

DETECCIÓN DE NEUMONÍA-COVID EN DATOS DE
TOMOGRAFÍA COMPUTACIONAL MEDIANTE TÉCNICAS DE
APRENDIZAJE PROFUNDO

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Especialidad en

Ciencias de la Computación y Matemáticas Industriales

Presenta

Henry David Ocampo Meraz

Director de Tesis:

Dr. Mariano José Juan Rivera Meraz

Autorización de la versión final

*A Francis Vasquez, la mejor profesora del Departamento de Matemática Pura en la
Universidad Nacional Autónoma de Honduras . . .*

"En vez de intentar producir un programa que simule la mente adulta, ¿por qué no tratar de producir uno que simule la mente del niño? Si ésta se sometiera entonces a un curso educativo adecuado, se obtendría el cerebro de adulto".

- Alan Turing

Resumen

La pandemia causada por el coronavirus 2019, más conocida como COVID-19, sigue teniendo un fuerte impacto en los pacientes y sistemas de atención médica a nivel mundial. En la lucha contra esta nueva enfermedad, existe la urgente necesidad de herramientas que faciliten a los expertos un diagnóstico eficiente para identificar pacientes infectados con COVID-19. Las tomografías computarizadas proporcionan imágenes detalladas de los pulmones y ayudan a los radiólogos a diagnosticar COVID-19. Sin embargo, el análisis de tomografías computarizadas (CTs, por sus siglas en inglés), rebanada a rebanada, demanda demasiado tiempo por los especialistas y puede provocar retrasos en los hospitales. Los métodos de inteligencia artificial podrían ayudar a los radiólogos a detectar de forma rápida y precisa la infección por COVID-19 a partir de estas exploraciones. En esta tesis se propone un diagnóstico automatizado de la infección por COVID-19 a partir de CTs de los pacientes mediante la técnica de aprendizaje profundo. El modelo propuesto, EffCOV-B0, utiliza tomografías computarizadas de tórax para clasificar volúmenes CT de pacientes entre quienes padecen COVID-19 y sanos. Para mejorar la precisión de la detección, las CTs se preprocesaron empleando segmentación e interpolación. El esquema propuesto se basa en la red neuronal convolucional (CNN) EfficientNet B0 que actualmente ha superado en precisión, eficiencia y reducción de parámetros a varios modelos CNN existentes. Mostramos que el modelo de aprendizaje profundo propuesto logra una exactitud del 94.05 %, valor F1 de 94.05 % y AUC del 97.5 %, para detectar COVID-19 en CTs.

Palabras clave: Tomografía computarizada, COVID-19, Redes neuronales convolucionales, Aprendizaje profundo

Agradecimientos

A ti, Dios de mis padres, te doy gracias y te alabo en primer lugar, porque me has dado fuerza y sabiduría, y ahora me has ayudado en todo lo que te he pedido.

Es imposible tratar de incluir en este pequeño espacio a todas aquellas personas que han contribuido a lo largo de mi vida en la realización de esta meta. Sepan que los aprecio mucho y en mi corazón tengo guardado cada hermoso recuerdo que nos une.

A mis padres, María Meraz y David Ocampo; y mi hermano, Omar. Gracias por todo su amor, paciencia, confianza y apoyo incondicional que he recibido, en todo momento, desde que tengo memoria.

A mis amigos de Honduras, que ofrecieron a mi disposición su ayuda y motivación en todo momento. En especial a mi novia Breydi Canales, gracias por estar siempre presente.

A todos los amigos que conocí en México. Gracias por toda la amabilidad y atención que recibí en estos dos años. Quiero mencionar especialmente a César Romero, un excelente amigo de quien aprendí muchas cosas y tuve la bendición de convivir como un hermano.

A mi asesor de tesis, el Dr. Mariano José Juan Rivera Meraz. Por los conocimientos que me brindó para llevar a cabo esta investigación, y sobretodo su gran paciencia para esperar a que este trabajo pudiera llegar a su fin.

Al Centro de Investigación en Matemáticas (CIMAT), por brindarme la oportunidad de continuar con mi preparación académica y así cumplir una meta más en mi vida profesional. A todos mis profesores, que contribuyeron a mi formación profesional. En especial al Dr. Oscar Dalmau por darme la oportunidad de ser su ayudante; y a mis profesores de inglés, Janet Izzo y Sigifredo Aguilar por toda la paciencia y dedicación que tuvieron en mi aprendizaje.

Finalmente, agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo económico brindado durante los últimos tres semestres de la maestría, para lograr este grado académico.

Índice general

Resumen	v
Agradecimientos	vii
Índice de figuras	xi
Índice de tablas	xvii
Introducción	1
1. La enfermedad de COVID-19	7
1.1. Descripción de la enfermedad	7
1.1.1. Origen	7
1.1.2. Síntomas	8
1.1.3. Diagnóstico	8
1.2. Hallazgos identificados en CT de tórax	8
1.2.1. Opacidad de vidrio esmerilado	9
1.2.2. Pavimentación loca	9
1.2.3. Signo de halo invertido	10
1.2.4. Consolidación	10
1.2.5. Broncograma aéreo	11
2. Arquitectura de la red	13
2.1. Componente base	15
2.1.1. Modelo EfficientNet B0	16
2.1.2. Bloque de clasificación	18
2.2. Complemento 3D	18
2.2.1. Bloque convolucional	19
2.2.2. Bloque de clasificación	19
2.3. Optimizador y función de pérdida	20
2.3.1. RMSProp	20
2.3.2. Entropía cruzada binaria con pesos balanceados	20
3. Datos	23
3.1. Conjunto de datos 2D	23
3.1.1. Preprocesamiento de datos	24

3.2. Conjunto de datos 3D	25
3.2.1. Preprocesamiento de datos	26
3.2.2. Segmentación pulmonar basada en CNN	27
4. Resultados	29
4.1. Modelo 2D	29
4.1.1. Clases: COVID-19, Neumonía y Normal	29
4.1.2. Clases: COVID-19 y Normal	41
4.2. Modelo 3D	50
4.2.1. Partición de los Conjuntos de Datos	50
4.2.2. Entrenamiento	51
4.2.3. Resultados con Segmentación del Cuerpo Humano	51
4.2.4. Resultados con Segmentación Pulmonar	59
4.2.5. Evaluaciones de los conjuntos de datos	67
4.3. Comparación con otros modelos	68
5. Conclusiones y trabajo futuro	71
Bibliografía	73
A. Anexo I	77
A.1. Hallazgos identificadas en CT de tórax	77
A.2. Descripción del hardware	82
A.3. Grad-CAM	82

Índice de figuras

1.1.	Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado en los lóbulos superiores. Imagen tomada de <i>COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay</i> [1]. La imagen en tamaño completo se encuentra en el apéndice A.1.	9
1.2.	Tomografía computarizada que muestra patrón de pavimentación loca en la periferia del lóbulo inferior derecho. Imagen tomada de <i>COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay</i> [1]. La imagen en tamaño completo se encuentra en el apéndice A.2.	9
1.3.	Tomografía computarizada que muestra signo de halo invertido en los lóbulos inferiores. Imagen tomada de <i>COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay</i> [1]. La imagen en tamaño completo se encuentra en el apéndice A.3.	10
1.4.	Tomografía computarizada que muestra consolidaciones periféricas en los lóbulos superiores. Imagen tomada de <i>COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay</i> [1]. La imagen en tamaño completo se encuentra en el apéndice A.4.	10
1.5.	Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado, consolidaciones posteriores con broncogramas aéreos. Imagen tomada de <i>COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay</i> [1]. La imagen en tamaño completo se encuentra en el apéndice A.5.	11
2.1.	Estructura general del modelo de red neuronal convolucional EffCOV-B0, para identificar casos de COVID-19 a través de imágenes de tomografías computarizadas.	13
2.2.	Componente base y Complemento 3D del modelo EffCOV-B0.	14
2.3.	Estructura de la Componente Base de EffCOV-B0.	16
2.4.	Arquitectura de la red EfficientNet B0. Imagen tomada de <i>EfficientNet: Improving Accuracy and Efficiency through AutoML and Model Scaling</i> [2].	16
2.5.	Bloque de convolución móvil invertida de cuello de botella (MBCConv). Imagen tomada de <i>COVID-19 Detection in CT Images with Deep Learning: A Voting-based Scheme and Cross-Datasets Analysis</i> [3].	17

2.6.	Gráfica comparativa del número de parámetros <i>vs</i> precisión (en el conjunto de datos ImageNet) de los modelos EfficientNet con respecto a otros modelos. Imagen tomada de <i>EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks</i> [4].	17
2.7.	Estructura general del Complemento 3D.	19
3.1.	Segmentación de imagen 2D, correspondiente a un corte de tomografía computarizada. Se retira todo indicador visual que se encuentra fuera del cuerpo del paciente.	24
3.2.	Método de segmentación de imagen 2D. a) Imagen original; b) Imagen suavizada con filtro Gaussiano; c) Imagen binaria construida a partir del umbral definido con la imagen suavizada; d) Imagen del mayor contorno de la imagen binaria (mascara del cuerpo del paciente); e) Imagen segmentada del cuerpo del paciente.	24
3.3.	Preprocesamiento de imagen 3D. a) Segmentación de los cortes, se retira todo indicador visual que se encuentra fuera del cuerpo del paciente; b) Selección de cortes basado en el elemento central y los cortes vecinos más cercanos; c) Cortes seleccionados; d) Dato procesado, los cortes seleccionados se evalúan en la Componente Base de EffCOV-B0.	26
3.4.	Segmentación pulmonar 3D, imagen correspondiente a un corte de tomografía computarizada. Se retira todo indicador visual que se encuentra fuera de la región pulmonar del cuerpo del paciente.	27
3.5.	Método de segmentación de imagen 3D, para un corte específico. a) Corte original del volumen; b) Mascara del corte original, el pulmón izquierdo y el pulmón derecho están separados; c) Corte segmentado.	28
4.1.	Gráfica de la pérdida (sección 2.3.2) con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	31
4.2.	Gráfica de la exactitud con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	31
4.3.	Gráfica de la función F1 con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	32
4.4.	Gráfica de la métrica AUC con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	32
4.5.	Gráfica de la matriz de confusión de la Componente Base en el subconjunto de prueba. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	33
4.6.	Gráfica de la matriz de confusión de la Componente Base en el subconjunto de entrenamiento. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	34

4.7. Gráfica de la matriz de confusión de la Componente Base en el subconjunto de validación. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.	34
4.8. Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9991, 0.0005 y 0.0004 para las clases COVID-19, Neumonía y Normal respectivamente.	36
4.9. Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9983, 0.0012 y 0.0006 para las clases COVID-19, Neumonía y Normal respectivamente.	37
4.10. Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9992, 0.0005 y 0.0003 para las clases COVID-19, Neumonía y Normal respectivamente.	37
4.11. Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9983 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.	38
4.12. Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9984 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.	38
4.13. Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9984 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.	39
4.14. Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0907, 0.0036 y 0.9057 para las clases COVID-19, Neumonía y Normal respectivamente.	39
4.15. Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0008, 0.0006 y 0.9987 para las clases COVID-19, Neumonía y Normal respectivamente.	40
4.16. Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0007, 0.0006 y 0.9988 para las clases COVID-19, Neumonía y Normal respectivamente.	40
4.17. Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 2D. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.	42
4.18. Gráfica de la exactitud con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.	42

4.19. Gráfica de la función F1 con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.	43
4.20. Gráfica de la métrica AUC-ROC con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.	43
4.21. Gráfica de la matriz de confusión de la Componente Base en el subconjunto de prueba. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.	44
4.22. Gráfica de la matriz de confusión de la Componente Base en el subconjunto de entrenamiento. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.	45
4.23. Gráfica de la matriz de confusión de la Componente Base en el subconjunto de validación. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.	45
4.24. Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.8 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.	47
4.25. Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.9 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.	47
4.26. Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.10 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.	48
4.27. Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.14 y el modelo retorna los resultados 0.0003 y 0.9997 para las clases COVID-19 y Normal respectivamente.	48
4.28. Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.15 y el modelo retorna los resultados 0.0002 y 0.9998 para las clases COVID-19 y Normal respectivamente.	49
4.29. Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.16 y el modelo retorna los resultados 0.0002 y 0.9998 para las clases COVID-19 y Normal respectivamente.	49
4.30. Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano. . .	53
4.31. Gráfica de la exactitud con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	54
4.32. Gráfica de la función F1 con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	54
4.33. Gráfica de la métrica AUC-ROC con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	55

4.34. Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano. . .	55
4.35. Gráfica de la exactitud con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	56
4.36. Gráfica de la función F1 con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	56
4.37. Gráfica de la métrica AUC con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	57
4.38. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de prueba. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	58
4.39. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	58
4.40. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de validación. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	59
4.41. Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	61
4.42. Gráfica de la exactitud con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	61
4.43. Gráfica de la función F1 con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	62
4.44. Gráfica de la métrica AUC con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	62
4.45. Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	63
4.46. Gráfica de la exactitud con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	63
4.47. Gráfica de la función F1 con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	64
4.48. Gráfica de la métrica AUC-ROC con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	64

4.49. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de prueba. El conjunto de datos tiene preprocesado de segmentación pulmonar.	65
4.50. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.	66
4.51. Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de validación. El conjunto de datos tiene preprocesado de segmentación pulmonar.	66
A.6. Mapeo de la activación de la clase: el puntaje de clase predicho se asigna a la capa convolucional anterior para generar los mapas de activación de clase (CAMs). Imagen tomada de <i>Learning Deep Features for Discriminative Localization</i> [5].	82

Índice de tablas

2.1. Arquitectura del modelo EffCOV-B0.	15
2.2. Arquitectura de la componente base.	18
2.3. Arquitectura del complemento 3D.	19
3.1. Distribución de datos del conjunto COVIDx CT-2.	23
3.2. Distribución de datos en el conjunto <i>MosMedData: Chest CT Scans with COVID-19 Related Findings</i>	27
4.1. Tiempo de entrenamiento-validación y prueba de los modelos bidimensionales.	29
4.2. Partición del conjunto de datos COVIDx CT-2.	30
4.3. Distribución de pacientes en la partición del conjunto de datos COVIDx CT-2.	30
4.4. Pérdida, exactitud, funciones F1 y AUC-ROC de la Componente Base en el conjunto de datos COVIDx CT-2.	35
4.5. <i>Recall</i> de la Componente Base en el conjunto de datos COVIDx CT-2.	35
4.6. <i>Precision</i> de la Componente Base en el conjunto de datos COVIDx CT-2.	35
4.7. Partición del conjunto de datos COVIDx CT-2. No se incluye la categoría Neumonía	41
4.8. Distribución de pacientes en la partición del conjunto de datos COVIDx CT-2. No se incluye la categoría Neumonía	41
4.9. Pérdida, exactitud, funciones F1 y AUC-ROC de la Componente Base. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.	44
4.10. <i>Recall</i> de la Componente Base en el conjunto de datos COVIDx CT-2. Únicamente se consideran las categorías COVID-19 y Normal.	46
4.11. <i>Precision</i> de la Componente Base en el conjunto de datos COVIDx CT-2. Únicamente se consideran las categorías COVID-19 y Normal.	46
4.12. Partición de los conjuntos de datos <i>COVID-19 CT-INER México</i> y <i>MosMedData</i> . Las categorías <i>CT-2</i> , <i>CT-3</i> , <i>CT-4</i> e <i>INER</i> corresponden a volúmenes clasificados en la categoría COVID-19, mientras que los datos en la categoría <i>CT-0</i> corresponden a volúmenes Normales.	50
4.13. Tiempo de entrenamiento-validación y prueba de los modelos tridimensionales.	51

4.14. Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la primera fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación del cuerpo del humano	52
4.15. Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la segunda fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación del cuerpo del humano	52
4.16. <i>Recall</i> del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	52
4.17. <i>Precision</i> del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.	52
4.18. Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la primera fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación pulmonar. . . .	59
4.19. Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la segunda fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación pulmonar. . . .	60
4.20. <i>Recall</i> del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación pulmonar.	60
4.21. <i>Precision</i> del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación pulmonar.	60
4.22. Resumen de resultados del modelo EffCOV-B0 (los datos tienen preprocesado de segmentación del cuerpo humano). Las categorías CT-0, CT-1, CT-2, CT-3 y CT-4 corresponden al conjunto de datos MosMedData.	67
4.23. Resultados de modelos bidimensionales.	68
4.24. Resultados de modelos tridimensionales.	69
4.25. Comparación de resultados de la métrica Recall.	70
4.26. Comparación de resultados de la métrica Precision.	70
A.1. Descripción del hardware utilizado para desarrollar modelos de redes neuronales.	82

Introducción

La enfermedad de Coronavirus de 2019, conocida como COVID-19, ha desafiado al mundo en todos sus niveles; desde la crisis en los sistemas de salud hasta la interrupción social y el colapso económico. Debido a la globalización, la velocidad de propagación del virus del Síndrome Respiratorio Agudo Severo (SARS-COV-2), causante de la pandemia del COVID-19, se ha superado continuamente las estimaciones en todo el mundo. Desde marzo del 2020, cuando el virus fue declarado una pandemia por la Organización Mundial de la Salud debido al número de muertos y casos positivos, los impactos económicos de esta enfermedad han sido enormes. El cierre de empresas, la reducción de la fuerza laboral, la interrupción del comercio y una fuerte disminución en la industria del turismo son algunos impactos de las consecuencias económicas de esta pandemia [6].

El creciente número de casos positivos de COVID-19 y el registro de pacientes que experimentan insuficiencia respiratoria grave, complicaciones cardíacas y otros síntomas son razones sustanciales para estar extremadamente preocupados por los resultados de esta infección viral [7]. El COVID-19 se propaga principalmente a través del aire y el contacto transmitido en forma de pequeñas gotas producidas por una persona infectada al estornudar y toser. Tocarse la nariz, la boca o los ojos después de entrar en contacto con la superficie expuesta al virus puede provocar una infección [7]. Las personas infectadas con COVID-19 muestran síntomas similares a los de la influenza, así como otras dolencias respiratorias. Los síntomas informados varían de síntomas leves a graves. Fiebre, tos, dificultad para respirar, dolor de cabeza y diarrea son algunos de los que ocurren en personas que tienen el virus COVID-19. Estos síntomas pueden tardar entre 2 y 14 días en aparecer después de la exposición al virus [8].

En la lucha contra esta pandemia, existe una necesidad apremiante de herramientas de detección rápidas y efectivas para identificar a los pacientes infectados con COVID-19 para garantizar un aislamiento y tratamiento oportuno. Actualmente, la prueba de reacción en cadena de la polimerasa con transcriptasa inversa en tiempo real (RT-PCR, por sus siglas en inglés), es el medio principal de detección de COVID-19 [1]. Si bien la prueba RT-PCR para COVID-19 es altamente específica, su sensibilidad es variable dependiendo del método de muestreo y del tiempo desde el inicio de los síntomas, y algunos estudios han reportado una sensibilidad de COVID-19 relativamente baja [1]. Además, las pruebas RT-PCR son

un proceso que requiere de equipo especializado, mucho tiempo y tienen una gran demanda. Esto conlleva a posibles retrasos en la obtención de resultados de las pruebas. Junto con estas pruebas de laboratorio, se ha comprobado que las exploraciones del área del tórax mediante una Radiografía o Tomografía Computarizada (CT por sus siglas en inglés *Computed Tomography*) ayudan en el diagnóstico de COVID-19 [1]. Las imágenes de tórax en los pacientes tienen afectaciones multilobares y opacidades redondeadas. Las CTs de tórax son más sensibles a las opacidades en vidrio esmerilado, que podrían pasarse por alto en las radiografías simples de tórax [9].

La CT de tórax se ha propuesto como una herramienta alternativa para detección de neumonía por covid-19 debido a su alta sensibilidad, y puede ser particularmente eficaz cuando se utiliza como complemento de las pruebas de RT-PCR [10]. Las imágenes de CT se utilizaron ampliamente durante las primeras etapas de la pandemia de COVID-19, particularmente en Asia. Aunque la prueba mediante análisis de CT no reemplaza la RT-PCR, es una excelente herramienta de diagnóstico en situaciones en las que se sospecha que los pacientes son positivos para COVID-19 pero sus pruebas de RT-PCR son negativas. Por ejemplo, los pacientes que han tenido contacto cercano con casos confirmados de COVID-19 y presentan síntomas de la enfermedad son altamente sospechosos, pero pueden tener resultados de RT-PCR negativos. En estos casos, se pueden utilizar imágenes de CT para confirmar sospechas de infección por COVID-19 en espera de resultados positivos de RT-PCR [11].

El creciente número de personas infectadas plantea un problema para los procedimientos de detección actuales, ya que están abrumados por la cantidad de personas examinadas. La evaluación del escaneo de una CT requiere mucho tiempo y supone una carga adicional para los médicos involucrados en la cuarentena y el tratamiento de los pacientes infectados. Para superar este problema, el proceso de diagnóstico a través de tomografías computarizadas se puede acelerar a través de técnicas de visión por computadora y aprendizaje profundo. Estos algoritmos se pueden utilizar para detectar COVID-19, ya que pueden identificar las lesiones y las variadas opacidades a un ritmo más rápido que los métodos existentes [12].

En los primeros estudios, se encontró que ciertas anomalías en las imágenes de CT de tórax son indicativas de infección por COVID-19, como opacidades en vidrio esmerilado, sombras irregulares, patrón de pavimentación loca y consolidaciones. Son algunas de las anomalías más comúnmente reportadas, típicamente con anomalías bilaterales [1]. Además, algunos estudios han encontrado que pueden estar presentes anomalías en la CT de tórax de un paciente debido a la infección por COVID-19 a pesar de una prueba de RT-PCR negativa [1]. Sin embargo, estas anomalías en las imágenes pueden no ser específicas de la infección por COVID-19, y las diferencias visuales entre las anomalías relacionadas con COVID-19 y otras anomalías pueden ser bastante sutiles. Como resultado, el desempeño de los radiólogos para distinguir las anomalías relacionadas con COVID-19 de las

anomalías de otra etiología puede variar considerablemente [13]. Para los radiólogos, el análisis visual de las CTs también es una tarea manual que requiere mucho tiempo, particularmente cuando el volumen del paciente contiene muchos cortes o el número de estudios es grande.

Recientemente, los algoritmos basados en el aprendizaje profundo han proporcionado un enorme progreso en la interpretación y automatización de datos médicos debido a su capacidad para obtener características valiosas de los conjuntos de datos médicos disponibles [12]. El presente trabajo considera una Red Neuronal Convolutiva (CNN por sus siglas en inglés *Convolutional Neural Network*) cuya arquitectura base es la red bidimensional *EfficientNet B0* [4]. Inicialmente, la red base es entrenada para detectar COVID-19 en imágenes CT-2D con el conjunto de datos *COVIDx CT-2* [11]. Luego, implementada la técnica de transferencia de aprendizaje, la red base es adaptada en una arquitectura tridimensional para entrenar con volúmenes de tomografías computarizadas CT-3D tomadas de los conjuntos de datos *MosMedData* [14] y *COVID-19 CT-INER México*.

Contribución

Los aportes del presente trabajo son los siguientes:

1. La creación de un modelo de red neuronal convolutiva de vanguardia para detectar la enfermedad de COVID-19 en los cortes de volúmenes de tomografía computarizada de personas, impulsado por técnicas como el aprendizaje por transferencia, ajuste fino, segmentación de imagen y una sólida línea de preprocesamiento que reduce el sesgo general que puede ocurrir cuando se trabaja con conjuntos de datos que son relativamente pequeños.
2. El desarrollo de un modelo base bidimensional que emplea la arquitectura de red *EfficientNet B0*, entrenado para estimar COVID-19 en imágenes de tomografías computarizadas.

Trabajos relacionados

Aunque la falta de datos ha sido el principal problema en el desarrollo de modelos de redes neuronales, se han desarrollado distintos trabajos que implementan diversas arquitecturas y técnicas que mejoran los resultados de los modelos bases que son presentados cuando se publica un determinado conjunto de datos. Dado que los conjuntos de datos más extensos corresponden a imágenes 2D aisladas de CT, la mayoría de los trabajos realizados consisten en modelos bidimensionales. Mientras se espera la publicación de un conjunto de datos de volúmenes de CT lo suficientemente grande, algunos trabajos proponen modelos tridimensionales entrenados en pequeños conjuntos de datos, implementando distintas técnicas de reducción de sesgo.

Modelos bidimensionales

Uno de los primeros trabajos realizados fue el modelo *Self-Trans* de He et al. [15], el cual es entrenado en el conjunto de datos *COVID-CT* que ellos prepararon y que contiene únicamente 746 imágenes con 349 cortes de pacientes diagnosticados con COVID-19 y 397 cortes con diagnósticos normales y otras enfermedades. Para abordar el problema de sobreajuste debido al pequeño tamaño del conjunto de datos, ellos proponen un enfoque de aprendizaje por transferencia auto-supervisado contrastivo. En el proceso de aprendizaje auto-supervisado, las etiquetas supervisadas de las imágenes de CT son únicamente de las propias imágenes, sin utilizar anotaciones humanas. Con este método, el modelo logra la mejor exactitud de 86 % ($F1 = 85\%$ y $AUC = 94\%$) en el conjunto de prueba.

Poco tiempo después, Soares et al. [16] reunieron los datos *SARS-CoV-2* para entrenar el modelo *xDNN*. El conjunto de datos está compuesto de 2482 imágenes de CT con 1252 cortes positivos de COVID-19 y 1230 cortes que no contienen dicha enfermedad. Ellos proponen un método de aprendizaje profundo explicable basado en cálculos recursivos y el uso de muestras de datos de entrenamiento reales que son muy representativas. Con este método, el modelo obtiene la mejor exactitud de 97.4 % ($F1 = 97.3\%$ y $AUC = 97.4\%$) en el conjunto de prueba.

