaracnoideas (HSA). A diferencia de los aneurismas congénitos, los saculares suelen desarrollarse a lo largo de la vida [34], [35].

7.5 Aneurisma fusiforme

Por otro lado, los aneurismas fusiformes representan un tipo menos común. Se caracterizan por un ensanchamiento difuso que afecta a toda la arteria, sin un “cuello” definido, y suelen involucrar a la arteria madre en su totalidad. Estos aneurismas pueden ser congénitos, adquiridos, o resultar de defectos iatrogénicos en la pared vascular. Su estructura alargada y la falta de un cuello definido los distinguen claramente de los aneurismas saculares [34].

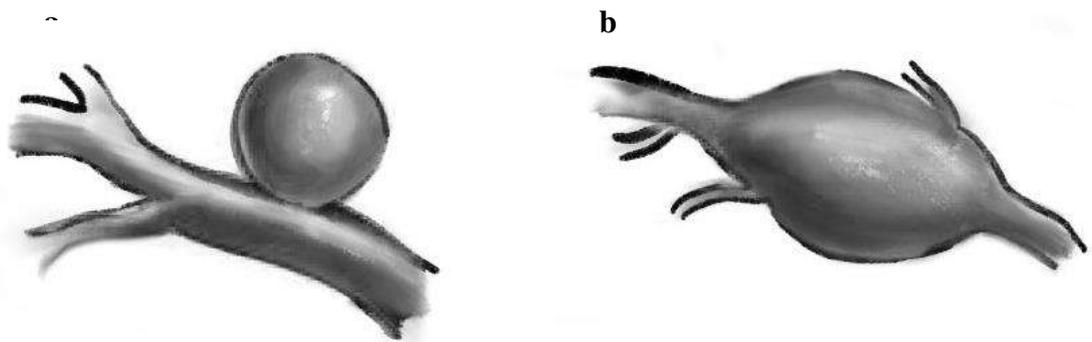


Figura 3.2. (a) aneurisma sacular, (b) aneurisma fusiforme. Imagen propia.

7.6 Procedimiento Clínico

El tratamiento de los aneurismas cerebrales puede incluir una intervención quirúrgica, especialmente en casos de ruptura o alto riesgo de ella. Para aneurismas con bajo riesgo, se opta generalmente por un seguimiento activo para monitorizar su evolución. En situaciones donde se decide proceder quirúrgicamente, existen dos técnicas principales: la colocación de clips intracraneales y la terapia endovascular.

Aunque la tendencia en las prácticas clínicas actuales se inclina hacia métodos mínimamente invasivos, el uso de clips intracraneales sigue siendo una técnica definitiva para el manejo de aneurismas cerebrales complejos. Este procedimiento implica realizar una incisión en el cráneo para exponer el cerebro, adaptándose el enfoque quirúrgico según la ubicación y el tamaño del aneurisma. Una vez localizado el aneurisma, se coloca un clip alrededor de su cuello. Este clip debe cerrar completamente las paredes del vaso sanguíneo, con el objetivo de mantener el flujo sanguíneo dentro del sistema circulatorio y prevenir que se dirija hacia el aneurisma [35].

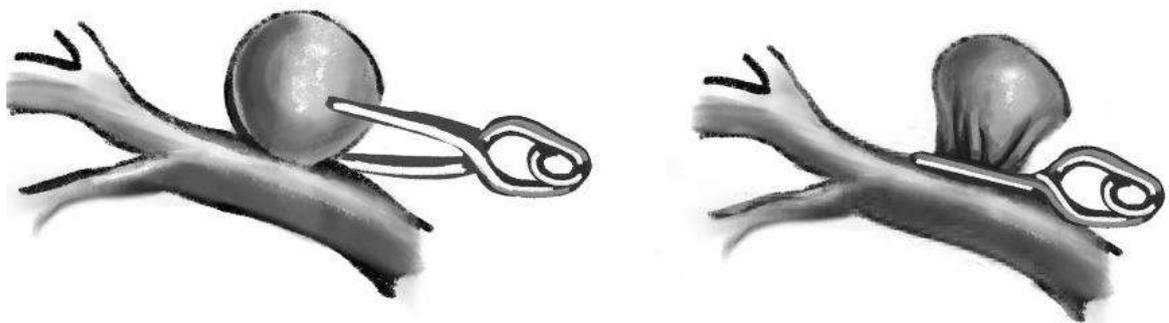


Figura 3.3. Clips implantables intracraneales sobre aneurismas. Imagen propia.

La segunda técnica de tratamiento para los aneurismas cerebrales implica un enfoque mínimamente invasivo. Este procedimiento comienza con la inserción de un catéter flexible a través de la arteria femoral. El catéter se navega cuidadosamente a través del sistema vascular hasta alcanzar el aneurisma dentro del cerebro. Una vez en posición, se pueden introducir dispositivos como mallas metálicas en forma de stents o espirales desmontables. Estos dispositivos están diseñados para bloquear el flujo de sangre hacia el aneurisma, asegurando que la circulación sanguínea se mantenga exclusivamente dentro del sistema circulatorio. Esta técnica, al ser mínimamente invasiva, ofrece ventajas en términos de recuperación y riesgos asociados en comparación con métodos más invasivos [35].



Figura 3.4.

Ejemplos de terapia endovascular para tratamiento de Aneurisma intracraneal. Imagen propia.

7.7 Imágenes médicas

Como se mencionó previamente, condiciones como las descritas en este estudio se detectan y examinan mediante estudios de imágenes médicas. En este campo, existe una diversidad de técnicas de adquisición, cada una con sus metodologías y enfoques propios. Entre ellas, la angiografía resalta por su relevancia en el diagnóstico y evaluación de aneurismas cerebrales. Esta técnica implica una visualización detallada de la vascularidad del cuerpo humano, centrándose en este caso en las angiografías cerebrales [36].

Esta técnica se divide en dos métodos principales. El primero es no invasivo, basado en las propiedades del flujo sanguíneo y su interacción con campos magnéticos, como en las resonancias magnéticas. Esta metodología distingue las señales del flujo sanguíneo de las emitidas por tejidos estacionarios. El segundo método, que implica la inyección intravenosa de un agente de contraste, mejora la diferenciación entre las estructuras vasculares y los tejidos circundantes, facilitando una visualización más clara de los aneurismas [36].

La Angiografía por Sustracción Digital (DSA) es un procedimiento notable en este campo. Se realiza en un entorno hospitalario o en un centro de radiología, donde el paciente se acuesta en una mesa de rayos X y su cabeza se inmoviliza para evitar movimientos. Después de sedar al paciente para su relajación y monitorear su actividad cardíaca mediante un electrocardiograma, se inserta un catéter en una arteria, generalmente en la ingle. Este catéter se guía cuidadosamente hasta una arteria en el cuello, utilizando imágenes de rayos X. Una vez en su lugar, se inyecta un tinte a través del catéter, y se toman radiografías para observar el desplazamiento del tinte a través de los vasos sanguíneos, destacando cualquier obstrucción del flujo sanguíneo [37].

La Angiografía de Tomografía Computarizada (CTA) es otra técnica relevante. Combina la tomografía computarizada (TC) con la inyección de un tinte de contraste para crear imágenes detalladas de los vasos sanguíneos. Durante este examen, el paciente se acuesta en una camilla que se desliza hacia el centro del escáner de TC. La máquina de rayos X rota alrededor del paciente, creando imágenes detalladas y, en algunos casos, modelos tridimensionales de los vasos sanguíneos [38].

Finalmente, la Angiografía por Resonancia Magnética (MRA) es una técnica valiosa, especialmente útil para pacientes con alergias a medios de contraste radiactivo. Utiliza ondas de radio y campos magnéticos para generar imágenes detalladas de los vasos sanguíneos [36].

Cada una de estas variantes de la angiografía ofrece ventajas únicas en la detección y análisis de aneurismas cerebrales, brindando a los médicos múltiples opciones para un diagnóstico preciso y adaptado a las necesidades individuales de cada paciente.

7.8 Visión por computadora

Afortunadamente, la visión por computadora en el ámbito médico representa un área de rápido crecimiento y potencial significativo, trascendiendo las capacidades humanas en varios aspectos cruciales. Gracias a los avances en inteligencia artificial y aprendizaje automático, se ha logrado una mejora sustancial en la efectividad del análisis de imágenes médicas como las previamente mencionadas. Estas tecnologías permiten no solo una identificación más rápida y precisa de patologías, sino también la posibilidad de detectar anomalías sutiles que a menudo pasan desapercibidas en las evaluaciones tradicionales. Además, la integración de la visión por computadora en la medicina ha abierto caminos para el desarrollo de sistemas de diagnóstico asistido por computadora (CAD), que ofrecen una segunda opinión valiosa para los profesionales de la salud. Estos sistemas están diseñados para proporcionar análisis cuantitativos, reduciendo la variabilidad y subjetividad inherentes a las interpretaciones humanas.

Otro avance significativo es la capacidad de estas tecnologías para manejar y analizar grandes volúmenes de datos, lo que es esencial dada la creciente cantidad de información generada en el cuidado de la salud. La automatización de tareas rutinarias y repetitivas, como la segmentación y clasificación de imágenes, no solo ahorra tiempo valioso, sino que también reduce la probabilidad de errores humanos.

La aplicación de la visión por computadora en la medicina no se limita solo al diagnóstico. También está revolucionando áreas como la cirugía asistida por robot, la planificación del tratamiento y el seguimiento del progreso del paciente. Por ejemplo, en la radioterapia, la visión por computadora permite un posicionamiento más preciso y una mejor adaptación del tratamiento a las necesidades individuales del paciente [39].

7.9 Preprocesamiento de imágenes médicas

Antes de sumergirnos en los intrincados detalles de los algoritmos aplicados en el ámbito de la visión por computadora para análisis de imágenes médicas, es esencial destacar la importancia del preprocesamiento de estas imágenes para entender cómo se llegan a los formatos de entrada de los modelos. Este paso crítico desempeña un papel fundamental en la optimización de los datos antes de que ingresen a los modelos de inteligencia artificial y aprendizaje profundo. Este proceso no solo implica la estandarización de formatos, sino también la transformación de datos crudos en representaciones que pudiesen resultar más significativas y manejables. En el contexto de las imágenes médicas, la estandarización de formatos se convierte en un paso esencial para garantizar la coherencia y compatibilidad entre diferentes conjuntos de datos, facilitando su procesamiento y análisis subsiguiente. Este enfoque no solo simplifica la entrada de datos para los modelos, sino que también optimiza la extracción de características relevantes, mejorando así la precisión y eficacia de los algoritmos de visión por computadora aplicados en tareas médicas críticas [7], [40]. En este contexto, exploraremos brevemente algunas de las técnicas clave de preprocesamiento utilizadas para preparar imágenes médicas, destacando cómo estas estrategias contribuyen a la excelencia y fiabilidad de los modelos de inteligencia artificial en el ámbito de la salud.

7.10 Transformación de histograma y filtrado de ruido

Uno de los procesos iniciales más fundamentales y esenciales que experimentan las imágenes médicas en su camino hacia el formato de entrada de los modelos es la aplicación de transformaciones de histogramas. La representación mediante histogramas proporciona una visualización efectiva del comportamiento de una imagen al presentar gráficamente el rango de valores de píxeles frente a su frecuencia. La capacidad de transformar una imagen mediante la aplicación de funciones específicas ofrece la oportunidad de modificar su comportamiento de manera significativa. Al elegir

cuidadosamente transformaciones adecuadas, es posible resaltar estructuras de interés, al mismo tiempo que se pueden atenuar o eliminar aquellas que no son relevantes.

En el ámbito de las manipulaciones posibles de las imágenes, la representación por histograma emerge como una herramienta valiosa. Esta representación visualiza la distribución de píxeles y sus respectivos valores dentro de un rango determinado por la profundidad del formato de la imagen.

De esta manera, se pueden llevar a cabo tareas como el aumento de contraste, brillo u otras transformaciones, como la segmentación de valores al asignar nuevos valores a cada píxel mediante una función de pertenencia [41]. Este proceso se apoya en una representación gráfica que proporciona una comprensión más clara del comportamiento interno de la imagen. Además, permite identificar patrones y comportamientos menos evidentes, contribuyendo así a la realización efectiva de la tarea específica en cuestión [40].

