

**SEGMENTACIÓN AUTOMÁTICA DE PLEXO BRAQUIAL PARA FACILITAR
EL BLOQUEO ECO DIRIGIDO UTILIZANDO REDES NEURONALES
CONVOLUCIONALES**

NICOLAS MAURICIO RAMIREZ TRIANA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2022

**SEGMENTACIÓN AUTOMÁTICA DE PLEXO BRAQUIAL PARA FACILITAR
EL BLOQUEO ECO DIRIGIDO UTILIZANDO REDES NEURONALES
CONVOLUCIONALES**

NICOLAS MAURICIO RAMIREZ TRIANA

**Una tesis presentada en cumplimiento de los requisitos para el grado de:
Ingeniería de Sistemas e Informática**

Directora:

Lola Xiomara Bautista Rozo

PhD en Automatic, Signal and Image Processing

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2022

AGRADECIMIENTOS

Agradezco profundamente a todas las personas que de una manera u otra han estado involucradas en este proceso.

A todos los profesores y académicos de quienes tuve oportunidad de aprender, especialmente a mi directora, Lola Xiomara Bautista, quien desde el inicio de mi carrera me orientó de manera correcta para mejorar como profesional.

Al Dr. Henry Jair Mayorga, quien fue pieza fundamental a la hora de concebir la idea de este proyecto.

A mis padres Ana Judith Triana y a Cesar Augusto Ramirez, principal apoyo, guía y motivación para seguir creciendo día a día.

A mis amigos, a todos y cada uno, quienes hicieron parte de mi paso por la universidad y de quienes sigo aprendiendo cada día.

Y finalmente al Coro de la UIS y a su director Juan Manuel Hernandez, a todos aquellos que hicieron y hacen parte de esta agrupación, en la cual viví experiencias inolvidables y entendí como la academia y el arte son parte fundamental de una formación profesional integral.

Índice general

	Pág
INTRODUCCIÓN	11
1. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA	13
2. OBJETIVOS	16
3. ANTECEDENTES	17
3.1. Marco Teórico	17
3.1.1. Aprendizaje profundo	17
3.1.2. Redes neuronales convolucionales	19
3.1.3. Segmentación de imágenes	23
3.1.4. Segmentación de instancias	23
3.1.5. Funciones de pérdida	24
3.1.6. Pérdida de datos	24
3.1.7. Plexo braquial	25
3.1.8. Bloqueo nervioso	26
3.1.9. Bloqueo eco dirigido de plexo braquial	27
3.2. Estado del Arte	30
3.2.1. Aprendizaje profundo	30
3.2.2. Funciones de activación:	31
3.2.3. Métodos de segmentación de imágenes clásicos	32
3.2.4. Kaggle	33
4. METODOLOGÍA	35
4.1. Comprensión de los datos	35

4.1.1. Descripción del conjunto de datos:	35
4.2. Preparación de los datos	36
4.3. Modelado	37
4.3.1. Optimización de parámetros	41
5. EVALUACIÓN Y RESULTADOS	44
5.1. Evaluación	44
5.2. Resultados	44
5.3. Curva del coeficiente de dados (DSC)	45
5.4. Predicción	45
6. CONCLUSIONES Y PERSPECTIVAS	47
6.1. Conclusión	47
6.2. Discusión	47
6.3. Trabajo futuro	48
BIBLIOGRAFÍA	49

Índice de figuras

	Pág
Figura 1. Anatomía del Plexo Braquial	13
Figura 2. Anatomía del Plexo Braquial y ubicación en el cuerpo humano	14
Figura 3. Red Neuronal Profunda.	18
Figura 4. Kernel ajustando valores aleatorios mediante backpropagation.	20
Figura 5. Ejemplo del kernel realizando producto matricial.	21
Figura 6. Subsampling: Aplicación de Max Pooling de 2x2 reduciendo la salida a la mitad.	21
Figura 7. Formato de un solo color para la segmentación semántica.	23
Figura 8. Coeficiente de dados (teoría de conjuntos).	25
Figura 9. Anatomía del Plexo Braquial.	26
Figura 10. Posición del paciente y los instrumentos para el bloqueo eco dirigido de plexo braquial.	29
Figura 11. M-Net con entradas muestreadas y supervisión multicapa.	31
Figura 12. Segmentación resultante de uno de los proyectos propuestos para el reto de Kaggle.	34
Figura 13. Ejemplo del conjunto de datos para entrenamiento. Panel a) muestra una imagen de ultrasonido aleatoria. Panel b) Corresponde a su respectiva máscara.	36
Figura 14. Visualización 3D de la arquitectura de la U-Net	38
Figura 15. Arquitectura de la U-Net (Visualización obtenida usando ¹).	40
Figura 16. Curva de aprendizaje del coeficiente de dados DSC experimentos con diferentes parámetros.	42

¹ *PlotNeuralNet*. <https://github.com/HarisIqbal88/PlotNeuralNet>. Accessed: 2022-02-23.

Figura 17. Curva de aprendizaje del coeficiente de datos DSC experimentos con diferentes parámetros.	42
Figura 18. Curva de aprendizaje del coeficiente de datos DSC durante el entrenamiento. .	45
Figura 19. Comparación del resultado del modelo con su respectiva máscara	46

Índice de cuadros

	Pág
Tabla 1. Red Neuronal Convolucional	22
Tabla 2. Parámetros de la U-Net	43
Tabla 3. Resultados del coeficiente de datos DSC y su respectiva pérdida	46
Tabla 4. Comparación de resultados con otros trabajos	46

RESUMEN

TÍTULO: SEGMENTACIÓN AUTOMÁTICA DE PLEXO BRAQUIAL PARA FACILITAR EL BLOQUEO ECO DIRIGIDO UTILIZANDO REDES NEURONALES CONVOLUCIONALES. *

AUTOR: NICOLAS MAURICIO RAMIREZ TRIANA **

PALABRAS CLAVE: Imágenes médicas, plexo braquial, aprendizaje profundo, redes neuronales convolucionales, segmentación.

DESCRIPCIÓN: El plexo braquial es una estructura nerviosa situada en el hombro, la cual transporta señales sensoriales y de movimiento desde la médula espinal hasta los brazos y las manos. El bloqueo del plexo braquial es el método anestésico más utilizado en las intervenciones quirúrgicas de los miembros superiores, el cual consiste en realizar una punción con una aguja especializada en el hombro del paciente hacia la sección del plexo que transportará la anestesia hacia el área deseada. Mediante el uso de imágenes de ultrasonido, los profesionales capacitados pueden realizar el procedimiento de bloqueo eco dirigido del plexo braquial, que consiste en observar la dirección y posición de la aguja con un ecógrafo a medida que esta se introduce. Sin embargo, la identificación y localización del plexo braquial mediante el ecógrafo presenta una dificultad considerable debido al ruido presente en las imágenes de ultrasonido, lo que puede implicar complicaciones en la salud del paciente en caso de realizarse de manera incorrecta. Las redes neuronales convolucionales podrían facilitar este procedimiento proporcionando una identificación y segmentación automática de la región de interés. Utilizando el conjunto de datos proporcionado por la competencia pública "Ultrasound Nerve Segmentation", se implementó una arquitectura de redes neuronales convolucionales denominada U-net. La red neuronal implementada localiza y segmenta automáticamente el plexo braquial en el conjunto de datos mencionado. El modelo alcanza una puntuación de coeficiente de datos DSC=0,86 lo que permite concluir que el método propuesto localiza y segmenta satisfactoriamente el plexo braquial. Dichos resultados están en línea con otros trabajos de aprendizaje profundo, lo que sugiere que las redes neuronales convolucionales (CNN) podrían facilitar el procedimiento de bloqueo eco dirigido del plexo braquial.

* Trabajo de investigación

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Lola Xiomara Bautista Roza, Ph.D.

ABSTRACT

TITLE: AUTOMATIC BRACHIAL PLEXUS SEGMENTATION TO FACILITATE TARGETED ECHO BLOCK USING CONVOLUTIONAL NEURAL NETWORKS . *

AUTHOR: NICOLAS MAURICIO RAMIREZ TRIANA **

KEYWORDS: Medical imaging, Brachial plexus, Deep Learning, Convolutional neuronal networks, Segmentation.

DESCRIPTION: The brachial plexus is a nerve structure located in the shoulder, which carries sensory and movement signals from the spinal cord to the arms and hands. The brachial plexus block is the most commonly used anesthetic method in upper limb surgery, which consists of puncturing the patient's shoulder with a specialized needle to the section of the plexus that will transport the anesthesia to the desired area. Using ultrasound imaging, trained professionals can perform the brachial plexus echo-targeted block procedure, which involves observing the direction and position of the needle with an ultrasound machine as it is introduced. However, the identification and localization of the brachial plexus by ultrasound presents considerable difficulty due to the noise present in the ultrasound images, which may involve complications to the patient's health if performed incorrectly. Convolutional neural networks could facilitate this procedure by providing automatic identification and segmentation of the region of interest. Using the dataset provided by the public competition "Ultrasound Nerve Segmentation", a convolutional neural network architecture called U-net was implemented. The implemented neural network automatically localizes and segments the brachial plexus in the aforementioned data set. The model achieves a Dice coefficient score $DSC=0.86$ which allows us to conclude that the proposed method successfully localizes and segments the brachial plexus. Such results are in line with other deep learning works, suggesting that convolutional neural networks (CNNs) could facilitate the brachial plexus echo-directed blocking procedure.

* Research work

** Faculty of Physical-Mechanical Engineering, School of Systems and Computer Engineering. Advisor: Lola Xiomara Bautista Roza, PhD

INTRODUCCIÓN

El plexo braquial es una estructura nerviosa ubicada en el hombro, la cual transporta el movimiento y las señales sensoriales desde la médula espinal hasta los brazos y las manos. Las lesiones del plexo braquial suelen derivarse de un traumatismo en el cuello y pueden causar dolor, debilidad y entumecimiento en el brazo y la mano. La anestesia local guiada por ultrasonido, también llamado bloqueo eco dirigido, se ha convertido en un procedimiento estándar en operaciones quirúrgicas y manejo del dolor, ya que este proporciona visualización de la estructura anatómica en la cual se realizará la punción. Esta técnica es la más utilizada y estudiada para la anestesia del miembro superior con una tasa de aplicación que oscila entre el 70 % y el 97 % de todas las anestесias del miembro superior que se realizan ¹. El seguimiento nervioso representa un momento crucial para la práctica del bloqueo eco dirigido y es útil e importante desarrollar una herramienta computacional para facilitar este paso. Las complicaciones asociadas al bloqueo ecodirigido del plexo braquial incluyen la inyección intraarterial o intravenosa, que puede provocar toxicidad anestésica local ² Esto puede caracterizarse por problemas graves del sistema nervioso central, como convulsiones epilépticas, depresión del sistema nervioso central y posible coma ³. En casos graves, pueden producirse disritmias cardíacas, paros cardíacos y la muerte. Los pacientes se beneficiarían de una intervención automática.

Recientemente, se ha descubierto que las redes neuronales convolucionales (CNN) tienen un

¹ Javier Lenis Chacón y Nelson Iván Rodríguez Castro y Lucas Cordoví de Armas y Idoris Cordero Escobar y Isabel Díaz Mora. *Bloqueo del plexo braquial por vía supraclavicular y axilar guiados por ultrasonido*. 2017. URL: <http://www.revanestesia.sld.cu/index.php/anestRean/article/view/180>.