Wang et al. [17] rediseñaron el modelo COVID-NET X-Ray de Lin et al. [18], desarrollado para el diagnóstico de COVID-19 en imágenes de rayos X, para desarrollar un modelo de diagnóstico de COVID-19 a partir de imágenes de CT. Ellos proponen una solución al problema de negligencia de la heterogeneidad de los datos combinando los conjuntos de datos *COVID-CT* y *SARS-CoV-2* antes mencionados. El nuevo modelo *COVID-NET* rediseñado incluye dos bloques (un bloque convolucional 2D y un bloque denso) que trabajan de forma paralela para generar el diagnóstico de la imagen de CT. La mejor exactitud de este modelo, en el conjunto de prueba, es $90.83 \pm 0.93\%$ para los datos *SARS-CoV-2* y $78.69 \pm 1.54\%$ para el conjunto *COVID-CT*.

Uno de los últimos trabajos realizados en el desarrollo de modelos 2D y recolección de datos es el modelo *COVIDNet-CT* y el conjunto de datos *COVIDx CT-2* de Gunraj et al. [11]. El conjunto de datos fue creado a partir de la selección de cortes 2D de volúmenes de distintos conjuntos de datos, de diferentes países. La versión *2A* del conjunto de datos *COVIDx CT-2*, el cual hemos utilizado para entrenar el modelo 2D presentado en este trabajo, comprende 194922 cortes de Tomografías Computarizadas clasificados en tres categorías: 94548 cortes en COVID-19, 40291 para Neumonía común y 60083 en categoría Normal. La arquitectura de COVIDNet-CT está inspirada en el modelo COVID-NET X-Ray de Lin et al. [18] y realiza diagnósticos considerando las tres categorías definidas en el conjunto de datos. La mejor exactitud del modelo, en la versión *2A* del conjunto de datos, es de 98.1 %

Modelos tridimensionales

En mayo del 2020, el Centro Nacional de BioInformación de China (CNCB por sus siglas en inglés *China National Center for Bioinformatics*) publicó el conjunto de datos CNCB Dataset el cual contiene 267 volúmenes de Tomografías Computarizadas de pacientes diagnosticados con COVID-19, clasificados en 6 categorías de severidad. Con este conjunto de datos, Aboutalebi et al. [19] desarrollaron un modelo que implementa bloques tridimensionales (convolucionales y residuales) para clasificar el nivel de severidad de COVID-19, agrupado en tres categorías. En los experimentos realizados sobre el conjunto de prueba, obtuvieron la mejor exactitud de 78.5 %.

Con la publicación del conjunto de datos MosMedData de Morozov et al. [14], el cual hemos tomado para complementar el conjunto de datos para el entrenamiento y evaluación del modelo 3D que se presenta en este trabajo, se desarrollaron diversos modelos de redes neuronales para diagnóstico de COVID-19. El modelo desarrollado por Serte y Demirel [20], adapta la arquitectura de la ResNet-50 en una red tridimensional para realizar diagnósticos de COVID-19 en volúmenes truncados de CT. En sus experimentos, ellos alcanzan la mejor exactitud de 84 % ($AUC = 96\%$).

Rohila et al. [12] proponen el modelo *ReCOV-101* para detección de COVID-19 y severidad de la enfermedad, utilizando el conjunto de datos MosMedData et al. [14]. Ellos utilizan el modelo ResNet-101, multiplicado en paralelo, para crear una red tridimensional que realiza diagnósticos de COVID-19 en volúmenes de CT y clasifica los casos positivos en tres categorías de severidad. Los resultados en sus evaluaciones alcanzan la mejor exactitud de 94.9 %, en el conjunto de prueba.

Organización de la tesis

El presente trabajo está organizado de la siguiente manera:

El capítulo 1 describe la enfermedad de COVID-19. Se explica el origen de la enfermedad, los síntomas que se pueden presentar en las personas infectadas y los diferentes métodos de diagnóstico que actualmente se están utilizando. Hacemos énfasis en el método de detección de la enfermedad basado en el análisis de imágenes de tomografía computarizada, describiendo de manera breve las lesiones identificadas por los radiólogos.

En el capítulo 2 se detalla la arquitectura de la red. Explicamos a detalle cada uno de los elementos que componen la red principal: la componente base y el complemento 3D. Al final, se describe la función de pérdida (o función de costo) y el método de optimización que se utilizó.

En el capítulo 3 se describen los conjuntos de datos que se utilizaron para entrenar los modelos presentados en este trabajo. Se explica a detalle cómo están

compuestos, cómo están clasificados cada uno de sus datos y se describe el preprocesamiento que se aplicó a cada uno de ellos.

En el capítulo 4 se muestran los resultados obtenidos para cada modelo. Abordamos primero los resultados que se obtuvieron con los distintos modelos bidimensionales. Luego, explicamos la partición del conjunto de datos y el método de entrenamiento que se implementó para desarrollar los modelos tridimensionales. Al final, se muestran los resultados obtenidos con los modelos tridimensionales.

Finalmente, comparamos los resultados obtenidos con los resultados del estado del arte, realizamos conclusiones y se discuten los posibles trabajos futuros.

Capítulo 1

La enfermedad de COVID-19

1.1. Descripción de la enfermedad

La enfermedad por Coronavirus de 2019, conocida como COVID-19, es una enfermedad causada por el virus del Síndrome Respiratorio Agudo Severo (SARS-COV-2) [7, 21]. Los primeros casos se presentaron en diciembre de 2019, en la ciudad de Wuhan, provincia de Hubei, en China; y el 11 de marzo del 2020, la Organización Mundial de la salud la declaró una pandemia [22].

1.1.1. Origen

Taxonómicamente hablando, el virus SARS-COV-2 se ubica en la familia Coronaviridae [23]. Esta familia se subdivide en cuatro géneros: Alphacoronavirus, Betacoronavirus, Gammacoronavirus y Deltacoronavirus [24, 25]. Muchos coronavirus son causantes de enfermedades en animales domésticos, y por lo tanto son de interés veterinario. Los coronavirus de importancia médica, actualmente conocidos, pertenecen a uno de los dos primeros géneros mencionados [24].

Desde una perspectiva ecoepidemiológica, los coronavirus se pueden clasificar en dos grupos: coronavirus humano y coronavirus zoonótico. Los coronavirus humanos circulan libremente en la población mundial y suelen causar enfermedades respiratorias leves [24]. Por otro lado, los coronavirus zoonóticos circulan transitoriamente y pueden generar grandes epidemias de enfermedades respiratorias graves [26].

El origen de los coronavirus de importancia médica parece ser zoonótico. En particular, los betacoronavirus zoonóticos están filogenéticamente relacionados con los coronavirus de murciélagos, los cuales podrían haber sido su fuente para el hombre, ya sea directamente o a través de un agente intermediario [24, 25]. Aunque es importante mencionar que, a la fecha, aún no es claro cuál pudo haber sido el intermediario para el SARS-COV-2, o si pasó directamente del murciélago al humano.

1.1.2. Síntomas

Los síntomas más habituales (fiebre, tos seca y cansancio) están presentes en la mayoría de los pacientes, pero no en todos los casos sintomáticos. Otros síntomas como la pérdida del gusto o el olfato, congestión nasal, dolor de cabeza, dolor de garganta, dolores musculares o articulares, náuseas o vómitos y diarrea, son menos frecuentes y pueden afectar a algunos pacientes. Cuando el paciente presenta un cuadro grave de COVID-19 se observan síntomas tales como disnea (dificultad respiratoria), dolor u opresión persistente en el pecho y temperatura alta (por encima de los 38°C) [8].

1.1.3. Diagnóstico

El diagnóstico de COVID-19 presenta algunas limitaciones. Actualmente, el diagnóstico se confirma mediante la prueba de reacción en cadena de la polimerasa con transcriptasa reversa en tiempo real (RT-PCR, por sus siglas en inglés). Sin embargo, el diagnóstico de los pacientes y el aislamiento temprano puede verse afectado por el retraso de los resultados y la falta de equipo de pruebas [1].

Algunos estudios han demostrado una considerable tasa inicial de falsos negativos, incluso en pacientes sintomáticos; algunos de ellos ya presentaban anomalías en la imagen de CT de tórax y luego se volvieron positivos en los resultados de laboratorio [1]. Aunque la prueba mediante análisis de CT no reemplaza la RT-PCR, es una excelente herramienta de diagnóstico para tales pacientes y de mucha utilidad para el monitoreo de la progresión de la enfermedad.

1.2. Hallazgos identificados en CT de tórax

El hallazgo más común consiste en múltiples opacidades de vidrio esmerilado. Suelen tener forma redondeada y se localizan principalmente en la periferia de los lóbulos pulmonares y en las regiones posteriores, algunas veces en las bases. La afección puede evolucionar a un patrón de pavimentación loca y consolidaciones coalescentes [1].

Hallazgos como afectación de la vía aérea, agrandamiento de los ganglios linfáticos, excavaciones, consolidaciones lobares, nódulos o cambios perihiliares no son frecuentes. Cuando estos hallazgos se encuentran en la CT, se debe considerar la co-infección u otros diagnósticos [1].

1.2.1. Opacidad de vidrio esmerilado

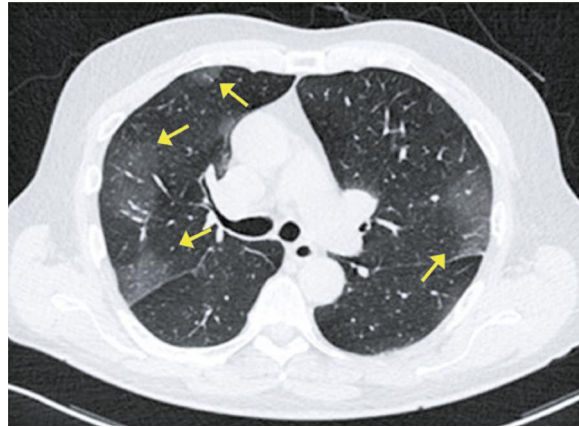


Figura 1.1: Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado en los lóbulos superiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]. La imagen en tamaño completo se encuentra en el [apéndice A.1](#).

Las opacidades en vidrio esmerilado se definen como un ligero aumento de la densidad pulmonar, sin oscurecer las paredes de los vasos y los bronquios. Es el hallazgo más común y temprano en pacientes con COVID-19; a menudo presentan distribución bilateral, periférica y subpleural en los lóbulos inferiores [1, 27].

1.2.2. Pavimentación loca

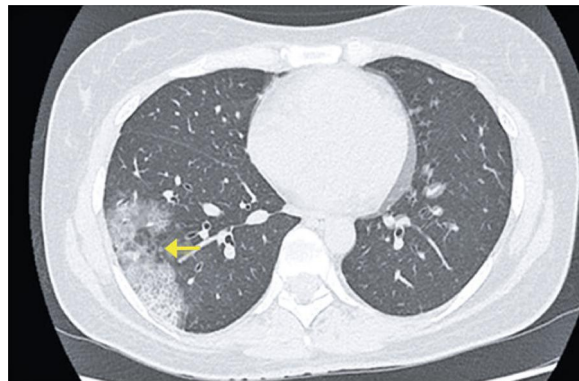


Figura 1.2: Tomografía computarizada que muestra patrón de pavimentación loca en la periferia del lóbulo inferior derecho. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]. La imagen en tamaño completo se encuentra en el [apéndice A.2](#).

Las opacidades en vidrio esmerilado a veces se superponen con un engrosamiento septal ubicado dentro o en medio de los lóbulos pulmonares secundarios. Estos hallazgos superpuestos se denominan patrón de pavimentación loca. En la

enfermedad de COVID-19, se caracteriza con mayor frecuencia algunos días después del inicio de los síntomas [1, 27].

1.2.3. Signo de halo invertido

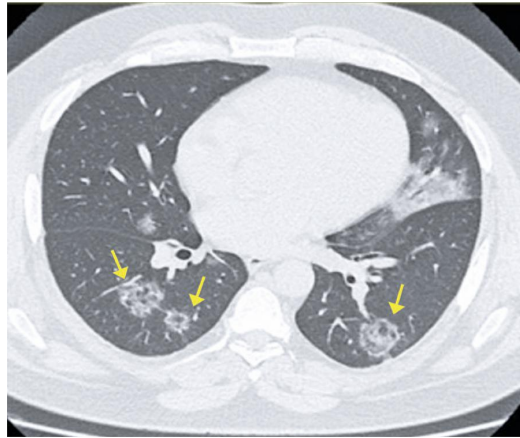


Figura 1.3: Tomografía computarizada que muestra signo de halo invertido en los lóbulos inferiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]. La imagen en tamaño completo se encuentra en el [apéndice A.3](#).

El signo del halo invertido se describe como un área central de opacidad en vidrio esmerilado, rodeada por un anillo de consolidación casi completo [1, 27].

1.2.4. Consolidación



Figura 1.4: Tomografía computarizada que muestra consolidaciones periféricas en los lóbulos superiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]. La imagen en tamaño completo se encuentra en el [apéndice A.4](#).

Después de la opacidad de vidrio esmerilado, la consolidación es el segundo hallazgo más frecuente encontrado en la neumonía causada por COVID-19. Tiende a

ocurrir en las últimas fases de la infección, principalmente después del décimo día. Este hallazgo radiológico se caracteriza por un aumento de la densidad pulmonar con vasos ocultos y líneas intersticiales, y a menudo presenta una forma redonda [1, 27].

1.2.5. Broncograma aéreo

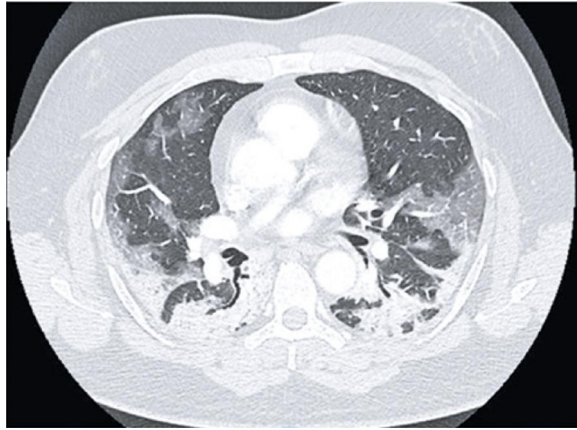


Figura 1.5: Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado, consolidaciones posteriores con broncogramas aéreos. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]. La imagen en tamaño completo se encuentra en el *apéndice A.5*.

El broncograma aéreo es la identificación de bronquios llenos de aire, dentro de una consolidación alveolar. Radiológicamente, se caracteriza por una estructura tubular hipoatenuante en medio de una parénquima pulmonar consolidada [1, 27].

Capítulo 2

Arquitectura de la red

En esta sección se describe el enfoque propuesto basado en Redes Neuronales Convolucionales (CNN por sus siglas en inglés *Convolutional Neural Networks*) que se utiliza para identificar casos de COVID-19 a través de imágenes CT de tórax. Por lo general, cuando el conjunto de datos está formado a partir de imágenes 3D, los tipos de redes neuronales artificiales que se suelen implementar son las Redes Convolucionales Tridimensionales 3D-CNN. Dichas redes, por muy pequeñas que sean, necesitan enormes cantidades de datos para generar excelentes resultados [12]. En nuestro caso, el conjunto de imágenes 3D no es lo suficientemente grande para desarrollar un modelo 3D-CNN con los rendimientos deseables. Este problema se resuelve mediante la implementación de una red 2D-CNN que recibe la imagen 3D como una muestra en sus componentes 2D, produce una salida para cada elemento, y luego, mediante una red 1D-CNN, procesa dichos resultados para generar una respuesta final.

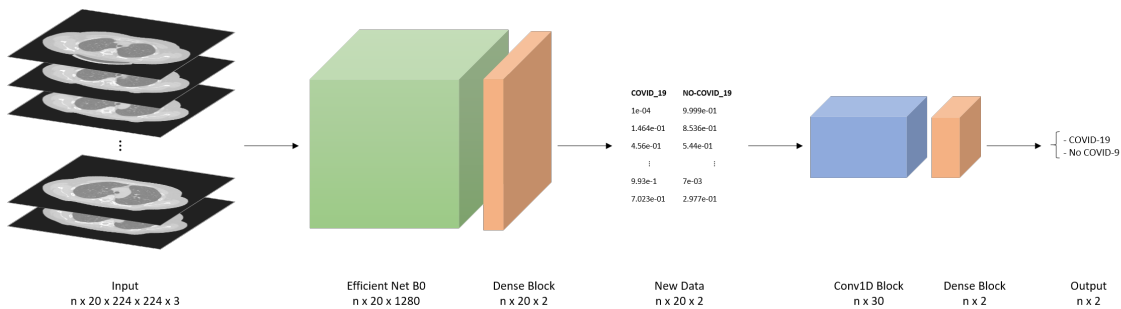


Figura 2.1: Estructura general del modelo de red neuronal convolucional EffCOV-B0, para identificar casos de COVID-19 a través de imágenes de tomografías computarizadas.

La estructura general del modelo CNN que se implementó en este trabajo se muestra en la *Figura 2.1*. El modelo recibe como entrada lotes (de tamaño n) de volúmenes CT de tórax compuestos por 20 cortes, cada uno de tamaño 224, 224, 3. La entrada es analizada por un sub-modelo 2D-CNN que se encarga de evaluar cada corte de manera independiente para generar una lista de diagnósticos (valor probabilístico para las categorías COVID-19 y Normal) que finalmente son

procesados por un sub-modelo 1D-CNN para generar el diagnóstico final.

La arquitectura del modelo EffCOV-B0 (*Figura 2.2*) que se propone está compuesta de dos sub-modelos CNN que fueron entrenados de manera independiente, pero trabajan de manera conjunta para generar el diagnóstico final de la CT del paciente. El primer sub-modelo de EffCOV-B0 es la *Componente Base* 2D-CNN cuya estructura principal es la red EfficientNet B0. Luego, de forma secuencial, se encuentra el segundo sub-modelo que corresponde al *Complemento 3D* que está basado en capas 1D-CNN. Los modelos EfficientNet poseen la cualidad de lograr la mayor precisión y la mejor eficiencia sobre las CNN existentes, reduciendo el tamaño de los parámetros y los FLOPS en un orden de magnitud [4]. La implementación de la componente base se puede interpretar mediante la analogía de un médico que analiza cada imagen 2D de una tomografía computarizada completa, generando una lista con los diagnósticos individuales de cada imagen. Dicha lista de diagnósticos se convierte en la entrada del complemento 3D que procesa dicha información y produce un diagnóstico final para la tomografía computarizada.

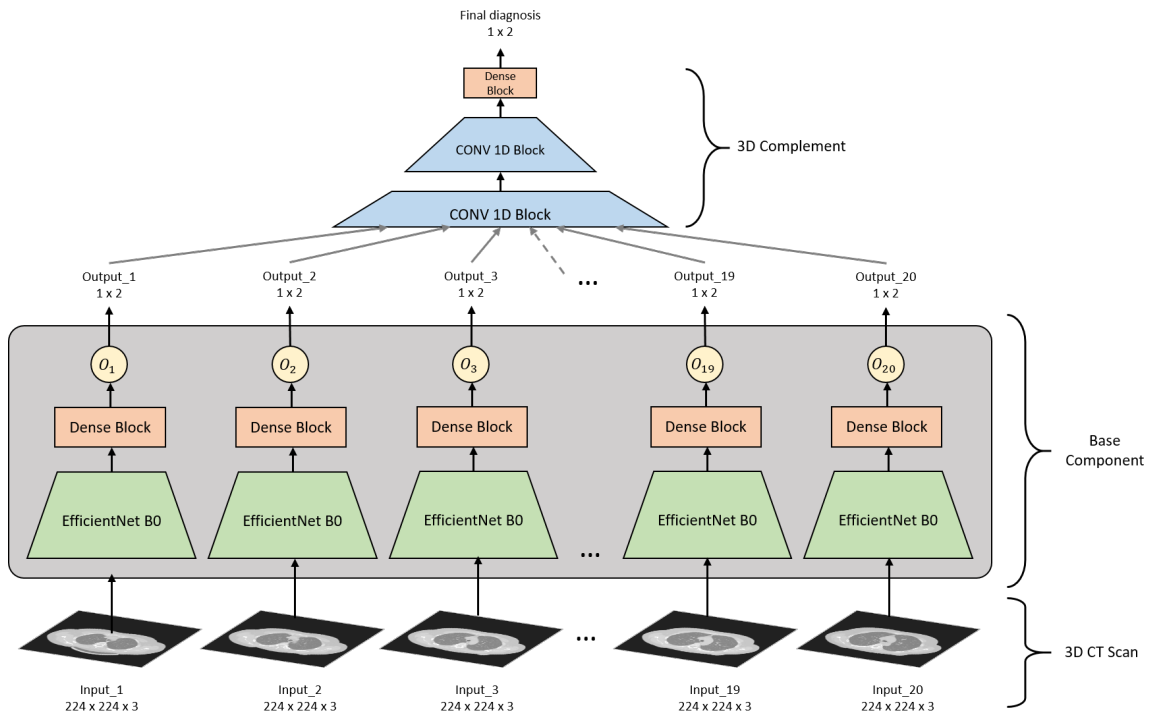


Figura 2.2: Componente base y Complemento 3D del modelo EffCOV-B0.

La incorporación de la componente base en el modelo EffCOV-B0 se logra mediante la implementación de la capa de tiempo distribuido *TimeDistributed*. Dicha capa puede aplicar la misma transformación (componente base) para una lista de datos de entrada (lotes de volúmenes de CT en nuestro caso). Posteriormente, puede dejar la capa distribuida en el tiempo para continuar trabajando con los datos, lo que equivale a tener múltiples capas que trabajan de forma paralela sin

extender la cantidad de parámetros que se podrían generar en este proceso. De esta forma, la entrada del modelo EffCOV-B0 se configura a $(n, 20, 224, 224, 3)$ en la capa de tiempo distribuido, asignando un tamaño de lote fijo (igual a 20) a la componente base, para generar una salida de tamaño $(n, 20, 2)$. Luego, después de procesar los lotes de volúmenes CT en la capa de tiempo distribuido, el complemento 3D se encarga de recibir dicha información y generar la salida de tamaño $(n, 2)$ con los diagnósticos de cada volumen. Cabe mencionar que el tamaño de la salida pudo haberse configurado a $(n, 1)$, definiendo el umbral a 0.5 para clasificar imágenes CT en la categoría COVID-19 si la salida es mayor que 0.5, ó en la categoría Normal en caso contrario. La *Tabla 2.1* resume la arquitectura del modelo EffCOV-B0. Los detalles de la componente base y el complemento 3D se explican en las siguientes sub-secciones y se resumen en las *Tabla 2.2* y *Tabla 2.3*, respectivamente.

Capa	Dimensión entrada	Dimensión salida	Parámetros
TimeDistributed(BaseComponent)	20, 224, 224, 3	20, 2	4, 385, 790
3D Complement	20, 2	2	257

Tabla 2.1: Arquitectura del modelo EffCOV-B0.

2.1. Componente base

La arquitectura de la componente base de EffCOV-B0 está basada en el modelo EfficientNet B0. Se utilizó dicha red, pre-entrenada en el conjunto de datos ImageNet [28] y eliminando su bloque de clasificación. El modelo recibe de entrada un lote de imágenes de tamaño $(n, 224, 224, 3)$ y genera una salida de tamaño $(n, 1280)$; donde n es el tamaño del lote, 224×224 las dimensiones en renglones y columnas, y 3 es el número de canales (colores). El número de canales debería ser uno para imágenes en tonos de gris, pero dado que usamos una red pre-entrenada en imágenes de color (canales Rojo-Verde-Azul) por construcción es tres. Lo que hacemos es replicar cada imagen en los tres canales. Una vez aclarado este punto, continuamos con la descripción del proceso. La salida de $(n, 1280)$ se pasa a un bloque de clasificación que se encarga de producir los diagnósticos individuales para cada imagen. La estructura de la componente base de EffCOV-B0, propuesta en este trabajo, se muestra en la *Figura 2.3* y en la *Tabla 2.2* se resume la arquitectura de dicha red.

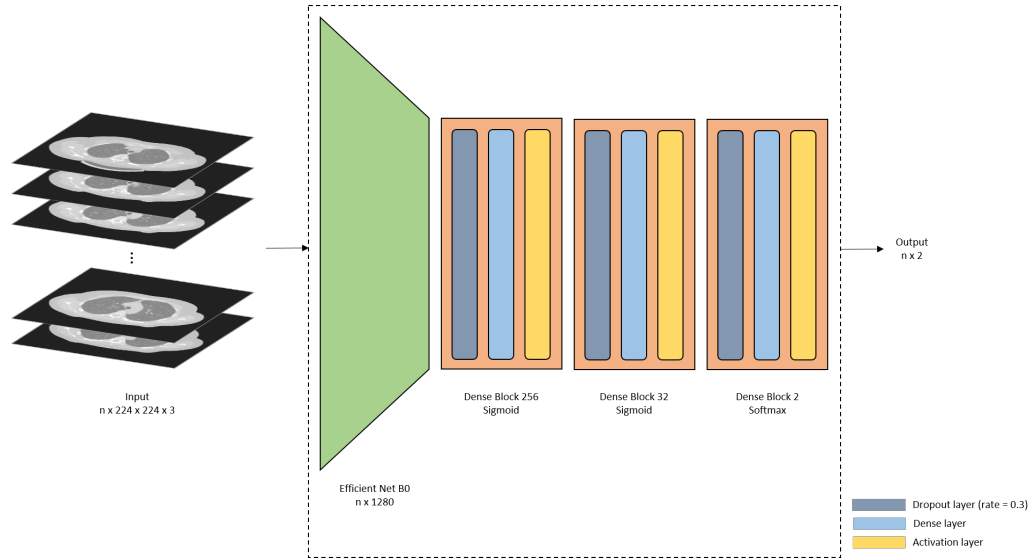


Figura 2.3: Estructura de la Componente Base de EffCOV-B0.