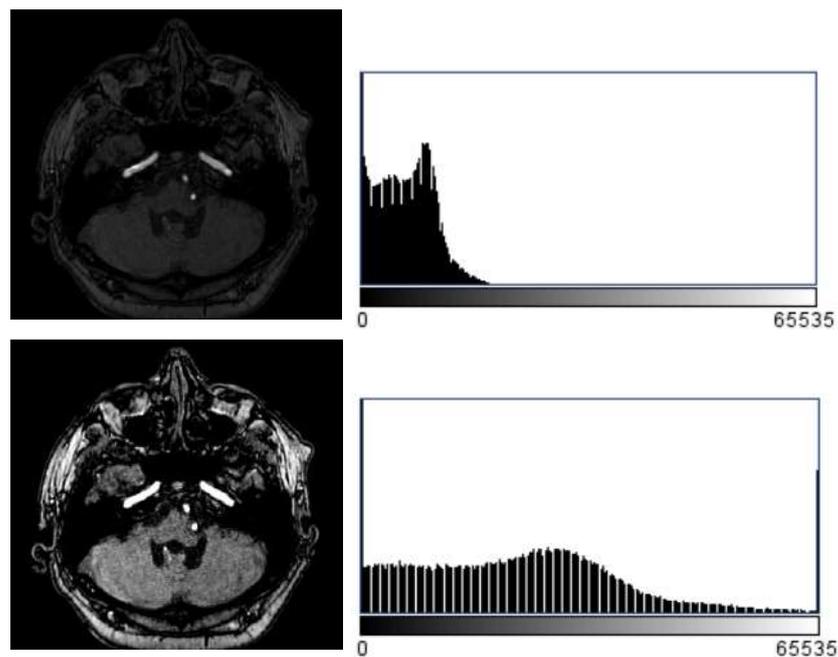


Figura 3.5. Transformación de histograma de imágenes de 16 bits. Imagen propia.

Otro aspecto del comportamiento que puede ser identificado en el histograma es la presencia de ruido. Como se mencionó anteriormente, estas variaciones aleatorias en los valores de la imagen pueden originarse durante el proceso de adquisición debido a fluctuaciones imprevistas en las señales utilizadas para generar las imágenes. Estas fluctuaciones pueden manifestarse como simples cambios de brillo o incluso como texturas no deseadas. Para abordar este comportamiento no deseado, se recurre a prácticas como el filtrado de ruido [41].

El filtrado de ruido se lleva a cabo considerando los píxeles vecinos al pixel que se está filtrando, formando así un kernel de $N \times N$. Este enfoque busca mejorar la calidad de la imagen al eliminar o reducir las variaciones no deseadas, al tiempo que facilita la posible separación de las estructuras de interés. La aplicación cuidadosa de técnicas de filtrado no solo contribuye a la mejora estética de la imagen, sino que también promueve una representación más precisa y fiable, esencial para el éxito de los modelos de inteligencia artificial y aprendizaje profundo en el posterior análisis de imágenes médicas [41].

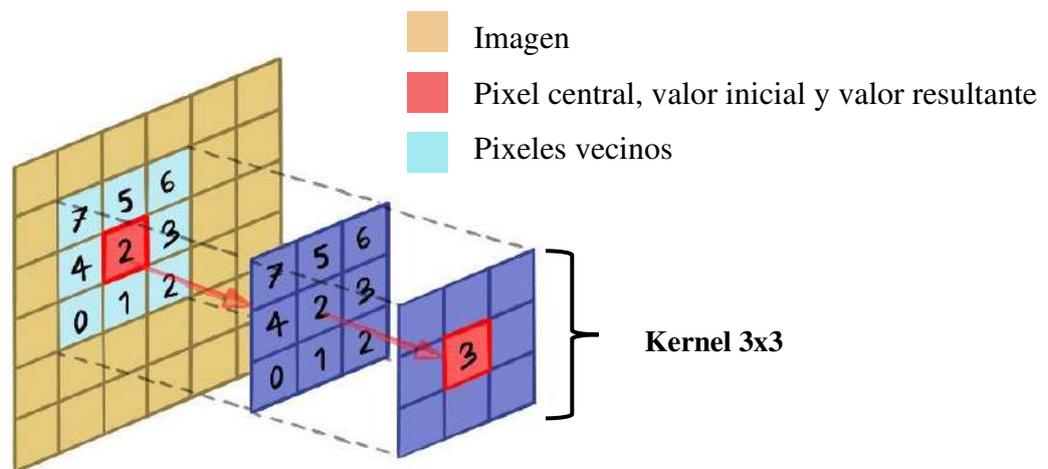


Figura 3.6. Kernel y visualización de concepto de vecindad por medio de un filtro.
Imagen propia.

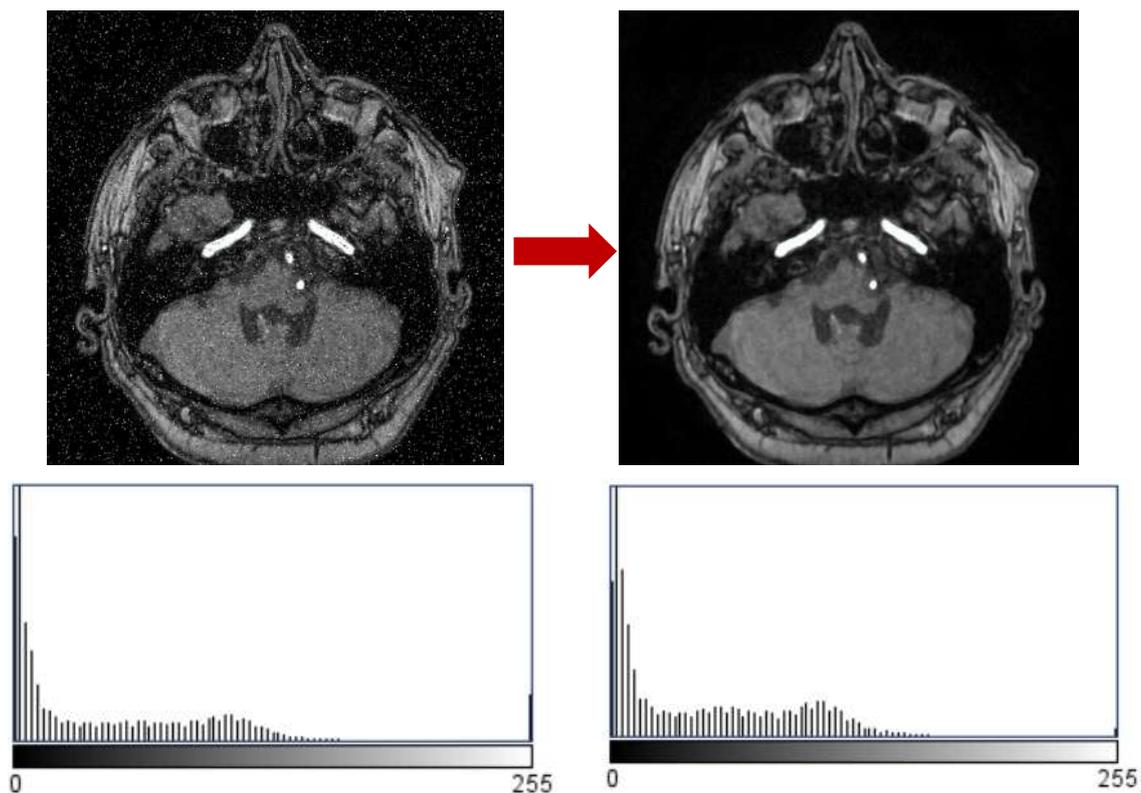


Figura 3.7. Ejemplo de filtrado de ruido en imagen de 8 bits. Imagen propia.

7.11 Reconstrucción de anatomía

Tras lograr la estandarización y el filtrado inicial de las imágenes, nuestros algoritmos avanzan hacia una etapa de mayor complejidad, adaptándose según sea necesario a diferentes formatos de entrada. En este nivel avanzado, se utilizan imágenes previamente filtradas y estandarizadas, las cuales se ajustan a una normativa dimensional y de realce específica, dependiendo del modelo y arquitectura utilizados. Estas imágenes pueden incluir, por ejemplo, versiones umbralizadas para aislar estructuras específicas de interés. Esto resulta crucial para la generalización de patrones relacionados con afecciones o anatomías específicas en aplicaciones médicas.

No obstante, este proceso no se limita a adaptaciones iniciales. Las imágenes pueden ser sometidas a una capa adicional de preprocesamiento, marcando el inicio de una transición hacia formatos más avanzados y detallados. Esta fase extendida puede incluir transformaciones que superan la mejora visual básica, orientándose hacia una representación más detallada y dimensional de estructuras anatómicas. Por ejemplo, se pueden aplicar técnicas de reconstrucción tridimensional y restauración de segmentos

anatómicos específicos, a través de algoritmos que convierten imágenes 2D estandarizadas en modelos 3D completos.

Este nivel de procesamiento abre caminos novedosos para la visualización y análisis de datos en el ámbito médico. Facilita una comprensión más profunda de las estructuras anatómicas y permite la creación de representaciones detalladas y manipulables, como la voxelización y las nubes de puntos. Estos formatos avanzados brindan una perspectiva más rica y detallada, fundamental para la generalización del modelo, ya que las relaciones espaciales se amplían más allá del formato base estándar. Esto es crucial para un análisis más profundo y preciso en la medicina moderna. Así, la reconstrucción anatómica utilizando estas técnicas avanzadas de procesamiento de imágenes emerge como una herramienta esencial en el sector de la salud. Su utilidad se extiende al diagnóstico, la investigación y el tratamiento, incluso sin la incorporación de modelos adicionales de inteligencia artificial. Estos procesos, por su eficacia inherente, se perfilan como métodos estándar potenciales en laboratorios médicos para el estudio y diagnóstico clínico.

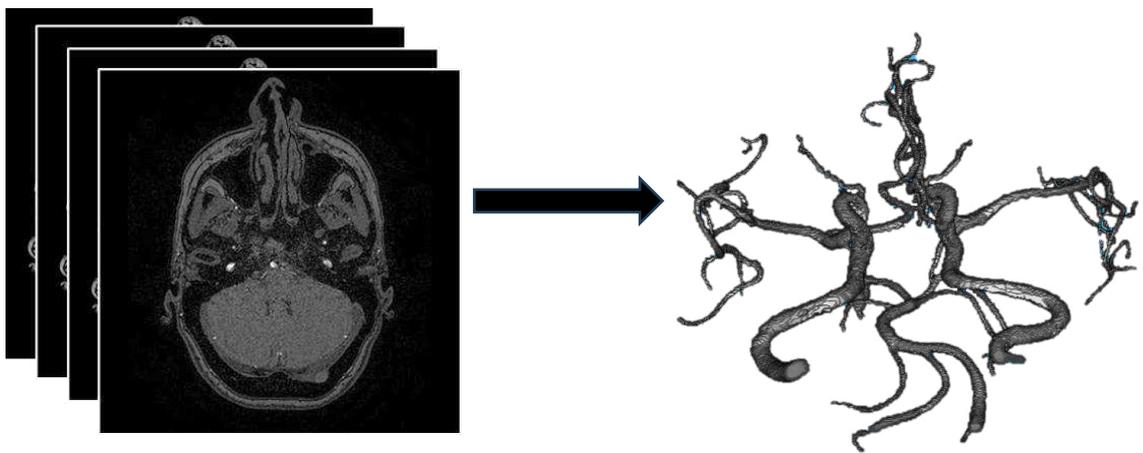


Figura 3.8. Ejemplo de reconstrucción anatómica de la estructura cerebro vascular a partir del filtrado de imágenes médicas tipo TOF-MRA. Imagen propia.

7.12 Inteligencia artificial en el área médica

Esta evolución en la reconstrucción anatómica adquiere una dimensión adicional con la integración de la inteligencia artificial. El avance en técnicas de aprendizaje automático y aprendizaje profundo ha revolucionado el ámbito médico, ampliando significativamente las posibilidades en el procesamiento de imágenes médicas y la visión por computadora. Estas innovaciones no solo complementan los métodos existentes, sino que también abren nuevos horizontes en la creación de herramientas de apoyo avanzadas. Dichas herramientas trascienden la simple visualización, tal como se emplea actualmente de manera más tradicional, contribuyendo sustancialmente en áreas como el análisis clínico, la toma de decisiones médicas y la asistencia en procedimientos quirúrgicos [20].