² Laurence E. Mather, Susan E. Copeland y Leigh A. Ladd. «Acute Toxicity of Local Anesthetics: Underlying Pharmacokinetic and Pharmacodynamic Concepts». En: *Regional Anesthesia & Pain Medicine* 30.6 (2005). DOI: 10.1016/j.rapm.2005.07.186.

³ Theodore E Rothman James C Crews 1. «Desarrollo de la Radiología. Centenario del descubrimiento de los Rayos X». En: *Anesth Analg* 26.4 (2003). DOI: 10.1213/01.ANE.0000053235.07186.39.

enorme potencial en el campo del análisis de las ecografías médicas ⁴. Se trata de modelos de aprendizaje profundo para localizar y segmentar automáticamente el plexo braquial. Estas implementaciones podrían asistir a los profesionales entrenados, ayudando a minimizar el riesgo que puede acarrear llevar acabo este procedimiento.

El objetivo de este proyecto es construir una herramienta computacional que permita la segmentación automática del plexo braquial en imágenes de ultrasonido para facilitar el proceso del bloqueo eco dirigido de dicha área usando redes neuronales convolucionales. Esto se realizará a través de la implementación de redes neuronales convolucionales, las cuales, en los últimos años han logrado grandes resultados en la clasificación de imágenes, la segmentación y la detección de objetos, incluso en imágenes de ultrasonido desafiantes.

El desarrollo de este proyecto se basó en una metodología utilizada en proyectos relacionados con el procesamiento de datos llamada CRISP-DM (*CRoss Industry Standard Process for Data Mining*) la cual proporciona una descripción normalizada del ciclo de vida de un proyecto estándar de análisis de datos, de forma análoga a como se hace en la ingeniería del software con los modelos de ciclo de vida de desarrollo de software. El modelo *CRISP-DM* cubre las fases de un proyecto, sus tareas respectivas, y las relaciones entre estas tareas.

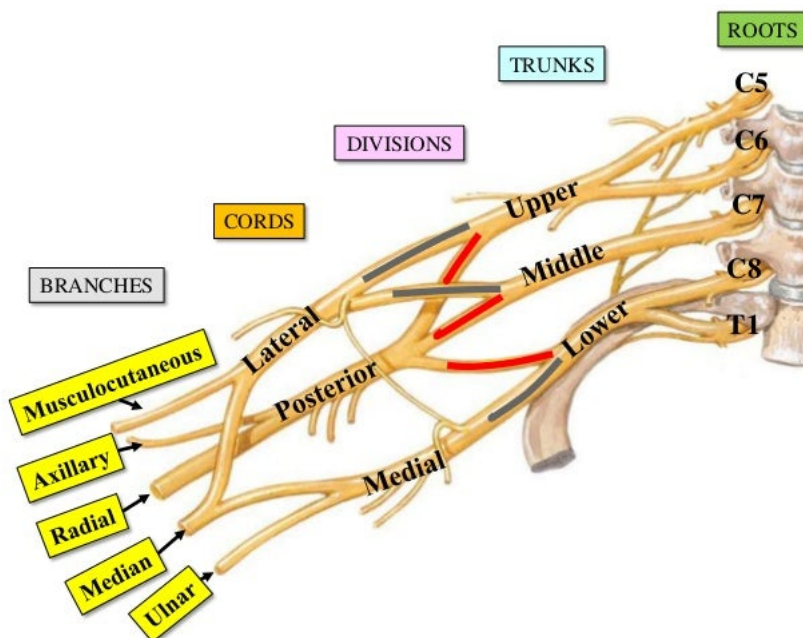
Entre los aportes de este proyecto, se encuentra desarrollar una alternativa viable y efectiva para la aplicación en el campo de la medicina, específicamente en el bloqueo eco dirigido de plexo braquial, así como generar un espacio para el desarrollo o continuación de proyectos similares, como un aporte desde la academia para la ciencia.

⁴ Sun YN Yang TH Horng MH Yang CW. «DeepNerve: A New Convolutional Neural Network for the Localization and Segmentation of the Median Nerve in Ultrasound Image Sequences». En: *Ultrasound Med Biol* 46(9).2439-2452 (2020). DOI: 10.1016/j.ultrasmedbio.2020.03.017.

1. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

El plexo braquial es una estructura nerviosa localizada en la base del cuello que conecta la médula espinal con todos los nervios periféricos del brazo. Está formado por cinco raíces espinales, las ramas ventrales de las cuatro raíces cervicales más bajas y por la primera raíz torácica (C5, C6, C7, C8 Y T1), estas raíces se unen formando tres troncos primarios (SUPERIOR: C5,C6; MEDIO: C7; INFERIOR: C8, T1), posteriormente, estos troncos se dividen en divisiones, fascículos, ramas y nervios terminales. El plexo braquial es el encargado de inervar por completo (a excepción de algunas zonas) el miembro superior.⁵ (Ver Figura 1)

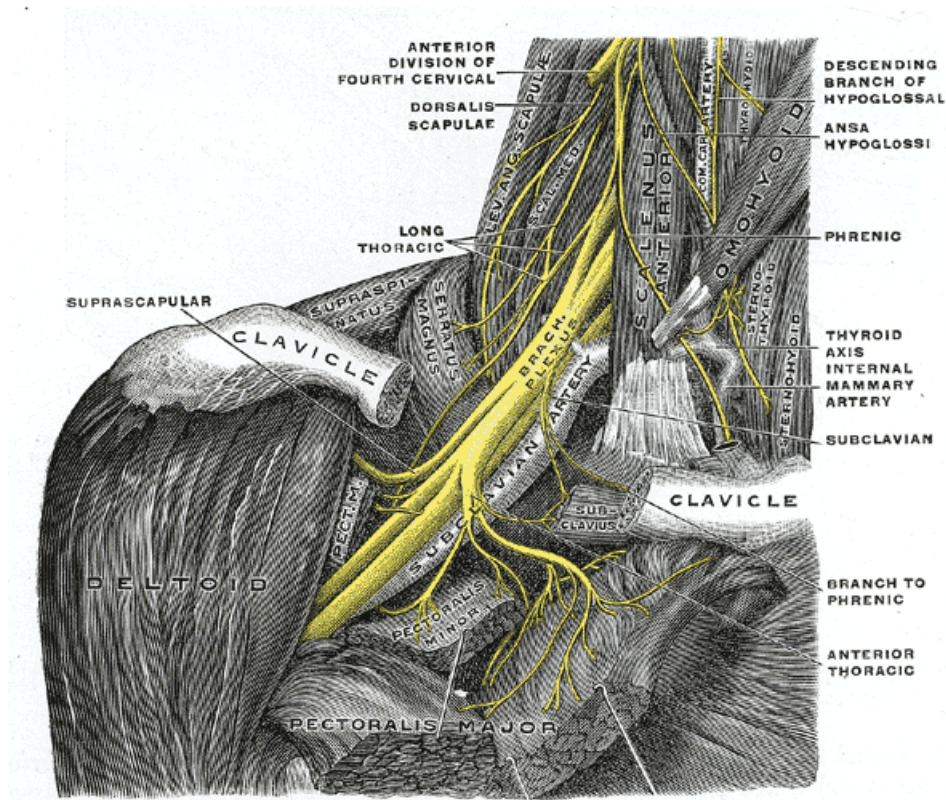
Figura 1. Anatomía del Plexo Braquial



Fuente: Restituyo, D. (2016). Resumen del Plexo Braquial. Disponible en: <https://www.acteaz.org/wp-content/uploads/2018/09/AT-More-than-Taping-Ankles.pdf>

⁵ Geena Matthew Sandeep Kusre Andrew McEwen. *Ultrasound-Guided Supraclavicular Brachial Plexus Block*. 2018. URL: <https://resources.wfsahq.org/atotw/ultrasound-guided-supraclavicular-brachial-plexus-block/>.

Figura 2. Anatomía del Plexo Braquial y ubicación en el cuerpo humano



Fuente: Herrero, B. (2012). Lesiones del Plexo Braquial. Disponible en: <https://rehabilitacionpremiummadrid.com/blog/beatriz-herrero/lesiones-del-plexo-braquial/>

Las complicaciones asociadas con el bloqueo del plexo braquial incluyen inyección intraarterial o intravenosa, que puede conducir a toxicidad anestésica local. Esto puede caracterizarse por graves problemas del sistema nervioso central como convulsiones epilépticas, depresión del sistema nervioso central y posible estado de coma. En casos graves, puede ocurrir disritmia cardíaca, paro cardíaco y muerte.²⁾ Debido a la proximidad del pulmón al plexo braquial a nivel de la clavícula, la complicación más frecuentemente asociada con este bloqueo es el neumotórax, con un riesgo de hasta el 6,1 %.⁶ Otras complicaciones del bloqueo supraclavicular incluyen la

⁶ Chan VW Brull R Macfarlane AJ Prasad GA. «Does regional anesthesia improve outcome after total knee arthroplasty?» En: *Clin Orthop Relat Res* 467 (2009). DOI: 10.1007/s11999-008-0666-9.

punción de la arteria subclavia, y la propagación del anestésico local para causar parálisis del ganglio estrellado, el nervio frénico y el nervio laríngeo recurrente. Recientemente, se ha encontrado que las redes neuronales convolucionales (CNN) tienen un enorme potencial en el campo del análisis de ultrasonidos médicos ⁴. La automatización del proceso de identificación de la ROI, (Region of Interest), utilizando segmentación de imágenes podría beneficiar el desarrollo de las prácticas de bloqueo eco dirigido y reducir los riesgos asociados con dichos tratamientos como muestran ⁴ al implementar segmentación automática en ecografías del nervio mediano usando redes neuronales convolucionales, mostrando resultados satisfactorios para la localización y segmentación del nervio mediano y también creando mediciones prometedoras para aplicaciones en el diagnóstico clínico del síndrome del túnel carpiano.

¿De qué manera se pueden implementar los algoritmos de aprendizaje profundo para realizar la segmentación automática del plexo braquial que permita facilitar el proceso de bloqueo eco dirigido de dicha área usando redes neuronales convolucionales?

2. OBJETIVOS

Objetivo general

Construir una herramienta computacional que permita la segmentación automática del plexo braquial en imágenes de ultrasonido para facilitar el proceso de bloqueo eco dirigido de dicha área usando redes neuronales convolucionales.

Objetivos específicos

- Construir un conjunto de datos a partir de ecografías de la región de interés para la identificación del plexo braquial y segmentación del mismo.
- Entrenar el modelo de segmentación con el conjunto de datos a partir de la implementación de la arquitectura de red neuronal convolucional.
- Evaluar la estrategia de segmentación propuesta mediante métricas de precisión.

3. ANTECEDENTES

3.1. Marco Teórico

El marco teórico que se describe a continuación pretende dar a conocer y orientar el enfoque que se desarrolla a lo largo de la investigación sobre la segmentación automática del plexo braquial usando aprendizaje profundo; en este marco se exponen diferentes teorías y desarrollos en distintas áreas académicas, además de establecer un hilo conductor para la ejecución del proyecto.