2.1.1. Modelo EfficientNet B0

EfficientNet es una familia de modelos que actualmente han superado la precisión de los modelos CNN, con una eficiencia hasta 10 veces mayor (más pequeña y más rápida) [2]. Utilizan un método de escalado de modelo novedoso que utiliza un coeficiente compuesto simple pero altamente efectivo para escalar las CNN de una manera más estructurada. A diferencia de los enfoques convencionales que escalan arbitrariamente las dimensiones de la red, como el ancho, la profundidad y la resolución; el método utilizado en EfficientNet escala uniformemente cada dimensión con un conjunto fijo de coeficientes de escala [2].

La arquitectura base (EfficientNet B0) utiliza convolución móvil invertida de cuello de botella (MBConv), similar a MobileNetV2 y MnasNet, pero es ligeramente más grande debido a un aumento del número de FLOPS [2]. En la *Figura 2.4* se muestra la estructura general del modelo base EfficientNet B0 y el resto de los modelos que forman la familia se obtienen de la ampliación del modelo base.

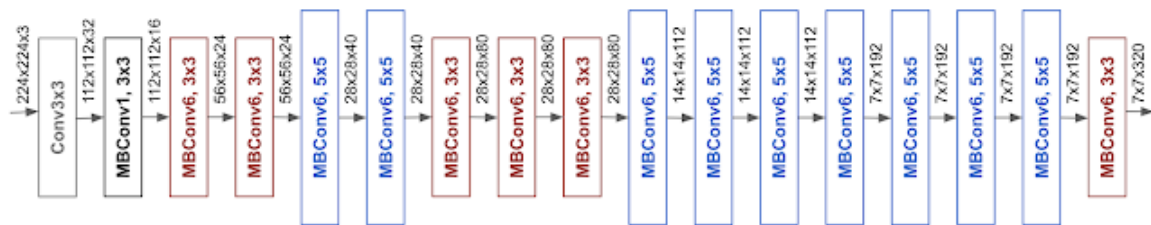


Figura 2.4: Arquitectura de la red EfficientNet B0. Imagen tomada de *EfficientNet: Improving Accuracy and Efficiency through AutoML and Model Scaling* [2].

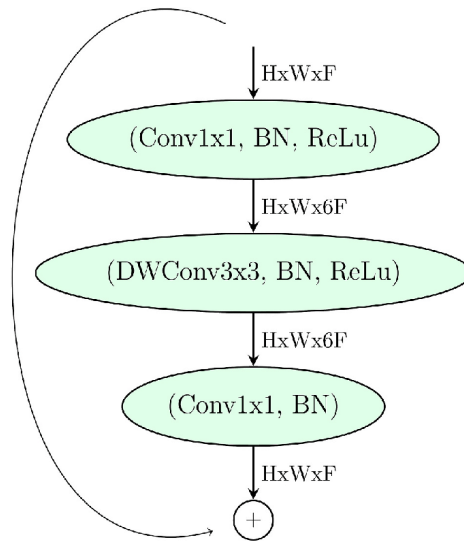


Figura 2.5: Bloque de convolución móvil invertida de cuello de botella (MBConv). Imagen tomada de *COVID-19 Detection in CT Images with Deep Learning: A Voting-based Scheme and Cross-Datasets Analysis* [3].

El bloque MBConv es una estructura residual que se compone de una secuencia de 6 capas convolucionales bidimensionales con activación ReLu y normalización por lote, seguidas de 6 capas de convolución profunda (cada una con activación ReLu y normalización por lote) y una capa de convolución 2D con normalización por lote (ver *Figura 2.5*).

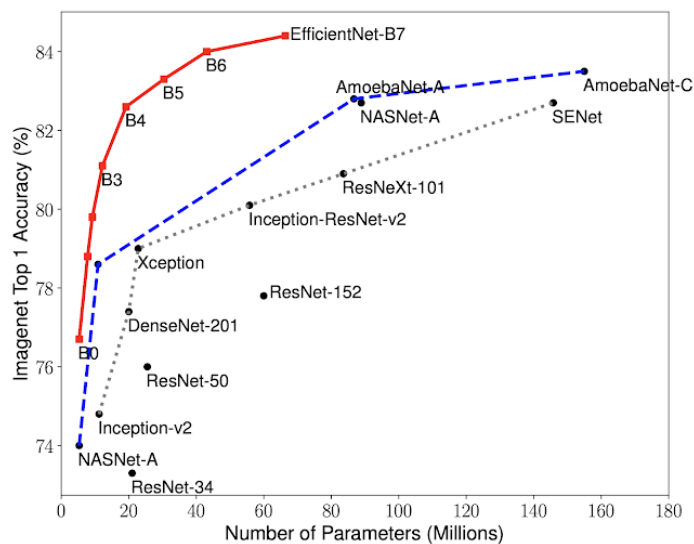


Figura 2.6: Gráfica comparativa del número de parámetros *vs* precisión (en el conjunto de datos ImageNet) de los modelos EfficientNet con respecto a otros modelos. Imagen tomada de *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks* [4].

Los modelos EfficientNet se han comparado con varios modelos CNN existentes

en el conjunto de datos ImageNet. En general, los modelos EfficientNet lograron una mayor precisión y una mejor eficiencia sobre las CNN existentes, reduciendo el tamaño de los parámetros y los FLOPS en, al menos, un orden de magnitud.

También se ha comprobado que los modelos EfficientNet pre-entrenados en ImageNet han obtenido buenos resultados al transferirse a otros conjuntos de datos de aprendizaje por transferencia ampliamente utilizados, como CIFAR-100 (*accuracy* 91.7%) y Flowers (*accuracy* 98.8%).

2.1.2. Bloque de clasificación

El bloque de clasificación en la componente base recibe como entrada el tensor de tamaño $(n, 1280)$ generado por la base convolucional del modelo Efficient-Net B0. En la *Figura 2.3* se puede visualizar la estructura de dicho bloque, el cual está formado por tres sub-bloques densos secuenciales los cuales se componen de una capa *Dropout* con tasa de 0.3 seguido de una capa densa. Los primeros dos sub-bloques tienen función de activación *sigmoide* y generan una salida de tamaño 256 y 32 respectivamente. El último sub-bloque tiene función de activación *softmax* y produce una salida de tamaño $(n, 2)$ la cual contiene los valores probabilísticos de cada clase (COVID-19, NORMAL).

Capa	Dimensión Entrada	Dimensión Salida	Parámetros
EfficientNet B-0	224, 224, 3	1280	4, 049, 564
Dropout 0.3	1280	1280	0
Dense (sigmoid)	1280	256	327, 936
Dropout 0.3	256	256	0
Dense (sigmoid)	256	32	8, 224
Dropout 0.3	32	32	0
Dense (softmax)	32	2	66

Tabla 2.2: Arquitectura de la componente base.

2.2. Complemento 3D

La salida generada por la componente base del modelo EffCOV-B0 que contiene los diagnósticos de cada imagen 2D de la tomografía computarizada, se transfiere al complemento 3D para producir el diagnóstico final. La entrada en el complemento 3D es un lote de tamaño $(n, 20, 2)$ el cual se procesa en un bloque convolucional unidimensional que produce una salida de tamaño $(n, 3, 10)$. La salida aplanada de tamaño $(n, 30)$ del bloque convolucional pasa a la capa de clasificación para obtener la salida de tamaño $(n, 2)$ que contiene el diagnóstico final en valores probabilísticos de cada clase (COVID-19, NORMAL). En la *Figura 2.7* se visualizan cada bloque

del Complemento 3D y en la *Tabla 2.3* se resume su arquitectura.

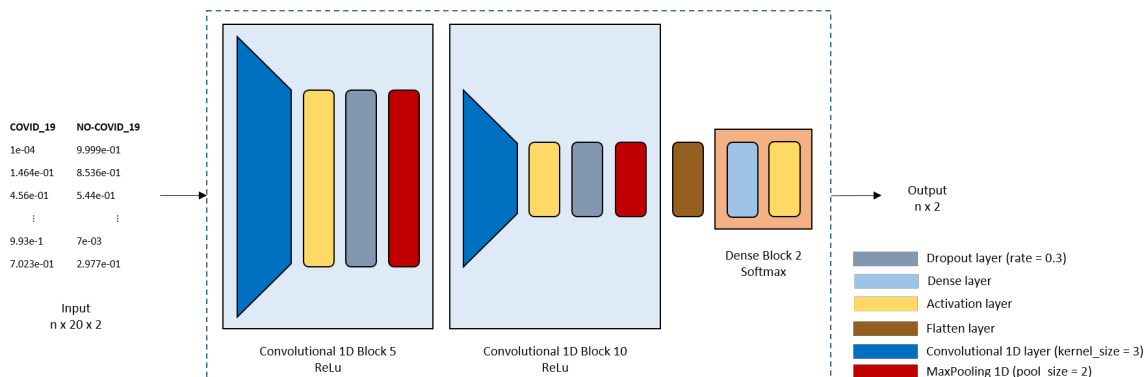


Figura 2.7: Estructura general del Complemento 3D.

2.2.1. Bloque convolucional

El bloque convolucional está compuesto de dos sub-bloques de convolución unidimensional. El primer sub-bloque contiene una capa convolucional unidimensional de 5 filtros, tamaño de kernel 3 y función de activación ReLu; una capa *Dropout* con tasa de 0.3 y una capa *MaxPooling* unidimensional de tamaño 2. El segundo sub-bloque convolucional genera una salida de tamaño $(n, 3, 10)$ y tiene la misma estructura del primer sub-bloque. La única diferencia se encuentra en el número de filtros de la capa convolucional, el cual se duplica para formar 10 filtros.

2.2.2. Bloque de clasificación

El bloque de clasificación recibe la salida aplanada de tamaño $(n, 30)$ del bloque convolucional y está formada por una única capa densa con función de activación *softmax* y salida de tamaño $(n, 2)$.

Capa	Dimensión Entrada	Dimensión Salida	Parámetros
Conv1D (relu)	20, 2	18, 5	35
Dropout 0.3	18, 5	18, 5	0
MaxPooling1D 2	18, 5	9, 5	0
Conv1D (relu)	9, 5	7, 10	160
Dropout 0.3	7, 10	7, 10	0
MaxPooling1D 2	7, 10	3, 10	0
Flatten	3, 10	30	0
Dense (softmax)	30	2	62

Tabla 2.3: Arquitectura del complemento 3D.

2.3. Optimizador y función de pérdida

En los modelos que componen la arquitectura de red que se propone en este trabajo (Componente Base y Complemento 3D) se utilizó la misma función de pérdida y el mismo método de optimización. La función de pérdida fue la Entropía Cruzada Binaria con pesos balanceado, mientras que, el optimizador fue el método de Propagación Cuadrática Media (RMSProp) con tasa de aprendizaje $\alpha = 1e - 5$.

2.3.1. RMSProp

Considere el problema de optimización sin restricciones mín $f(x)$, donde

$$f(x) = \frac{1}{n} \sum_{i=1}^n f_i(x). \quad (2.1)$$

Cada función $f_i(x)$ se asocia con la i -ésima observación en el conjunto de datos (conjunto de entrenamiento).

El método RMSProp (contracción del inglés *Root Mean Square Propagation*) es una variante del método del Gradiente Descendente Estocástico en el que la tasa de aprendizaje se adapta para cada uno de los parámetros. La tasa de aprendizaje α de un peso se divide por un promedio móvil de las magnitudes de gradientes recientes para ese peso. Inicialmente se calcula el promedio móvil en términos de medias cuadradas:

$$v(x, t) = \gamma v(x, t - 1) + (1 - \gamma) (\nabla f_i(x))^2, \quad (2.2)$$

donde γ es el factor de olvido. Finalmente, los parámetros se actualizan uno a uno con

$$x_i = x_i - \frac{\alpha}{\sqrt{v_i(x, t)}} \nabla f_i(x). \quad (2.3)$$

RMSProp ha demostrado una buena adaptación de la tasa de aprendizaje en diferentes aplicaciones y es capaz de trabajar con lotes pequeños [29].

2.3.2. Entropía cruzada binaria con pesos balanceados

Cuando la distribución de valores en las clases del conjunto de datos son muy diferentes, los algoritmos de aprendizaje automático tienden a sesgarse hacia las categorías con valores mayoritarios, realizando un mal desempeño en la categoría de menor valor.

La mayoría de los modelos de redes neuronales no son útiles con datos de clases sesgadas. Cuando este problema ocurre, es necesario modificar la función de costo para considerar la distribución sesgada de las clases. Este problema se puede resolver asignando diferentes pesos a las clases dependiendo inversamente del

número de casos.

Cuando se asignan pesos a las clases, la función de pérdida de entropía binaria cruzada se define por:

$$\mathcal{L}(\theta) = -\frac{1}{n} \sum_{i=1}^n [w_1 y_i \log(h_\theta(x_i)) + w_2 (1 - y_i) \log(1 - h_\theta(x_i))] \quad (2.4)$$

donde

- n : número de elementos en el conjunto de entrenamiento
- w_j : pesos $j = 1, 2$
- y_i : etiqueta del i -ésimo elemento de entrenamiento
- x_i : entrada del i -ésimo elemento de entrenamiento
- h_θ : modelo de red neuronal con pesos θ .

Los pesos se pueden configurar para ajustar la importancia de las etiquetas positivas. El propósito es penalizar la clasificación errónea hecha por la clase minoritaria estableciendo un peso de clase más alto y al mismo tiempo reduciendo el peso para la clase mayoritaria. [30].

Cuando se utilizan pesos balanceados, el modelo asigna automáticamente los pesos de clase inversamente proporcionales a sus respectivas frecuencias. La fórmula para calcular esto es:

$$w_j = \frac{n}{k \bar{n}_j} \quad (2.5)$$

donde

- w_j : peso de la j -ésima clase
- n : número de elementos en el conjunto de entrenamiento
- k : número total de clases únicas
- \bar{n}_j : número de elementos de la j -ésima clase (en el conjunto de entrenamiento)

Capítulo 3

Datos

3.1. Conjunto de datos 2D

Para entrenar y validar la Componente Base del modelo EffCOV-B0, para la detección de COVID-19 a partir de imágenes 2D de CT, se utilizó el conjunto de datos **COVIDx CT-2** de *Kaggle* [31]. Únicamente se consideró la variante “A” del conjunto de datos, ya que dicha variante consta de casos con diagnósticos confirmados por medio de RT-PCR y radiólogos.

Actualmente, el conjunto de datos comprende 194,922 cortes de CTs de 3,745 pacientes. Los datos están clasificados en tres categorías: Neumonía por Nuevo Coronavirus (NCP por sus siglas en inglés: *Novel Coronavirus Pneumonia*) debido a infección viral del SARS-CoV-2, Neumonía Común (CP por sus siglas en inglés: *Common Pneumonia*) y Controles Normales. Para los volúmenes de NCP y CP, se aprovecharon los cortes marcados que contenían anomalías pulmonar. [11]. La distribución de los datos se muestra en la *Tabla 3.1*.

Categoría	# Datos
COVID-19	94,548
Pneumonia	40,291
Normal	60,083

Tabla 3.1: Distribución de datos del conjunto COVIDx CT-2.

Basado en los experimentos iniciales del modelo COVIDNet-CT [11]. Se descubrió que los indicadores erróneos en las imágenes, como son, la base de los escáneres de CT donde reposa el paciente, artefactos de imagen, entre otros, estaban siendo aprovechados por la red para hacer predicciones. Para prevenir este problema, se realizó segmentación de imagen en el conjunto de datos, para retirar todo indicador visual que se encontrara fuera del cuerpo del paciente, como se muestra en la *Figura 3.1*.

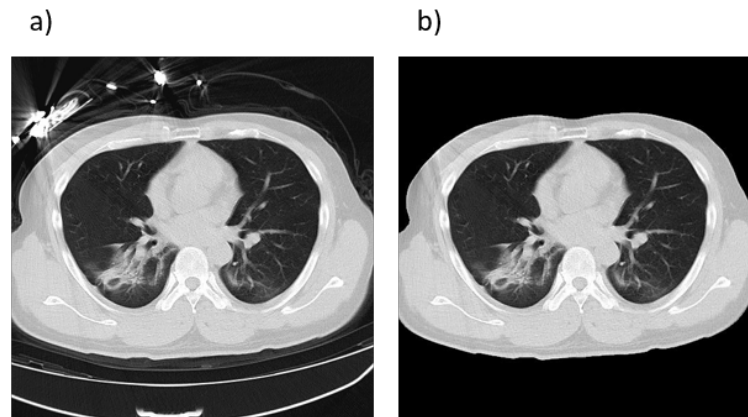


Figura 3.1: Segmentación de imagen 2D, correspondiente a un corte de tomografía computarizada. Se retira todo indicador visual que se encuentra fuera del cuerpo del paciente.

3.1.1. Preprocesamiento de datos

Se redimensionó y reescaló de imagen para adaptar el conjunto de datos a los requerimientos de la red. Se cambió el tamaño de cada imagen a $(224, 224, 3)$, que corresponde al tamaño de entrada de datos en la arquitectura de la Componente Base. Los píxeles de cada imagen se reescalaron del intervalo de números enteros $[0, 255]$ al intervalo de números reales $[0, 1]$, para el correcto entrenamiento de la red.

El método de segmentación de imagen que se implementó utiliza funciones de la librería *OpenCV* de Python [11]. El proceso inicia reduciendo el ruido de la imagen mediante la aplicación de un filtro Gaussiano (con tamaño de kernel $(5, 5)$). Luego, con la imagen suavizada, se calcula el umbral binario ($THRESH_BINARY + THRESH_OTSU$) para clasificar los píxeles de la imagen original y obtener una nueva imagen binaria definida únicamente con dos tipos de píxeles (1 ó 0). Finalmente, se obtiene el mayor contorno de la imagen binaria para definir la máscara del cuerpo del paciente. Los resultados de cada etapa del proceso de segmentación de imagen se muestran en la *Figura 3.2*.

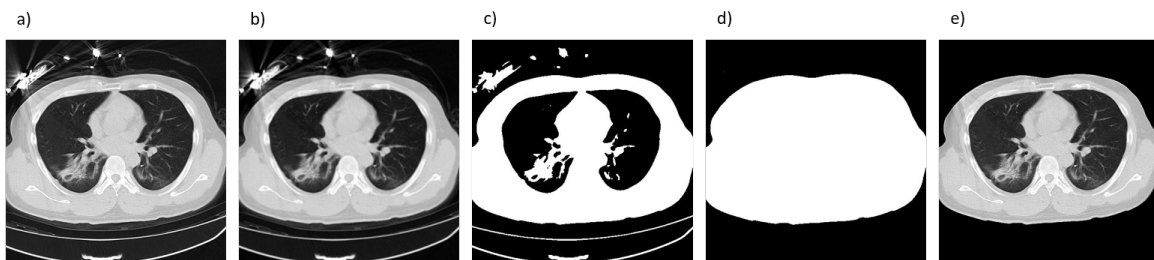


Figura 3.2: Método de segmentación de imagen 2D. a) Imagen original; b) Imagen suavizada con filtro Gaussiano; c) Imagen binaria construida a partir del umbral definido con la imagen suavizada; d) Imagen del mayor contorno de la imagen binaria (máscara del cuerpo del paciente); e) Imagen segmentada del cuerpo del paciente.

Finalmente, para mejorar el rendimiento y evitar el sobreajuste de la Componente Base del modelo EffCOV-B0, se implementó una estrategia de “aumento de datos” (*data augmentation*) con los siguientes parámetros:

- Rotación de imagen: $rotation_range \in [-15, 15]$ grados .
- Ancho de desplazamiento de imagen: $width_shift_range \in [-0.15, 0.15]$ relativos a la dimensión de la imagen.
- Alto de desplazamiento de imagen: $height_shift_range \in [-0.15, 0.15]$ relativos a la dimensión de la imagen.
- Inclinación de imagen: $shear_range \in [-0.15, 0.15]$ relativos a la dimensión de la imagen.
- Enfoque de imagen: $zoom_range \in [-0.15\%, 0.15\%]$.

3.2. Conjunto de datos 3D

El Complemento 3D de EffCOV-B0 se entrenó y se validó con los conjuntos de datos **COVID-19 CT-INER México** y **MosMedData**. El conjunto COVID-19 CT-INER México consta de 200 CT de tórax, de pacientes que fueron diagnosticados con COVID-19. En el conjunto MosMedData: *Chest CT Scans with COVID-19 Related Findings*, las tomografías computarizadas se obtuvieron entre el 1 de marzo de 2020 y 25 de abril de 2020, y fueron proporcionadas por hospitales municipales en Moscú, Rusia [14]. El conjunto de datos contiene escáneres de tomografía computarizada de 1,110 pacientes, clasificados en cinco categorías:

- CT-0: tejido pulmonar normal, sin signos de neumonía viral.
- CT-1: algunas opacidades de vidrio esmerilado, la afectación de la parénquima pulmonar es inferior al 25 %.
- CT-2: opacidades de vidrio esmerilado, la afectación de la parénquima pulmonar está entre el 25 y el 50 %.
- CT-3: opacidades de vidrio esmerilado y regiones de consolidación, la afectación de la parénquima pulmonar está entre el 50 y el 75 %.
- CT-4: opacidades de vidrio esmerilado y consolidación, así como cambios reticulares en los pulmones. La afectación de la parénquima pulmonar supera el 75 %.

La *Tabla 3.2* contiene la distribución de los datos para cada una de las categorías. Los pacientes con CT en las categorías CT-1 y CT-2 fueron remitidos a casa. El seguimiento de los pacientes en CT-1 se realizó con tecnologías de telemedicina, mientras que, los pacientes en CT-2 el seguimiento fue realizado por un médico de atención primaria. En las categorías CT-3 y CT-4, los pacientes requirieron transferencia inmediata a la unidad de cuidados intensivos en hospitales especializados para pacientes diagnosticados con COVID-19 [14].

3.2.1. Preprocesamiento de datos

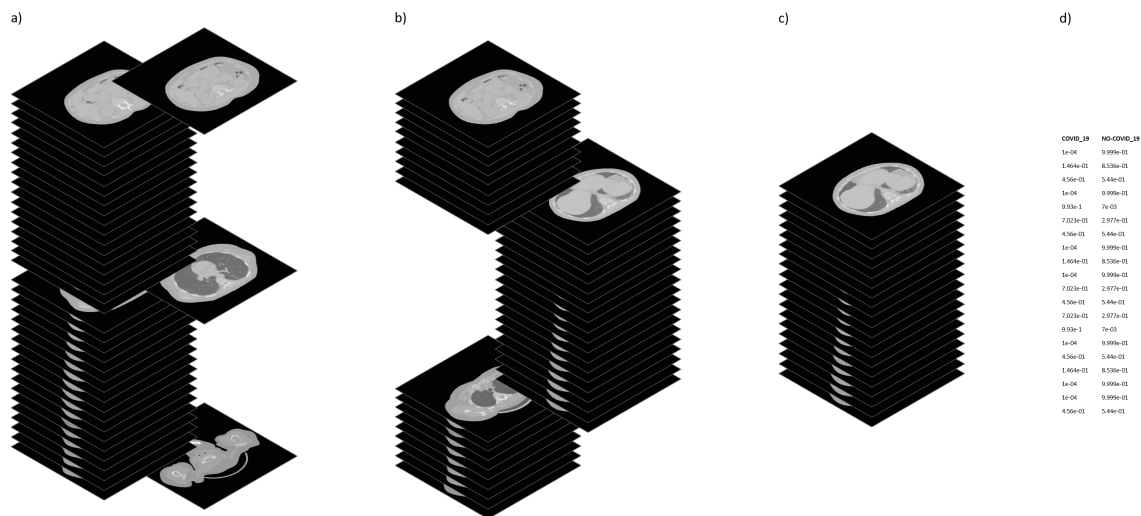


Figura 3.3: Preprocesamiento de imagen 3D. a) Segmentación de los cortes, se retira todo indicador visual que se encuentra fuera del cuerpo del paciente; b) Selección de cortes basado en el elemento central y los cortes vecinos más cercanos; c) Cortes seleccionados; d) Dato procesado, los cortes seleccionados se evalúan en la Componente Base de EffCOV-B0.

Para cada volumen, se obtienen los cortes de manera individual y se realiza el preprocesamiento de datos que se explica en la sección 3.1.1. Dado que los cortes ubicados en los extremos del volumen no contienen región pulmonar, se seleccionaron únicamente los cortes centrales. Para esto, se tomó el corte central del volumen y se seleccionaron los cortes vecinos de cada extremo (10 cortes en el extremo superior y 9 cortes en el extremo inferior). Finalmente, los cortes seleccionados se evalúan en la Componente Base para formar el dato de entrada en el Complemento 3D de EffCOV-B0.

Categoría	# Datos
CT-0	254
CT-1	684
CT-2	125
CT-3	45
CT-4	2

Tabla 3.2: Distribución de datos en el conjunto *MosMedData: Chest CT Scans with COVID-19 Related Findings*.

3.2.2. Segmentación pulmonar basada en CNN

Además del método de segmentación de imagen que se explica en la sección 3.1.1 y se aplica en el conjunto de datos 3D, también se utilizó el método de segmentación que se describe en el artículo titulado *Automated lung segmentation in CT under presence of severe pathologies* [32]. Dicho método consiste en el modelo de red U-Net entrenado para realizar la tarea de segmentación pulmonar, como se muestra en la *Figura 3.4*, en imágenes de tomografías computarizadas 3D.

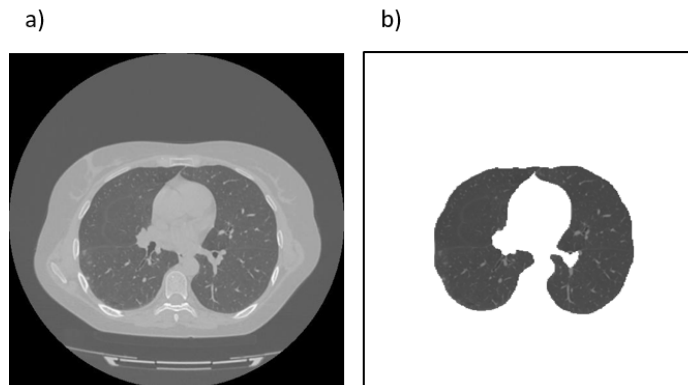


Figura 3.4: Segmentación pulmonar 3D, imagen correspondiente a un corte de tomografía computarizada. Se retira todo indicador visual que se encuentra fuera de la región pulmonar del cuerpo del paciente.