Los modelos de inteligencia artificial, aplicados a las imágenes ya optimizadas y detalladas obtenidas en fases previas, posibilitan interpretaciones más precisas y exhaustivas de las estructuras anatómicas. Por lo tanto, la integración de la reconstrucción anatómica avanzada con la inteligencia artificial no solo mejora la eficacia en el diagnóstico y la investigación, sino que también podría convertirse en un pilar fundamental en el tratamiento y manejo clínico, reflejando la curva de adaptación tecnológica mencionada anteriormente. En consecuencia, estas técnicas no solo emergen como complementos valiosos, sino también como componentes integrales en el futuro de la medicina diagnóstica y terapéutica, estableciendo un nuevo estándar en la práctica médica [42].

7.13 Inteligencia artificial en planeación preoperatoria

Como se ha comentado, en el ámbito de la planificación quirúrgica, el uso de imágenes médicas es un pilar fundamental para diversas prácticas rutinarias, tales como la clasificación, detección, segmentación y registro de datos cruciales en el proceso de toma de decisiones médicas. Por esta razón, resulta natural observar una progresiva incorporación y desarrollo de modelos de inteligencia artificial diseñados para desempeñar estas tareas. Estos modelos no solo aportan eficiencia temporal, sino que también ofrecen mayor facilidad al cirujano en la formulación de propuestas de intervenciones quirúrgicas, las cuales son presentadas a los pacientes.

Es importante destacar que estos modelos de inteligencia artificial facilitan, además, la creación de aplicaciones preoperatorias, como la visualización de patologías en modelos anatómicos completos antes de su segmentación. Esto permite basar la toma de decisiones

no solo en el tipo y características específicas de la afección, sino también en su interacción con estructuras adyacentes, ofreciendo así una visión más integral. Este enfoque contribuye a una mejor comprensión y visualización del cuadro clínico, mejorando de este modo el proceso quirúrgico en su conjunto [20].

7.14 Inteligencia artificial como guía intraoperativa

Continuando con la exploración del impacto transformador de la inteligencia artificial en la medicina, es crucial destacar su creciente interés en la aplicación en el ámbito quirúrgico más allá de la planificación preoperatoria. La integración de la inteligencia artificial no se limita solo a la preparación y análisis de imágenes médicas, sino que también se extiende a la ejecución misma de las cirugías. En este contexto, el siguiente aspecto a considerar es el uso de la computación asistida por inteligencia artificial como guía durante las intervenciones quirúrgicas, direccionando su uso hacia intervenciones cada vez menos invasivas y seguras [20].

7.15 Machine learning

En el apasionante campo de la medicina moderna, el Machine Learning (ML) emerge como un protagonista clave, particularmente en el análisis y reconocimiento de patrones en imágenes médicas [43]. Esta disciplina se bifurca en dos vertientes fundamentales: (I) el aprendizaje supervisado y (II) el aprendizaje no supervisado. El aprendizaje supervisado implica el aprender la relación entre las entradas (datos de características) y las salidas (etiquetas o respuestas deseadas). Este proceso se realiza mediante el uso de un conjunto de datos de entrenamiento donde cada entrada ya tiene una salida asociada conocida. La clave de este enfoque radica en la manera en que el algoritmo ajusta sus parámetros internos para poder predecir la salida más precisa para nuevas entradas, un proceso que se logra mediante la minimización de una función de pérdida [44].

Es importante destacar que el aprendizaje supervisado no solo se trata de ajustar los parámetros para ajustarse al conjunto de datos de entrenamiento; también implica la generalización. Esto significa que el algoritmo debe ser capaz de aplicar lo que ha aprendido a nuevos datos que nunca ha visto antes. Para garantizar esto, se utiliza un conjunto de datos de prueba separado del de entrenamiento. Este conjunto de datos de prueba ayuda a evaluar qué tan bien el algoritmo ha aprendido a generalizar más allá de los ejemplos específicos con los que fue entrenado [44].

En contraste, el aprendizaje no supervisado se centra en algoritmos que operan únicamente con datos de entrada, sin ninguna información previa sobre los resultados esperados. A pesar de las numerosas aplicaciones exitosas de estos métodos, suelen ser más complejos de comprender y evaluar en comparación con el aprendizaje supervisado [44]. Esta modalidad permite descubrir patrones y relaciones ocultas en los datos sin la guía explícita de un resultado predeterminado, lo que resulta útil en situaciones donde las respuestas correctas no son conocidas de antemano o son demasiado complejas para ser definidas de forma explícita.

7.16 Redes neuronales artificiales

Habiendo explorado los algoritmos de aprendizaje supervisado en la sección anterior, es esencial ahora adentrarnos en uno de los enfoques más influyentes y versátiles en la inteligencia artificial moderna, las redes neuronales, que no son solo fundamentales en el procesamiento de problemas complejos, sino que también forman la base de numerosos algoritmos avanzados utilizados actualmente. Estas estructuras, adaptadas de manera versátil, desempeñan un papel crucial en el manejo de una amplia gama de tareas. Su uso se extiende desde el procesamiento de imágenes, tanto en video como en formatos estáticos, donde son fundamentales para el reconocimiento de patrones y la interpretación visual. Esta capacidad abarca desde la identificación de objetos hasta la segmentación, así como el manejo de datos estructurados en formas tabulares, procesamiento de audio, texto, y señales eléctricas. Estos ejemplos solo raspan la superficie de las posibilidades, ya que el límite real lo impone la naturaleza del problema a abordar. Si se requiere una complejidad que las redes puedan manejar y se puede estandarizar el tipo de datos para su entrada, estas redes neuronales subrayan su centralidad tanto en técnicas novedosas como en aquellas ya establecidas, revolucionando campos que a menudo eran inimaginables en sus inicios.

El funcionamiento de las redes neuronales se inspira en el cerebro humano, donde cada 'neurona' en una red neuronal es un nodo matemático que imita la función de una neurona biológica, interconectadas entre sí y cada conexión tiene un 'peso' que determina la importancia y la influencia de una neurona sobre otra [45].

La unidad básica de este sistema es la neurona, que recibe datos de entrada ya sea del mundo externo o de las neuronas anteriores en la red, y estas entradas se ponderan según la importancia asignada a cada conexión, luego una función de activación determina si la neurona se 'activa' o no, basándose en la suma ponderada de sus entradas, lo cual ocurre en distintas capas de neuronas, empezando por la capa de entrada que recibe los datos iniciales, seguida por las capas ocultas que realizan la mayoría de los cálculos a través de sus conexiones ponderadas y finalmente la capa de salida que produce el resultado final, ya sea una clasificación, una predicción u otro tipo de salida [45].

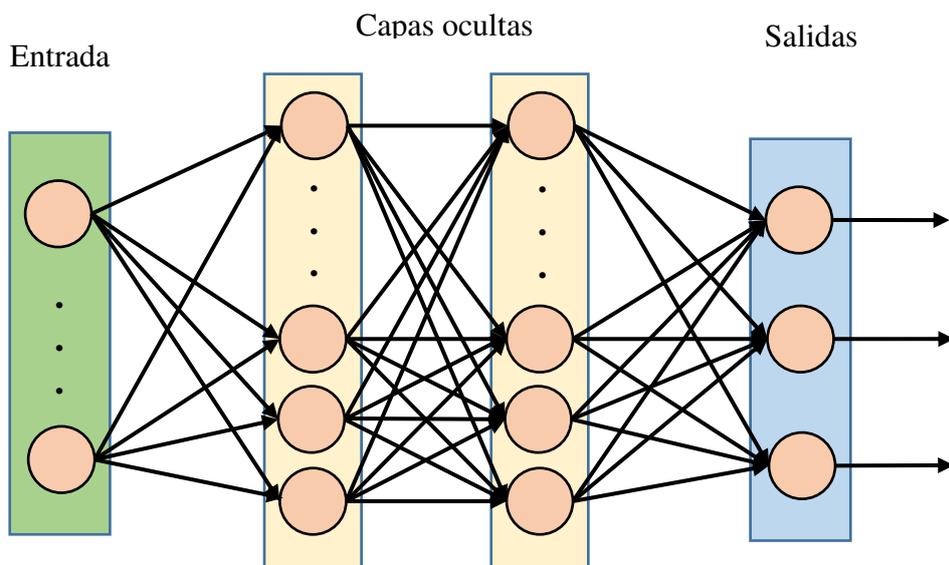


Figura 3.9. Estructura básica de una red neuronal. Imagen propia.

7.17 Deep learning

En el fascinante mundo de las redes neuronales, una evolución significativa ha llevado al surgimiento de arquitecturas cada vez más especializadas. Dentro de este espectro, el Deep Learning o aprendizaje profundo emerge como una subcategoría vital del aprendizaje automático, representando un salto cualitativo en la forma en que las máquinas aprenden y procesan la información.

El aprendizaje profundo se fundamenta en la utilización de redes neuronales con múltiples capas ocultas. A diferencia de las redes neuronales tradicionales, que pueden tener una sola capa oculta, las redes de deep learning incorporan varias de estas capas. Este diseño es lo que justifica su calificativo de "profundas".

La presencia de múltiples capas ocultas es crucial, ya que cada capa está diseñada para extraer y procesar características específicas y simples de los datos de entrada. A medida que la información avanza a través de estas capas, el nivel de abstracción aumenta, permitiendo a la red identificar y aprender patrones cada vez más complejos. Este proceso es análogo a cómo los humanos aprendemos y procesamos la información, comenzando por conceptos simples y construyendo gradualmente hacia comprensiones más complejas. Un aspecto crucial del deep learning es su dependencia y eficacia en relación con el volumen de datos utilizados para el entrenamiento. En esencia, mientras mayor sea el conjunto de datos de entrenamiento, más precisas y efectivas se vuelven estas redes en sus tareas. Esto se debe a que con una mayor cantidad de datos, la red tiene más oportunidades de aprender y ajustar sus parámetros internos, lo que resulta en un mejor reconocimiento y procesamiento de patrones complejos [45].

7.18 Transformers

En la continuación de nuestra exploración sobre el aprendizaje profundo en inteligencia artificial, nos adentramos en un concepto revolucionario: la atención [4]. La atención en inteligencia artificial se puede comprender mejor mediante una analogía avanzada de llaves y candados, donde cada dato es simultáneamente una llave y un candado, capaz de desbloquear y ser desbloqueado por otros.

En esta analogía, imagine que cada pieza de información en un conjunto de datos representa un candado que necesita ser abierto, así como una llave que puede abrir otros candados.

Aquí, el algoritmo de atención funciona cual experto cerrajero, decidiendo meticulosamente para cada dato, que funge tanto como candado como llave, qué otros elementos de información son esenciales (identificando qué llaves se acoplan mejor con qué candados) y de qué manera estos contribuyen significativamente al sentido y al contexto global.

Cada elemento de datos tiene el potencial de revelar el significado de otros elementos, y al mismo tiempo, su significado puede ser revelado por otros. El algoritmo de atención aprende estas interconexiones dinámicas, asignando pesos a las diferentes interacciones entre los datos. Estos pesos indican cuánta 'atención' debe prestar el modelo a cada interacción, permitiendo un enfoque más preciso en las conexiones relevantes y una menor atención a las menos significativas.

Esta habilidad para discernir y priorizar conexiones relevantes es lo que eleva a los modelos de atención por encima de enfoques más convencionales y los vuelve de los algoritmos más potentes en la actualidad. Al identificar qué partes de los datos son cruciales para comprender y predecir otras partes, estos modelos manejan tareas complejas con mayor eficiencia y precisión, descubriendo patrones y relaciones que serían difíciles de detectar de otra manera.

7.19 Aprendizaje contrastivo

Indudablemente, existe una amplia gama de algoritmos y técnicas avanzadas mencionables dentro de la construcción del contexto del área. Sin embargo, en lugar de intentar abordarlas todas, resulta crucial, para el marco de investigación, explorar el concepto del aprendizaje contrastivo. Aunque este concepto pueda parecer complejo, se fundamenta en una idea simple e intuitiva.