3.1.1. Aprendizaje profundo El aprendizaje profundo es una rama del aprendizaje automático (machine learning), caracterizado por construir redes compuestas de un determinado número de niveles jerárquicos donde el nivel inicial recibe y procesa información o “aprende” para posteriormente enviar la información procesada al siguiente nivel. La complejidad del aprendizaje aumenta a medida que se aumenta el nivel. Por otra parte,⁷, expresan que, aunque es todavía complicado definir concretamente los términos sobre el aprendizaje automático y profundo, lo definen de la siguiente manera: “Las redes neuronales profundas o DNN (Deep Neuronal Network) transforman los datos en múltiples etapas de procesamiento para producir representaciones que se pueden usar para resolver tareas complejas, como el reconocimiento de objetos en imágenes. Sin embargo, la comprensión completa de cómo logran esto sigue siendo difícil de alcanzar. La complejidad de las redes neuronales biológicas supera sustancialmente la complejidad de las DNN, lo que dificulta aún más la comprensión de las representaciones que aprenden.”

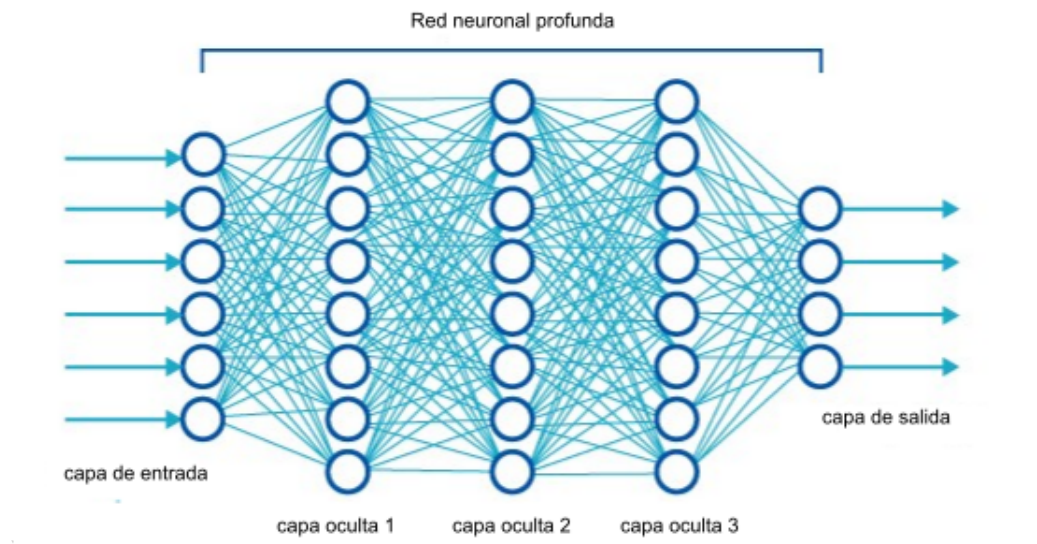
Profundizando en las Redes Neuronales Profundas o DNN, se tienen los siguientes conceptos,

⁷ Ari S.; Macke Jakob H Barrett David Gt; Morcos. «Analyzing biological and artificial neural networks: challenges with opportunities for synergy.» En: *Current opinion in neurobiology* (2019). DOI: 10.1016/j.conb.2019.01.007.

los cuales son necesarios para la implementación de modelos y entrenamiento de estos:

- **Neurona o nodo:** Es la unidad básica en una red neuronal, la cual recibe información de otros nodos para calcular una salida.
- **Nodos de entrada:** Proporcionan información del mundo exterior a la red, en conjunto son llamados capas de entrada. Su función es pasar información a los nodos ocultos.
- **Nodos ocultos:** Son los encargados de realizar cálculos y transferir información a los nodos de salida, en conjunto se les conoce como capa oculta.
- **Nodos de salida:** Son los responsables de los cálculos y transferencia de datos al exterior, en conjunto son llamados capa de salida.
- **Funciones de activación:** Introducen la no linealidad a las salidas de las neuronas, son importantes porque los datos en el mundo real no son lineales.

Figura 3. Red neuronal Profunda.



Fuente: Fuente: Sachdeva, A (2017) Deep Learning for Computer Vision for the average person. Disponible en internet: <https://medium.com/diaryofawannapreneur/deep-learning-for-computer-vision-for-the-average-person-861661d8aa61>

3.1.2. Redes neuronales convolucionales Las redes neuronales convolucionales (Convolutional neural networks), también llamadas CNN por sus siglas en inglés, son un tipo de red neuronal artificial con aprendizaje supervisado que procesa sus capas imitando al cortex visual del ojo humano para identificar distintas características en las entradas que en definitiva hacen que pueda identificar objetos y “ver” ⁸. Para ello, la CNN contiene varias capas ocultas especializadas y con una jerarquía: esto quiere decir que las primeras capas pueden detectar líneas, curvas y se van especializando hasta llegar a capas más profundas que reconocen formas complejas como un rostro o la silueta de un animal. El proceso que distingue las CNN, las llamadas “convoluciones”, ⁹ consiste en una operación entre dos funciones f y g , que produce una tercera función la cual se puede interpretar como una versión modificada ("filtrada") de f . En esta interpretación se denomina a g como el filtro. Si f se define en una variable espacial como x en lugar de una variable de tiempo como t , se llama a la operación convolución espacial. Aplicado a funciones bidimensionales como imágenes, es útil para la búsqueda de bordes, detección de características, detección de movimiento, coincidencia de imágenes, entre otros. Para las funciones $f(x)$ y $g(x)$ de una variable continua x , la convolución se define como: 1):

$$f(x, y) * g(x, y) = \int_{\tau_1=-\infty}^{\infty} \int_{\tau_2=-\infty}^{\infty} f(\tau_1, \tau_2) \cdot g(x - \tau_1, y - \tau_2) d\tau_1 d\tau_2 \quad (1)$$

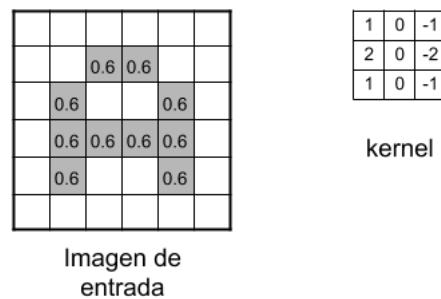
donde $*$ significa convolución y $\cdot$ significa multiplicación ordinaria para las funciones de dos variables x e y (por ejemplo imágenes). Cuando se realiza el proceso de convolución en imágenes, las cuales se pueden analizar como conjuntos ordenados de píxeles, el procedimiento consiste

⁸ N.N.; Lyakhov P.A.; Valuev G.V.; Chervyakov N.I Valueva M.V.; Nagornov. «Application of the residue number system to reduce hardware costs of the convolutional neural network implementation». En: *Mathematics and Computers in Simulation* 117 (2020). DOI: 10.1016/j.matcom.2020.04.031.

⁹ *Ultrasound Nerve Segmentation dataset*. <https://www.kaggle.com/c/ultrasound-nerve-segmentation>. 2016.

en tomar “grupos de píxeles cercanos” de la imagen de entrada e ir operando matemáticamente (producto escalar) contra una matriz denominada kernel. Este kernel, que puede ser de tamaño 3×3 píxeles, “recorre” todas las neuronas de entrada (de izquierda-derecha, de arriba-abajo) y genera una nueva matriz de salida que representará la nueva capa de neuronas ocultas. Al conjunto de kernels se les denomina filtros. Retomando el ejemplo específico presentado por ¹⁰ en una primera convolución se pueden tener 32 filtros, con lo cual realmente se obtendrán 32 matrices de salida (este conjunto se conoce como “feature mapping”), cada una de $28 \times 28 \times 1$ dando un total del 25.088 neuronas para la primer capa oculta de neuronas.

Figura 4. Kernel ajustando valores aleatorios mediante backpropagation.

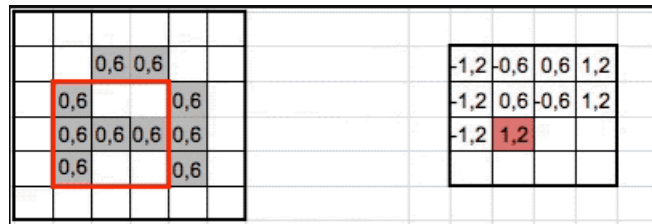


Fuente: Fuente: Na8 (2018). ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador. Disponible en internet: <https://www.aprendemachinellearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>

En la figura 4 se puede evidenciar como el kernel toma inicialmente valores aleatorios que se irán ajustando mediante backpropagation. Posteriormente, en la figura 6 se puede apreciar al kernel realizando el producto matricial con la imagen de entrada y desplazando de a 1 píxel de izquierda a derecha y de arriba a abajo para así ir generando una nueva matriz que compone al mapa de características. A medida que se va desplazando el kernel se va obteniendo una “nueva

¹⁰ Na8. ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador? <https://www.aprendemachinellearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>. 2012.

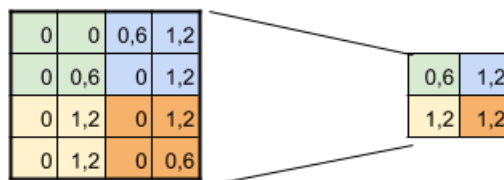
Figura 5. Ejemplo del kernel realizando producto matricial.



Fuente: Fuente: Na8 (2018). ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador. Disponible en internet: <https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>

imagen” filtrada por el mismo. En esta primera convolución y siguiendo con el ejemplo anterior, es como si se obtuvieran 32 “imágenes filtradas nuevas”. Estas imágenes nuevas aparentemente están “dibujando” ciertas características de la imagen original. Esto ayudará en el futuro a poder distinguir un objeto de otro. Luego, se procede a realizar un procedimiento donde se reducirá la cantidad de neuronas antes de hacer una nueva convolución, para ello, se realiza un proceso conocido como subsampling en el que se reduce el tamaño de las imágenes filtradas, pero en donde deberán prevalecer las características más importantes que detectó cada filtro. El tipo de subsampling más usado se denomina Max Pooling.

Figura 6. Subsampling: Aplicación de Max Pooling de 2x2 reduciendo la salida a la mitad.



Fuente: Fuente: Na8 (2018). ¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador. Disponible en internet: <https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>

Hasta ahora se ha descrito el procedimiento para realizar una primera convolución, el cual consiste en una entrada, un conjunto de filtros, generar un mapa de características y luego realizar un subsampling, obteniendo lo siguiente:

	1)Entrada: Imagen	2)Se aplica el kernel	3)Se obtiene Feature Mapping	4)Se aplica Max Pooling	5)Salida de la convolución
Primera convolución	28x28x1	32 filtros de 3×3	28x28x32	de 2×2	14x14x32
Segunda convolución	14x14x32	64 filtros de 3×3	14x14x64	de 2×2	7x7x64
Tercera convolución	7x7x64	128 filtros de 3×3	7x7x128	de 2×2	3x3x128

Tabla 1. Ejemplo de procesamiento CNN, primeras 3 convoluciones.

Como se observa en la Tabla 1, en la primera convolución la cantidad de capas es menor, debido a esto, el mapa de características solo es capaz de detectar detalles primitivos como líneas o curvas. A medida que se implementen más capas con las convoluciones, los mapas de características serán capaces de reconocer formas más complejas, y el conjunto total de capas de convoluciones podrá “ver”. La salida de la segunda convolución tendrá un tamaño 7×7 píxeles, al realizar posteriormente el Max pooling quedará en 3×3 con lo cual se puede hacer sólo 1 convolución más. En este ejemplo se partió de una imagen de 28×28 píxeles y se hicieron 3 convoluciones. Si la imagen inicial hubiese sido mayor (de 224×224 píxeles) se hubiese podido seguir haciendo convoluciones.