Dado que el modelo de red se entrenó en un conjunto de datos que contiene volúmenes con diversas patologías (incluyendo la enfermedad COVID-19), este método posee la ventaja de realizar la segmentación pulmonar en volúmenes que contiene cortes con lesiones causadas por COVID-19. La única desventaja que presenta el modelo es que sólo funciona en cortes completos. El corte a procesar debe mostrar el pulmón completo y el pulmón debe estar rodeado de tejido para segmentarse.

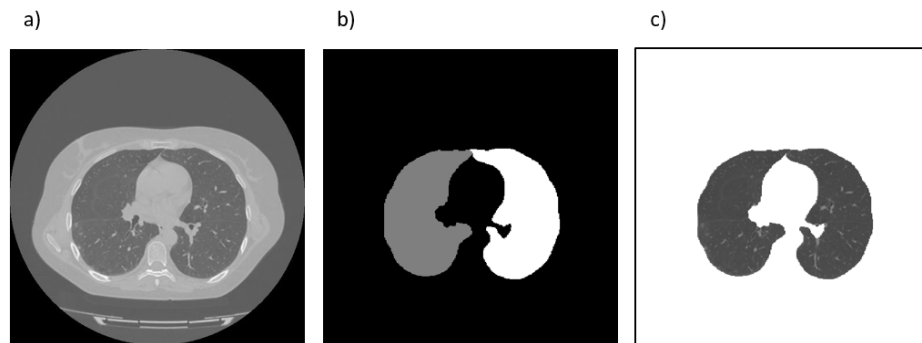


Figura 3.5: Método de segmentación de imagen 3D, para un corte específico. a) Corte original del volumen; b) Mascara del corte original, el pulmón izquierdo y el pulmón derecho están separados; c) Corte segmentado.

El método no requiere preprocesamiento de la imagen 3D. Recibe como entrada el volumen a segmentar (en formato *Nifti* o *DICOM*) y lo procesa para generar la máscara 3D que contiene los cortes de la región pulmonar (en caso de existir dicha región). Cada corte incluye el pulmón izquierdo y el pulmón derecho por separado, como se muestra en la *Figura 3.5*.

Capítulo 4

Resultados

4.1. Modelo 2D

Los resultados que se presentan en esta sección se obtuvieron con el conjunto de datos **COVIDx CT-2** de *Kaggle* [31], el cual se describe en la sección 3.1 del capítulo anterior. Los modelos se desarrollaron utilizando la librería *Keras* de *Tensorflow*, en el lenguaje de programación *Python*.

Dado que el conjunto de datos está compuesto de imágenes clasificadas en tres categorías, se desarrollaron dos modelos de redes utilizando la misma arquitectura. El primer modelo fue entrenado considerando todas las categorías del conjunto de datos, mientras que, el segundo modelo fue entrenado utilizando únicamente las categorías COVID-19 y Normal. La *Tabla 4.1* contiene información sobre el número de épocas, el tiempo de entrenamiento-validación y tiempo de evaluación del conjunto de prueba de cada modelo.

Modelo	Épocas	Tiempo (horas) entrenamiento-validación	Tiempo (minutos) prueba
Modelo, 3 clases	10	4.24	1.30
Modelo, 2 clases	15	6.34	1.29

Tabla 4.1: Tiempo de entrenamiento-validación y prueba de los modelos bidimensionales.

La descripción del hardware utilizado (memoria, procesador y GPU) para desarrollar los modelos de redes neuronales se encuentra en el *Apéndice A.1*.

4.1.1. Clases: COVID-19, Neumonía y Normal

El modelo que se describe en esta sección fue entrenado utilizando todas las categorías del conjunto de datos (COVID-19, Neumonía y Normal). Dicho proceso de entrenamiento, y de validación, se realizó con 10 épocas y para medir el desempeño de la red se utilizaron las métricas:

- Función de pérdida (*loss*)
- Exactitud (*accuracy*, *umbral* = 0.5)
- Función F1 (*F1*, *umbral* = 0.5)
- Área bajo la curva de ROC (*AUC*)

Dichas métricas también se utilizaron para evaluar el desempeño de la red en el conjunto de prueba.

En las siguientes secciones se describe la partición del conjunto de datos, se visualizan gráficas de las métricas en el proceso de entrenamiento, se tabulan las evaluaciones de las métricas y matrices de confusión en cada subconjunto de la partición, y finalmente, se muestran resultados de predicciones de la red y visualizaciones de mapa de calor con la técnica GradCAM.

Partición del Conjunto de Datos

En esta sección se describe la partición del conjunto COVIDx CT-2 en los subconjuntos de entrenamiento, validación y prueba. Se consideró la partición de datos predeterminada que ya se incluye en dicho conjunto:

Partición	COVID-19	Neumonía	Normal	Total
entrenamiento	88,467	25,496	35,996	159,959
validación	6,244	7,400	11,842	25,486
prueba	6,018	7,395	12,245	25,658

Tabla 4.2: Partición del conjunto de datos COVIDx CT-2.

Partición	COVID-19	Neumonía	Normal	Total
entrenamiento	2714	558	321	3593
validación	166	190	126	482
prueba	175	125	126	426

Tabla 4.3: Distribución de pacientes en la partición del conjunto de datos COVIDx CT-2.

Gráficas

Para medir el aprendizaje en el proceso de entrenamiento del modelo de red, se registraron los resultados obtenidos de las métricas en cada época. Las *Figura 4.1*, *Figura 4.2*, *Figura 4.3* y *Figura 4.4* muestran gráficas comparativas de cada métrica con respecto al número de época, para los subconjuntos de entrenamiento y validación.

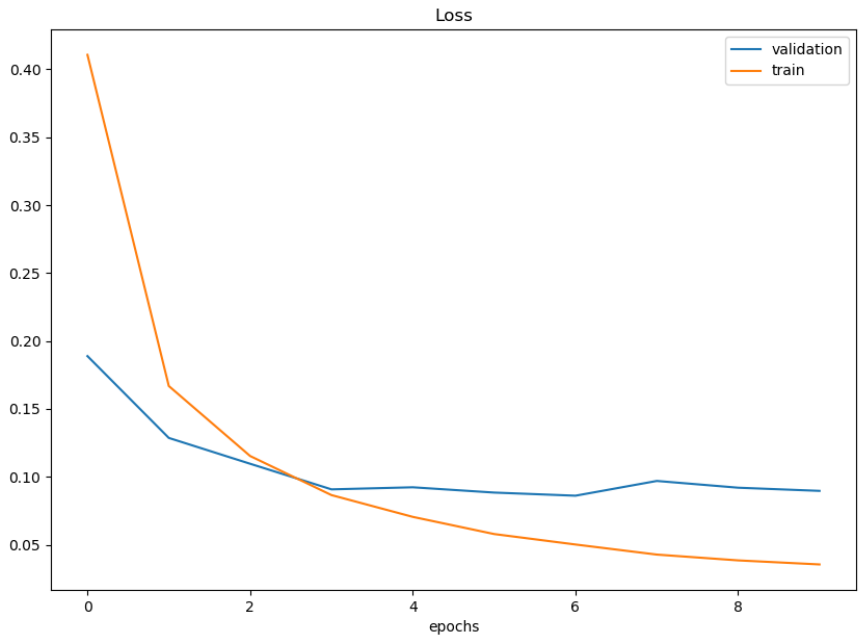


Figura 4.1: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

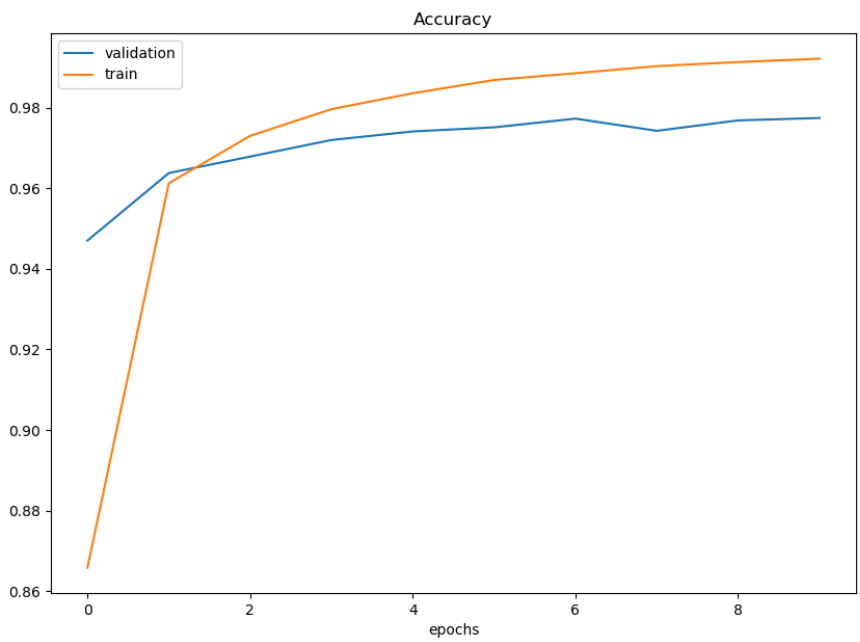


Figura 4.2: Gráfica de la exactitud con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

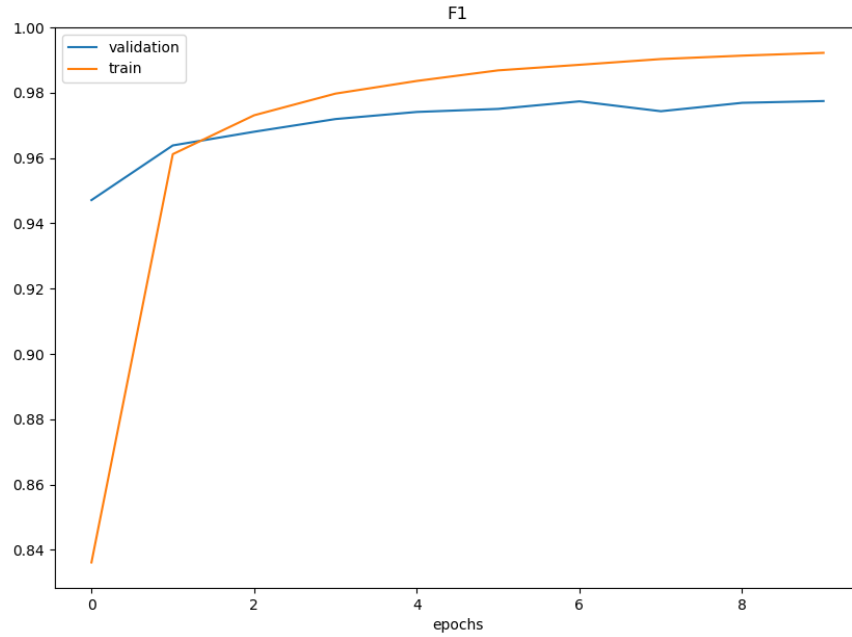


Figura 4.3: Gráfica de la función F1 con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

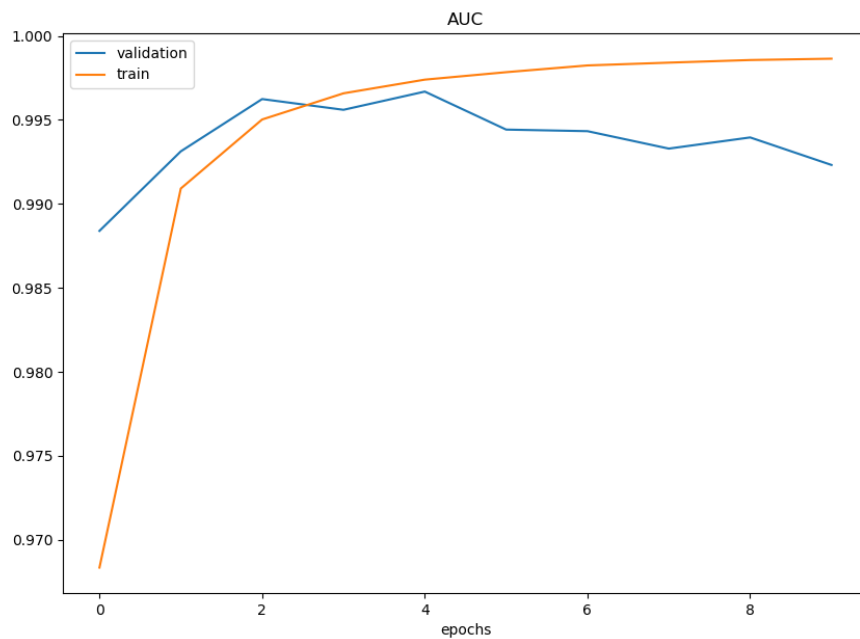


Figura 4.4: Gráfica de la métrica AUC con respecto al número de época de la Componente Base. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

En las gráficas de entrenamiento-validación se puede observar que la curva de entrenamiento se aplana a partir de la octava época, y en la quinta época para la curva de validación. Inicialmente se realizó un entrenamiento con 20 épocas y en las gráficas se observó que a partir de la décima época, las mejoras en el modelo no eran significativas. Por tal razón, para reducir el tiempo de entrenamiento y para evitar el sobreajuste del modelo; se redujo el número de épocas a la mitad.

Evaluaciones y Matrices de Confusión

Los resultados finales de las métricas que miden el desempeño de la red se muestran en esta sección. La *Tabla 4.4* contiene los resultados de las funciones de pérdida, exactitud, F1 y AUC-ROC (Área bajo la curva *Receiver Operating Characteristic, ROC*) en los subconjuntos de entrenamiento, validación y prueba. Adicionalmente se incluyen los resultados *Recall* y *Precision*, de cada categoría del conjunto de datos, en las tablas *Tabla 4.5* y *Tabla 4.6* respectivamente. Finalmente, las *Figura 4.5*, *Figura 4.6* y *Figura 4.7* muestran las matrices de confusión para los subconjuntos de prueba, entrenamiento y validación, respectivamente. Para cada una de las matrices de confusión de esta tesis, el eje vertical representa los valores de predicción, mientras que el eje horizontal representa los valores reales.

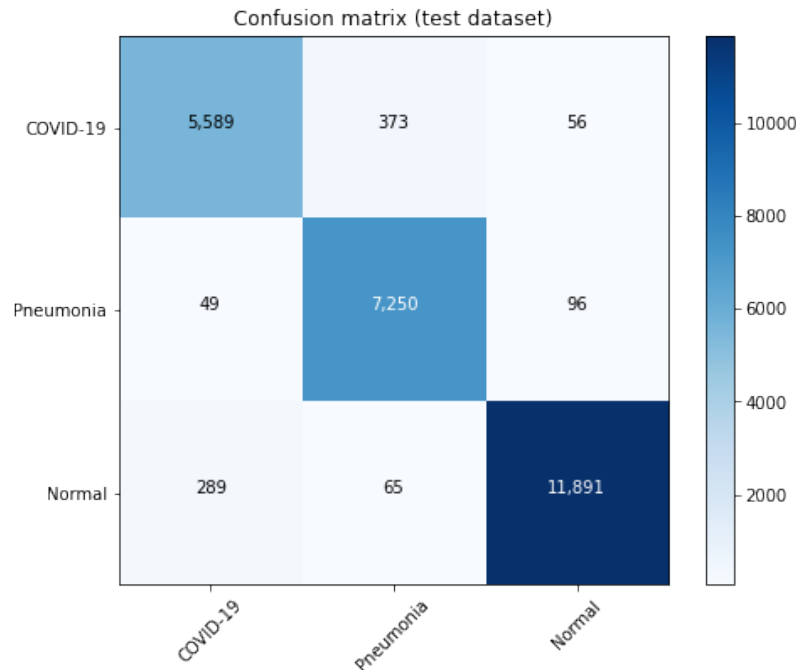


Figura 4.5: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de prueba. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

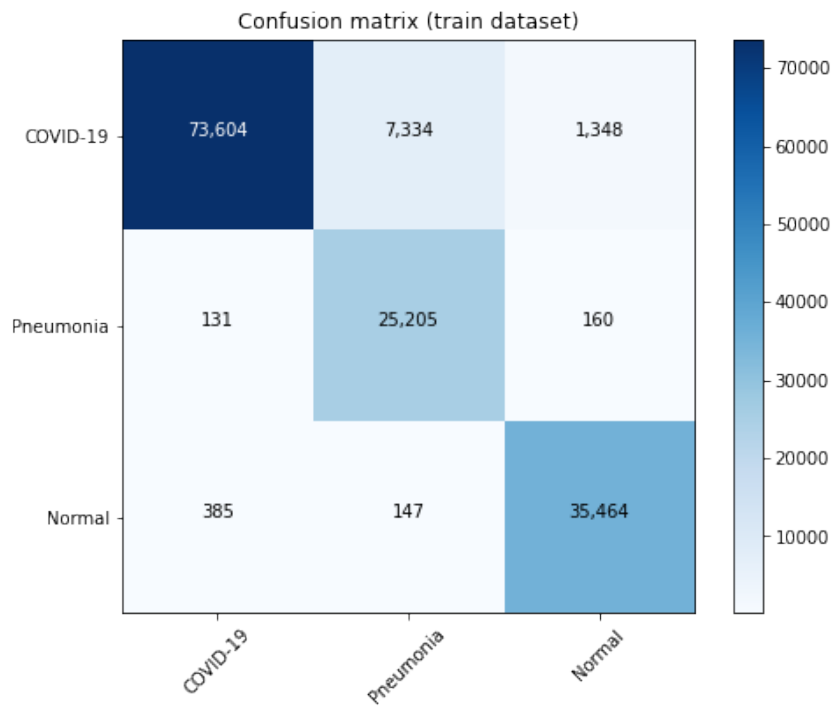


Figura 4.6: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de entrenamiento. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

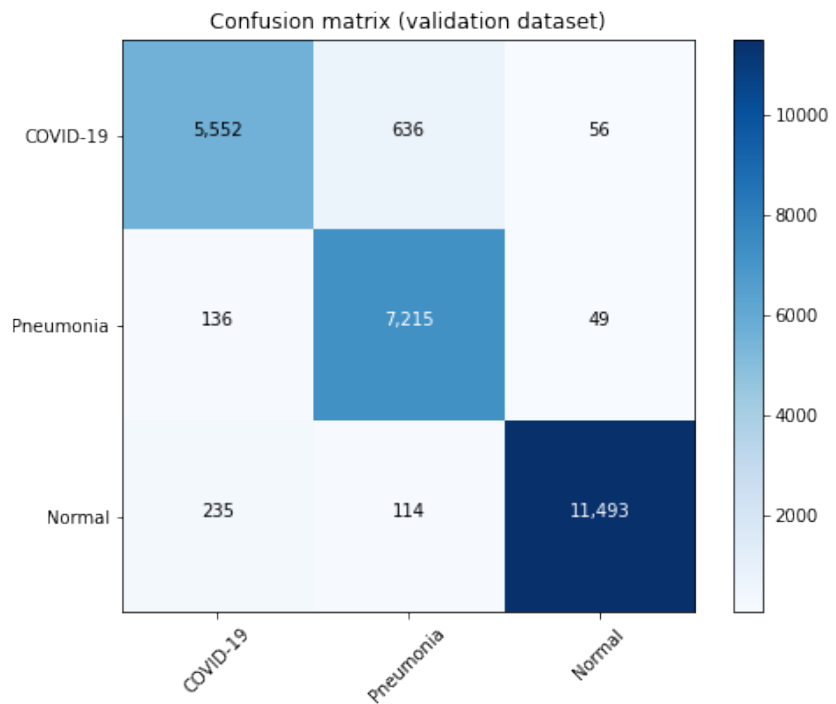


Figura 4.7: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de validación. Se consideran las tres categorías del conjunto de datos COVIDx CT-2.

Para cada partición del conjunto de datos, el mayor valor de cada fila en las matrices de confusión se encuentra en la diagonal principal. Esto nos indica que el modelo trabaja de manera correcta. No obstante, en la matriz del conjunto de entrenamiento (*Figura 4.6*) existe una notable confusión de la categoría Neumonía con respecto a la categoría COVID-19 y esto repercute en las demás particiones. En la misma figura también se puede observar una ligera confusión de las categorías Normal y COVID-19, sin embargo, esto no genera confusiones en las otras particiones.

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.042	99.1	99.1	99.8
validación	0.111	97.7	97.7	99.0
prueba	0.142	96.4	96.4	99.2

Tabla 4.4: Pérdida, exactitud, funciones F1 y AUC-ROC de la Componente Base en el conjunto de datos COVIDx CT-2.

Partición	COVID-19 %	Neumonía %	Normal %
entrenamiento	89.4	98.9	98.5
validación	88.9	97.5	97.1
prueba	92.9	98.0	97.1

Tabla 4.5: *Recall* de la Componente Base en el conjunto de datos COVIDx CT-2.

Partición	COVID-19 %	Neumonía %	Normal %
entrenamiento	99.3	77.1	95.9
validación	93.7	90.6	99.1
prueba	94.3	94.3	98.7

Tabla 4.6: *Precision* de la Componente Base en el conjunto de datos COVIDx CT-2.

La métrica de precisión nos permite analizar la calidad del modelo para predecir una categoría. En el conjunto de entrenamiento (ver *Tabla 4.6*), el modelo se limita a reconocer el 77.1 % de las imágenes con Neumonía. Sin embargo, en el conjunto de prueba el porcentaje en esta categoría aumenta a 94.3%. Para el resto de las categorías, en cada partición, los resultados son prometedores.

En el caso de la exhaustividad (*recall*), con ella podemos medir la cantidad de casos que se pueden identificar para una categoría específica. Los resultados de la *Tabla 4.5* nos muestra los porcentajes más bajos de identificación de datos en la categoría COVID-19 (para cada partición). Para el resto de las categorías, los porcentaje de identificación están por arriba del 97 % lo cual es muy bueno.

Predicciones

Finalmente, se incluyen algunas predicciones de imágenes que fueron seleccionadas de forma arbitraria del conjunto de prueba. Adicionalmente se visualizan las regiones de la imagen que fueron más relevantes durante la predicción, mediante la técnica de Grad-CAM [33]. Las *Figura 4.8*, *Figura 4.9* y *Figura 4.10* corresponden a predicciones de imágenes clasificadas en la categoría *COVID-19*; mientras que, las *Figura 4.11*, *Figura 4.12* y *Figura 4.13* corresponden a la categoría de *Neumonía* y las *Figura 4.14*, *Figura 4.15* y *Figura 4.16* son predicciones de imágenes clasificadas en la categoría *Normal*.

El mapa de activación de clase ponderada por el gradiente (Grad-CAM) produce un mapa de calor que destaca las regiones importantes de una imagen utilizando los gradientes del objetivo de la capa convolucional final [33]. Es una técnica útil para comprender cómo se ha conducido a una red neuronal convolucional a tomar una decisión de clasificación. Es específica para cada clase, lo que significa que puede producir una visualización diferente para cada clase presente en la imagen. El mapa de activación de clase atribuye una importancia a cada posición (x, y) en la última capa convolucional calculando la combinación lineal de las activaciones, ponderadas por los pesos de salida correspondientes para la clase observada. La cartografía de activación de clase resultante después se vuelve a muestrear según el tamaño de la imagen de entrada [33]. Esto se ilustra con el mapa térmico (heatmap) de la *Figura A.6*.

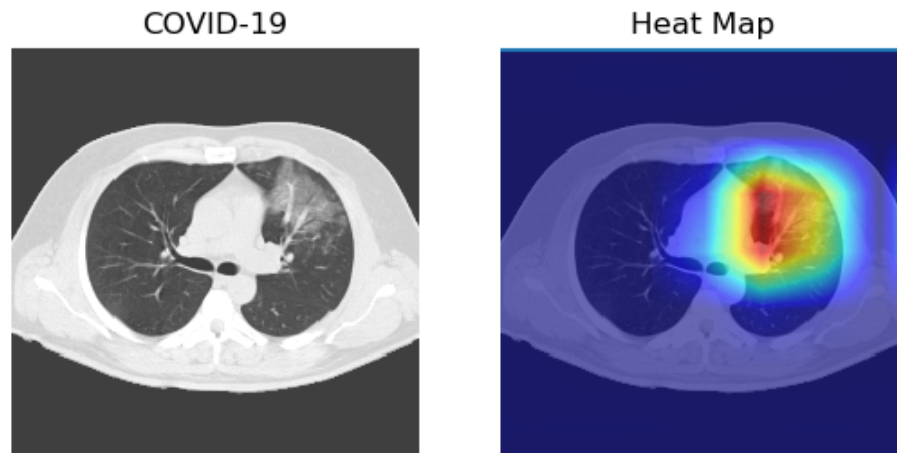


Figura 4.8: Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9991, 0.0005 y 0.0004 para las clases COVID-19, Neumonía y Normal respectivamente.

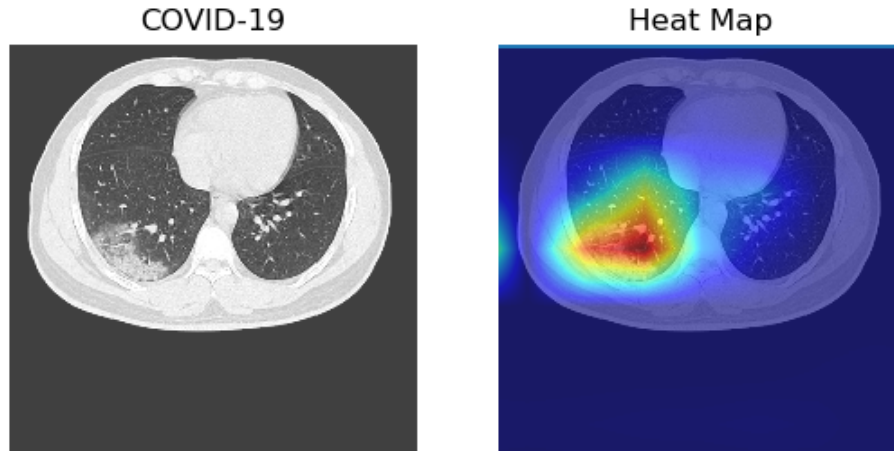


Figura 4.9: Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9983, 0.0012 y 0.0006 para las clases COVID-19, Neumonía y Normal respectivamente.