El aprendizaje contrastivo es una técnica que busca incrementar la similitud de características de elementos que deberían formar parte de un mismo espacio incrustado de características. Funciona al comparar diferentes ejemplos para aprender las características compartidas dentro de las clases de datos y los rasgos distintivos que diferencian una clase de otra [46].

Esta aproximación de aprendizaje, inspirada en la aproximación humana de comprensión por asociación, ha demostrado un notable éxito en el ámbito del Aprendizaje Profundo, convirtiéndose en un enfoque significativo en la investigación de Visión por Computadora.

En esencia, el aprendizaje contrastivo implica acercar muestras similares cerca unas de otras en un espacio semántico dado por la caracterización, mientras que se alejan las muestras de distribuciones diferentes generando una claridad más delimitada y diferenciable entre clases. Esta estrategia mejora la capacidad del modelo para discernir patrones y características dentro de los datos, contribuyendo a un rendimiento mejorado en diversas del uso de algoritmos que creen y extraigan características de los datos [47].

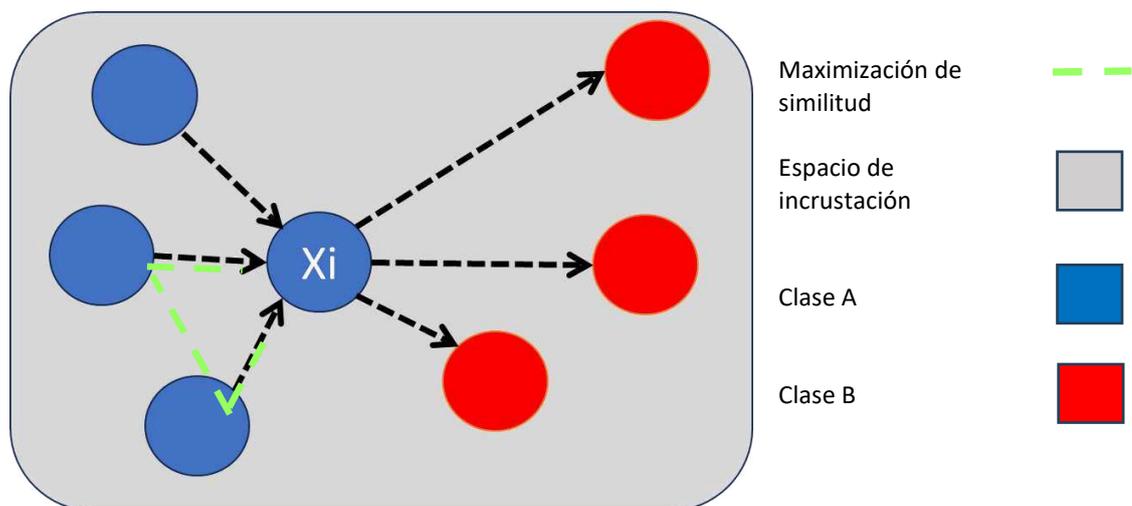


Figura 3.10. Aprendizaje contrastivo para la maximización de similitud de clases.

Imagen propia.

7.20 Función de pérdida

Como ya se mencionó antes es extensivo el número de técnicas de las que podemos hablar, dicho esto, es imperativo abordar un componente esencial que opera en el corazón de estos avanzados sistemas: la función de pérdida previamente ya mencionada inevitablemente dentro de la introducción al mundo del ML. Esta función, vital en el proceso de aprendizaje, se convierte en un elemento fundamental para la eficiencia y precisión de los algoritmos.

En el contexto del aprendizaje supervisado, del cual ya hemos hablado, la función de pérdida desempeña un rol crucial. Sirve como un barómetro que mide qué tan bien el algoritmo puede predecir resultados en base a los datos que se le presentan. En esencia, esta función calcula la discrepancia entre las predicciones del modelo y los valores reales o deseados. El objetivo principal durante el proceso de entrenamiento es minimizar esta discrepancia, optimizando así el rendimiento del modelo [48].

En las redes neuronales y el DL la función de pérdida adquiere una relevancia aún mayor. Dada la complejidad de estas estructuras, con múltiples capas y una gran cantidad de parámetros, la función de pérdida se convierte en una guía indispensable. Facilita la adaptación de los modelos para que puedan aprender de manera efectiva, identificando patrones y realizando predicciones precisas según la dirección indicada por la pérdida. En muchos casos, esta pérdida se calcula como el error directo entre la predicción y el valor real, durante su uso en el proceso de aprendizaje el modelo trata minimizar este error. Sin embargo, en otros contextos, como en el uso del algoritmo mencionado previamente, el aprendizaje contrastivo como pérdida tiene el objetivo de incrementar las probabilidades en lugar de eliminar las diferenciaciones, de cierto modo es abordar el proceso de generalización desde diferentes enfoques o incluso mediante funciones híbridas que engloben varios enfoques a la vez [46], [47] [48].

El proceso de minimización de la función de pérdida se lleva a cabo mediante técnicas como el descenso de gradiente, que ajusta iterativamente los parámetros del modelo en dirección a reducir el error. Este proceso es análogo a afinar un instrumento musical, donde cada ajuste acerca más el sonido a la nota deseada.

Además, en el ámbito del aprendizaje profundo, con arquitecturas como las redes de atención, la función de pérdida es igualmente crucial. Aquí, la capacidad del modelo para discernir y ponderar la importancia de diferentes partes de los datos es esencial. La

función de pérdida guía este proceso, asegurando que el modelo preste "atención" a los aspectos más relevantes de los datos para realizar predicciones o análisis eficientes.

7.21 Segmentación

La segmentación, un desafío esencial en el campo de la visión por computadora, busca alcanzar un entendimiento profundo de las imágenes más allá de la mera clasificación o detección de estructuras. Esta técnica involucra la habilidad de dividir una imagen o representación visual de múltiples dimensiones en distintos objetos, mediante el agrupamiento de clases o categorías, presentado en forma de máscaras. Este enfoque es crucial para generar una comprensión visual más detallada, permitiendo abordar problemas que requieren un análisis espacial más refinado, algo que las técnicas básicas de clasificación o detección no pueden proporcionar. En este contexto, la aplicación de arquitecturas avanzadas de Deep Learning y modelos de atención en la segmentación abre nuevas posibilidades. Por ejemplo, las redes neuronales profundas, con su capacidad para aprender y extrapolar patrones complejos, pueden identificar de manera más efectiva las distintas características de los objetos en las imágenes. Además, al incorporar mecanismos de atención, estos sistemas pueden focalizarse en las áreas más relevantes de la imagen, permitiendo una segmentación más precisa y contextual [49]. Esto resulta particularmente útil en aplicaciones como el diagnóstico médico por imagen, donde la precisión en la segmentación puede ser crucial, donde la correcta identificación y segmentación de elementos en el entorno es fundamental para la seguridad y eficacia del sistema.

7.22 Segmentación 3D

Extendiendo los conceptos esenciales de segmentación en el campo de la visión por computadora, nos adentramos en el desafiante territorio de la segmentación tridimensional. Mientras que la segmentación tradicional se enfoca en planos bidimensionales, la segmentación 3D lleva la comprensión visual a un nivel superior. Esta técnica no solo se ocupa de la clasificación y agrupamiento en representaciones tridimensionales, sino que también introduce la segmentación por partes, elevando nuestra capacidad de interpretar y desglosar cada componente de un objeto dentro de un espacio volumétrico. Dicha profundidad en el análisis permite una comprensión más detallada de la escena, abarcando desde las geometrías complejas hasta las sutilezas de

las texturas y formas. Este enfoque avanzado es particularmente crucial en aplicaciones como la interpretación de imágenes médicas, donde la precisión tridimensional puede ser decisiva para verdaderamente volverse un facilitador en la interpretación. La combinación de arquitecturas de Deep Learning y modelos de atención, aplicada a la segmentación 3D, explota una riqueza de información mucho mayor en comparación con el procesamiento de imágenes bidimensionales. Cada tipo de segmentación, ya sea semántica, por instancias o por partes, aporta distintos niveles de claridad y precisión según el objetivo perseguido, evidenciando la evolución y el potencial creciente de la visión por computadora [9].

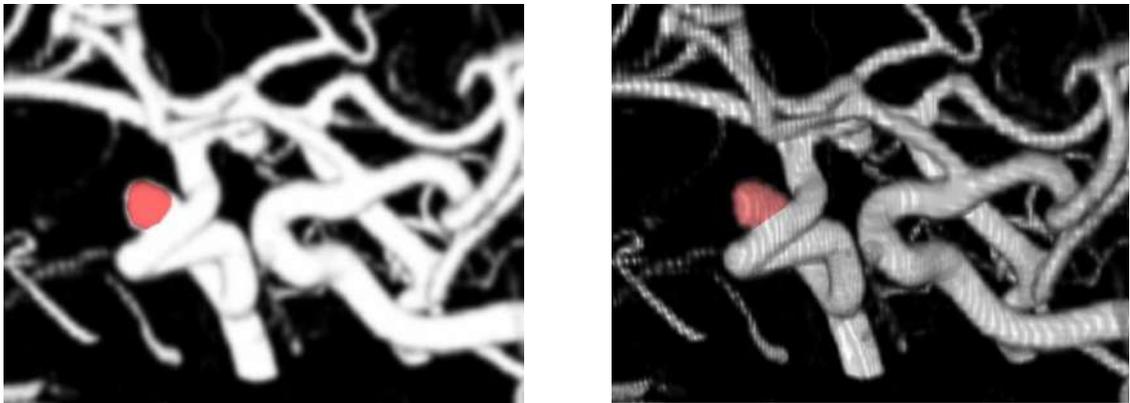


Figura 3.11. Segmentación bidimensional y tridimensional de un aneurisma intracraneal. Imagen propia.

7.23 Métricas de evaluación

Tras abordar las complejidades de los modelos en visión por computadora y la segmentación tridimensional, resulta imperativo enfocarse en la evaluación de dichos modelos. La importancia de las métricas de evaluación es crucial, pues sin ellas, discernir entre un modelo eficiente y uno deficiente sería un desafío considerable [50]. Estas métricas proporcionan una base cuantitativa para juzgar el desempeño de los sistemas de inteligencia artificial y son esenciales en el desarrollo de cualquier método predictivo.

Una selección adecuada de criterios de evaluación es primordial. Dentro del abanico de métricas habituales, encontramos la sensibilidad y la especificidad, que determinan qué proporción de casos positivos reales (como aneurismas) son identificados correctamente, y qué fracción de evaluaciones negativas son acertadas. Otras métricas populares incluyen la precisión, que mide la exactitud de las predicciones positivas realizadas por el modelo.

Además, es fundamental conocer las cuatro categorías de clasificación de predicciones: verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos, elementos clave para calcular la mayoría de estas métricas, que se pueden apreciar de forma más intuitiva dentro del cuadro 3.1.

Cuadro 3.1. Clasificación de valores predictivos.

		Presencia real de la afección	
		POSITIVA (+)	NEGATIVA (-)
Predicción	+	VP	FP
	-	FN	VN

Categorías:
 VP: Verdadero positivo
 FP: Falso positivo
 FN: Falso negativo
 VN: Verdadero negativo

Sin embargo, la evaluación de la segmentación de imágenes, especialmente en 3D, va más allá de la simple categorización de presencia o ausencia de afecciones. Se requieren métricas que aprecien directamente la calidad de las máscaras de segmentación.

La métrica esencial para la evaluación de las máscaras en la segmentación de imágenes y la más popular es la Intersección sobre Unión (IOU) [51]. Esta métrica ofrece una valoración cuantitativa del rendimiento de un modelo de segmentación, proporcionando un valor entre 0 y 1. Un valor más alto en el IOU indica una mayor concordancia entre la región predicha por el modelo y la región real, destacando así su eficacia en la precisión de la superposición y la cobertura de áreas relevantes. Esta idea se ilustra de manera clara y sencilla en la Figura 3.12, facilitando su comprensión y análisis.