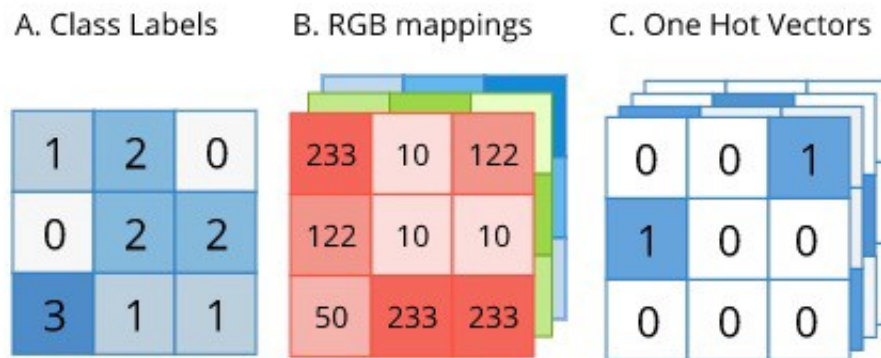
Luego de la última convolución se realiza el último paso, el cual consiste en conectar con una red neuronal “tradicional”. Para finalizar, se toma la última capa oculta a la que se le hizo subsampling, que se dice que es “tridimensional” por tomar la forma -en el ejemplo actual - $3 \times 3 \times 128$ (alto, ancho, mapas) y se realiza un proceso de “aplanamiento”, esto es que deja de ser tridimensional, y pasa a ser una capa de neuronas “tradicionales”. Por ejemplo, se podría aplanar (y conectar) a una nueva capa oculta de 100 neuronas feedforward (red neuronal recurrente).

El proceso que se realiza a partir de este momento determina si la red neuronal soluciona un problema de segmentación o de clasificación, para lo cual se expondrá el concepto de segmentación y posteriormente las funciones de pérdida.

3.1.3. Segmentación de imágenes La segmentación de imágenes es un proceso crítico en la visión por computador. Implica dividir una entrada visual en segmentos para simplificar el análisis de imágenes. Los segmentos representan objetos o partes de objetos y comprenden conjuntos de píxeles. La segmentación de imágenes ordena los píxeles en componentes más grandes, lo que elimina la necesidad de considerar píxeles individuales como unidades de observación. Existen tres niveles de análisis de imágenes, los cuales son:

- Clasificación: categorizar toda la imagen en una clase como personas, animales, exteriores.
- Detección de objetos: detectar objetos dentro de una imagen y dibujar un rectángulo a su alrededor, por ejemplo, una persona o una oveja.
- Segmentación: identificar partes de la imagen y comprende a qué objeto pertenecen. La segmentación establece la base para realizar la detección de objetos y clasificación.

Figura 7. Formato de un solo color para la segmentación semántica.



Fuente: Fuente: Raj. B (2019). Una guía simple sobre segmentación semántica. Disponible en internet: <https://medium.com/beyondminds/a-simple-guide-to-semantic-segmentation-effcf83e7e54>

3.1.4. Segmentación de instancias La segmentación de instancias es el proceso mediante el cual se busca detectar un objeto en una escena y generar una máscara que permita extraer con mayor precisión el objeto detectado, puede verse como el conjunto de dos procesos; primero,

detectar el área rectangular que contiene el objeto y luego obtener la máscara que segmenta dicho objeto.

3.1.5. Funciones de pérdida Cuando se trabaja en un problema de aprendizaje automático o de aprendizaje profundo, las funciones de pérdida / costo se utilizan para optimizar el modelo durante el entrenamiento. El objetivo es casi siempre minimizar la función de pérdida. Cuanto menor sea la pérdida, mejor será el modelo. La pérdida de entropía cruzada (Crossentropy loss) es una función de costo muy importante. Se utiliza para optimizar los modelos de segmentación. A diferencia de los clasificadores normales, se debe seleccionar una función de pérdida diferente para la segmentación.

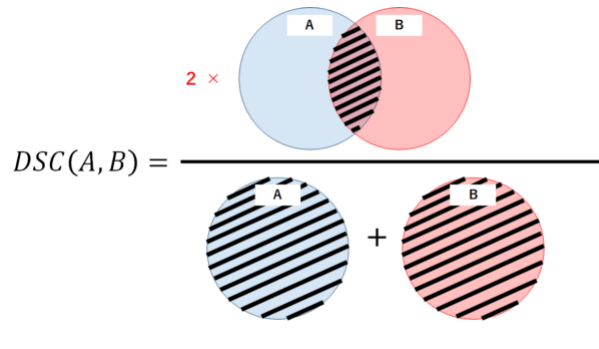
3.1.6. Pérdida de dados La pérdida de dados o Dice Loss tiene su origen en el coeficiente de S-ensen-Dice, proceso estadístico desarrollado en la década de 1940 para medir la similitud entre dos muestras referencia. ¹¹ lo lleva a la comunidad de visión por computadora para la segmentación de imágenes médicas.

$$D = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \quad (2)$$

La ecuación 2 muestra la ecuación del coeficiente de dados, en la que p_i y g_i representan pares de valores de píxeles correspondientes de predicción y exactitud, respectivamente. En el escenario de detección de límites, los valores de p_i y g_i son 0 o 1, lo que representa si el píxel es límite (valor de 1) o no (valor de 0). Por lo tanto, el denominador es la suma de los píxeles de límite total de la predicción y la exactitud, y el numerador es la suma de los píxeles de límite predichos correctamente porque la suma se incrementa solo cuando p_i y g_i coinciden (ambos de valor 1).

¹¹ Milletari. *Redes neuronales totalmente convolucionales para la segmentación de imágenes médicas volumétricas*. <https://arxiv.org/abs/1606.04797>. 2016.

Figura 8. Coeficiente de dados (teoría de conjuntos).



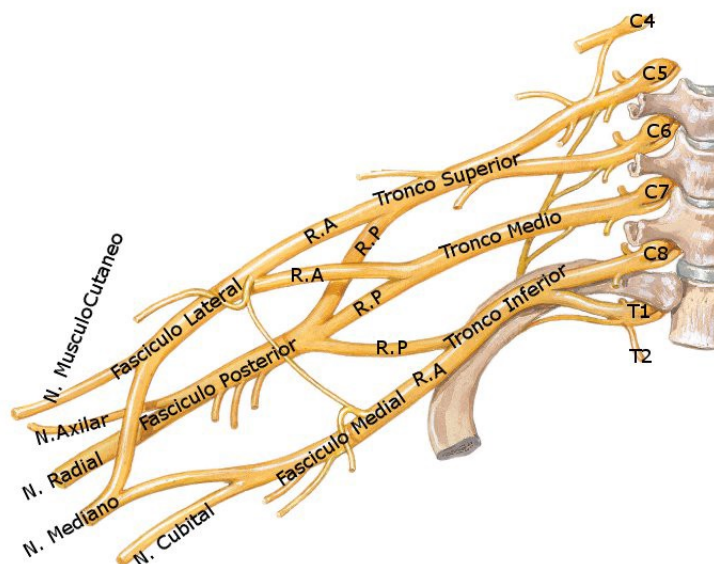
Fuente: Fuente: Du. S, 2020. Comprender la pérdida de datos para la detección de límites. Disponible en internet: <https://medium.com/ai-salon/understanding-dice-loss-for-crisp-boundary-detection-bb30c2e5f62b#:~:text=Dice%20loss%20originates%20from%20S%C3%B8rensen,between%20two%20samples%20%5BWikipedia%5D.&text=3%20from%20the%20perspective%20of,of%20overlap%20between%20two%20sets.>

La Figura 8 es otra visión de la Figura 7 desde la perspectiva de la teoría de conjuntos, en la que el coeficiente de datos (DSC) es una medida de superposición entre dos conjuntos. Por ejemplo, si dos conjuntos A y B se superponen perfectamente, DSC obtiene su valor máximo en 1. De lo contrario, DSC comienza a disminuir, llegando a su valor mínimo a 0 si los dos conjuntos no se superponen en absoluto. Por lo tanto, el rango de DSC está entre 0 y 1, cuanto mayor sea mejor. Por lo tanto, es posible usar $1 - DSC$ como pérdida de datos para maximizar la superposición entre dos conjuntos. En las tareas de detección de límites, los píxeles de límite objetivo y los píxeles de límite predichos se pueden ver como dos conjuntos. Al aprovechar la pérdida de datos, los dos conjuntos se entrenan para superponerse poco a poco. Como se muestra en la ecuación 2, el denominador considera el número total de píxeles de límite a escala global, mientras que el numerador considera la superposición entre los dos conjuntos a escala local. Por lo tanto, la pérdida de datos considera la información de pérdida tanto a nivel local como global, lo que es fundamental para una alta precisión.

3.1.7. Plexo braquial El plexo braquial es una estructura nerviosa localizada en la base del cuello que conecta la médula espinal con todos los nervios periféricos del brazo, está formado

por cinco raíces espinales, las ramas ventrales de las cuatro raíces cervicales más bajas y por la primera raíz torácica (C5, C6, C7, C8 Y T1), estas raíces se unen formando tres troncos primarios (SUPERIOR: C5,C6; MEDIO: C7; INFERIOR: C8, T1), posteriormente, estos troncos se dividen en divisiones, fascículos, ramas y nervios terminales. El plexo braquial es el encargado de inervar por completo (a excepción de algunas zonas) el miembro superior. ⁵

Figura 9. Anatomía del Plexo Braquial.



Fuente: Fuente: Restituyo, D. (2016). Resumen del Plexo Braquial. Disponible en internet: <https://medium.com/@danielrestituyo/resum%C3%A9n-del-plexo-braquial-5c406bd9eb75>

3.1.8. Bloqueo nervioso El bloqueo nervioso es un proceso que se utiliza en anestesiología en donde se inyecta un anestésico o un antiinflamatorio en un nervio o un grupo de nervios específicos con el objetivo de controlar el dolor en una zona específica del cuerpo, evitar la inflamación o eliminar las respuestas reflejas, así como interrumpir los impulsos simpáticos de

una zona en específico ¹². Este método de bloqueo del dolor es altamente eficaz y puede usarse la mayoría de los casos que requieran un bloqueo temporal de una zona específica del cuerpo. Los bloqueos nerviosos pueden ser de tipo terapéutico, para controlar el dolor por un tiempo limitado o prolongado como parte de una terapia, así mismo, en estados avanzados también puede ser usado para destruir el tejido nervioso y eliminar por completo el dolor, similar a lo que se produciría con una sección quirúrgica; diagnóstico, para conseguir información del origen del dolor y plantear un futuro tratamiento; y pronóstico, para predecir el efecto de un bloqueo antes de una intervención ¹³. Por otra parte, es posible clasificar los bloqueos de acuerdo con el lugar en el que actúan ¹².

- Bloqueo de nervios periféricos
- Bloqueo de sistema nervioso autónomo (bloqueos simpáticos)
- Bloqueo raquídeo

Para finalizar, cabe resaltar que el bloqueo del plexo braquial se clasifica como un bloqueo de nervios periféricos.