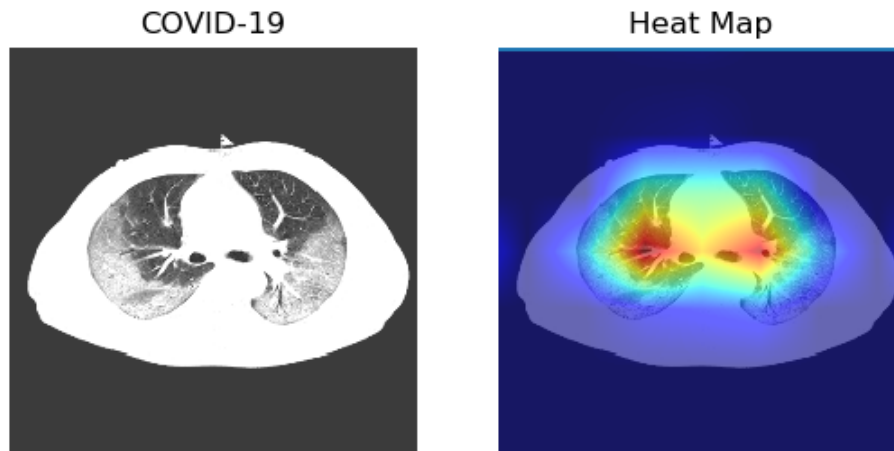


Figura 4.10: Grad-CAM de imagen CT categoría COVID-19. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.9992, 0.0005 y 0.0003 para las clases COVID-19, Neumonía y Normal respectivamente.

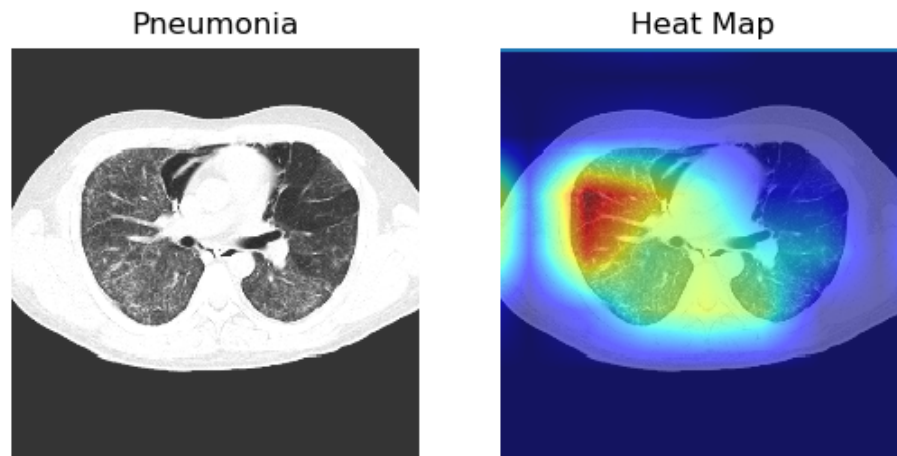


Figura 4.11: Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9983 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.

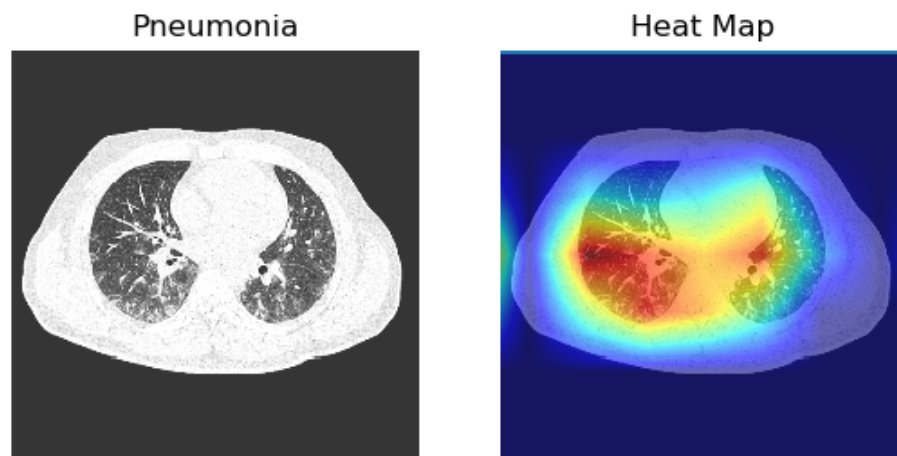


Figura 4.12: Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9984 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.

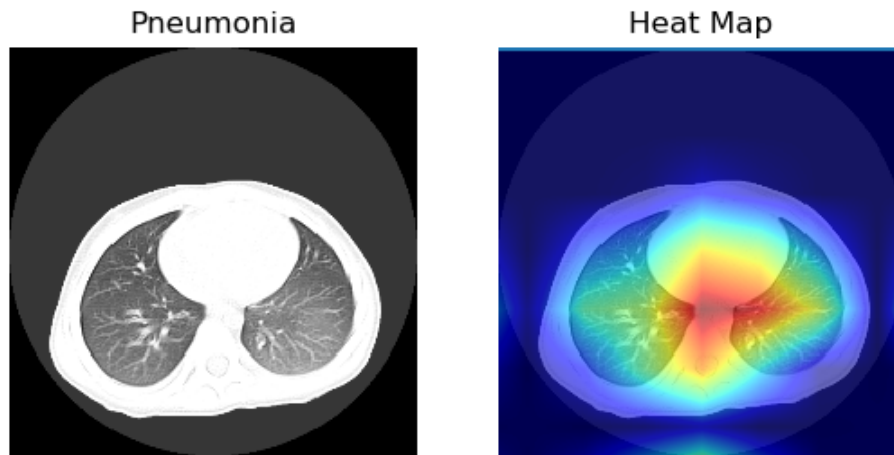


Figura 4.13: Grad-CAM de imagen CT categoría Neumonía. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0009, 0.9984 y 0.0007 para las clases COVID-19, Neumonía y Normal respectivamente.

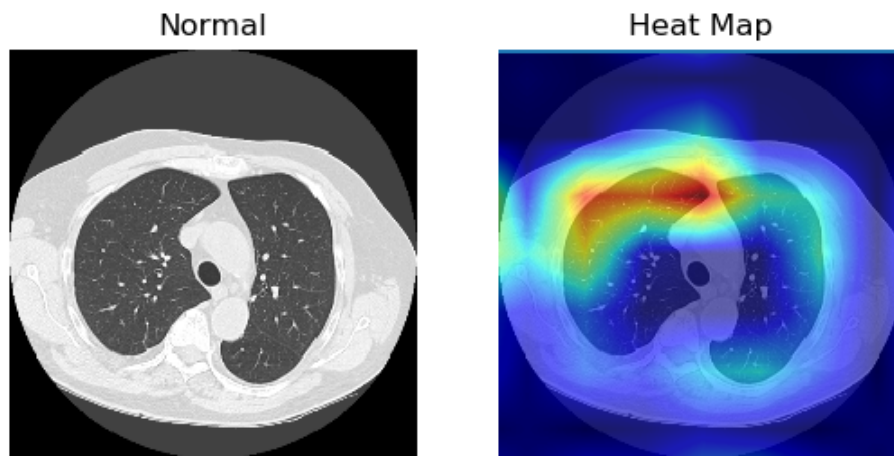


Figura 4.14: Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0907, 0.0036 y 0.9057 para las clases COVID-19, Neumonía y Normal respectivamente.

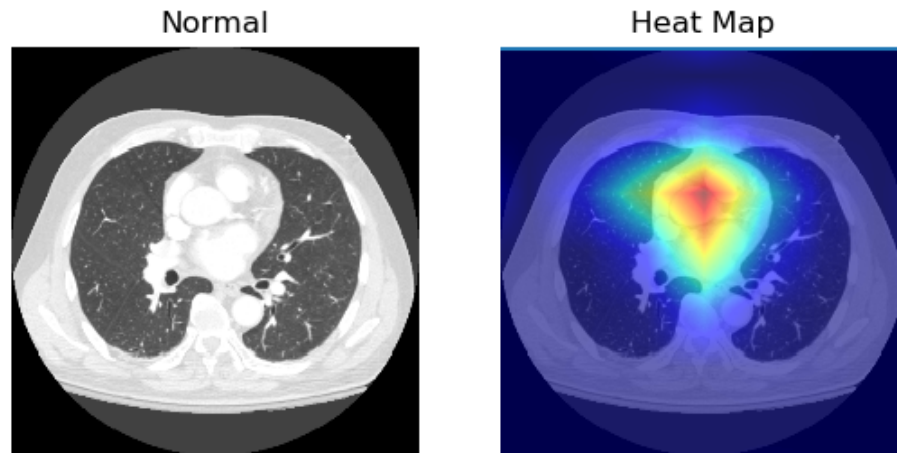


Figura 4.15: Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0008, 0.0006 y 0.9987 para las clases COVID-19, Neumonía y Normal respectivamente.

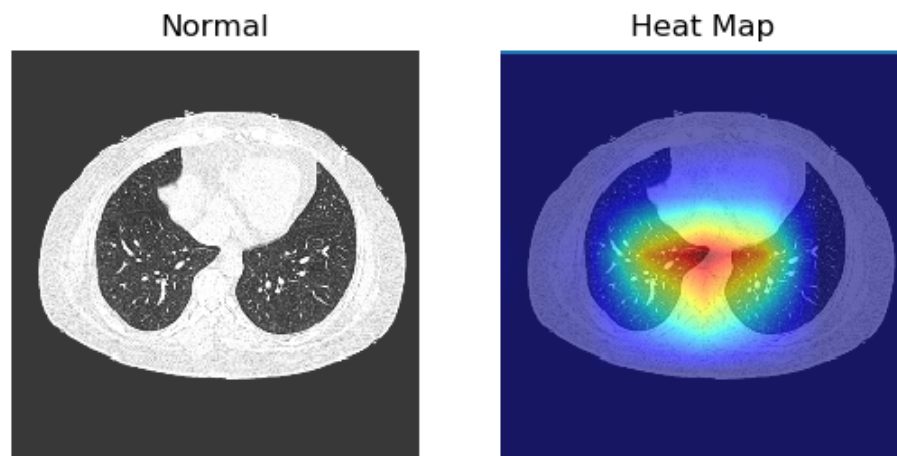


Figura 4.16: Grad-CAM de imagen CT categoría Normal. La imagen fue seleccionada de forma arbitraria y el modelo retorna los resultados 0.0007, 0.0006 y 0.9988 para las clases COVID-19, Neumonía y Normal respectivamente.

En general, los resultados de predicción en este modelo son muy prometedores. Los mapas de calor en las imágenes de categoría COVID-19 son asombrosos, en ellos se puede observar que las regiones relevantes para la predicción se ubican sobre la lesión causada por neumonía COVID-19. En el caso de las imágenes CT de categoría Neumonía, las regiones relevantes para la predicción abarcan ambos pulmones, esto se debe a que las lesiones de neumonía abarcan toda la región pulmonar en la mayoría de datos, en esta categoría. En el caso de los cortes normales, el modelo se limita a considerar pequeñas regiones relevantes de la imagen, en algunos casos afuera de los pulmones.

4.1.2. Clases: COVID-19 y Normal

Al igual que en la *Sección 4.1.1*, en esta última sección se muestran resultados correspondientes a la Componente Base del modelo EffCOV-B0. En este caso, se utilizaron únicamente las categorías COVID-19 y Normal de todo el conjunto de datos (ver *Tabla 4.7* y *Tabla 4.8*). El entrenamiento del modelo se realizó con 15 épocas utilizando las métricas de evaluación de desempeño en el aprendizaje que se describen en el modelo anterior (pérdida, exactitud, funciones F1 y AUC-ROC). La partición del conjunto de datos, así como las gráficas de las métricas durante el entrenamiento y los resultados de las predicciones del modelo, se describen en las siguientes secciones.

Partición del Conjunto de Datos

Partición	COVID-19	Normal	Total
entrenamiento	88467	35996	124463
validación	6244	11842	18086
prueba	6018	12245	18263

Tabla 4.7: Partición del conjunto de datos COVIDx CT-2. No se incluye la categoría Neumonía

Partición	COVID-19	Normal	Total
entrenamiento	2714	321	3035
validación	166	126	292
prueba	175	126	301

Tabla 4.8: Distribución de pacientes en la partición del conjunto de datos COVIDx CT-2. No se incluye la categoría Neumonía

Gráficas

Al igual que en la sección anterior, para medir el aprendizaje en el proceso de entrenamiento se registraron los resultados obtenidos de las métricas en cada época. Las *Figura 4.17*, *Figura 4.18*, *Figura 4.19* y *Figura 4.20* muestran gráficas comparativas de cada métrica con respecto al número de época, para los subconjuntos de entrenamiento y validación. En cada gráfica se puede observar que la curva de entrenamiento se aplatina a partir de la décima época. En el caso de la curva de validación, esta se eleva a partir de la décima época para la función de pérdida y sigue decreciendo para la métrica AUC-ROC. Para el resto de las métricas, las curva de validación se aplatina en la décima época. Para reducir el tiempo de entrenamiento y para evitar el sobreajuste del modelo; se definió el entrenamiento con 15 épocas y SE configuró para guardar el mejor modelo (usando *callbacks*) a medida se reducía el valor de la función de pérdida.

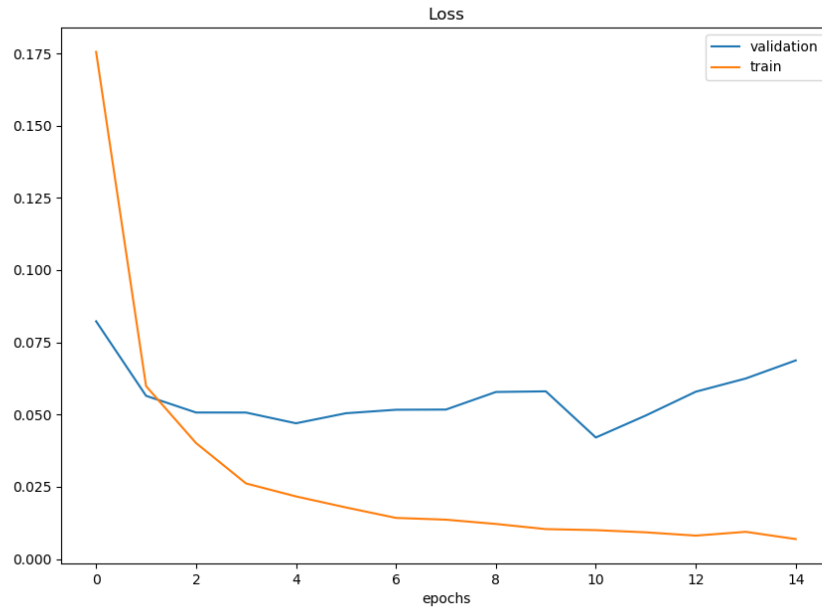


Figura 4.17: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 2D. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.

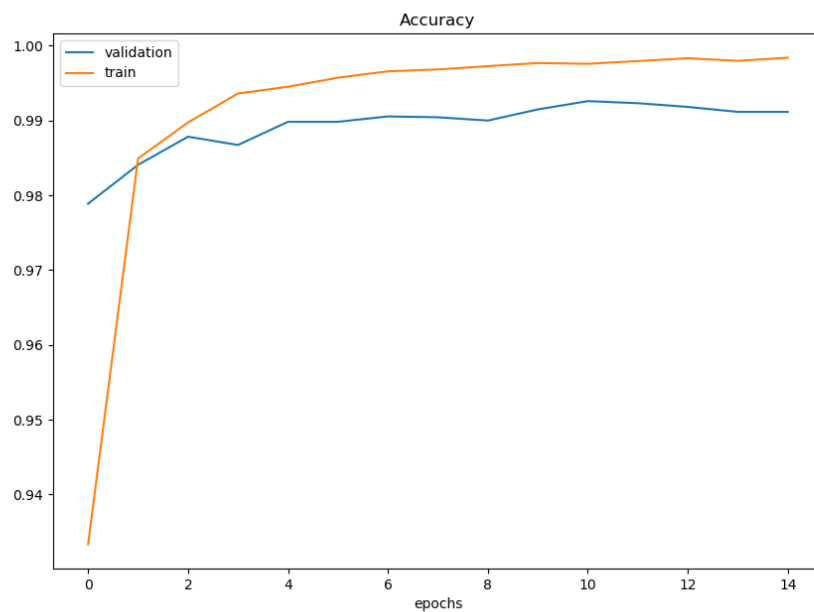


Figura 4.18: Gráfica de la exactitud con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.

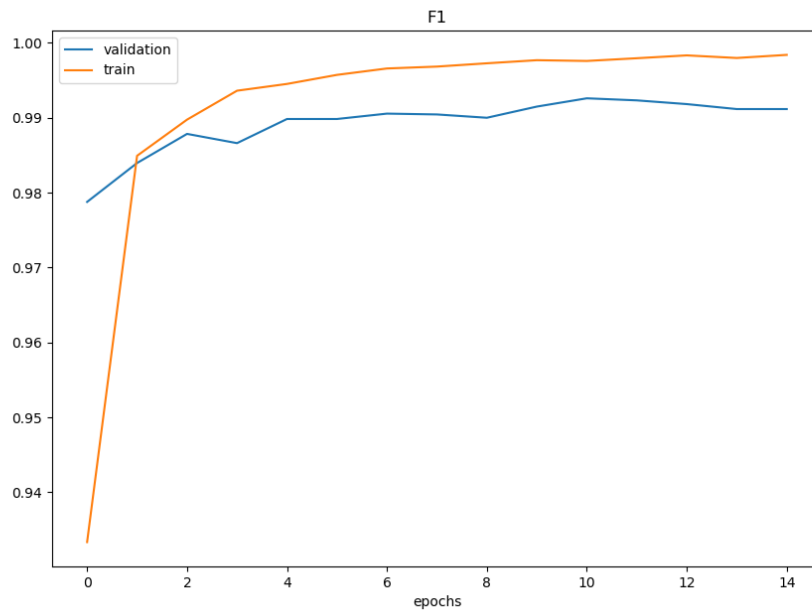


Figura 4.19: Gráfica de la función F1 con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.

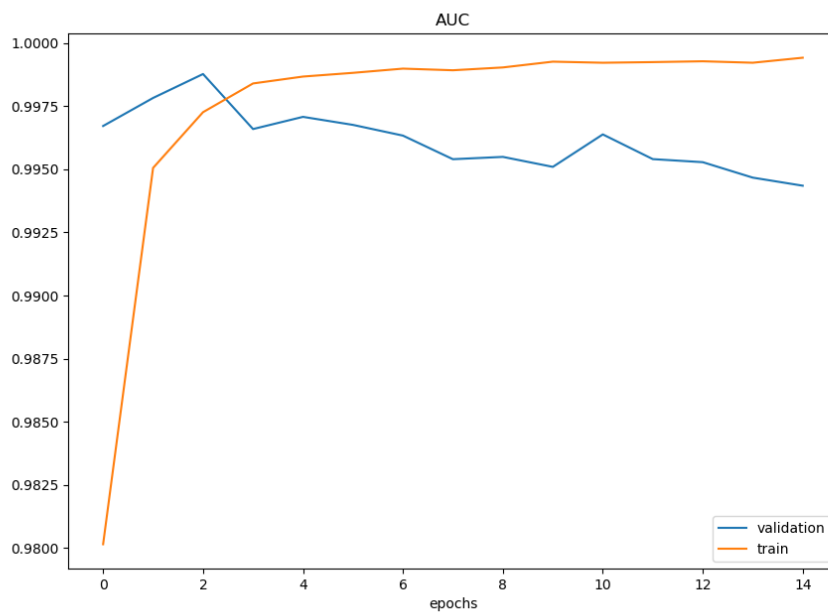


Figura 4.20: Gráfica de la métrica AUC-ROC con respecto al número de época de la Componente Base. Se consideran las categorías COVID-19 Y Normal del conjunto de datos COVIDx CT-2.

Evaluaciones y Matrices de Confusión

Mostramos en esta sección los resultados finales de las métricas que miden el desempeño de la red. La *Tabla 4.9* contiene los resultados de las funciones de pérdida, exactitud, F1 y AUC-ROC (Área bajo la curva *Receiver Operating Characteristic, ROC*) en los subconjuntos de entrenamiento, validación y prueba. Las *Tabla 4.10* y *Tabla 4.11* contienen resultados de las métricas *Recall* y *Precision*, respectivamente. Y finalmente, las *Figura 4.21*, *Figura 4.22* y *Figura 4.23* muestran las matrices de confusión para los subconjuntos de prueba, entrenamiento y validación, respectivamente.

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.01	99.8	99.8	99.9
validación	0.067	99.1	99.1	99.4
prueba	0.123	98.3	98.3	98.9

Tabla 4.9: Pérdida, exactitud, funciones F1 y AUC-ROC de la Componente Base. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.

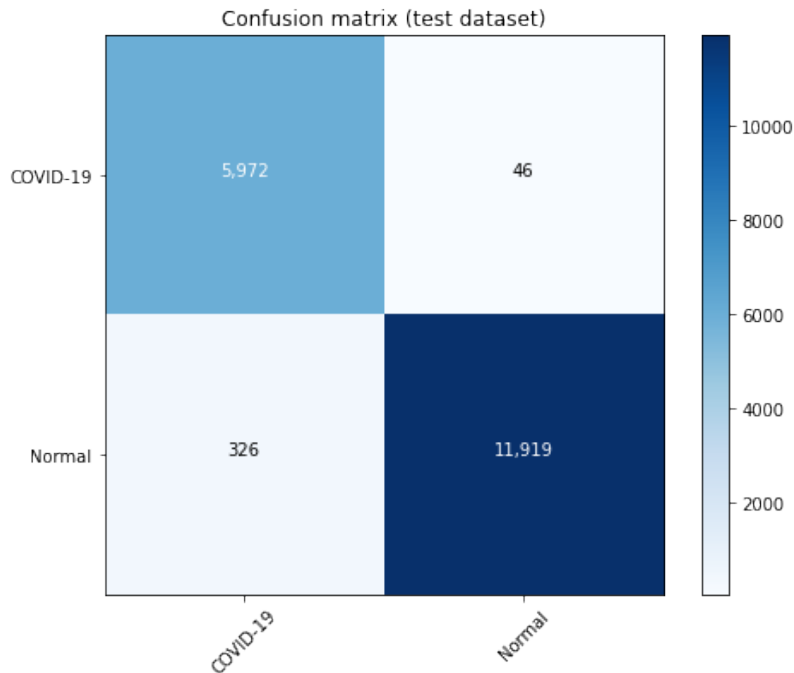


Figura 4.21: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de prueba. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.

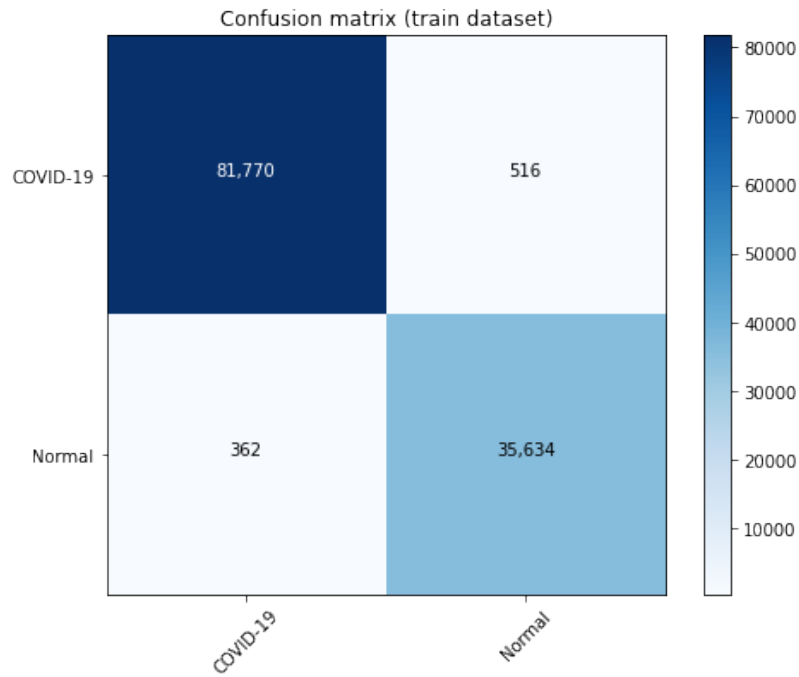


Figura 4.22: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de entrenamiento. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.

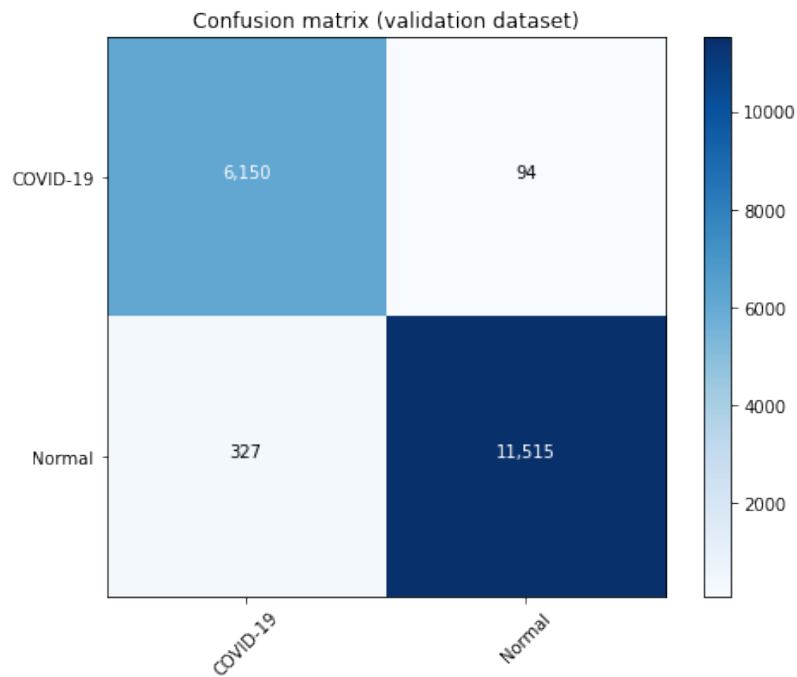


Figura 4.23: Gráfica de la matriz de confusión de la Componente Base en el subconjunto de validación. Se consideran las categorías COVID-19 y Normal del conjunto de datos COVIDx CT-2.