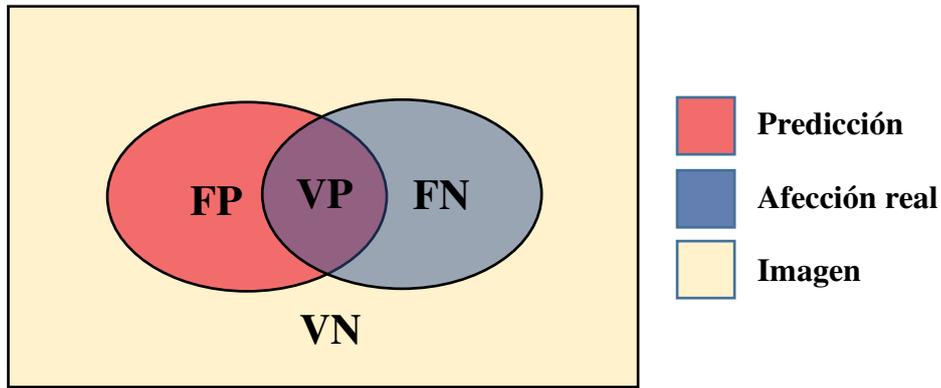


Figura 3.12. Valores predictivos en la segmentación de imágenes. Imagen propia.

La fórmula para calcular el IOU es intuitiva: se toma el número unidades de Verdaderos Positivos (VP) - aquellos correctamente identificados por el modelo - y se divide por la suma de las unidades de Falsos Positivos (FP), Verdaderos Positivos (VP) y Falsos Negativos (FN). Matemáticamente, se expresa como:

$$IOU = \frac{VP}{FP+VP+FN} \quad (1)$$

Esta ecuación proporciona un índice que mide la proporción de la intersección (área común identificada correctamente) sobre la unión (total del área cubierta por ambas la predicción y la referencia) de los conjuntos de unidades. Un valor elevado de IOU indica una alta precisión en la segmentación, reflejando una gran superposición entre la predicción del modelo y la segmentación de referencia. Por el contrario, un valor bajo sugiere una discrepancia significativa entre la segmentación predicha y la real, implicando una menor exactitud del modelo.

El IOU, por tanto, proporciona una evaluación integral y robusta del rendimiento de un modelo de segmentación, permitiendo una apreciación objetiva y cuantitativa de su precisión y eficacia en diversas aplicaciones dentro del campo de la visión por computadora.

VIII. METODOLOGIA

Esta sección se presenta el modelo Nested Contrastive Boundary Learning Point Transformer (NCBL-PT) diseñado para segmentaciones de nubes de puntos 3D, como se muestra en la Figura 6.3. Este marco utiliza el aprendizaje contrastivo para mejorar las representaciones aprendidas, maximizando la discriminación de características de clase en puntos del límite. Esta mejora permite una predicción más precisa de la clase de los puntos en proximidad a la intersección de clases, un área crítica donde las definiciones de clase pueden volverse ambiguas.

Mientras que el aprendizaje contrastivo típicamente evalúa la similitud de clase para establecer pares de relaciones positivas o negativas, este enfoque supera esto al no solo distinguir diferencias de clase, sino también identificar la pertenencia dentro o fuera del límite, añadiendo un nivel mejorado de aprendizaje consciente del borde. En consecuencia, para mejorar significativamente el rendimiento del modelo en el manejo de este desafío, se activa el marco de aprendizaje dentro de cada nube de puntos submuestreada generada en cada etapa de los bloques del codificador, refinando el proceso de aprendizaje a través de la arquitectura en términos de extracción de características.

8.1 Conjunto de Datos

En la investigación, se aprovecha del conjunto de datos Intra de acceso público, creado por Xi Yang et al. [52]. Este conjunto de datos comprende 1909 nubes de puntos de vasos sanguíneos generadas a partir de modelos 3D reconstruidos de pacientes, capturando meticulosamente las sutilezas de las estructuras vasculares. Dentro de este conjunto, se encuentran 1694 segmentos de vasos sanos distintos, proporcionando una base fundamental para análisis comparativos y evaluaciones diagnósticas, junto con 215 segmentos de aneurismas, facilitando el estudio en profundidad de condiciones vasculares patológicas.

Un aspecto destacable es que 116 de los 215 segmentos de aneurisma han sido sometidos a una subdivisión exhaustiva y anotación meticulosa por expertos médicos, mejorando significativamente la idoneidad del conjunto de datos para segmentación e investigaciones diagnósticas, un enfoque clave del trabajo. Además, la investigación propone MeshLab para generar el vector normal para cada punto. En cada subdivisión de aneurisma/vaso, cada punto se caracteriza por sus coordenadas, el vector normal

correspondiente y una etiqueta binaria que indica si es un punto de aneurisma o no. Además, se incluyen matrices de distancia geodésica pensadas para cada uno de estos segmentos 3D anotados. Esta mejora es particularmente valiosa, ya que la distancia geodésica ofrece una medición más precisa que la distancia euclidiana, adaptándose a las formas intrincadas de las estructuras vasculares. Estas matrices están formateadas como $N \times N$, optimizando la eficiencia computacional para modelos basados en puntos.

8.2 Aprendizaje contrastivo

Con el conjunto de datos Intra como piedra angular de la investigación, ahora nos centramos en la metodología que subyace al enfoque investigativo. Fundamental en la estrategia es el concepto de regularización de la clasificación categórica, que juega un papel crucial en discernir características semánticas entre datos etiquetados. Esto, a su vez, mejora la capacidad para lograr una categorización precisa y conduce notablemente a mejores resultados de segmentación.

En el corazón de este enfoque yace la técnica de aprendizaje contrastivo, que nos permite descubrir percepciones más profundas del conjunto de datos. Al aprovechar las relaciones intrínsecas incrustadas en los datos, profundizamos en la ampliamente utilizada pérdida InfoNCE [53]. Junto con su generalización [47] y el aprendizaje de límites contrastivos (CBL) propuesto por Liyao Tang et al. [54], en este contexto, el aprendizaje contrastivo está orientado a fomentar el desarrollo de representaciones de características, motivando que estas representaciones aprendidas muestren una mayor similitud con puntos cercanos dentro de la misma clase mientras que, al mismo tiempo, sean más distinguibles de aquellos que pertenecen a una clase diferente. Este delicado equilibrio está encapsulado por la Ecuación 2, que sirve como un marco fundamental que encapsula el concepto discutido como una forma de la probabilidad media de que los vecinos de un punto, basados en características aprendidas, muestren similitud dentro de su respectiva clase.

Es importante destacar que este concepto se aplica dentro de la ecuación no para promediar las probabilidades en toda la nube de puntos, sino específicamente para los puntos asociados con intersecciones de clases. Estos puntos también pueden describirse como el límite entre las clases.

$$CBL = -\frac{1}{|Bl|} \sum_{xi \in Bl} \log \frac{\sum_{xk \in Ni \wedge li=lk} \exp(-d(fi, fk))}{\sum_{xk \in Ni} \exp(-d(fi, fk))} \quad (2)$$

donde:

- $|Bl|$ número de puntos del borde Bl
- fk características aprendidas del vecino xk que pertenece al vecindario Ni
- fi características aprendidas del punto central xi .
- $-d$ distancia euclidiana negativa.

8.3 Definición de Borde

Habiendo establecido la base para el concepto de aprendizaje contrastivo con puntos característicos a lo largo del borde, el imperativo subsiguiente es localizar e identificar con precisión estos cruciales puntos de límite para permitir su incorporación en los cálculos. La metodología define estos límites como los puntos distintivos que delimitan las transiciones entre clases dentro de la nube de puntos. Para lograr esto, empleamos un método de k-vecinos más cercanos (KNN), con k establecido en 8, representando los ocho vecinos más cercanos a cada punto. En particular, incorporamos distancias geodésicas para determinar estos vecinos más cercanos. Se emplean distancias geodésicas para capturar caminos a lo largo de la superficie de la nube de puntos, proporcionando una representación más precisa de la proximidad y conectividad real entre puntos. Incluir distancias geodésicas es una elección deliberada para abordar posibles problemas de sesgo direccional inherentes a las nubes de puntos 3D irregulares. A diferencia de la distancia euclidiana, que puede introducir sesgos en una dirección particular, las distancias geodésicas representan los caminos más cortos a lo largo de la superficie de la nube de puntos. Esta consideración garantiza que los vecinos seleccionados reflejen verdaderamente las relaciones espaciales en el espacio 3D irregular. Para cada punto xi en la nube de puntos X , evaluamos si cualquiera de sus ocho vecinos más cercanos Ni exhibe una etiqueta de clase diferente del punto en cuestión. Si se cumple esta condición, clasificamos el punto como perteneciente al borde Bl . Esta clasificación se puede expresar matemáticamente de la siguiente manera:

$$Bl = \{xi \in X | \exists xk \in Ni \wedge li \neq lk\}$$

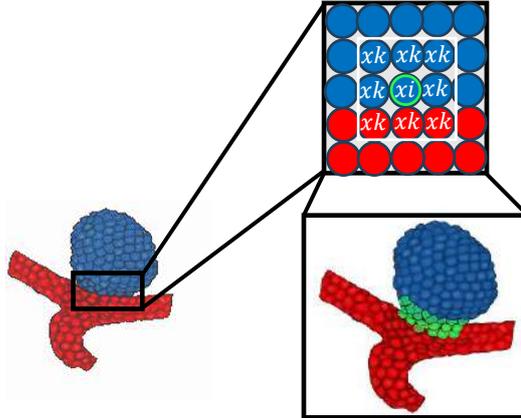


Figura 6.1. Definición de Borde. Imagen propia.

donde:

- l_i denota la etiqueta de verdad fundamental del punto x_i en la nube de puntos X .
- l_k denota la etiqueta de verdad fundamental del vecino x_k que está dentro del vecindario N_i .
- x_i es el punto central en consideración.
- x_k es un punto vecino de x_i dentro del vecindario N_i .
- N_i es el conjunto de vecindarios que contiene k vecinos más cercanos de x_i en la nube de puntos X .

Este mecanismo de etiquetado identifica efectivamente puntos asociados con transiciones de clase, formando la base fundamental para análisis de límites posteriores.

8.4 Particionamiento de los Puntos Vecinos al Borde

La segmentación puede representar una tarea desafiante al intentar discernir las diferentes partes de un solo objeto, particularmente en proximidad a intersecciones de clases. Los objetos con geometrías bien definidas pueden tener partes más fácilmente discernibles para aprender. Sin embargo, en el caso de anatomías variables, como aneurismas dentro de estructuras vasculares, que a menudo exhiben formas y patrones irregulares y difíciles de generalizar, definir las partes que constituyen la clase de aneurisma versus la clase de vaso sanguíneo se convierte en una empresa compleja. Una vez identificada la presencia de un aneurisma, distinguir sus partes del vaso sanguíneo se vuelve relativamente sencillo, al menos hasta el punto de su intersección. A medida que uno se acerca a esta intersección, definir clases se vuelve más desafiante debido a la ambigüedad semántica generada en esa región.

Por lo tanto, una vez establecido un marco de referencia, en este caso, el etiquetado de la intersección de clases se puede argumentar, incluso desde una perspectiva de aprendizaje humano, que los puntos definidos como parte del borde podrían exhibir una relación semántica más ambigua con puntos vecinos dentro del borde en comparación con aquellos fuera del borde y sus vecinos alrededor de los cuales no están en esta intersección de clases. Estos últimos generalmente muestran un menor grado de ambigüedad semántica a medida que se alejan del borde.

Para estudiar el marco de aprendizaje propuesto, es imperativo profundizar en las relaciones espaciales entre puntos dentro del mismo vecindario de clase de los puntos de borde propuestos. Esta exploración se puede dividir en dos grupos distintos, una división conceptual que se alinea con la metodología introducida por Lixia He et al [55], donde se emplean cuatro agrupaciones discretas con fines analíticos, principalmente dependiendo de si un punto reside dentro del borde o fuera de él. El marco adopta esta división fundamental de conceptos vecinos dentro del mismo vecindario de clase. Esta división potencia la capacidad del modelo para discernir si cada punto vecino del borde cae dentro de una zona caracterizada por una mayor ambigüedad semántica. Esta distinción permite que el modelo se enfoque en aprender no solo de características de puntos cercanos dentro de esta intersección de clases, sino también de puntos que existen más allá de esta intersección. El objetivo es evitar enredar al modelo en establecer relaciones de características entre puntos de diferentes clases en un área que puede exhibir ambigüedad.