3.1.9. Bloqueo eco dirigido de plexo braquial El bloqueo de plexo braquial es un procedimiento médico perteneciente a la anestesiología en donde se anestesia el miembro superior desde la parte distal hasta el hombro, esta anestesia se hace directamente al plexo braquial. Esta técnica es la más utilizada y estudiada para la anestesia del miembro superior con una tasa de aplicación que oscila entre el 70 % y 97 % del total de anestesias de miembro superior suministradas ¹. Este procedimiento tiene 4 diferentes abordajes, supraclavicular, infraclavicular,

¹² M. S. Acedo Gutiérrez. *Tratamiento del dolor: técnicas locorregionales. Bloqueos. Oncología*. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0378-48352005000300007. 2005.

¹³ T Wheeler. *Manejo del dolor y bloqueos nerviosos*. <https://www.webmd.com/pain-management/guide/nerve-blocks#2>. 2020.

axilar e Inter escaleno. Para el procedimiento de bloqueo de plexo braquial, es necesario contar con:

- Transductor lineal de alta frecuencia (10 a 15MHz)
- Aguja aislada punta roma para neuroestimulación de 50 mm (si la constitución del paciente lo determina puede necesitarse un tamaño diferente)
- Estimulador de nervio periférico y electrodo de superficie
- Agujas de infiltración local
- Gasas, guantes estériles, lápiz y regla para dibujar las referencias de superficie jeringas preparadas con la solución a utilizar

Posición del paciente: Como se muestra en la Figura 10 , el paciente debe estar sentado con una inclinación entre 30° y 45° con la cabeza elevada y girada levemente hacia el lado contrario del bloqueo, apoyada en una almohada. El profesional debe estar parado del lado donde se va a practicar la anestesia, con el monitor de ultrasonido del lado contrario para permitir una mejor visualización. Esta posición permite un mejor control del paciente mientras se ejecuta el procedimiento.

Realización del bloqueo: El bloqueo se ejecuta en 4 pasos. Primero, debe realizarse el escaneo con el transductor lineal en diferentes ángulos hasta obtener la mejor imagen posible; Segundo, se debe localizar los nervios por correlación, es importante identificar la arteria subclavia, la primera costilla y el plexo braquial; Tercero, la inserción de la aguja, este paso se debe insertar la aguja a lo largo del eje longitudinal del transductor, con el fin de visualizar de mejor manera la aguja durante todo el proceso, se debe avanzar la aguja hasta la intersección entre la arteria subclavia y la primera costilla hasta la zona conocida como “Bolsillo de la esquina”; Para concluir, la aplicación del anestésico en la zona, lo ideal es inyectar la mitad del anestésico debajo del plexo braquial y la otra mitad se aplica en el centro del racimo del plexo, por lo

cual la aguja debe ser redirigida a esa posición. Al finalizar, se espera que el plexo braquial este completamente rodeado de anestésico ⁵.

Figura 10. Posición del paciente y los instrumentos para el bloqueo eco dirigido de plexo braquial.



Fuente: Disponible en internet: <https://www.anestesiologoemonterrey.com/uncategorized/beneficios-de-la-anestesia-regional-por-ultrasonido/>

3.2. Estado del Arte

Los antecedentes expuestos a continuación, son estudios o aplicaciones que han girado en torno al desarrollo de técnicas de segmentación en imágenes médicas.

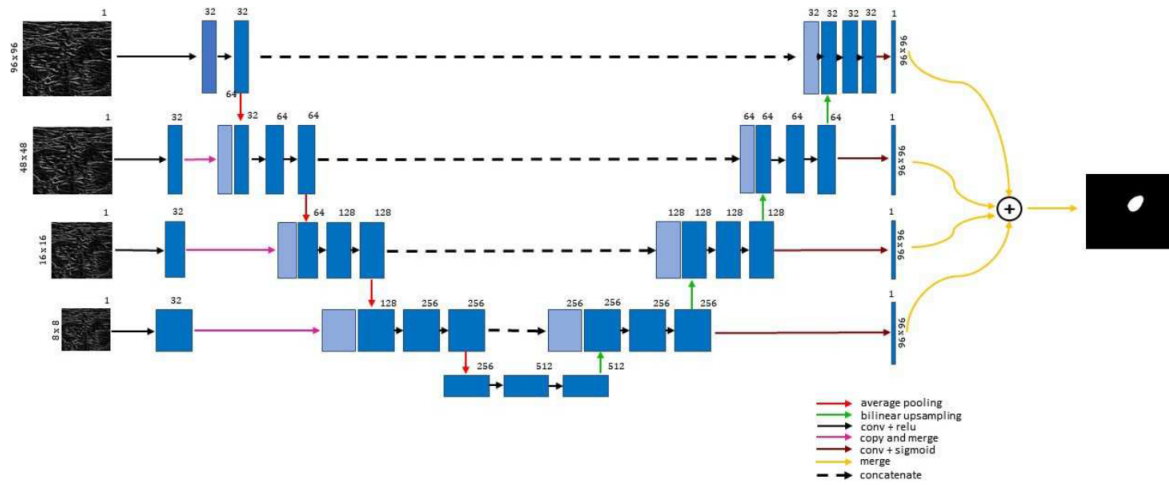
3.2.1. Aprendizaje profundo

U-Net: ¹⁴ Entre los grandes avances que se han realizado en la última década en cuanto a la segmentación de imágenes médicas, se presenta el caso de la arquitectura U-Net, la cual fue desarrollada en el Departamento de Ciencias de la Computación de la Universidad de Freiburg en Alemania. A nivel general, esta consiste en una red totalmente convolucional conformada por una red encoder y seguida por una red decoder, que consiste en la repetición de convoluciones acompañadas de una unidad lineal rectificada y una operación de agrupación máxima. Durante este evento la información espacial se reduce mientras que las características aumentan. Desde su creación, U-Net ha sido la base para muchos proyectos de investigación sobre segmentación de imágenes médicas, tal es el caso de la implementada por ¹⁴ donde se utiliza una variación de la anterior mencionada U-Net. Este artículo exploró la segmentación de los nervios del plexo braquial a partir de ecografías. Se concluyó que U-Net puede delinear los límites del nervio con ayuda de un filtro basado en bordes que pre procesa las imágenes. Un segundo experimento con el novedoso M-Net produce resultados prometedores en la misma prueba datos sobre los que se evaluó U-Net sin acceso a datos de prueba etiquetados.

Otro estudio relacionado donde también se utiliza una variación de U-Net llamada Deep Nerve ⁴ propone un marco de red neuronal convolucional para la localización y segmentación del nervio mediano. DeepNerve integra las características de MaskTrack y la memoria convolucional a corto plazo para localizar y segmentar eficazmente el nervio mediano con el objetivo de identi-

¹⁴ Khan Androutsos Abraham Kandasamy. «Deep Learning for Semantic Segmentation of Brachial Plexus Nerves in Ultrasound Images Using U-Net and M-Net». En: (). DOI: 10.1109/ICISPC.2019.8935668.

Figura 11. M-Net con entradas muestreadas y supervisión multicapa.



Fuente: Nabila Abraham, Kandasamy Illanko, Naimul Khan, Dimitri Androutsos. (2019). Deep Learning for Semantic Segmentation of Brachial Plexus Nerves in Ultrasound Images using U-Net and M-Net Disponible en: <https://doi.org/10.1109/ICISPC.2019.8935668>

ficar el síndrome del túnel carpiano en un determinado paciente. Sobre la base de los resultados experimentales, el modelo propuesto logró un alto rendimiento y generó valores promedio de coeficiente de dados (DSC), precisión, recuperación y puntuación 0.8975, 0.8912, 0.9119 y 0.9015, respectivamente.

3.2.2. Funciones de activación: Una función de activación devuelve la salida que será generada por la neurona a partir de una o varias entradas. Generalmente se buscan funciones cuyas derivadas sean simples para minimizar el coste computacional. A cada una de las capas de la red neuronal le corresponde una función de activación, la cual le permitirá predecir o reconstruir.

Función de activación ELU: La Unidad Lineal Exponencial (ELU) Eq. 3 es una función de activación para redes neuronales. A diferencia de las ReLU, las ELU tienen valores negativos que permiten acercar las activaciones unitarias medias más cerca de cero, como la normalización por lotes, pero con una menor complejidad computacional. Los cambios medios hacia cero aceleran

el aprendizaje al acercar el gradiente normal al gradiente natural unitario debido a un efecto de desplazamiento de sesgo reducido. Las ELU se tienden a saturarse a un valor negativo con entradas más pequeñas y, por lo tanto, disminuyen la variación y la información propagadas hacia adelante.

$$\begin{cases} f(x); & \text{si } x > 0 \\ \alpha(\exp(x) - 1); & \text{si } x \leq 0 \end{cases} \quad (3)$$

BPMSegnet: ¹⁵ propone otro método para segmentar múltiples instancias en una imagen de ultrasonido. Específicamente, el primer conjunto de datos de imágenes de ultrasonido del plexo braquial (UBPD) y una red (BPMSegNet) se proponen para la segmentación del plexo braquial y su anatomía circundante. Se realizan experimentos cuantitativos para verificar la superioridad de la red propuesta en el conjunto de prueba de UBPD, y la red logra el objetivo de manera consistente tanto en tareas de detección como de segmentación.

3.2.3. Métodos de segmentación de imágenes clásicos Existen técnicas adicionales de segmentación de imágenes que se utilizaban comúnmente en el pasado, pero son menos eficientes que sus contrapartes de aprendizaje profundo porque utilizan algoritmos rígidos y requieren intervención humana y experiencia, entre los cuales se encuentran:

Umbral: Divide una imagen en un primer plano y fondo. Un valor de umbral especificado separa los píxeles en uno de los dos niveles para aislar objetos. El umbral convierte imágenes en escala de grises en imágenes binarias o distingue los píxeles más claros y oscuros de una imagen en color.

¹⁵ Guozheng Wu Jian Zhang Zhiguang Qin Yi Ding Qiqi Yang. «Multiple Instance Segmentation in Brachial Plexus Ultrasound Image Using BPMSegNet». En: (2020). DOI: <https://doi.org/10.48550/arXiv.2012.12012>.

Clústeres de medios K: Un algoritmo identifica los grupos en los datos, con la variable K que representa el número de grupos. El algoritmo asigna cada punto de datos (o píxel) a uno de los grupos en función de la similitud de la entidad. En lugar de analizar grupos predefinidos, la agrupación en clústeres funciona de forma iterativa en grupos de forma orgánica.

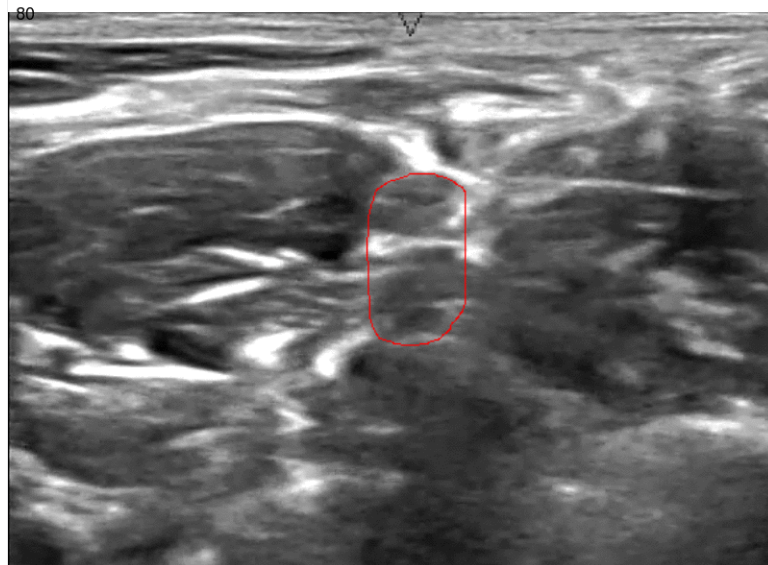
Segmentación de imagen basada en histograma: Utiliza un histograma para agrupar píxeles en función de los "niveles grises". Las imágenes simples consisten en un objeto y un fondo. El fondo suele ser un nivel gris y es la entidad más grande. Por lo tanto, un pico grande representa el nivel de gris de fondo en el histograma. Un pico más pequeño representa el objeto, que es otro nivel gris.