Al igual que los resultados de la sección anterior, el mayor valor de cada fila en las matrices de confusión se encuentra en la diagonal principal y esto nos indica que el modelo trabaja de manera correcta. Al retirar la categoría Neumonía del conjunto de datos, se puede observar poca confusión para distinguir cortes de COVID-19 con respecto a casos normales.

Los porcentajes de las métricas de precisión y exhaustividad son mejores que los resultados del modelo anterior. La precisión (ver *Tabla 4.11*), que nos permite analizar la calidad del modelo para predecir una categoría, muestra resultados casi perfectos en la categoría Normal y los porcentajes en la categoría COVID-19 (para cada partición del conjunto de datos) se encuentran arriba del 94%. En el caso de la exhaustividad (*recall*, *Tabla 4.10*), que nos ayuda a medir la cantidad de casos que se pueden identificar para una categoría específica, sus porcentajes están arriba del 97% lo cual es muy bueno.

Partición	COVID-19 %	Normal %
entrenamiento	99.4	98.9
validación	98.5	97.2
prueba	99.2	97.3

Tabla 4.10: *Recall* de la Componente Base en el conjunto de datos COVIDx CT-2. Únicamente se consideran las categorías COVID-19 y Normal.

Partición	COVID-19 %	Normal %
entrenamiento	99.6	98.6
validación	95.0	99.2
prueba	94.8	99.6

Tabla 4.11: *Precision* de la Componente Base en el conjunto de datos COVIDx CT-2. Únicamente se consideran las categorías COVID-19 y Normal.

Predicciones

Finalizamos esta sección incluyendo resultados de predicciones y visualización con Grad-CAM [33] de las imágenes que se utilizaron en la sección anterior, considerando únicamente las categorías COVID-19 y Normal. Las *Figura 4.24*, *Figura 4.25* y *Figura 4.26* corresponden a predicciones de imágenes clasificadas en la categoría *COVID-19*; mientras que, las *Figura 4.27*, *Figura 4.28* y *Figura 4.29* son predicciones de imágenes clasificadas en la categoría *Normal*.

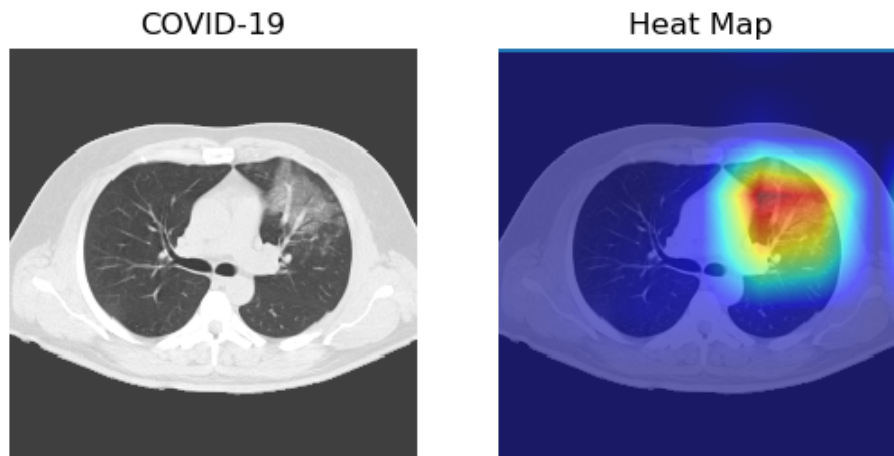


Figura 4.24: Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.8 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.

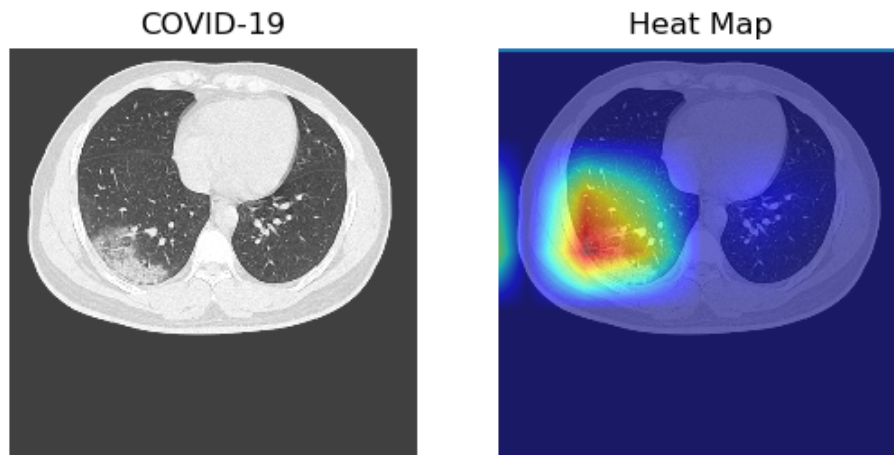


Figura 4.25: Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.9 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.

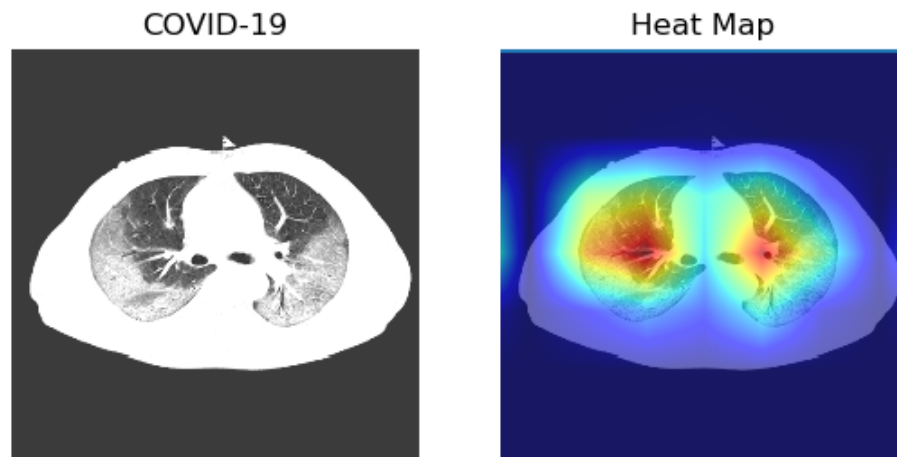


Figura 4.26: Grad-CAM de imagen CT categoría COVID-19. Se utilizó la imagen de la Figura 4.10 y el modelo retorna los resultados 0.9999 y 0.0001 para las clases COVID-19 y Normal respectivamente.

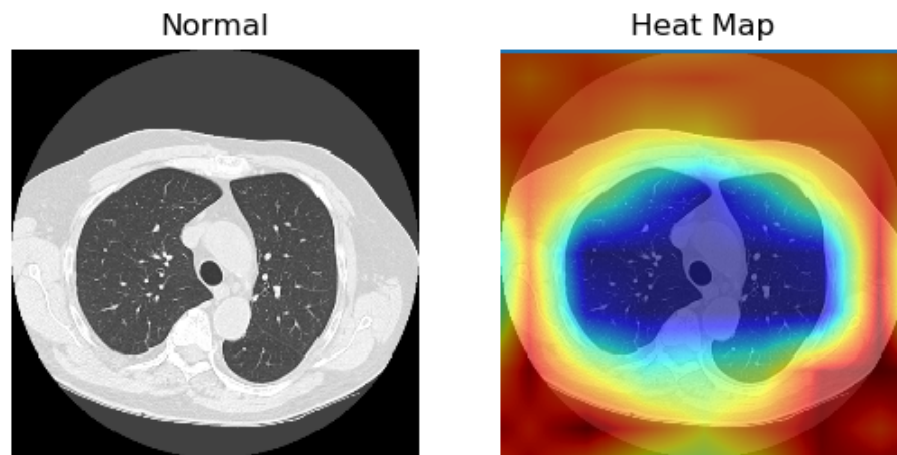


Figura 4.27: Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.14 y el modelo retorna los resultados 0.0003 y 0.9997 para las clases COVID-19 y Normal respectivamente.

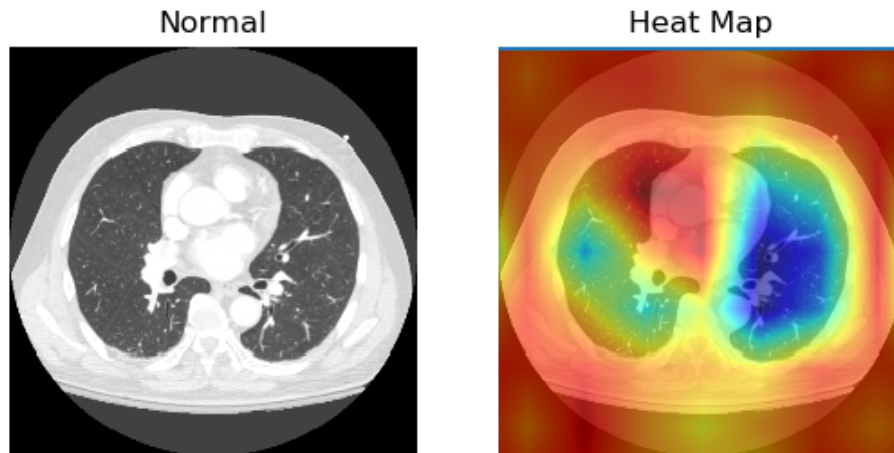


Figura 4.28: Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.15 y el modelo retorna los resultados 0.0002 y 0.9998 para las clases COVID-19 y Normal respectivamente.

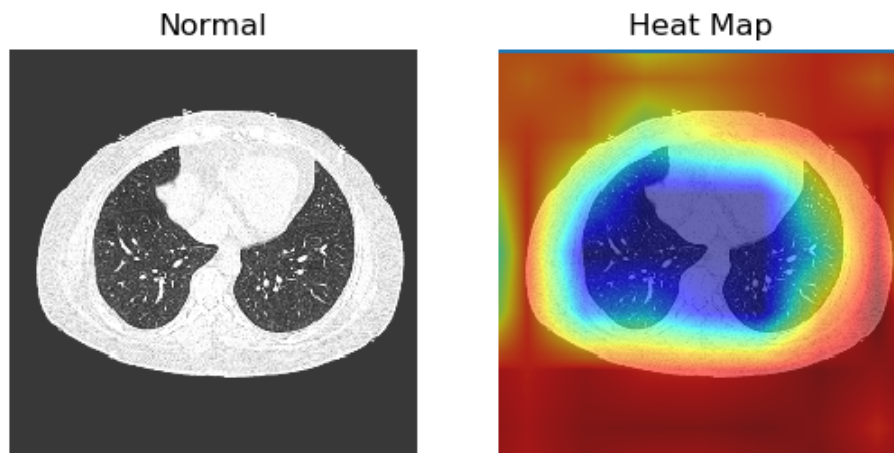


Figura 4.29: Grad-CAM de imagen CT categoría Normal. Se utilizó la imagen de la Figura 4.16 y el modelo retorna los resultados 0.0002 y 0.9998 para las clases COVID-19 y Normal respectivamente.

En este modelo, los resultados de predicción siguen siendo prometedores. Los mapas de calor en las imágenes de categoría COVID-19 muestran que el modelo trabaja de manera correcta al considerar como regiones relevantes los lugares donde se encuentran las lesiones causadas por neumonía COVID-19. En el caso de los cortes normales, las regiones relevantes abarcan gran parte de la imagen (casi en su totalidad) y esto nos indica que el modelo realiza una evaluación exhaustiva de la imagen para asignar una predicción de esta categoría.

4.2. Modelo 3D

En esta sección se presentan los resultados obtenidos con el modelo EffCOV-B0, en el conjunto de datos que se describe en la sección 3.2, el cual contiene volúmenes de los conjuntos de datos **COVID-19 CT-INER México** y **MosMedData**. Los modelos fueron desarrollados utilizando la librería *Keras* de *Tensorflow*, en el lenguaje de programación *Python*. La componente base que se utilizó en el desarrollo de EffCOV-B0 fue el modelo bidimensional que se entrenó para clasificar imágenes CT en las categorías COVID-19 y Normal (ver sección 4.1.2).

Puesto que en el preprocesado de datos se consideró dos tipos diferentes de segmentación de imagen, se desarrollaron dos modelos de redes utilizando la misma arquitectura. El primer modelo, al que llamaremos EffCOV-B0 SCH, se entrenó con los datos preprocesados con segmentación del cuerpo humano, mientras que el segundo modelo, al cual llamaremos EffCOV-B0 SP, fue entrenado con los datos preprocesados con segmentación pulmonar. Las secciones 4.2.3 y 4.2.4 contienen la información de todos los resultados obtenidos con estos dos modelos, respectivamente.

4.2.1. Partición de los Conjuntos de Datos

Basado en los entrenamientos iniciales del modelo EffCOV-B0, se comprobó que el modelo no estaba realizando la clasificación correcta de la categoría *CT-1* la cual corresponde a casos de COVID-19 en el conjunto de datos *MosMedData*. Los volúmenes de dicha categoría muestran una afectación de la parénquima pulmonar inferior al 25 % y los pacientes que recibieron este diagnostico fueron remitidos a sus casa con seguimiento mediante tecnologías de telemedicina [14]. Dado que dichos volúmenes no representan casos graves de COVID-19, y que la mayoría de los cortes no contienen lesiones causadas por la enfermedad, se decidió retirar los volúmenes correspondientes a la categoría *CT-1*. La *Tabla 4.12* muestra la distribución de los conjuntos de datos (CT-INER y MosMedData) para los subconjuntos de entrenamiento, validación y prueba.

Partición	CT-0	CT-2	CT-3	CT-4	INER	Total
entrenamiento	152	75	27	0	80	334
validación	50	25	9	0	40	124
prueba	52	25	9	2	80	168

Tabla 4.12: Partición de los conjuntos de datos *COVID-19 CT-INER México* y *MosMedData*. Las categorías *CT-2*, *CT-3*, *CT-4* e *INER* corresponden a volúmenes clasificados en la categoría COVID-19, mientras que los datos en la categoría *CT-0* corresponden a volúmenes Normales.

4.2.2. Entrenamiento

El entrenamiento del modelo EffCOV-B0 se realizó en dos etapas. La primer etapa corresponde al entrenamiento de la componente base; la cual se realizó con 15 épocas, utilizando el conjunto de datos COVIDx CT-2 y cuyos resultados se muestran en la sección 4.1.2. La segunda etapa del entrenamiento se realizó para el complemento 3D; definimos una estrategia de entrenamiento que combina las técnicas de transferencia de aprendizaje y ajuste fino, utilizando las mismas métricas de desempeño que se utilizaron en la Componente Base (pérdida, exactitud, funciones F1 y AUC).

El entrenamiento del complemento 3D se desarrolló en dos fases. La primera fase se realizó implementando la técnica de transferencia de aprendizaje (tomando la componente base previamente entrenada para clasificar imágenes 2D en categorías COVID-19 y Normal) con 500 épocas e inhabilitando el entrenamiento de la componente base. La segunda fase del entrenamiento se llevó a cabo con un ajuste completo de 100 épocas, permitiendo el entrenamiento en la Componente base. La *Tabla 4.13* contiene información sobre los tiempos de entrenamiento-validación (para cada fase) y evaluación del conjunto de prueba de cada modelo.

Modelo	Épocas	Tiempo (horas) entrenamiento-val.	Tiempo (segundos) prueba
EffCOV-B0 SCH <i>fase1</i>	500	2.0	—
EffCOV-B0 SCH <i>fase2</i>	100	12.87 (sin GPU)	15.2
EffCOV-B0 SP <i>fase1</i>	500	2.1	—
EffCOV-B0 SP <i>fase2</i>	100	12.86 (sin GPU)	15.2

Tabla 4.13: Tiempo de entrenamiento-validación y prueba de los modelos tridimensionales.

La descripción del hardware utilizado (memoria, procesador y GPU) para desarrollar los modelos de redes neuronales se encuentra en el *Apéndice A.1*. Por limitaciones computacionales, la segunda fase del entrenamiento se llevó a cabo sin GPU.

4.2.3. Resultados con Segmentación del Cuerpo Humano

Evaluaciones

Los resultados que se muestran en esta sección siguen el mismo esquema de los resultados presentados en los modelos bidimensionales y corresponden al modelo que se entrenó con datos preprocesados con segmentación del cuerpo humano. Las *Tabla 4.14* y *Tabla 4.15* contienen los resultados de la función de pérdida y las métricas de exactitud, F1 y AUC-ROC para cada fase de entrenamiento. También se consideran los resultados de las métricas *Recall* y *Precision*, los cuales se encuentran en las

Tabla 4.16 y Tabla 4.17, respectivamente.

La estrategia de entrenamiento que se implementó nos permite observar una mejora en los resultados obtenidos de cada fase entrenamiento. Se puede observar que la exactitud mejora, más del 10 %, en la partición de entrenamiento; y casi el 3 % en la partición de prueba. Sin embargo, también se puede observar un sobreajuste de aprendizaje al comparar los resultados de exactitud en la Tabla 4.15.

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.513	87.0	87.0	93.0
validación	0.467	95.0	95.0	98.0
prueba	0.493	91.7	91.7	95.6

Tabla 4.14: Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la primera fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación del cuerpo del humano

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.309	98.5	98.5	99.9
validación	0.354	94.2	94.2	99.4
prueba	0.376	94.0	94.0	97.5

Tabla 4.15: Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la segunda fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación del cuerpo del humano

Partición	COVID-19 %	Normal %
entrenamiento	98.4	98.7
validación	94.6	96.0
prueba	93.1	96.2

Tabla 4.16: *Recall* del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

Partición	COVID-19 %	Normal %
entrenamiento	98.9	98.0
validación	97.2	92.3
prueba	98.2	86.2

Tabla 4.17: *Precision* del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

Gráficas

Al igual que en los modelos bidimensionales; se registró el resultado de las métrica de desempeño (función de pérdida, exactitud, función F1 y AUC-ROC), en cada época, para visualizar el aprendizaje en la etapa de entrenamiento. En este caso, dada la estrategia de entrenamiento que se llevó a cabo, la cual consiste en dos fases, se dividieron las gráficas en dos categorías. La primer categoría contiene las *Figura 4.30*, *Figura 4.31*, *Figura 4.32* y *Figura 4.33*, y corresponden de la primera fase de entrenamiento. Mientras que las *Figura 4.34*, *Figura 4.35*, *Figura 4.36* y *Figura 4.37* se encuentran en la segunda categoría las cuales corresponden a los resultados de la segunda fase del entrenamiento.

Dada la limitante del pequeño tamaño del conjunto de datos, se puede observar un comportamiento oscilatorio en cada una de las gráficas. Dicho comportamiento es más notable en la primera fase del entrenamiento, la cantidad de parámetros habilitados para el entrenamiento también es causa relevante de este comportamiento. Las gráficas de la segunda fase del entrenamiento muestran un comportamiento similar a las gráficas de los modelos bidimensionales. No obstante, también se observan oscilaciones pero con menor periodicidad que en la primera fase. Cabe mencionar que en la segunda fase de entrenamiento, todos los parámetros del modelo están habilitados para ser entrenados.

Gráficas primera fase de entrenamiento

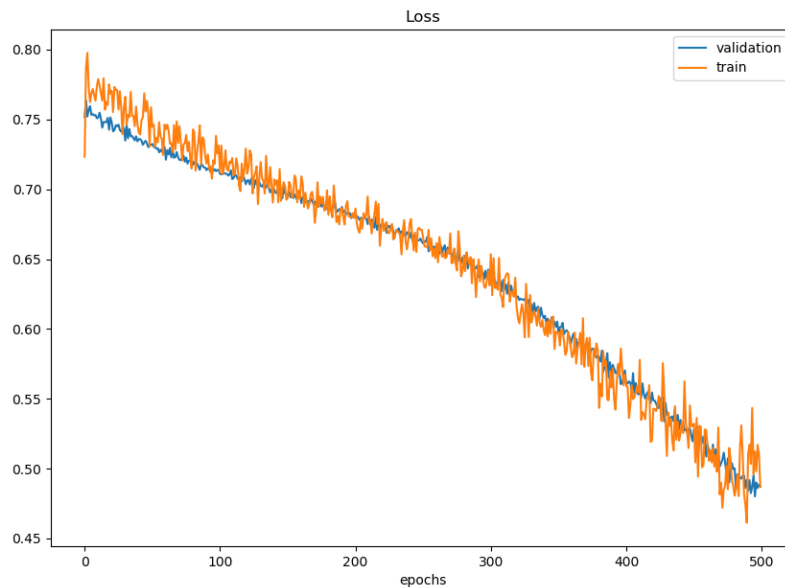


Figura 4.30: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

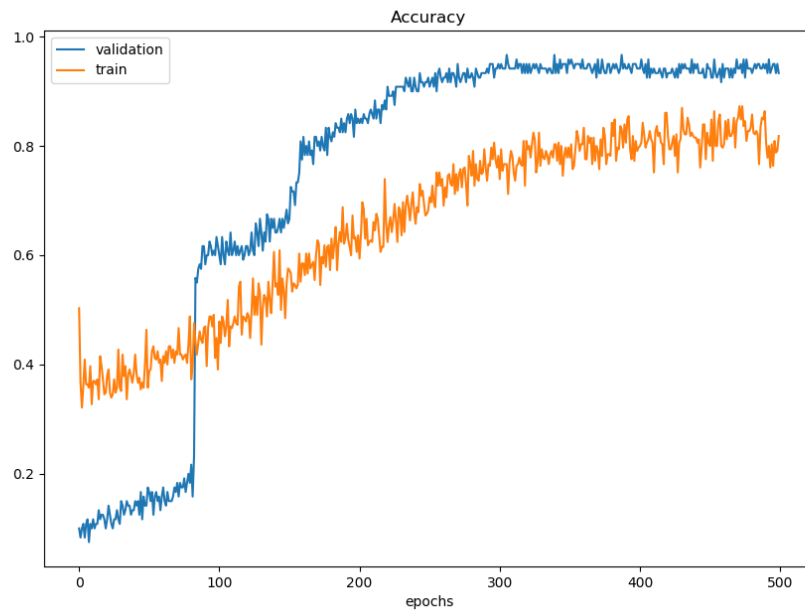


Figura 4.31: Gráfica de la exactitud con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

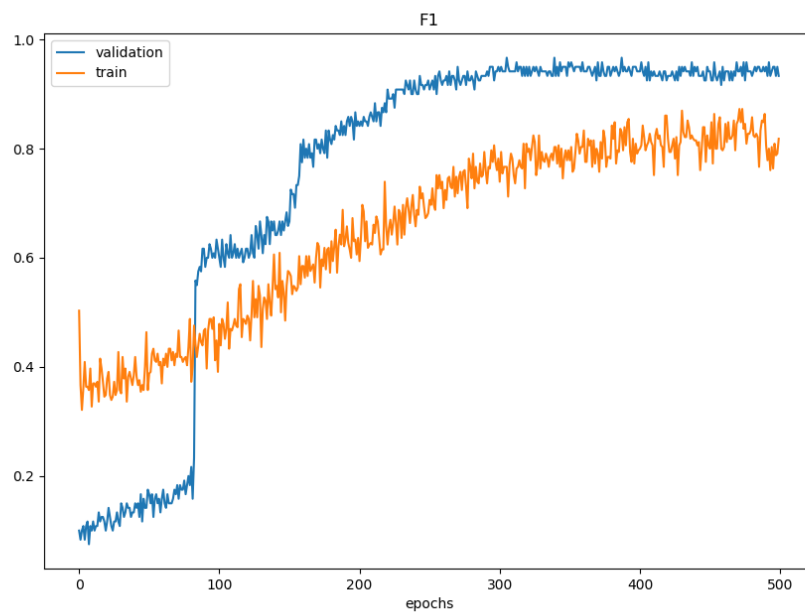


Figura 4.32: Gráfica de la función F1 con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

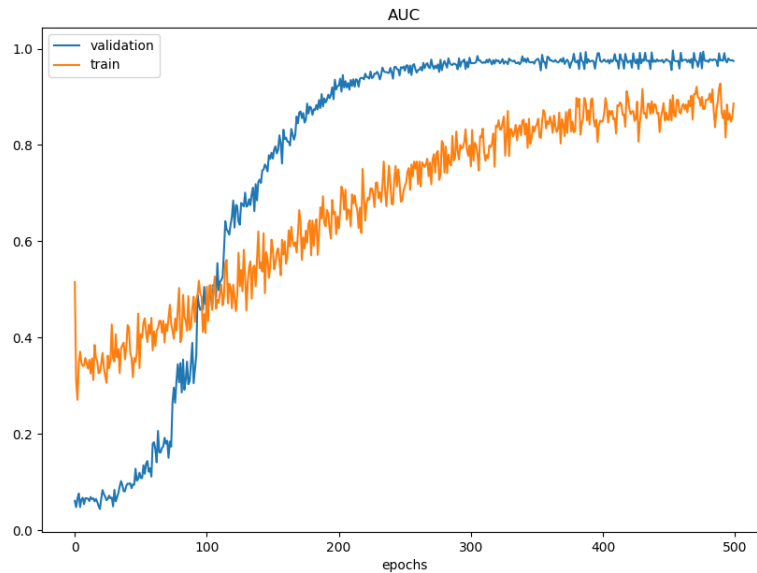


Figura 4.33: Gráfica de la métrica AUC-ROC con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