Este escenario es donde los puntos de diferentes clases dentro de la intersección pueden asociarse erróneamente debido a sus similitudes de características y proximidad.

Como lo dictan las anotaciones de verdad fundamental, los puntos situados a lo largo del borde requieren un grado elevado de relevancia semántica, particularmente en relación con las características compartidas con otros puntos de la misma clase. Como resultado, utilizamos el concepto de dos grupos distintos para categorizar el vecindario de la misma clase de cada punto de borde: aquellos que comprenden el borde en sí y aquellos que se encuentran fuera de él. Esta categorización nos permite fomentar el aprendizaje de similitudes a través de dos enfoques distintos.

Por lo tanto, es imperativo que el modelo no solo aprenda a establecer similitudes con puntos cercanos de la misma clase dentro de esta intersección, sino que también priorice el aprendizaje de puntos más distantes de la misma clase que residen fuera del límite. Estos puntos distantes ofrecen un mayor grado de certeza semántica con respecto a su afiliación de clase, ya que no están situados dentro de la región ambigua de intersección de clases. En consecuencia, el modelo obtiene una base más confiable para identificarlos como miembros de la misma clase.

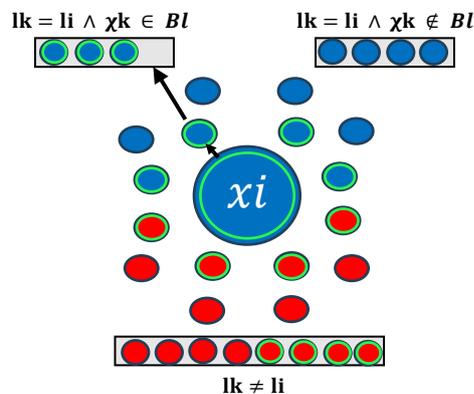


Figura 6.2. partición de vecindad del borde. Imagen propia.

Esta agrupación se puede expresar matemáticamente de la siguiente manera

$$P1 = \{\chi^k \in Ni \mid lk = li \wedge \chi^k \in Bl\}; \text{ puntos del borde con la misma categoría.}$$

$$P2 = \{\chi^k \in Ni \mid lk = li \wedge \chi^k \notin Bl\}; \text{ puntos fuera del borde con la misma categoría.}$$

8.5 Nested Contrastive Boundary Learning (NCBL)

En este contexto, aplicamos los principios del aprendizaje contrastivo para mitigar la ambigüedad semántica que puede surgir con las características aprendidas en las intersecciones de clases. Esto se logra particionando el vecindario de puntos de borde que pertenecen a la misma clase en los dos grupos distintos establecidos. Definimos el vecindario como los 16 vecinos. Aprovechamos la generalización del aprendizaje contrastivo desde las dos perspectivas distintas. Nos esforzamos por mejorar la similitud de las características para cada punto de borde con pares de clases tanto dentro como fuera del borde. Este enfoque nos permite lograr una forma de aprendizaje contrastivo anidado, como se muestra en la Ecuación 3. En esta ecuación, el logaritmo se aplica directamente a la suma de los cálculos de ambas probabilidades condicionales, probabilidades que forman la unificación de una sola probabilidad completa, la de que la vecindad en términos de características sea puramente de puntos de la misma clase, pero al anidarse se le otorga la oportunidad al algoritmo de matizarla.

Como se mencionó anteriormente, estas probabilidades significan la probabilidad de que el vecindario definido por la proximidad de las características aprendidas para cada punto de borde pertenezca a la misma clase. Ahora, refinamos este concepto distinguiendo la probabilidad de vecinos basados en características fuera del borde de aquellos dentro del borde. Este enfoque proporciona el enriquecimiento de la similitud de clases basadas en la proximidad.

$$NCBL = -\frac{1}{|Bl|} \sum_{x_i \in Bl} \log \left(\frac{\sum_{x_k \in N_i \wedge x_k \in P_1} \exp(-d(f_i, f_k))}{\sum_{x_k \in N_i} \exp(-d(f_i, f_k))} + \frac{\sum_{x_k \in N_i \wedge x_k \in P_2} \exp(-d(f_i, f_k))}{\sum_{x_k \in N_i} \exp(-d(f_i, f_k))} \right) \quad (3)$$

Esta adaptación otorga al algoritmo la capacidad de aprender de una manera específica a través del aprendizaje contrastivo, proporcionando conocimientos complementarios a partir de la información consciente del borde.

8.6 Nested Contrastive Boundary Learning Point Transformer (NCBL-PT)

El enfoque de aprendizaje del Aprendizaje Contrastivo Anidado de Límites (NCBL, por sus siglas en inglés) se integra perfectamente en la arquitectura propuesta por Hengshuang Zhao et al. [56], conocida como el Transformador de Puntos, como se ilustra en la Figura 6.3. Para adaptar el aprendizaje dentro de la arquitectura, se realizaron modificaciones a la lógica del bloque de transición hacia abajo de la arquitectura. Como se mencionó antes, NCBL requiere la utilización de características aprendidas por el modelo. Específicamente, las extracciones de características definidas por los módulos de atención del Transformador de Puntos se emplean en este caso. Además, son necesarias las etiquetas de clase de verdad fundamental para cada punto y los puntos que pertenecen al borde.

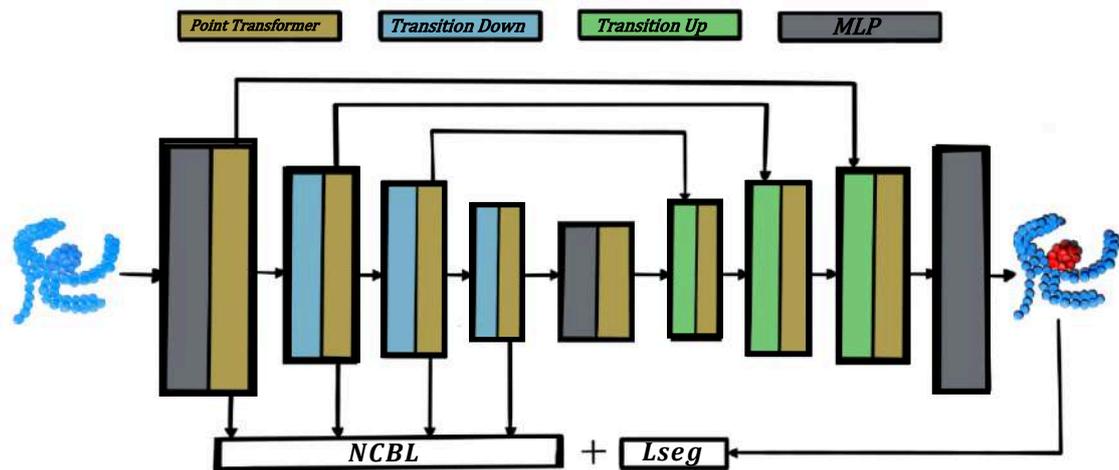


Figura 6.3. Nested Contrastive Boundary Learning Point Transformer (NCBL-PT). Imagen propia.

La arquitectura del Transformador de Puntos sigue un diseño de codificador-decodificador al estilo U-net. Sus componentes centrales incluyen el módulo Transformador de Puntos, que reemplaza las convoluciones convencionales empleando auto-atención vectorial. Para cada subconjunto $X(i) \subseteq X$, que representa un conjunto de puntos dentro de un vecindario local de x_i , aplicamos auto-atención. Específicamente, utilizamos $k=16$ vecinos más cercanos, manteniendo la consistencia con el mismo vecindario definido en los cálculos de NCBL para cada punto x_i . Esto asegura la regularización del contexto de auto-atención del vecindario local, ya que el valor de NCBL también puede considerarse como el grado de atención entre el punto x_i y los puntos vecinos dentro de la misma clase que se establecieron en las agrupaciones. Esto

reduce efectivamente la ambigüedad de las características entre puntos de la misma clase en el borde, describiendo la intersección de clases e introduciendo este nivel adicional de regularización.

La arquitectura además abarca módulos de transición hacia abajo y hacia arriba junto con un perceptrón multicapa (MLP). Como se mencionó, adaptamos el módulo de transición hacia abajo para cumplir con los requisitos del marco de trabajo de NCBL. Este módulo opera a través del muestreo de punto más lejano (FPS) y k-vecinos más cercanos (kNN) para transformar características de espacio de nube de puntos de alta dimensión C1 en un espacio de menor dimensión C2. El algoritmo FPS juega un papel crucial en el proceso de transformación del espacio C1 al C2. Inicialmente, crea un conjunto vacío para los puntos seleccionados y selecciona un punto aleatorio de C1 como punto inicial del subconjunto. Luego, el algoritmo calcula las distancias entre este punto y los demás en C1. El punto más alejado de la selección actual se agrega al subconjunto. Este procedimiento se repite iterativamente, ampliando cada vez el subconjunto con el punto más distante de los ya seleccionados. Este método tiene como objetivo preservar la estructura general y la dispersión de los datos, lo cual es crucial para mantener la fidelidad de los datos en el espacio dimensional reducido C2. El algoritmo FPS mantiene eficazmente distancias relativas y relaciones espaciales, que son esenciales para la integridad de los datos transformados. Este proceso iterativo continúa hasta que el subconjunto alcanza un número predefinido de puntos que, en este contexto, corresponde a una tasa de muestreo de reducción de 0,5 en cada bloque codificador de la arquitectura. Este enfoque es eficaz para reducir la dimensionalidad y al mismo tiempo capturar la esencia de la estructura de datos original.

Esta reducción de muestreo resulta en una reducción de la cardinalidad del conjunto de puntos en cada etapa. Durante este proceso, de acuerdo con el marco de NCBL, también transferimos las etiquetas de verdad fundamental de clase y borde a este espacio de menor dimensión, como se ilustra en las Figuras 6.4 y 6.5.

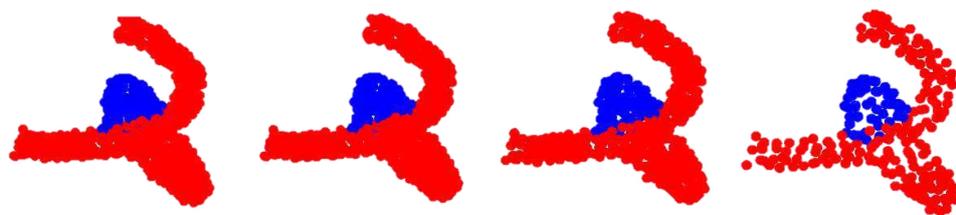


Figura 6.4. Propagación de etiquetas de clase a través de la reducción de la dimensionalidad del espacio del codificador. Imagen propia.

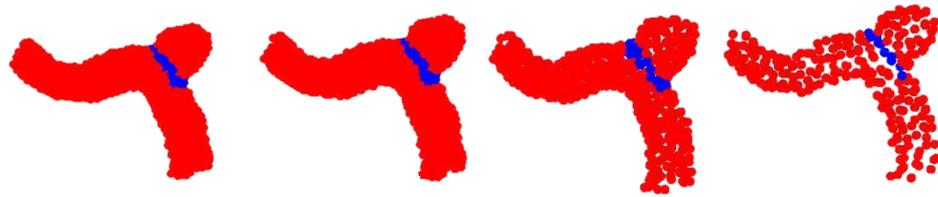


Figura 6.5. Propagación de etiquetas de clase “borde” a través de la reducción de la dimensionalidad del espacio del codificador. Imagen propia.

Para lograr esto, después de obtener un espacio de menor dimensión $C2$ a partir de $C1$, calculamos los tres vecinos más cercanos para cada punto en $C2$ de $C1$. Esto se facilita porque $C2 \subset C1$. Para las etiquetas de clase, asignamos la clase mayoritaria entre estos tres vecinos al punto correspondiente en $C2$ como se muestra en la figura 6.6. De manera similar, consideramos solo el vecino más cercano de cada punto para las etiquetas de borde. Si este vecino en $C1$ formaba parte del borde, entonces en el espacio $C2$, ese punto también pertenece al borde.