Detección de bordes: Identifica cambios bruscos o discontinuidades de brillo. La detección de aristas suele implicar la organización de puntos de discontinuidad en segmentos de línea curva o aristas. Por ejemplo, el borde entre un bloque de rojo y un bloque de azul.

3.2.4. Kaggle Kaggle es una comunidad en línea de científicos de datos y profesionales del aprendizaje automático que permite a los usuarios encontrar y publicar conjuntos de datos, explorar y crear modelos en un entorno de ciencia de datos basado en la web, trabajar con otros científicos de datos e ingenieros de aprendizaje automático, y participar en concursos para resolver desafíos de ciencia de datos. Uno de estos concursos denominado, segmentación de nervios por ultrasonido, presenta un conjunto de datos debidamente clasificado y etiquetado con ecografías de plexo braquial sobre el cual los participantes deben trabajar para encontrar la herramienta de clasificación más eficiente, sin embargo, debido a las condiciones del concurso, ninguna de las herramientas presentadas por los participantes implementa el aprendizaje profundo, estos se limitan a utilizar métodos de clásicos de tratamiento de imágenes como lo son la segmentación por intensidad de color. Todos los trabajos propuestos en el mencionado concurso toman como base una máscara inicial o semilla por lo tanto el proceso de segmentación

es semiautomático.

Figura 12. Segmentación resultante de uno de los proyectos propuestos para el reto de Kaggle



Fuente: Hefe, C. (2016). Animated Images with Outlined Nerve Area. Disponible en: <https://www.kaggle.com/chefe/animated-images-with-outlined-nerve-area>

Como se puede observar en la Figura 12, el resultado de la segmentación semiautomática lograda en los proyectos propuestos para el reto se asemeja a el objetivo de este proyecto, sin embargo, se tomarán como referencia otros artículos en donde se parta de la lógica del aprendizaje profundo.

4. METODOLOGÍA

El desarrollo de este proyecto se basó en una metodología utilizada en proyectos relacionados con el procesamiento de datos llamada CRISP-DM *Cross Industry Standard Process for Data Mining* la cual describe 6 fases relacionadas con el proceso del ciclo de vida de los proyectos de ciencia de datos. Estas fases son:

1. Comprensión de los datos
2. Preparación de los datos
3. Modelado
4. Evaluación
5. Despliegue

La propuesta de este proyecto se basa en implementar una arquitectura basada en aprendizaje profundo que permita realizar segmentación automática del plexo braquial utilizando redes neuronales convolucionales. Teniendo en cuenta las fases mencionadas anteriormente, la metodología del presente proyecto se apoya en los primeros 5 pasos.

4.1. Comprensión de los datos

Para comprender los datos es necesario hacer una descripción y exploración de los mismos.

4.1.1. Descripción del conjunto de datos: El dataset utilizado para entrenar y evaluar el modelo fue obtenido de la competición de *Kaggle* denominada *Ultrasound Nerve Segmentation*, el cual está compuesto de un total de 16788 imágenes obtenidas de 47 personas diferentes, de las cuales 5635 imágenes corresponden a ecografías del plexo braquial y otras 5635 corresponden a

sus respectivas máscaras debidamente anotadas, para un total de 11270. Las 5508 imágenes restantes no contienen máscaras, por lo cual no serán utilizadas para el entrenamiento y evaluación del modelo.

La figura 13 muestra un ejemplo aleatorio del conjunto de entrenamiento. El panel a) muestra la imagen de la ecografía. El panel b) corresponde a su respectiva máscara.

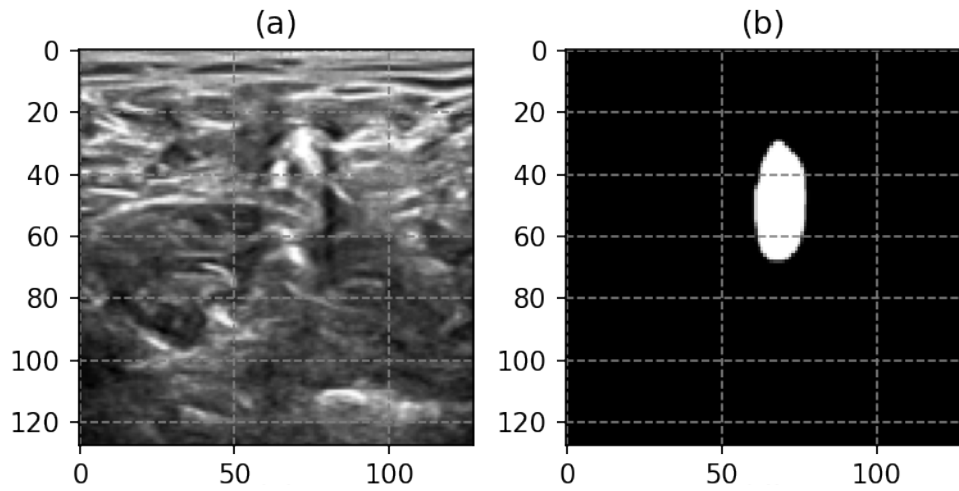


Figura 13. Ejemplo del conjunto de datos para entrenamiento. Panel a) muestra una imagen de ultrasonido aleatoria. Panel b) Corresponde a su respectiva máscara.

4.2. Preparación de los datos

En primer lugar, las imágenes se redimensionan a una escala de 128x128. Después las 11270 imágenes utilizadas en este trabajo se dividen en conjuntos de entrenamiento, validación y prueba, para ello fue utilizada la función *sklearn.model.selection.train_test_split* la cual particiona el dataset a las proporciones deseadas, a demás de realizar un barajado de los datos por medio del parámetro *random.state*. El dataset fue particionado utilizando una proporción de 70 %, 20 % y 10 %, respectivamente, es común utilizar distribuciones de 70/30 o 80/20 para entrenamiento

y validación, como lo muestran ¹⁶, por lo cual se determinó utilizar un híbrido entre las dos distribuciones designando por lo menos un 70 % para datos de entrenamiento y al menos un 20 % para validación, teniendo así un 10 % restante para la realización de pruebas como se muestra a continuación:

- *Conjunto de entrenamiento*: 4057 imágenes de ecografías del plexo braquial y sus respectivas máscaras anotadas. Para un total de 8114 imágenes.
- *Conjunto de validación*: 1127 ecografías del bucle braquial y sus respectivas máscaras anotadas. Para un total de 2254 imágenes.
- *Conjunto de pruebas*: 451 ecografías del plexo braquial y sus respectivas máscaras anotadas. Para un total de 902 imágenes.

4.3. Modelado

Para el desarrollo del proyecto fue implementada la arquitectura U-net. La cual está compuesta por una ruta de contracción (codificador) y una ruta de expansión (decodificador). La parte del codificador del modelo consta de cuatro capas convolucionales con función de activación ELU. Posteriormente, cada convolución va seguida de una capa de normalización de grupos con un tamaño de grupo de 8. En la primera capa, el número de filtros es de 32. Este número se multiplica por un factor dos en cada capa inferior.

La figura 14 presenta una visualización tridimensional de la arquitectura U-net utilizada en el proyecto, donde se visualiza como los caminos de codificación y decodificación aumentan y disminuyen su dimensión de manera simétrica.

¹⁶ Juul Van Boxtel y col. «Hybrid Deep Neural Network for Brachial Plexus Nerve Segmentation in Ultrasound Images». En: *2021 29th European Signal Processing Conference (EUSIPCO)* (2021). DOI: 10.23919/eusipco54536.2021.9616329.

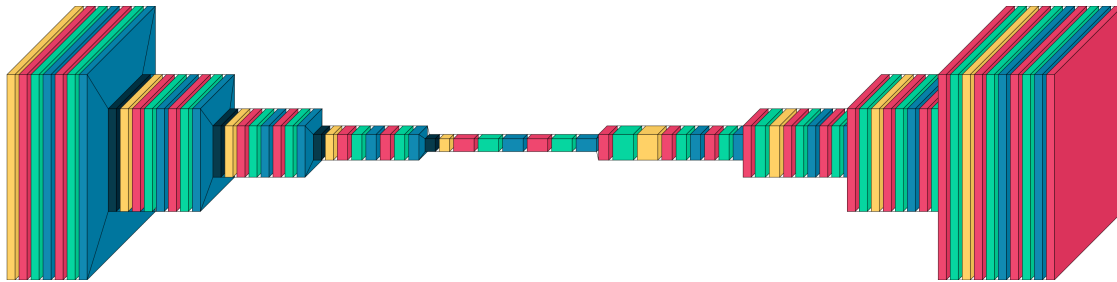


Figura 14. Visualización 3D de la arquitectura de la U-Net

A continuación se presenta el código de la arquitectura U-Net implementada en este proyecto, presentado cada una de sus capas.

```
#capa numero 1, inicio del camino de codificacion
def U_Net(img_tensor , n_filters = 32):
    conv1 = Conv2D_Block(img_tensor , n_filters * 1)
    pool1 = tf.keras.layers.MaxPooling2D((2 , 2))(conv1)
    pool1 = tf.keras.layers.Dropout(0.05)(pool1)

    #capa numero 2
    conv2 = Conv2D_Block(pool1 , n_filters * 2) #los filtros van aumentando en un
        factor exponencial de 2 para ir reduciendo la dimensin de la imagen por medio
        del Max Pooling
    pool2 = tf.keras.layers.MaxPooling2D((2 , 2))(conv2)
    pool2 = tf.keras.layers.Dropout(0.05)(pool2)

    #capa numero 3
    conv3 = Conv2D_Block(pool2 , n_filters * 4)
    pool3 = tf.keras.layers.MaxPooling2D((2 , 2))(conv3)
    pool3 = tf.keras.layers.Dropout(0.05)(pool3)

    #capa numero 4
```

```

conv4 = Conv2D_Block(pool3 , n_filters * 8)
pool4 = tf.keras.layers.MaxPooling2D((2 , 2))(conv4)
pool4 = tf.keras.layers.Dropout(0.05)(pool4)

#capa numero 5, fin del camino de codificacion e inicio del camino de decodificacion
conv5 = Conv2D_Block(pool4 , n_filters * 16)
#en este punto la imagen alcanza su minima dimension

#capa numero 6
pool6 = tf.keras.layers.Conv2DTranspose(n_filters * 8 , (2 , 2) , (2, 2) , padding
    = 'same')(conv5)
pool6 = tf.keras.layers.concatenate([pool6 , conv4])
pool6 = tf.keras.layers.Dropout(0.05)(pool6)
conv6 = Conv2D_Block(pool6 , n_filters * 8)

#capa numero 7
pool7 = tf.keras.layers.Conv2DTranspose(n_filters * 4 , (2 , 2) , (2 , 2) , padding
    = 'same')(conv6)
pool7 = tf.keras.layers.concatenate([pool7 , conv3])
pool7 = tf.keras.layers.Dropout(0.05)(pool7)
conv7 = Conv2D_Block(pool7 , n_filters * 4)

#capa numero 8
pool8 = tf.keras.layers.Conv2DTranspose(n_filters * 2 , (2 , 2) , (2 , 2) , padding
    = 'same')(conv7)
pool8 = tf.keras.layers.concatenate([pool8 , conv2])
pool8 = tf.keras.layers.Dropout(0.05)(pool8)
conv8 = Conv2D_Block(pool8 , n_filters * 2)