Gráficas segunda fase de entrenamiento

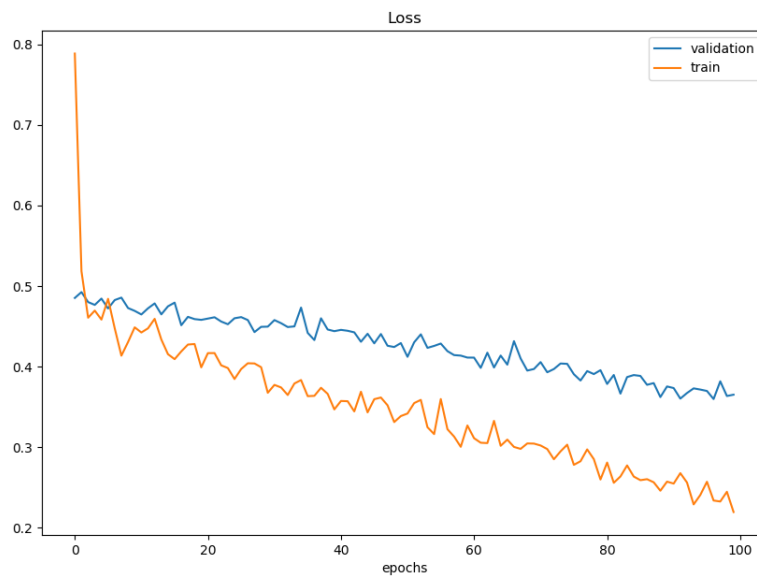


Figura 4.34: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

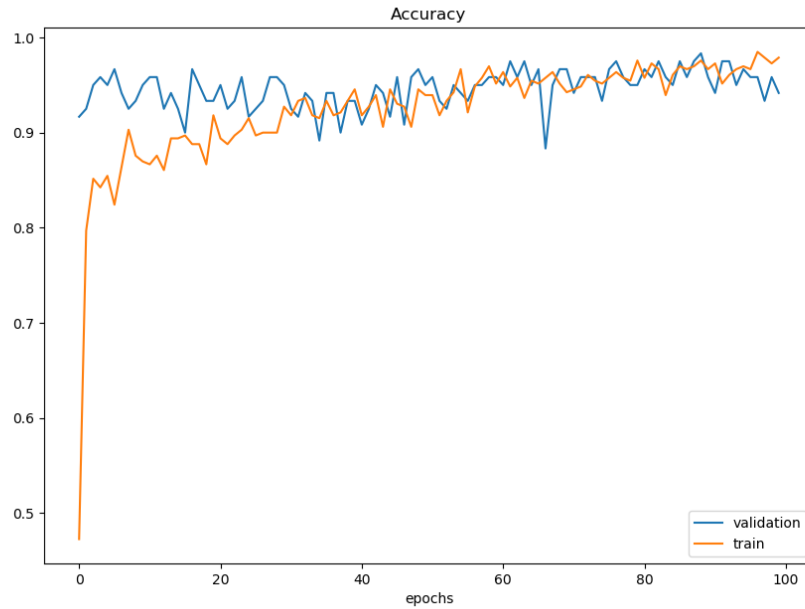


Figura 4.35: Gráfica de la exactitud con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

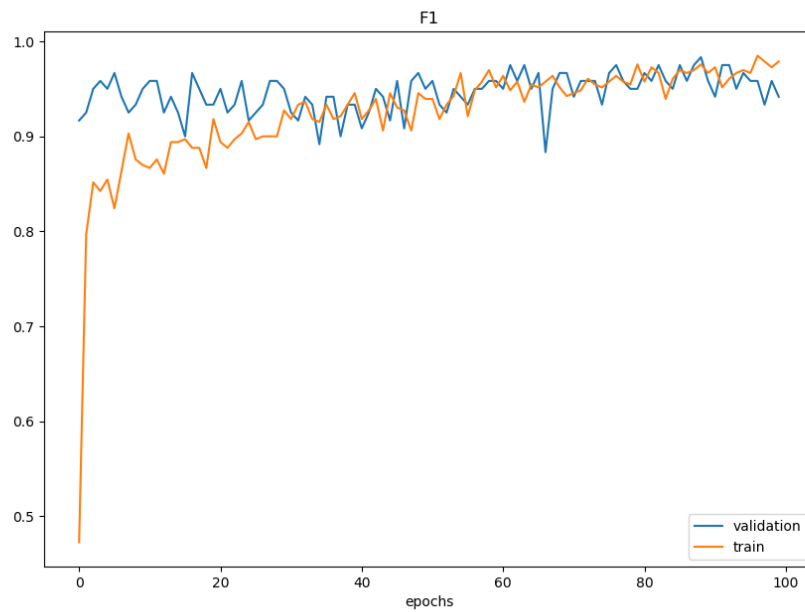


Figura 4.36: Gráfica de la función F1 con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

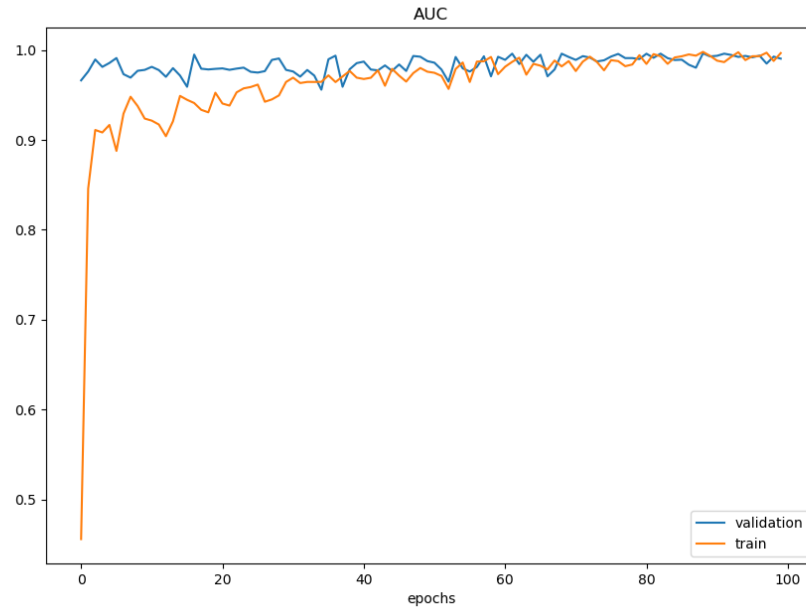


Figura 4.37: Gráfica de la métrica AUC con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

Matrices de Confusión

Finalizamos los resultados de esta sección con las matrices de confusión para cada partición del conjunto de datos. Las *Figura 4.38*, *Figura 4.39* y *Figura 4.40* muestran dichas matrices, para los subconjuntos de prueba, entrenamiento y validación, respectivamente. En cada una de las matrices, el eje vertical representa los valores de predicción y el eje horizontal representa los valores reales.

Se puede observar la certeza del modelo, en cuanto a la predicción, al verificar que el mayor valor de cada casilla de la matriz de confusión se encuentra en la diagonal principal (para cada partición del conjunto de datos). No obstante, en cada una de las gráficas, se observa que la mayor confusión del modelo se encuentra al predecir la categoría Normal, siendo mínima en el conjunto de validación pero extendiéndose notablemente en el conjunto de prueba.

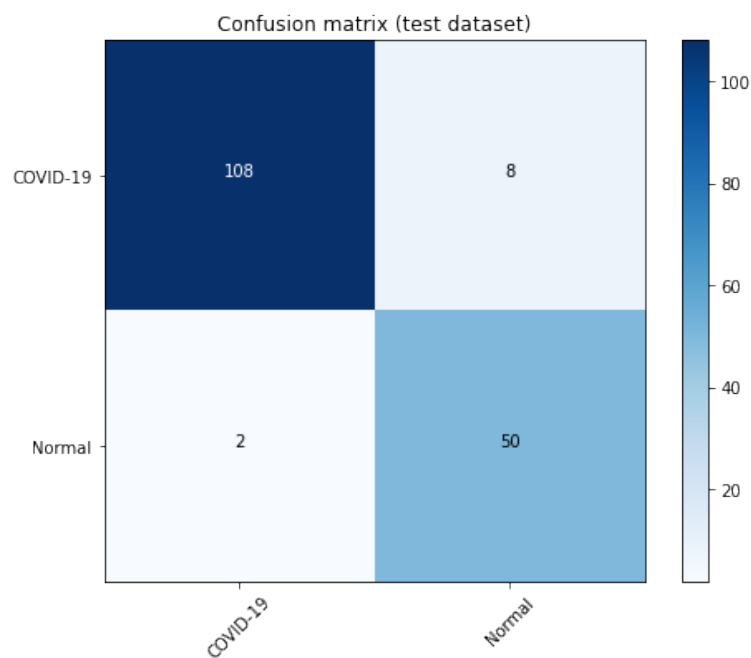


Figura 4.38: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de prueba. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

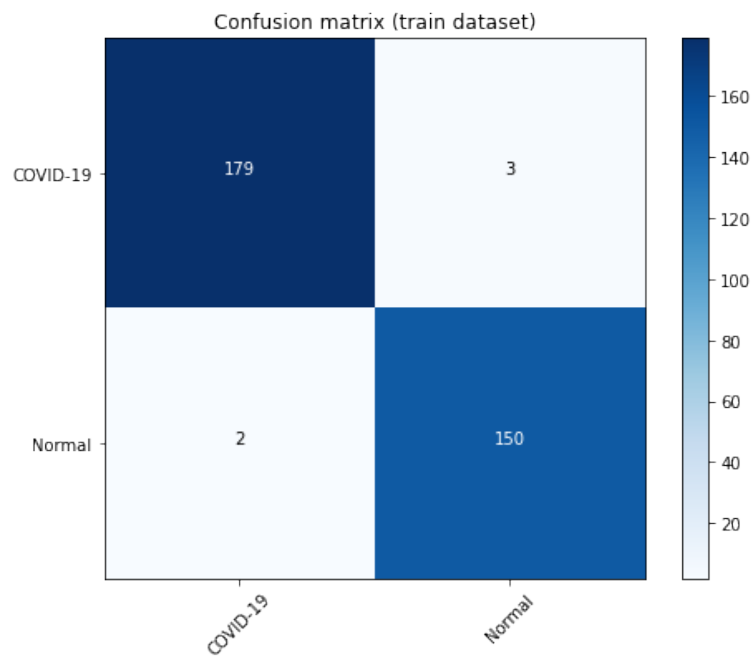


Figura 4.39: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de entrenamiento. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

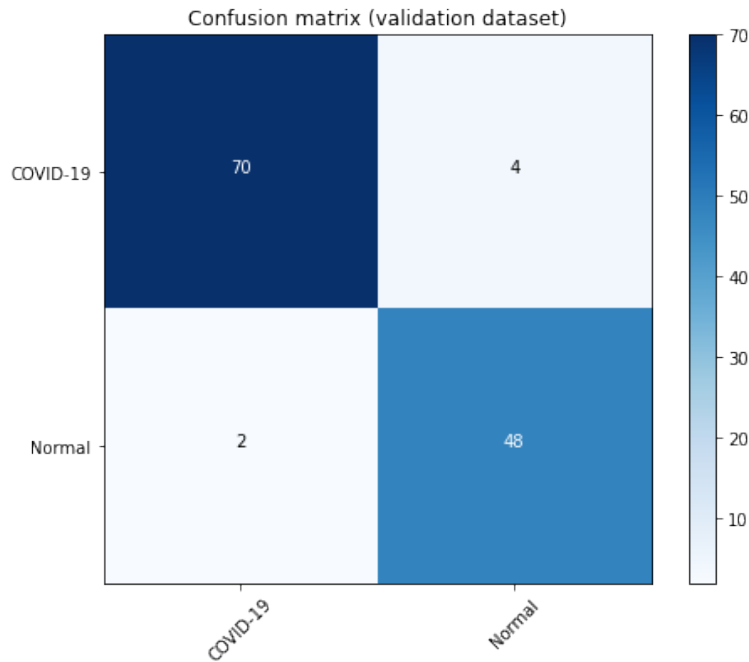


Figura 4.40: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de validación. El conjunto de datos tiene preprocesado de segmentación del cuerpo del humano.

4.2.4. Resultados con Segmentación Pulmonar

Evaluaciones

En esta sección se muestran los resultados del modelo entrenado con datos preprocesados con segmentación pulmonar. Las *Tabla 4.18* y *Tabla 4.19* contienen los resultados de la función de pérdida y las métricas de exactitud, F1 y AUC-ROC para cada fase de entrenamiento. Mientras que en las *Tabla 4.20* y *Tabla 4.21* se observan los resultados de las métricas *Recall* y *Precision*, respectivamente. Al igual que en los resultados del modelo anterior, se observa una mejora en cada fase entrenamiento, aunque dichos resultados no superan los resultados del modelo anterior. Se puede observar que la exactitud mejora, aproximadamente 9.5%, en la partición de entrenamiento; y casi el 6% en la partición de prueba. También se puede observar un sobreajuste de aprendizaje al comparar los resultados de exactitud en la *Tabla 4.19*.

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.632	77.9	77.9	85.0
validación	0.629	78.3	78.3	87.7
prueba	0.605	84.0	84.0	92.7

Tabla 4.18: Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la primera fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación pulmonar.

Partición	Loss	Accuracy %	F1 %	AUC-ROC %
entrenamiento	0.497	96.4	96.4	99.1
validación	0.496	95.0	95.0	99.3
prueba	0.509	89.9	89.9	97.0

Tabla 4.19: Pérdida, exactitud, funciones F1 y AUC-ROC del Complemento 3D. Los resultados corresponden a la segunda fase del entrenamiento y el conjunto de datos tiene preprocesado de segmentación pulmonar.

Partición	COVID-19 %	Normal %
entrenamiento	96.7	96.1
validación	93.2	92.0
prueba	93.2	82.7

Tabla 4.20: *Recall* del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación pulmonar.

Partición	COVID-19 %	Normal %
entrenamiento	96.7	96.1
validación	94.5	90.2
prueba	92.4	84.3

Tabla 4.21: *Precision* del Complemento 3D. El conjunto de datos tiene preprocesado de segmentación pulmonar.

Gráficas

De la misma forma que en el modelo anterior, se registró el resultado de las métricas de desempeño para visualizar el aprendizaje en la etapa de entrenamiento (en cada época y para cada fase). Las *Figura 4.41*, *Figura 4.42*, *Figura 4.43* y *Figura 4.44* corresponden de la primera fase de entrenamiento; mientras que las *Figura 4.45*, *Figura 4.46*, *Figura 4.47* y *Figura 4.48* corresponden a los resultados de la segunda fase del entrenamiento.

El comportamiento oscilatorio en cada una de las gráficas se observa con mayor frecuencia con respecto a las gráficas del modelo anterior. Dicho comportamiento también es más notable en la primera fase del entrenamiento, donde además de las oscilaciones se observa un salto brusco, en las gráficas de exactitud y función F1 (en los datos de validación), a un tercio del número de épocas. La diferencia del tipo de segmentación de datos en el entrenamiento de la componente base y el pequeño tamaño del conjunto de datos son factores fundamentales de este extraño comportamiento. Las gráficas de la segunda fase del entrenamiento muestran un comportamiento similar a las gráficas del modelo anterior, también se observan oscilaciones con menor periodicidad que en la primera fase.

Gráficas primera fase de entrenamiento

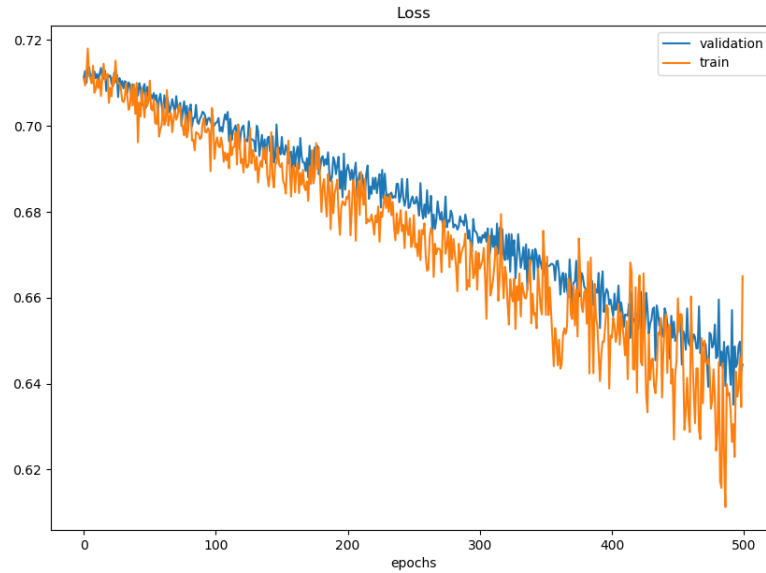


Figura 4.41: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

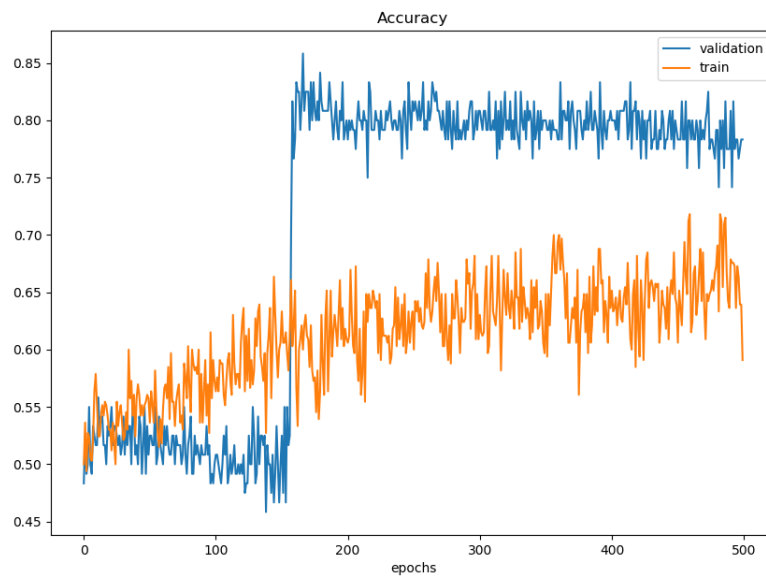


Figura 4.42: Gráfica de la exactitud con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

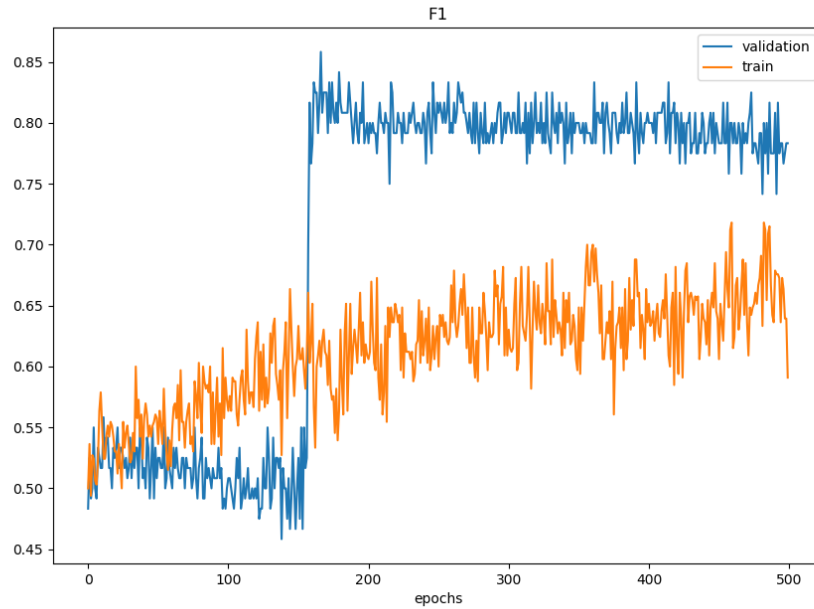


Figura 4.43: Gráfica de la función F1 con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

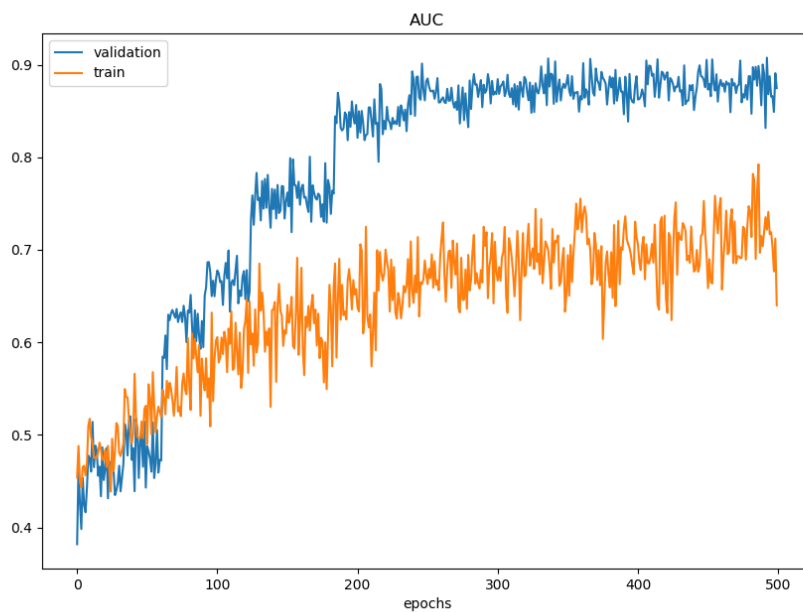


Figura 4.44: Gráfica de la métrica AUC con respecto al número de época del modelo 3D en la primera fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

Gráficas segunda fase de entrenamiento

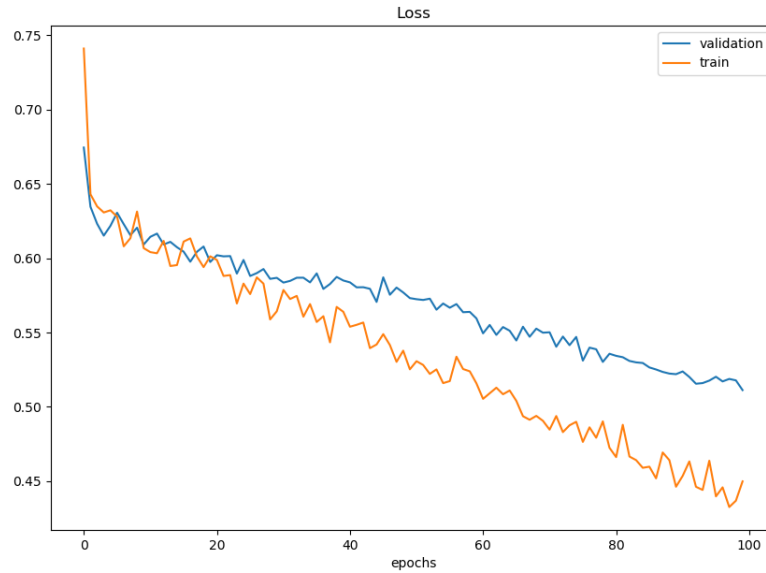


Figura 4.45: Gráfica de la pérdida (sección 2.3.2) con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

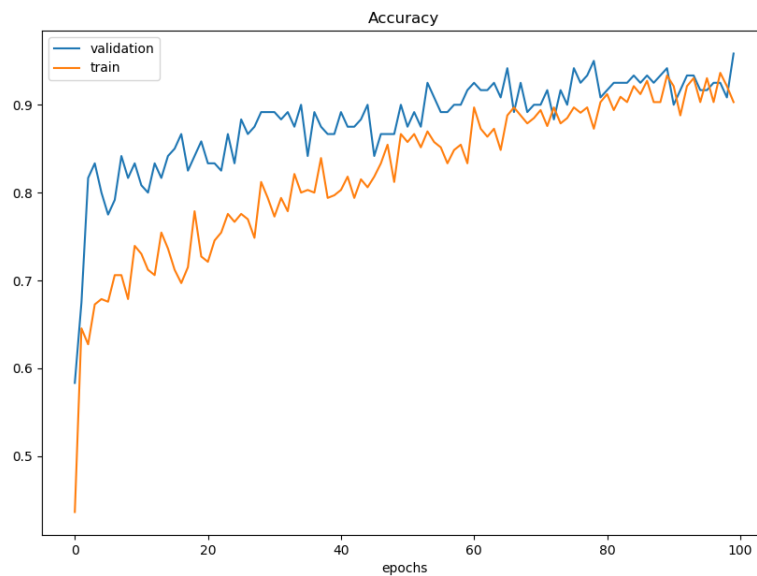


Figura 4.46: Gráfica de la exactitud con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

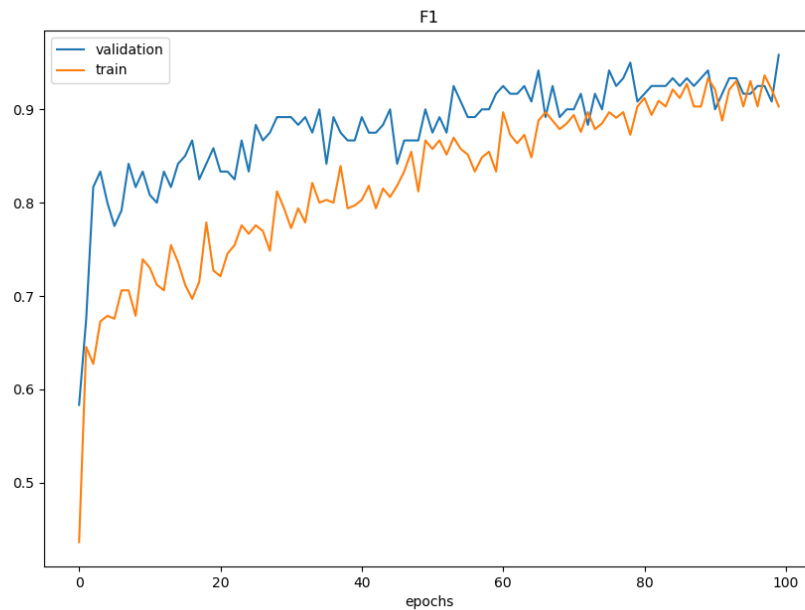


Figura 4.47: Gráfica de la función F1 con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

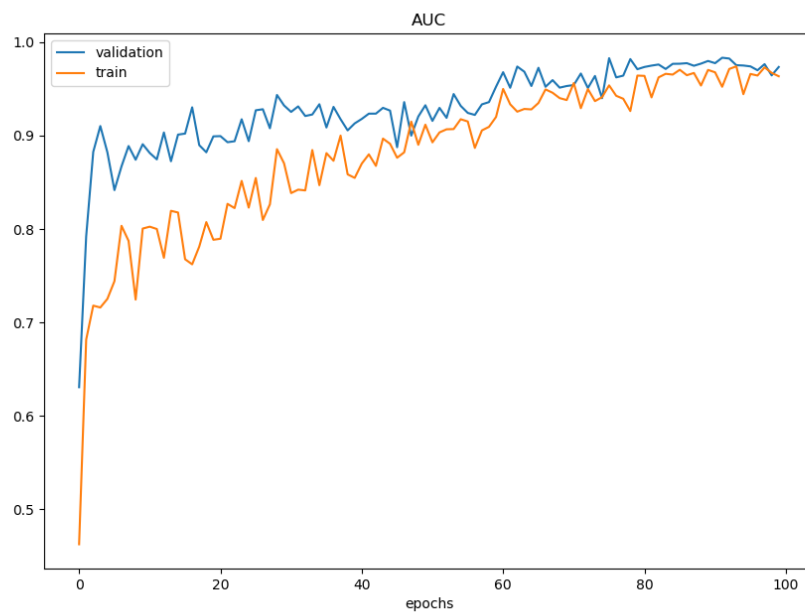


Figura 4.48: Gráfica de la métrica AUC-ROC con respecto al número de época del modelo 3D en la segunda fase del entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

Matrices de Confusión

Finalmente, se muestran las matrices de confusión para cada partición del conjunto de datos. Las *Figura 4.49*, *Figura 4.50* y *Figura 4.51* muestran dichas matrices; para los subconjuntos de prueba, entrenamiento y validación, respectivamente. En cada una de las matrices, el eje vertical representa los valores de predicción y el eje horizontal representa los valores reales.