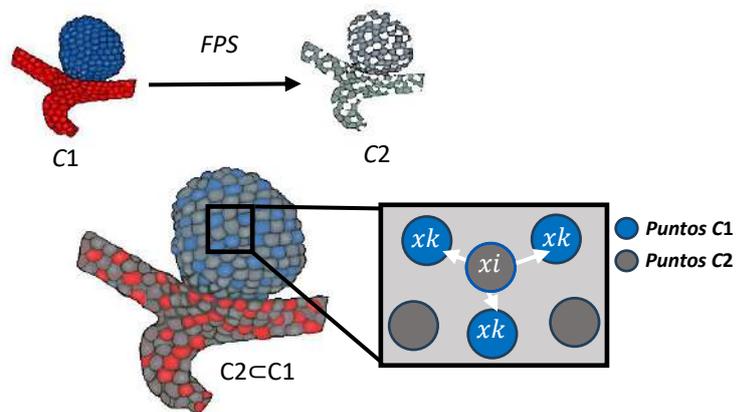


Figura 6.6. Propagación de las etiquetas de clase a través de Farthest Point Sampling (FPS) y KNN. Imagen propia.

Este proceso garantiza que, dentro de cada paso de extracción de características y reducción de dimensionalidad, cumplimos con los requisitos previos para calcular NCBL para cada módulo del codificador, aplicando efectivamente una regularización de aprendizaje para prevenir ambigüedades semánticas entre las características de los puntos de borde.

Como resultado de este enfoque integral, presentamos el Transformador de Puntos de Aprendizaje Contrastivo Anidado de Límites (NCBL-PT, por sus siglas en inglés). Este aprendizaje contrastivo anidado abarca varios niveles de la arquitectura, desde extracciones de características de bajo nivel en espacios de alta dimensión hasta el bloque final del codificador, que maneja características de alto nivel en espacios de menor dimensión.

Para navegar eficazmente por la integración de NCBL, la Figura 6.7 incluye un diagrama de flujo que aclara el flujo de trabajo integral para la integración de NCBL. Destaca los procesos específicos dentro de cada bloque codificador necesarios para la integración NCBL. Es importante destacar que el cuadro separa claramente las operaciones estándar necesarias de la arquitectura del transformador puntual de las nuevas integraciones necesarias para NCBL. Esta separación demuestra cómo los módulos nativos del transformador puntual se adaptan y utilizan dentro del marco NCBL. Dicha integración es crucial, ya que permite que los módulos NCBL operen dentro del entorno nativo del transformador puntual. Esto permite el uso de partes específicas de los procesos del transformador puntual para recopilar los datos necesarios para, en última instancia, obtener los datos necesarios para terminar con cada bloque codificador mediante el cálculo de NCBL.

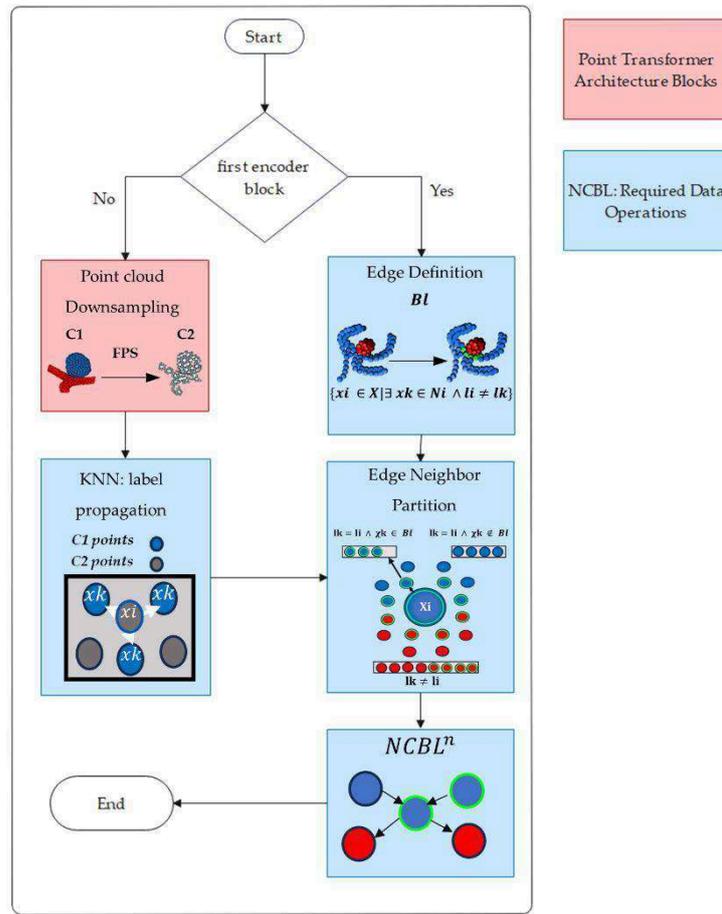


Figura 6.7. Nested Contrastive Boundary Learning Point Transformer (NCBL-PT): Diagrama de flujo que muestra lo que sucede en cada bloque para llegar a los cálculos del NCBL. Imagen propia.

Pérdida Global

La optimización de los pesos del transformador del modelo se logra a través de la función de pérdida global. Esta función de pérdida integral integra un componente tradicional orientado al rendimiento general de la segmentación semántica, junto con la metodología de aprendizaje contrastivo previamente explicada, que busca maximizar las similitudes de las características y reducir la ambigüedad semántica alrededor de la intersección de clases. La Ecuación 4 representa visualmente esta función de pérdida, combinando una pérdida de entropía cruzada supervisada basada en etiquetas semánticas (L_{seg}) con la pérdida NCBL inspirada en el aprendizaje contrastivo. aquí, $NCBL^n$ representa la pérdida de NCBL en la etapa n , agregando los resultados de cada bloque del codificador. Esta configuración de doble pérdida apunta a enriquecer el proceso de aprendizaje, capitalizando las fortalezas de los datos conscientes de la frontera etiquetados y los principios del aprendizaje contrastivo.

$$Loss = L_{seg} + \sum NCBL^n \quad (4)$$

8.7 Entrenamiento

Durante la fase de entrenamiento, utilizamos Google Colab Pro+ con sus recursos de GPU A100 para manejar eficientemente las exigentes tareas computacionales. Se empleó un tamaño de lote de 8, y el entrenamiento se llevó a cabo durante 200 épocas con una tasa de aprendizaje constante de 0.001. Para mitigar el sobreajuste, se aplicó una tasa de abandono del 0.5 durante el entrenamiento, asegurando que el rendimiento del modelo se extendiera bien a datos no vistos. Además, la densidad de la nube de puntos se estableció en 2048 puntos, proporcionando una representación adecuada para las estructuras 3D.

IX. RESULTADOS

En este estudio, se evaluó rigurosamente el rendimiento del método de segmentación propuesto comparándolo con el método base entrenado, utilizando la Intersección sobre la Unión (IoU) como la principal métrica de evaluación.

El análisis cuantitativo del estudio, detallado en el cuadro 7.1, destaca los avances del método propuesto. Muestra un notable aumento del 3,31 % en la intersección sobre la unión (IoU) para la segmentación de aneurismas en comparación con el modelo base. Esta mejora demuestra un mayor detalle de segmentación sin comprometer el rendimiento general. En la segmentación de embarcaciones, el IoU mantiene un desempeño consistente con una ligera mejora del 0,32%, equilibrando las mejoras en un área sin deterioro en otra. La Figura 7.1 ilustra visualmente estas mejoras, centrándose en la comparación con el modelo base. Esta descripción específica ayuda a comprender las mejoras del método.

El cuadro 7.1 también incluye un análisis comparativo más amplio con otros métodos de última generación que utilizan el mismo conjunto de datos y esquema de densidad de puntos. Esta comparación subraya la superioridad del enfoque sobre varios modelos líderes, incluidos Point Net++ [28], PointCNN [29], RandLA-Net [30] y el modelo de aprendizaje de rama dual propuesto recientemente [13]. El método muestra mejoras en el IoU medio del 5,04%, 6,19%, 9,13% y 22,72% con respecto a estos modelos, respectivamente. Además, en el IoU de aneurisma, nuestro método logra ganancias aún más significativas del 8,21 %, 10,84 %, 15,21 % y 32,97 %, respectivamente.

Es importante tener en cuenta el rendimiento de EPT-Net [18], que logra una mejora media general del IoU del 3,63 % con respecto al mismo modelo base. Si bien EPT-Net emplea una estrategia de modularización múltiple utilizando dos modelos en el mismo proceso, la fortaleza de nuestro método radica en su simplicidad y eficiencia. Utilizamos un modelo único con un módulo adicional dentro de la arquitectura principal, lo que ofrece ventajas significativas en cuanto a facilidad de implementación y escalabilidad potencial.

Cuadro 7.1. Comparativa del desempeño del método propuesto contra el estado del arte y el modelo base bajo el esquema de 2048 puntos.

Metodo	Vessel-IoU	Aneurysm-IoU	Mean-IoU
PointNet++	93.24%	76.21%	84.73%
PointCNN	93.59%	73.58%	83.58%
RandLA-Net	92.06%	69.21%	80.64%
Dual-branch Learning	82.65%	51.45%	67.05%
EPT-Net	95.37%	87.81%	91.59%
PT	94.8%	81.11%	87.96%
NCBL-PT	95.12%	84.42%	89.77%

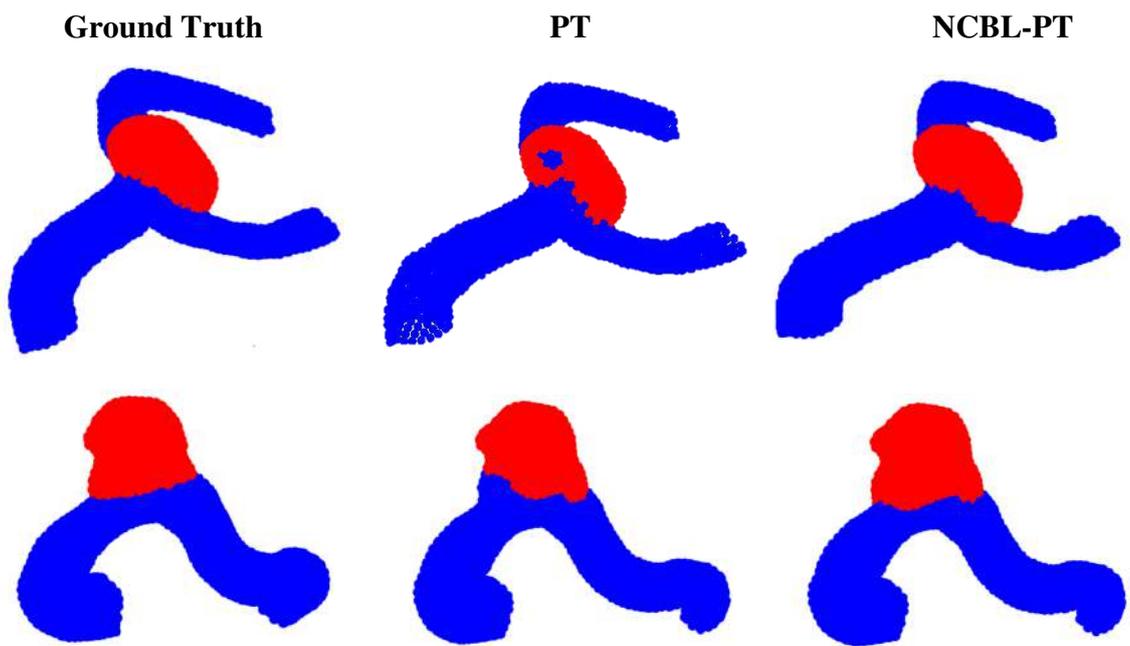


Figura 7.1. Explorando los resultados de la segmentación: análisis comparativo del etiquetado manual contra el modelo base y el método propuesto.

X. CONCLUSIONES

En este estudio, presentamos el marco de Aprendizaje de Fronteras Contrastivas Anidadas para el Transformador de Puntos (NCBL-PT), diseñado específicamente para la segmentación 3D de aneurismas basada en puntos. Hemos enfatizado la importancia de lograr una mejor similitud de las características semánticas sin necesidad de módulos complejos y enrevesados que podrían obstaculizar el proceso de aprendizaje. Al abordar dos desafíos importantes, el marco simplifica la dificultad inherente a la segmentación alrededor de los bordes, un problema que a menudo empeora por el potencial del modelo para generar características semánticamente ambiguas cuando diferentes clases están cerca. Este problema es particularmente pronunciado dadas las anatomías variables y complejas de los aneurismas. Además, nuestro enfoque reduce la dependencia de una modularización extensa para manejar esto, lo que a menudo complica el proceso de aprendizaje y da como resultado soluciones que son más difíciles de adaptar a otras aplicaciones.