```

```

#capa numero 9
pool9 = tf.keras.layers.Conv2DTranspose(n_filters * 1 , (2 , 2) , (2 , 2) , padding
    = 'same')(conv8)
pool9 = tf.keras.layers.concatenate([pool9 , conv1])
pool9 = tf.keras.layers.Dropout(0.05)(pool9)
conv9 = Conv2D_Block(pool9 , n_filters * 1)

output = tf.keras.layers.Conv2D(1 , (1 , 1) , activation = 'sigmoid')(conv9)

u_net = tf.keras.Model(inputs = [img_tensor] , outputs = [output])

return u_net

```

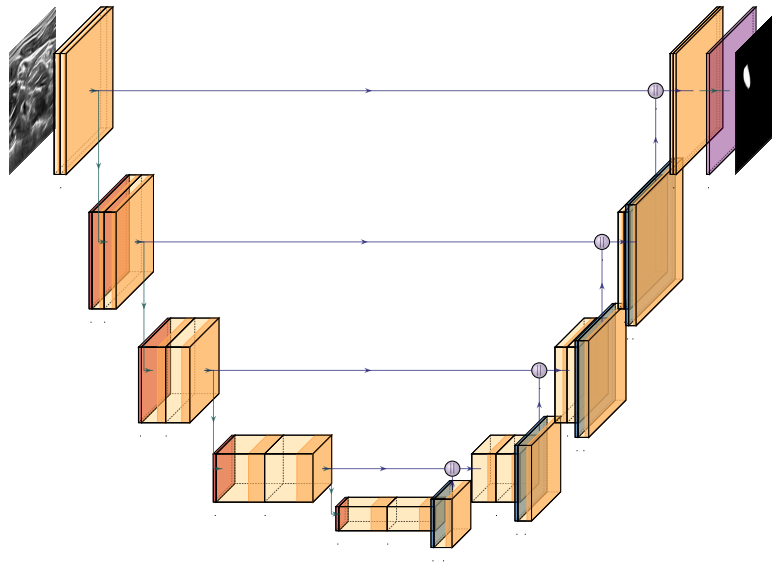


Figura 15. Arquitectura de la U-Net (Visualización obtenida usando ¹⁷).

Como se muestra en la Fig. 15 de izquierda a derecha, se pueden contar 9 bloques correspondientes a las capas de la red neuronal. Desde la primera a la cuarta capa hacen parte del camino de codificación (encoder) y de la sexta a la novena el camino de decodificación (decoder). La

dimensión de las imágenes se va reduciendo a medida que pasa por cada capa de la red hasta llegar a la quinta capa, en donde la imagen alcanza sus mínimas dimensiones y posteriormente se realiza el proceso inverso para recuperar la dimensión de la imagen de entrada. Las líneas azules horizontales que conectan cada bloque del camino de codificación con su equivalente en el camino de decodificación representan la función *tf.keras.layers.concatenate* la cual debe recibir en la capa de decodificación las mismas dimensiones que la imagen tenía al momento de pasar por las capas de codificación y así asegurar la integridad del proceso y la correcta utilización de los filtros.

La función de activación de la novena capa corresponde a la función sigmoidea, la cual satura y minimiza el gradiente, toma la entrada y da el valor de probabilidad entre 0 y 1 lo cual es ideal para la última capa.

Dentro del bloque convolucional se utiliza una capa Dropout del 50 %, seguida de una capa de Max Pooling. Entre los bloques, el muestreo ascendente se realiza mediante convoluciones transpuestas. La salida final se obtiene mediante una convolución 1x1 con activación sigmoidea.

4.3.1. Optimización de parámetros Para realizar la compilación y entrenamiento de la red U-NET se realizaron múltiples experimentos para verificar el comportamiento de la red, como se puede apreciar en las figuras 16 y 17, al utilizar diferentes optimizadores y rangos de aprendizaje se pueden lograr valores altos de la métrica objetivo, sin embargo la curva de aprendizaje tiene comportamientos anormales y tiende a disminuir o no mantenerse, lo que en ciertos casos podía reflejar sobre-entrenamiento, lo que provocaba que el modelo arrojara resultados erróneos como falsos positivos en donde se marcaba una ubicación de la región de interés sin que necesariamente el plexo estuviese presente en esa imagen en particular.

Optimizador y rango de aprendizaje: Al realizar experimentos con los optimizadores más utilizados en el campo de la segmentación en imágenes médicas como SGD, RMSprop, Adam y Adagrad; se encontró que existía un mejor comportamiento con el optimizador Adam y al variar el rango de aprendizaje entre 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , el modelo parecía tener mejores

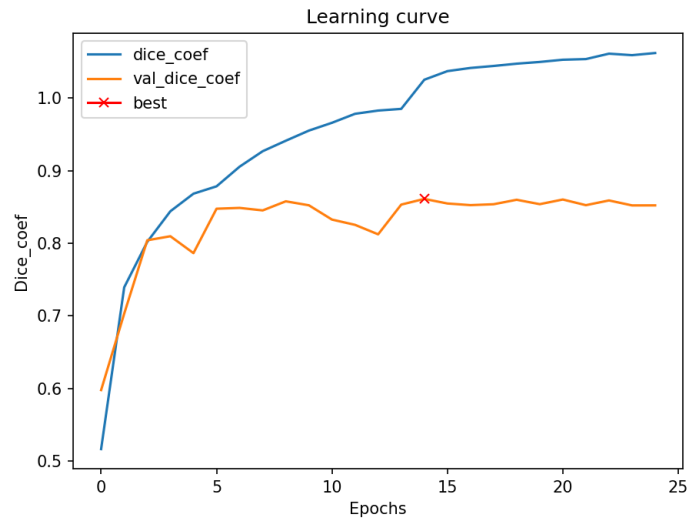


Figura 16. Curva de aprendizaje del coeficiente de datos DSC experimentos con diferentes parámetros.

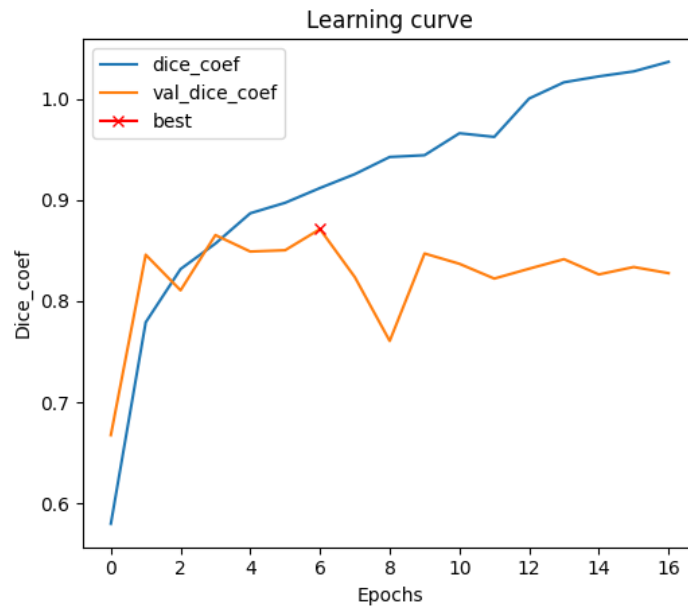


Figura 17. Curva de aprendizaje del coeficiente de datos DSC experimentos con diferentes parámetros.

resultados a medida que el rango de aprendizaje se reducía, sin embargo, valores en el rango de aprendizaje demasiado bajos causaban sobre-ajuste, por lo cual, se identificó que un rango

de aprendizaje de 10^{-3} permitiría aumentar la precisión del modelo sin causar un sobre-ajuste considerable.

Número de épocas El número de épocas para el modelo de cada arquitectura se definió con un valor de 50 debido a que para ese número de iteraciones la métrica estabiliza su valor. El uso de la función *earlyStopping* o "parada temprana" permite detener la ejecución del modelo al no detectar una mejora del mismo según una tolerancia determinada a la hora de definir dicha función.

Tamaño de imagen Las imágenes se redimensionaron a 128x128, esto se decidió con el fin de minimizar los tiempos de ejecución y recursos computacionales empleados a la hora de entrenar el modelo, estas dimensiones permiten reducir la dimensión de las imágenes originales manteniendo las características necesarias para segmentar correctamente el plexo braquial, experimentos realizados con menores dimensiones demostraron ser menos precisos a la hora de realizar segmentación.

La tabla 2 presenta los parámetros que arrojaron mejores resultados a raíz de los múltiples experimentos realizados.

Parámetro	Valor
Optimizador	Adam
Dropout	0,5
Épocas	50
Tamaño del Batch	32
Rango de aprendizaje	10^{-3}
Num. Filtros iniciales	32
Función de activación	ELU

Tabla 2. Parámetros de la U-Net

5. EVALUACIÓN Y RESULTADOS

5.1. Evaluación

El coeficiente de dados (DSC) se utiliza como métrica en el modelo (Eq.4). El DSC mide el solapamiento espacial entre el conjunto de los píxeles de la máscara etiquetada y el conjunto de los píxeles de los límites predichos. En la Ec. 4, X y Y son las regiones objetivo y la predicción, respectivamente. El valor de DSC oscila entre 0 y 1, cuanto mayor sea la puntuación, mejor. Su denominador tiene en cuenta el número total de píxeles de frontera a escala global, mientras que el numerador considera el solapamiento entre las dos áreas a escala local.

La pérdida se define en la Ec. 5 Aprovechando $Loss(DSC)$, los dos conjuntos se entrenan para maximizar gradualmente su solapamiento como se explica en la Ec. 4. Por lo tanto, $Loss(DSC)$ tiene en cuenta tanto la información de pérdida local como la global, lo que es fundamental para obtener una alta precisión.

$$DSC = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad (4)$$

$$Loss(DSC) = 1 - DSC \quad (5)$$

En el proceso de ajuste, el modelo toma como entrada los conjuntos de datos de entrenamiento y validación. Además, se ha implementado una parada temprana con una tolerancia de 10 épocas. Por último, una vez calculado el modelo, se evalúa en el conjunto de datos de prueba.

5.2. Resultados

En esta sección presentamos los resultados siguiendo la metodología propuesta en la sección 4. En la sección 6.3 se discuten los resultados.

5.3. Curva del coeficiente de datos (DSC)

La figura 18 muestra la evolución de la curva de aprendizaje de la métrica DSC . En las primeras 7 épocas, la curva de aprendizaje de validación (naranja) tiene un incremento pronunciado, como era de esperar. Después, la curva de aprendizaje sigue aumentando pero a un ritmo más lento hasta que alcanza la estabilidad. El modelo se detiene en torno a la época 43 debido a el uso de la función de parada temprana. El mejor modelo se obtiene en la época 32.