Esto se logró mediante la implementación de un marco de aprendizaje de fronteras contrastivas anidadas (NCBL). El enfoque está completamente integrado en cada etapa de generación espacial y de características del bloque codificador, que se esfuerza por extraer de manera integral la caracterización de nuestra aflicción objetivo. Esta metodología no sólo aborda con eficacia los desafíos mencionados anteriormente, sino que también mejora la versatilidad del método. Facilita la integración en varios modelos para aplicaciones médicas y más, al mismo tiempo que es potencialmente adaptable a diferentes tipos de imágenes más allá de las simples nubes de puntos. Además, la estrategia reconoce el problema de la ambigüedad de las características causada por la proximidad de diferentes clases. Esto se aborda dividiendo el proceso de aprendizaje en dos enfoques distintos: uno para aprender de secciones con clases situadas muy cerca que podrían plantear ambigüedad semántica, y otro para comprender secciones con características semánticas relacionadas más claras. Este enfoque da como resultado un proceso de aprendizaje de múltiples etapas, consciente de los límites y de la ambigüedad semántica dentro de un único algoritmo flexible.

Los resultados indican firmemente que la integración del NCBL, basado en principios de aprendizaje contrastivo, reduce significativamente la ambigüedad semántica de las

características aprendidas en el contexto del aprendizaje consciente de fronteras. Esto es evidente por una mejora del 3.31% en la Intersección sobre la Unión (IoU) para la segmentación de aneurismas en comparación con el modelo base de Transformador de Puntos. Esta implementación del aprendizaje abre posibilidades emocionantes para mejorar la segmentación semántica en varios análisis médicos, refinando la delineación de intersecciones de clases y avanzando potencialmente hacia aplicaciones en cirugías asistidas por computadora.

XI. REFERENCIAS

1. D. J. Nieuwkamp et al. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis, *Lancet Neurol.* pp. 635-642. Jul 2009.
2. M. D. K. Toumeh. Detección, Diagnóstico, Tratamiento y Pronóstico del Aneurisma Cerebral sin Ruptura. Evidencias y recomendaciones. Catálogo maestro de guías de práctica clínica: IMSS-432-11.
3. Y. LeCun et al, Gradient-Based Learning Applied to Document Recognition. in *Proceedings of the IEEE*, vol. 86. pp. 2278-2324. Nov 1998.
4. A. Vaswani et al. Attention is All you Need. in *NIPS*. 2017.
5. S. Islam et al. A Comprehensive Survey on Applications of Transformers for Deep Learning Tasks. Jun 2023.
6. S. Khan et al. Transformers in Vision: A Survey. *ACM Comput.* vol. 54. pp. 1-41. Ene 2022.
7. P. Vasuki, J. Kanimozhi, y M. B. Devi. A survey on image preprocessing techniques for diverse fields of medical imagery, en *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, Karur: IEEE. pp. 1-6. Abr 2017.
8. R. Yang y Y. Yu. Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis. *Front. Oncol.*, vol. 11. p. 638182. Mar 2021.
9. Y. He et al. Deep Learning Based 3D Segmentation: A Survey. *CoRR*. 2021.
10. D. Ueda et al. Deep Learning for MR Angiography: Automated Detection of Cerebral Aneurysms. *Radiology*, vol. 290. pp. 187-194. Ene 2019.
11. J. N. Stemmer et al. Convolutional Neural Networks for the Detection and Measurement of Cerebral Aneurysms on Magnetic Resonance Angiography. *J. Digit. Imaging*, vol. 32. pp. 808-815. Oct 2019.
12. T. Sichertmann et al. Deep Learning-Based Detection of Intracranial Aneurysms in 3D TOF-MRA. *Am. J. Neuroradiol.* vol. 40. pp. 25-32. Ene 2019.
13. D. Shao, X. Lu, y X. Liu. 3D Intracranial Aneurysm Classification and Segmentation via Unsupervised Dual-Branch Learning. *IEEE J. Biomed. Health Inform.*, vol. 27. Abr 2023.
14. R. Shahzad et al. Fully automated detection and segmentation of intracranial aneurysms in subarachnoid hemorrhage on CTA using deep learning. *Sci. Rep.* vol. 10. p. 21799. Dic 2020.

15. I. Rahmany, R. Guetari, y N. Khelifa, A Fully Automatic based Deep Learning Approach for Aneurysm Detection in DSA Images, en 2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS), Sophia Antipolis, France: IEEE. pp. 303-307. Dic 2018.
16. A. Park et al. Deep Learning–Assisted Diagnosis of Cerebral Aneurysms Using the HeadXNet Model. *JAMA Netw. Open*, vol. 2. p. e195600. Jun 2019.
17. T. Nakao et al. Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. *J. Magn. Reson. Imaging*, vol. 47. pp. 948-953, Abr 2018.
18. Y. Liu, J. Liu, y Y. Yuan. Edge-Oriented Point-Cloud Transformer for 3D Intracranial Aneurysm Segmentation. en *Lecture Notes in Computer Science*, vol. 13435. Cham: Springer Nature Switzerland. 2022.
19. A. Alaraj et al. Virtual Reality Cerebral Aneurysm Clipping Simulation With Real-Time Haptic Feedback. *Oper. Neurosurg.* vol. 11. pp. 52-58. Mar 2015.
20. X.-Y. Zhou et al. Application of artificial intelligence in surgery. *Front. Med.* vol. 14 pp. 417-430. Ago 2020.
21. A. Keedy. An overview of intracranial aneurysms. *McGill J. Med.*, vol. 9. Dic 2020.
22. J. Long, E. Shelhamer, y T. Darrell. Fully Convolutional Networks for Semantic Segmentation. en 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3431-3440. Boston, MA, USA, 2015.
23. O. Ronneberger, P. Fischer, y T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. en *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. en *Lecture Notes in Computer Science*, vol. 9351. Cham: Springer International Publishing, 2015, pp. 234-241.
24. W. Hou et al. 1D CNN-Based Intracranial Aneurysms Detection in 3D TOF-MRA. *Complexity*, vol. 2020, pp. 1-13, Nov 2020.
25. K. Kamnitsas et al. DeepMedic for Brain Tumor Segmentation», en *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. en *Lecture Notes in Computer Science*, vol. 10154. Cham: Springer International Publishing, 2016, pp. 138-149.
26. F. Claux, M. Baudouin, C. Bogey, y A. Rouchaud, Dense, deep learning-based intracranial aneurysm detection on TOF MRI using two-stage regularized U-Net. *J. Neuroradiol.* vol. 50. pp. 9-15, Feb 2023.
27. C. R. Qi, H. Su, K. Mo, y L. J. Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017.
28. C. R. Qi, L. Yi, H. Su, y L. J. Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Conference on Neural Information Processing Systems (NIPS)*. 2017.
29. Y. Li, R. Bu, M. Sun, W. Wu, X. Di, y B. Chen, PointCNN: Convolution On X-Transformed Points. *Advances in Neural Information Processing Systems 31 (NeurIPS)*. pp. 820-830. 2018.
30. Q. Hu et al., RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. pp. 11105 - 11114. 2020.
31. A. K. Afifi, *Neuroanatomía funcional: texto y atlas*. 2006.
32. J. M. González-Darder, V. Quilis Quesada y Capilla Guasch. *Aneurismas Cerebrales: Tratamiento Microquirúrgico*. Barcelona: Elsevier, 2022.

33. M. Forsting y C. Cognard. Eds. Intracranial vascular malformations and aneurysms: from diagnostic work-up to endovascular therapy; 9 tables, Nachdr. en Medical radiology / ed. board L. W. Brady. Berlin Heidelberg: Springer, 2006.
34. J. Frösen et al., Saccular intracranial aneurysm: pathology and mechanisms. *Acta Neuropathol. Berl.*, vol. 123. pp. 773-786. Jun 2012.
35. A. Ringer. *Intracranial Aneurysm*. (1st ed.). Elsevier. 2018.
36. P. Suetens, *Fundamentals of Medical Imaging*, Third Edition. Cambridge University Press, 2017.
37. D. Harrington, L. Boxt, y P. Murray. Digital subtraction angiography: overview of technical principles. *Am. J. Roentgenol.* vol. 139. pp. 781-786. Oct 1982.
38. V. Baliyan, K. Shaqdan, S. Hedgire, y B. Ghoshhajra. Vascular computed tomography angiography technique and indications. *Cardiovasc. Diagn. Ther.* vol. 9. pp. S14-S27, Ago 2019.
39. A. Anaya-Isaza, L. Mera-Jiménez, y M. Zequera-Diaz. An overview of deep learning in medical imaging. *Inform. Med. Unlocked*, vol. 26, p. 100723, 2021.
40. Y. Tachibana et al. The Utility of Applying Various Image Preprocessing Strategies to Reduce the Ambiguity in Deep Learning-based Clinical Image Diagnosis. *Magn. Reson. Med. Sci.*, vol. 19. pp. 92-98, 2020.
41. P Gupta, J S Kumare, U P Singh, y R K Singh. Histogram Based Image Enhancement Techniques: A Survey. *International Journal of Computer Sciences and Engineering*. vol. 5. 2017.
42. V. Kaul, S. Enslin, y S. A. Gross. History of artificial intelligence in medicine. *Gastrointest. Endosc.* vol. 92. pp. 807-812. Oct 2020.
43. A. El-Baz, G. Gimel'farb, y K. Suzuki. Machine Learning Applications in Medical Image Analysis. *Comput. Math. Methods Med.* vol. 2017. pp. 1-2. 2017.
44. C. Shah. *A Hands-On Introduction to Machine Learning*. Cambridge University, 2022.
45. J. Schmidhuber. Deep Learning in Neural Networks: An Overview. *Neural Netw.* vol. 61. pp. 85-117. Ene 2015.
46. T. Chen, S. Kornblith, M. Norouzi, y G. Hinton. A Simple Framework for Contrastive Learning of Visual Representations. *Proceeding of the 37th International Conference on Machine Learning (ICML)*. Article No. 149. pp. 1597 - 1607. Jul 2020.
47. N. Frosst, N. Papernot, y G. Hinton. Analyzing and Improving Representations with the Soft Nearest Neighbor Loss. *Proceeding of the 36th International Conference on Machine Learning (ICML)*. Feb 2019.
48. Q. Wang, Y. Ma, K. Zhao, y Y. Tian. A Comprehensive Survey of Loss Functions in Machine Learning. *Ann. Data Sci.* vol. 9. pp. 187-212. Abr 2022.
49. S. Minaee et al. Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. vol. 44. pp. 3523 - 3542. Jul 2022.
50. D. Müller, I. Soto-Rey, y F. Kramer. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Res. Notes*, vol. 15. p. 210. Dic 2022.
51. H. Rezaatofghi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, y S. Savarese. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. en 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA: IEEE. pp. 658-666. Jun 2019.
52. X. Yang, D. Xia, T. Kin, y T. Igarashi. *IntraA: 3D Intracranial Aneurysm Dataset for Deep Learning*. April 2020. Available online: <https://github.com/intra3d2019/IntraA> (accessed on 7 July 2022).

53. A. van den Oord, Y. Li, y O. Vinyals, «Representation Learning with Contrastive Predictive Coding», ene. 2019, Accedido: 8 de octubre de 2023. Disponible en: <http://arxiv.org/abs/1807.03748>
54. L. Tang, Y. Zhan, Z. Chen, B. Yu, y D. Tao. Contrastive Boundary Learning for Point Cloud Segmentation. en 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA: IEEE, Jun 2022, pp. 8479-8489.
55. L. He, J. She, Q. Zhao, X. Wen, y Y. Guan. Boundary–Inner Disentanglement Enhanced Learning for Point Cloud Semantic Segmentation. Appl. Sci. vol. 13. p. 4053. Mar 2023.
56. H. Zhao, L. Jiang, J. Jia, P. Torr, y V. Koltun. Point Transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 97- 106. 2021