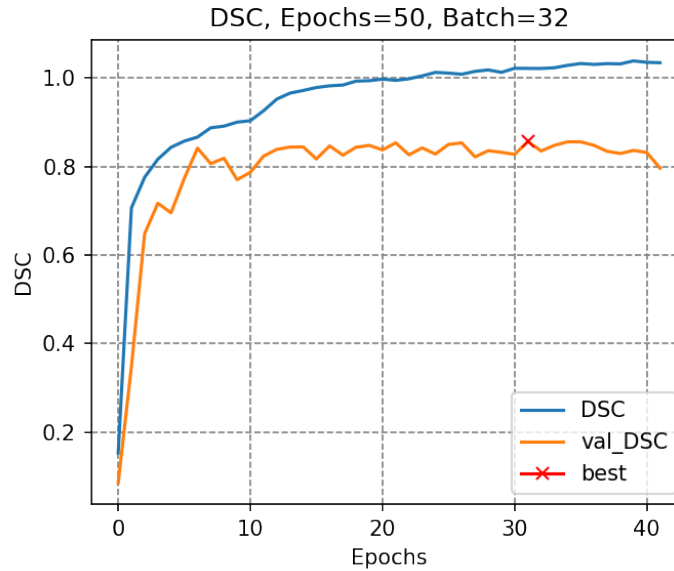


Figura 18. Curva de aprendizaje del coeficiente de datos DSC durante el entrenamiento.

5.4. Predicción

Una vez obtenido el mejor modelo, se proceder a visualizarlo. La figura 19 muestra el proceso paso a paso para un caso aleatorio del conjunto de pruebas. El panel a) muestra la ecografía. El panel b) corresponde a su respectiva máscara anotada proporcionada por ⁹, igual que la figura 13 pero para el conjunto de prueba. El panel c) es la combinación de los paneles a) y b) de la misma figura. El panel d) muestra la ecografía (es decir, el panel a) superpuesta a su máscara predicha obtenida por el mejor modelo.

La tabla 3 resume los valores de las ecuaciones 4 y 5. En el mejor modelo se obtiene un coeficiente de datos de $DSC = 0,86$ y una pérdida del coeficiente de datos de $Loss(DSC) = 0,14$.

Parámetro	Valor
DSC	0,86
$Loss(DSC)$	0,14

Tabla 3. Resultados del coeficiente de datos DSC y su respectiva pérdida

Método	Valor (DSC)
Abraham et al.[14] M-net	0,82
Boxtel et al. [16] U-net hibrida	0,83
Propio	0,86

Tabla 4. Comparación de resultados con otros trabajos

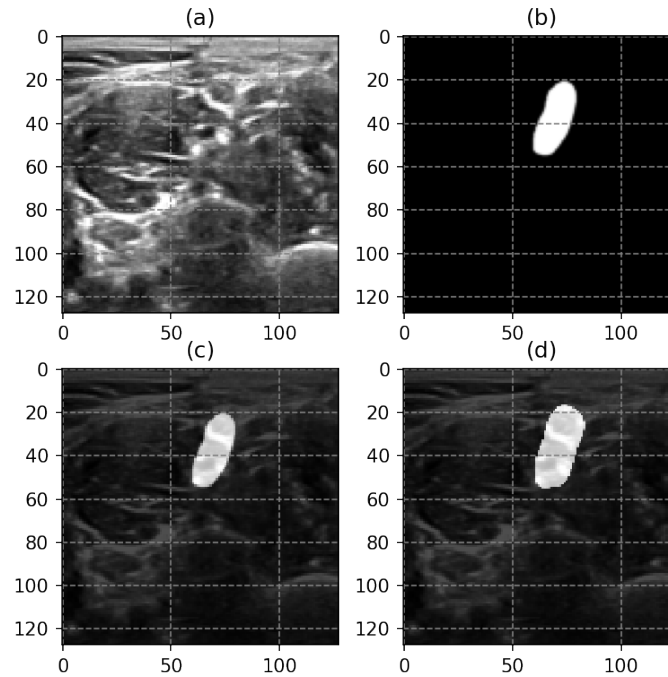


Figura 19. Comparación del resultado del modelo con su respectiva máscara

6. CONCLUSIONES Y PERSPECTIVAS

6.1. Conclusión

Utilizando el conjunto de datos ⁹, se implementó una arquitectura U-net para localizar y segmentar automáticamente el plexo braquial en imágenes de ultrasonido. El modelo resultante alcanza un buen rendimiento con una puntuación de coeficiente de datos de $DSC = 0,86$. El trabajo realizado alcanza resultados similares a otros trabajos recientes en la métrica (DSC) como ¹⁶ y ¹⁴, la cual es la mas utilizada en este tipo de problemas, esto se logró implementando un único modelo de la U-net. Los demás trabajos usados como referencia para este proyecto utilizan modificaciones de la U-net o arquitecturas híbridas que requieren una capacidad computacional mayor a la implementada en el presente proyecto, esto es beneficioso para la motivación inicial, la cual es que en un futuro esta herramienta pueda ser implementada en tiempo real asistiendo el proceso de bloqueo ecodirigido de plexo braquial en el Hospital Universitario de Santander u otras unidades hospitalarias del sector y para ello se requiere que la herramienta implementada utilice la menor cantidad de recursos posible.

6.2. Discusión

Según trabajos realizados anteriormente como ¹⁶, quienes utilizan el mismo dataset, si la puntuación de DSC es superior a 0,8; es posible considerar que el modelo ofrece una buena predicción de la ubicación del plexo braquial, lo cual sugiere que este tipo de aplicación podría apoyar a los profesionales capacitados en el procedimiento de bloqueo ecodirigido del plexo braquial desde el enfoque académico y práctico, ya sea usándose como método de aprendizaje para especialistas en formación o como herramienta práctica que asista en tiempo real el procedimiento.

Los trabajos de investigación anteriores, como ⁴, también han abordado este problema utilizando redes neuronales (véase la sección 3). En ¹⁶, los autores obtienen una puntuación DSC de 0,83 en su implementación U-net. Sin embargo, respecto a ¹⁶ el modelo implementado en el presente

proyecto no es tan robusto debido a que no se implementan la misma cantidad de experimentos con diversas arquitecturas, se podría realizar una mejor validación cruzada y no se posee la misma capacidad computacional, el proyecto presentado en este documento ofrece resultados comparables.

6.3. Trabajo futuro

Se busca que a futuro otro estudiante o grupo de estudiantes continúen con el desarrollo de una interfaz que permita a profesionales o aprendices del área de anestesiología utilizar esta herramienta.

Se tiene proyectado implementar métodos de validación cruzada mas robustos para aumentar la precisión del modelo a la hora de realizar el proceso de segmentación.

Se espera poder trabajar en conjunto con profesionales del Hospital Universitario para poder crear datasets del plexo braquial a partir de las imágenes recogidas por ecógrafos del hospital.

BIBLIOGRAFÍA

- Abraham Kandasamy, Khan Androustos. «Deep Learning for Semantic Segmentation of Brachial Plexus Nerves in Ultrasound Images Using U-Net and M-Net». En: (). DOI: 10.1109/ICISPC.2019.8935668 (vid. págs. 30, 47).
- Acedo Gutiérrez, M. S. *Tratamiento del dolor: técnicas locorregionales. Bloqueos. Oncología*. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0378-48352005000300007. 2005 (vid. pág. 27).
- Barrett David Gt; Morcos, Ari S.; Macke Jakob H. «Analyzing biological and artificial neural networks: challenges with opportunities for synergy.» En: *Current opinion in neurobiology* (2019). DOI: 10.1016/j.conb.2019.01.007 (vid. pág. 17).
- Horng MH Yang CW, Sun YN Yang TH. «DeepNerve: A New Convolutional Neural Network for the Localization and Segmentation of the Median Nerve in Ultrasound Image Sequences». En: *Ultrasound Med Biol* 46(9).2439-2452 (2020). DOI: 10.1016/j.ultrasmedbio.2020.03.017 (vid. págs. 12, 15, 30, 47).
- Isabel Díaz Mora, Javier Lenis Chacón y Nelson Iván Rodríguez Castro y Lucas Cordoví de Armas y Idoris Cordero Escobar y. *Bloqueo del plexo braquial por vía supraclavicular y axilar guiados por ultrasonido*. 2017. URL: <http://www.revanestesia.sld.cu/index.php/anestRean/article/view/180> (vid. págs. 11, 27).
- James C Crews 1, Theodore E Rothman. «Desarrollo de la Radiología. Centenario del descubrimiento de los Rayos X». En: *Anesth Analg* 26.4 (2003). DOI: 10.1213/01.ANE.0000053235.07186.39 (vid. pág. 11).

Macfarlane AJ Prasad GA, Chan VW Brull R. «Does regional anesthesia improve outcome after total knee arthroplasty?» En: *Clin Orthop Relat Res* 467 (2009). DOI: 10.1007/s11999-008-0666-9 (vid. pág. 14).

Mather, Laurence E., Susan E. Copeland y Leigh A. Ladd. «Acute Toxicity of Local Anesthetics: Underlying Pharmacokinetic and Pharmacodynamic Concepts». En: *Regional Anesthesia & Pain Medicine* 30.6 (2005). DOI: 10.1016/j.rapm.2005.07.186 (vid. págs. 11, 14).

Milletari. *Redes neuronales totalmente convolucionales para la segmentación de imágenes médicas volumétricas*. <https://arxiv.org/abs/1606.04797>. 2016 (vid. pág. 24).

Na8. *¿Cómo funcionan las Convolutional Neural Networks? Visión por Ordenador?* <https://www.aprendemachinelearning.com/como-funcionan-las-convolutional-neural-networks-vision-por-ordenador/>. 2012 (vid. pág. 20).

PlotNeuralNet. <https://github.com/HarisIqbal88/PlotNeuralNet>. Accessed: 2022-02-23 (vid. pág. 40).

Sandeep Kusre Andrew McEwen, Geena Matthew. *Ultrasound-Guided Supraclavicular Brachial Plexus Block*. 2018. URL: <https://resources.wfsahq.org/atotw/ultrasound-guided-supraclavicular-brachial-plexus-block/> (vid. págs. 13, 26, 29).

Ultrasound Nerve Segmentation dataset. <https://www.kaggle.com/c/ultrasound-nerve-segmentation>. 2016 (vid. págs. 19, 45, 47).

Valueva M.V.; Nagornov, N.N.; Lyakhov P.A.; Valuev G.V.; Chervyakov N.I. «Application of the residue number system to reduce hardware costs of the convolutional neural network implementation». En: *Mathematics and Computers in Simulation* 117 (2020). DOI: 10.1016/j.matcom.2020.04.031 (vid. pág. 19).

- Van Boxtel, Juul y col. «Hybrid Deep Neural Network for Brachial Plexus Nerve Segmentation in Ultrasound Images». En: *2021 29th European Signal Processing Conference (EUSIPCO)* (2021). DOI: 10.23919/eusipco54536.2021.9616329 (vid. págs. 37, 47).
- Wheeler, T. *Manejo del dolor y bloqueos nerviosos*. <https://www.webmd.com/pain-management/guide/nerve-blocks#2>. 2020 (vid. pág. 27).
- Yi Ding Qiqi Yang, Guozheng Wu Jian Zhang Zhiguang Qin. «Multiple Instance Segmentation in Brachial Plexus Ultrasound Image Using BPMSegNet». En: (2020). DOI: <https://doi.org/10.48550/arXiv.2012.12012> (vid. pág. 32).