La cantidad de falsos positivos y falsos negativos es casi balanceada pero se observa con mayor frecuencia con respecto al modelo anterior. No obstante, aunque este modelo no supera los resultados del modelo anterior, la mayor parte de la predicción es correcta (ver valores en la diagonal principal de cada matriz).

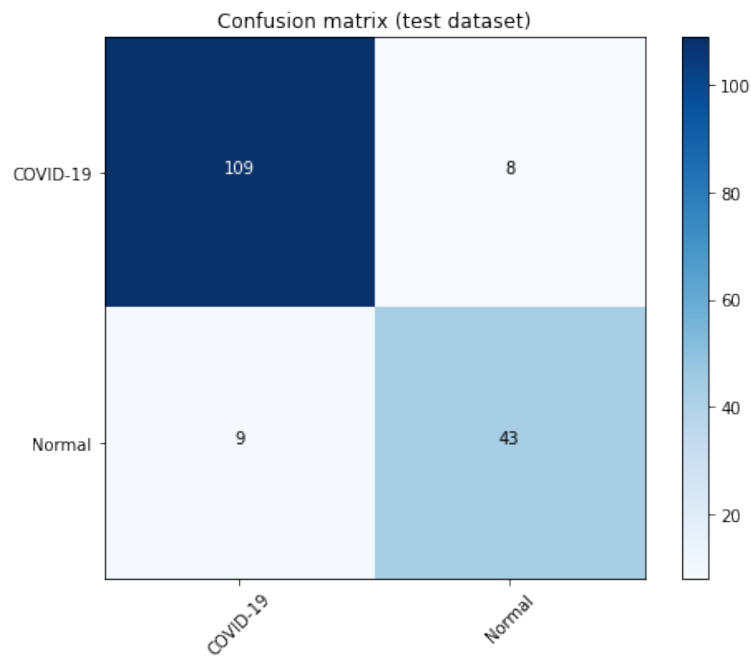


Figura 4.49: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de prueba. El conjunto de datos tiene preprocesado de segmentación pulmonar.

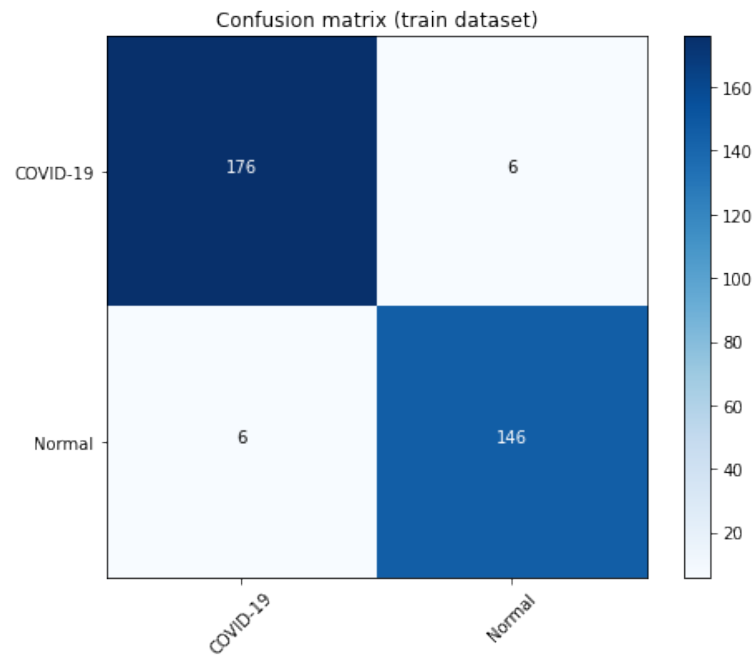


Figura 4.50: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de entrenamiento. El conjunto de datos tiene preprocesado de segmentación pulmonar.

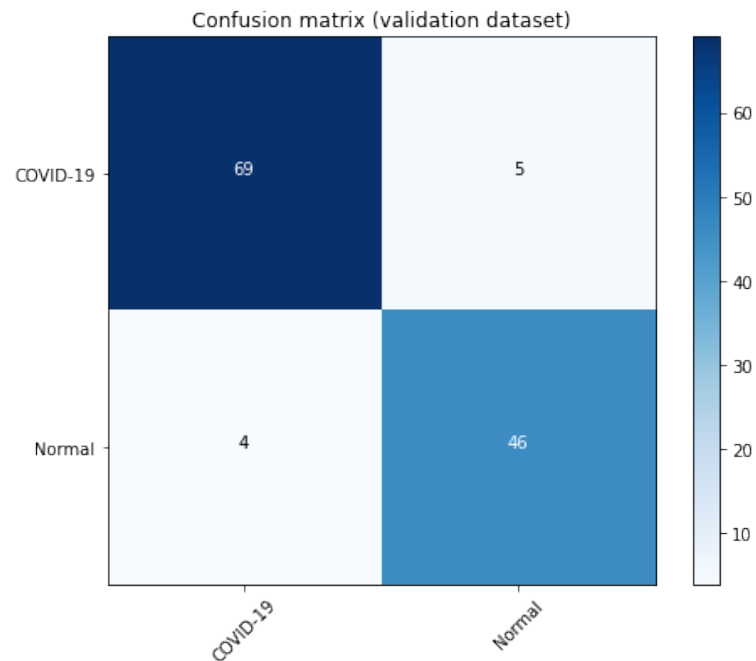


Figura 4.51: Gráfica de la matriz de confusión del Complemento 3D en el subconjunto de validación. El conjunto de datos tiene preprocesado de segmentación pulmonar.

4.2.5. Evaluaciones de los conjuntos de datos

Como hemos explicado en la *Sección 4.2.1*, el conjunto de datos, con el cual se han entrenado los modelos tridimensionales, está compuesto por los conjuntos MosMedData de Rusia [14] y COVID-19 CT-INER de México. También se mencionó que los datos correspondientes a la categoría CT-1 COVID-19, del conjunto MosMedData, se descartó del proceso de entrenamiento y validación, dado que la mayoría de los cortes no contienen lesiones causadas por la enfermedad. Aunque ya se han mostrado todos los resultados de los dos modelos tridimensionales realizados en este trabajo, en esta sección presentamos un resumen, de manera individual para conjunto de datos, de los resultados obtenidos con el modelo EffCOV-B0 y cuyos datos fueron preprocesados con segmentación del cuerpo humano (al cual llamaremos EffCOV-B0 SCH).

En la *Tabla 4.22* se resume los resultados del modelo EffCOV-B0 SCH. Las categorías CT-0, CT-1, CT-2, CT-3 y CT-4 corresponden al conjunto de datos MosMedData, mientras que la última categoría corresponde al conjunto COVID-19 CT-INER. Las columnas COVID-19 y Normal indican la clasificación de los datos para cada categoría, los datos CT-0 corresponden a volúmenes CT de pacientes sanos, mientras que el resto de las categorías contienen volúmenes de pacientes que fueron diagnosticados con COVID-19. Las últimas dos columnas, Predicción COVID-19 y Predicción Normal, contienen los resultados de las predicciones del modelo EffCOV-B0 SCH para cada categoría.

Categoría datos	COVID-19	Normal	Predicción COVID-19	Predicción Normal
CT-0	0	254	6	248
CT-1	684	0	307	377
CT-2	125	0	114	11
CT-3	45	0	41	4
CT-4	2	0	2	0
INER	200	0	200	0

Tabla 4.22: Resumen de resultados del modelo EffCOV-B0 (los datos tienen preprocesado de segmentación del cuerpo humano). Las categorías CT-0, CT-1, CT-2, CT-3 y CT-4 corresponden al conjunto de datos MosMedData.

Es importante analizar el comportamiento de EffCOV-B0 SCH en estos dos conjuntos de datos. La *Tabla 4.22* nos permite observar la perfecta exactitud de EffCOV-B0 SCH en el conjunto COVID-19 CT-INER y en la categoría CT-4 de MosMedData. También se puede ver la confusión que se produce al predecir la categoría CT-1 de COVID-19. Para el resto de las categorías de MosMedData, el número de predicciones erróneas es pequeño.

4.3. Comparación con otros modelos

En esta última sección se discuten los resultados obtenidos en este trabajo y se comparan con algunos resultados del estado del arte. Se consideraron las métricas de desempeño *accuracy*, *F1* y *AUC-ROC* dado que fueron las métricas más comunes en la mayoría de los trabajos estudiados. Adicionalmente se muestran resultados de las métricas *recall* y *precision*, de las categorías COVID-19 y Normal, de los modelos presentados en este trabajo (bidimensional y tridimensional) y se comparan con el modelo COVIDNet-CT de Gunraj et al. [11], puesto que dicho trabajo fue la base en el desarrollo del modelo bidimensional.

Cabe mencionar que no todos los modelos estudiados resuelven la tarea de clasificación binaria en las imágenes CT. En los modelos bidimensionales, el trabajo de Gunraj et al. [11] considera la categoría *Neumonía*, adicional a las categorías de COVID-19 y Normal. En el caso de los modelos tridimensionales, el trabajo realizado por Aboutalebi et al. [19] se limita a clasificar niveles de severidad de COVID-19 ignorando por completo las exploraciones normales; mientras que el modelo presentado por Rohila et al. [12], además de resolver la tarea de clasificación binaria, también clasifica los niveles de severidad de COVID-19.

Otro detalle a considerar en los resultados que se muestran en las siguientes tablas es la diversidad de los conjuntos de datos y la variabilidad en la distribución de datos en las particiones de entrenamiento, validación y prueba. En los resultados de la *Tabla 4.23*, cada uno de los modelos estudiados se desarrollaron con distintos conjuntos de datos. Los modelos EffCOV-B0 CB2 (componente base; clases COVID-19 y Normal) y EffCOV-B0 CB3 (componente base; clases COVID-19, Neumonía y Normal) fueron desarrollados con el mismo conjunto de datos del modelo COVIDNet-CT, también se utilizó la misma distribución de datos que ellos proponen (descartando la clase Neumonía para en modelo EffCOV-B0 CB2).

Modelo	Accuracy %	F1 %	AUC-ROC %	Dataset
Self-Trans [15]	86	85	94	COVID-CT
xDNN [16]	97.4	97.3	97.4	SARS-CoV-2
COVID-NET X-Ray [17]	91.8	92.2	96.6	COVID-CT + SARS-CoV-2
COVIDNet-CT [11]	98.1	—	—	COVIDx CT-2
EffCOV-B0 CB2	98.3	98.3	98.9	COVIDx CT-2
EffCOV-B0 CB3	96.4	96.4	99.2	COVIDx CT-2

Tabla 4.23: Resultados de modelos bidimensionales.

En los resultados de la *Tabla 4.24*, los modelos ReCOV-101 y Resnet50MV comparten el mismo conjunto de datos (MosMedData [14]) pero difieren en su distribución debido a las distintas tareas que resuelven cada modelo. Para los

modelos EffCOV-B0 SCH (segmentación del cuerpo humano en los datos) y EffCOV-B0 SP (segmentación pulmonar en los datos) también se consideró dicho conjunto de datos, la distribución también difiere en los modelos mencionados debido a la inclusión del conjunto de datos CT-INER México. Cabe mencionar que para ambos modelos se consideró la componente base CB2, la principal motivación de esta decisión radica en los buenos resultados que mostró dicha red en los experimentos.

Modelo	Accuracy %	F1 %	AUC-ROC %	Dataset
COVID-NET CT-S [19]	78.5	—	—	CNCB
ReCOV-101 [12]	94.9	—	—	MosMedData
Resnet50+MV [20]	84	—	96	MosMedData
EffCOV-B0 SP (CB2)	89.9	89.9	97.0	MosMedData + CT-INER MX
EffCOV-B0 SCH (CB2)	94.0	94.0	97.5	MosMedData + CT-INER MX

Tabla 4.24: Resultados de modelos tridimensionales.

Considerando cada uno de los factores mencionados, es difícil definir cuál es el modelo que presenta los mejores resultados. En la *Tabla 4.23* se puede observar que EffCOV-B0 CB2 es el modelo que muestra los mejores resultados en las métricas de desempeño. Sin embargo, EffCOV-B0 CB2 se limita a clasificar dos categorías de imágenes CT, mientras que el modelo COVIDNet-CT casi obtiene el mismo resultado de exactitud y, adicionalmente, considera la categoría neumonía junto con las categorías COVID-19 y normal. Con respecto a los modelos tridimensionales (ver *Tabla 4.24*), aunque no se logra superar los resultados de todos los modelos estudiados, EffCOV-B0 SCH muestra resultados de altura y su exactitud apenas está por debajo del modelo ReCOV-101.

Finalmente, considerando que nuestros modelos bidimensionales fueron entrenados con el mismo conjunto de datos del modelo COVIDNet-CT. En las *Tabla 4.25* y *Tabla 4.26* se muestran resultados de la métricas *recall* y *precision* de los modelos presentados en este trabajo (bidimensionales y tridimensionales), y se comparan con los resultados del modelo COVIDNet-CT de Gunraj et al. [11]. En dichas tablas podemos observar que los modelos desarrollados en este trabajo muestran resultados a la altura de COVIDNet-CT.

Modelo	COVID-19 %	Normal %	Neumonía %
COVIDNet-CT [11]	96.2	99.0	98.2
EffCOV-B0 CB2	99.2	97.3	—
EffCOV-B0 CB3	92.9	97.1	98.0
EffCOV-B0 SCH (CB2)	93.1	96.2	—
EffCOV-B0 SP (CB2)	93.2	82.7	—

Tabla 4.25: Comparación de resultados de la métrica Recall.

Modelo	COVID-19 %	Normal %	Neumonía %
COVIDNet-CT [11]	96.7	99.4	98.2
EffCOV-B0 CB2	94.8	99.6	—
EffCOV-B0 CB3	94.3	98.7	97.2
EffCOV-B0 SCH (CB2)	98.2	86.2	—
EffCOV-B0 SP (CB2)	92.4	84.3	—

Tabla 4.26: Comparación de resultados de la métrica Precision.

Capítulo 5

Conclusiones y trabajo futuro

En este trabajo, el modelo de red neuronal convolucional EffCOV-B0 se ha presentado y estudiado para diagnosticar COVID-19 en volúmenes de tomografías computarizadas. Se utilizó la red EfficientNet B0 para desarrollar nuestro modelo base bidimensional, el cual fue entrenado (con el conjunto de datos COVIDx CT-2) para detectar COVID-19 en imágenes de CT. Con el modelo base se desarrolló una red tridimensional para clasificar volúmenes de pacientes diagnosticados con COVID-19 así como volúmenes de pacientes sanos. Los resultados obtenidos muestran que el modelo propuesto tiene un rendimiento comparable con el estado del arte.

El principal objetivo de este trabajo es detectar COVID-19 en volúmenes de tomografías computarizadas. Actualmente, la mayoría de los trabajos realizados proponen modelos bidimensionales que no son útiles en el análisis de volúmenes CT. Los pocos modelos tridimensionales que se han desarrollados, en su mayoría, no muestran buenos resultados en sus experimentos. Esto puede ser debido a los pequeños tamaños en los conjuntos de datos de los cuales se puede tener acceso. En este trabajo se presentó el modelo de red neuronal convolucional 3D, EffCOV-B0, para detectar COVID-19 en volúmenes CT. En los experimentos realizados se observa que EffCOV-B0 tiene un desempeño a la altura del estado del arte y puede ser una herramienta de mucha utilidad para los profesionales encargados del diagnóstico temprano de dicha enfermedad.

Para resolver la limitante del pequeño tamaño en el conjunto de datos, el método utilizado para desarrollar el modelo de red neuronal 3D combina la técnica de transferencia de aprendizaje y el uso de una red bidimensional como base. Se construyó una red base CNN-2D utilizando el modelo EfficientNet B0 preentrenado en ImageNet como columna vertebral y se entrenó por completo con el conjunto de datos COVIDx CT-2 para clasificar imágenes CT de COVID-19 y cortes normales. Implementando la técnica de transferencia de aprendizaje, se utilizó el modelo base para adaptarlo a una red 3D la cual fue entrenada en un conjunto de datos que combina volúmenes de las base de datos ModMedData y COVID-19 CT-INER México. Después de aplicar transferencia de aprendizaje, se realizó una segunda fase

de entrenamiento en la cual se incluyeron las capas del modelo base (esta segunda fase de entrenamiento se denominó *ajuste completo*).

Para mejorar la robustez del modelo se implementó una sólida línea de preprocesamiento de datos. El trabajo realizado en el preprocesamiento de los datos incluye dos métodos avanzados de segmentación de imágenes. El primer método que se aplicó fue la segmentación del cuerpo humano y se llevó a cabo utilizando funciones de la librería OpenCV de Python, este método es una extensión a volúmenes CT del método que desarrollaron Gunraj et al. [11] para imágenes CT. El segundo método consiste en segmentación pulmonar 3D y está basado en redes neuronales convolucionales, este método se llevó a cabo con el paquete Lungmask [32] de Python. En los experimentos realizados se puede observar que los mejores resultados se obtuvieron con el método de segmentación del cuerpo humano.

Como trabajo futuro se pueden considerar otras técnicas de preprocesado de volúmenes CT, como el aumento de datos 3D para la generación de datos artificiales y así mejorar el rendimiento de los modelos. Otro trabajo que se puede considerar es la creación de un modelo CNN para detección de lesiones pulmonares causadas por COVID-19. En el futuro, cuando la base de datos para el análisis de COVID-19 sea lo suficientemente grande, se espera que el modelo EffCOV-B0 supere los resultados actualmente presentados.

Bibliografía

- [1] M. Rosa, M. Matos, R. Furtado, V. Brito, L. Amaral, G. Beraldo, E. Fonseca, R. Chate, R. Passos, G. Teles, M. Silva, P. Yokoo, E. Yanata, H. Shoji, G. Szarf, and M. Funari, “COVID-19 findings identified in chest computed tomography: a pictorial essay,” *Einstein (São Paulo)*, vol. 18, 2020.
- [2] “EfficientNet: Improving Accuracy and Efficiency through AutoML and Model Scaling,” 05 2019.
- [3] P. Silva, E. Luz, G. Silva, G. Moreira, R. Pedrosa Silva, D. R. Lucio, and D. Menotti, “Covid-19 detection in ct images with deep learning: A voting-based scheme and cross-datasets analysis,” *Informatics in Medicine Unlocked*, vol. 20, p. 100427, 09 2020.
- [4] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning*, vol. 5, p. 11, 2019.
- [5] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” 2015.
- [6] A. Pak, O. A. Adegboye, A. I. Adekunle, K. M. Rahman, E. S. McBryde, and D. P. Eisen, “Economic Consequences of the COVID-19 Outbreak: the Need for Epidemic Preparedness,” *Frontiers in Public Health*, vol. 8, 2020.
- [7] Z. Zu, M. Jiang, P. Xu, W. Chen, Q. Ni, G. Lu, and L. Zhang, “Coronavirus Disease 2019 (COVID-19): A Perspective from China,” *Radiology*, vol. 296, no. 2, pp. E15–E25, 2020.
- [8] OMS, “Preguntas y respuestas,” 2021.
- [9] W. Kong and P. P. Agarwal, “Chest Imaging Appearance of COVID-19 Infection,” *Radiology: Cardiothoracic Imaging*, vol. 2, no. 1, p. e200028, 2020.
- [10] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, “Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases,” *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.

- [11] H. Gunraj, L. Wang, and A. Wong, “Covidnet-ct: A tailored deep convolutional neural network design for detection of covid-19 cases from chest ct images,” *Frontiers in Medicine*, vol. 7, p. 1025, 2020.
- [12] V. Rohila, N. Gupta, A. Kaul, and D. Sharma, “Deep learning assisted COVID-19 detection using full CT-scans,” *Internet of Things*, vol. 14, p. 100377, 2021.
- [13] H. X. Bai, B. Hsieh, Z. Xiong, K. Halsey, J. W. Choi, T. M. L. Tran, I. Pan, L.-B. Shi, D.-C. Wang, J. Mei, X.-L. Jiang, Q.-H. Zeng, T. K. Egglin, P.-F. Hu, S. Agarwal, F.-F. Xie, S. Li, T. Healey, M. K. Atalay, and W.-H. Liao, “Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT,” *Radiology*, vol. 296, no. 2, pp. E46–E54, 2020.
- [14] S. Morozov, A. Andreychenko, I. Blokhin, P. Gelezhe, A. Gonchar, A. Nikolaev, N. Pavlov, V. Chernina, and V. Gombolevskiy, “MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic,” *Digital Diagnostics*, vol. 1, no. 1, pp. 49–59, 2020.
- [15] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing, and P. Xie, “Sample-efficient deep learning for covid-19 diagnosis based on ct scans,” *medRxiv*, 2020.
- [16] E. Soares, P. Angelov, S. Biaso, M. H. Froes, and D. K. Abe, “Sars-cov-2 ct-scan dataset: A large dataset of real patients ct scans for sars-cov-2 identification,” *medRxiv*, 2020.
- [17] Z. Wang, Q. Liu, and Q. Dou, “Contrastive Cross-Site Learning With Redesigned Net for COVID-19 CT Classification,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2806–2813, 2020.
- [18] L. Wang, Z. Q. Lin, and A. Wong, “COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images,” *Scientific Reports*, vol. 10, no. 1, 2020.
- [19] H. Aboutalebi, S. Abbasi, M. Javad Shafiee, and A. Wong, “COVID-Net CT-S: 3D Convolutional Neural Network Architectures for COVID-19 Severity Assessment using Chest CT Images,” *arXiv e-prints*, p. arXiv:2105.01284, May 2021.
- [20] S. Serte and H. Demirel, “Deep learning for diagnosis of covid-19 using 3d ct scans,” *Computers in Biology and Medicine*, vol. 132, p. 104306, 2021.
- [21] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R. Hesketh, L. Yang, and C. Zheng, “Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19),” *Radiology*, vol. 295, no. 3, pp. 715–721, 2020.
- [22] A. Ghebreyesus T, “WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020,” 03 2020.

- [23] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, “The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2,” *Nature Microbiology*, vol. 5, no. 4, pp. 536–544, 2020.
- [24] M. Pal, G. Berhanu, C. Desalegn, and V. Kandi, “Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update,” *Cureus*, 2020.
- [25] J. Chan, K. Kok, Z. Zhu, H. Chu, K. To, S. Yuan, and K. Yuen, “Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan,” *Emerging Microbes & Infections*, vol. 9, no. 1, pp. 221–236, 2020.
- [26] A. Khaitovich, “CORONAVIRUS (TAXONOMY, VIRUS STRUCTURE),” *Crimea Journal of Experimental and Clinical Medicine*, vol. 10, no. 3, pp. 69–81, 2021.
- [27] C. Silva, E. Marchiori, A. Souza Júnior, and N. Müller, “Consenso brasileiro ilustrado sobre a terminologia dos descritores e padrões fundamentais da TC de tórax,” *Jornal Brasileiro de Pneumologia*, vol. 36, no. 1, pp. 99–123, 2010.
- [28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
- [29] G. Hinton, N. Srivastava, and K. Swersky, “Neural Networks for Machine Learning Lecture 6a Overview of mini-batch gradient descent.”
- [30] Y. Ho and S. Wookey, “The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling,” *IEEE Access*, vol. 8, pp. 4806–4813, 2020.
- [31] H. Gunraj, A. Sabri, D. Koff, and A. Wong, “Covid-net ct-2: Enhanced deep neural networks for detection of covid-19 from chest ct images through bigger, more diverse learning,” 2021.
- [32] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs, “Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem,” *European Radiology Experimental*, vol. 4, no. 1, 2020.
- [33] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, 2019.

Apéndice A

Anexo I

A.1. Hallazgos identificadas en CT de tórax

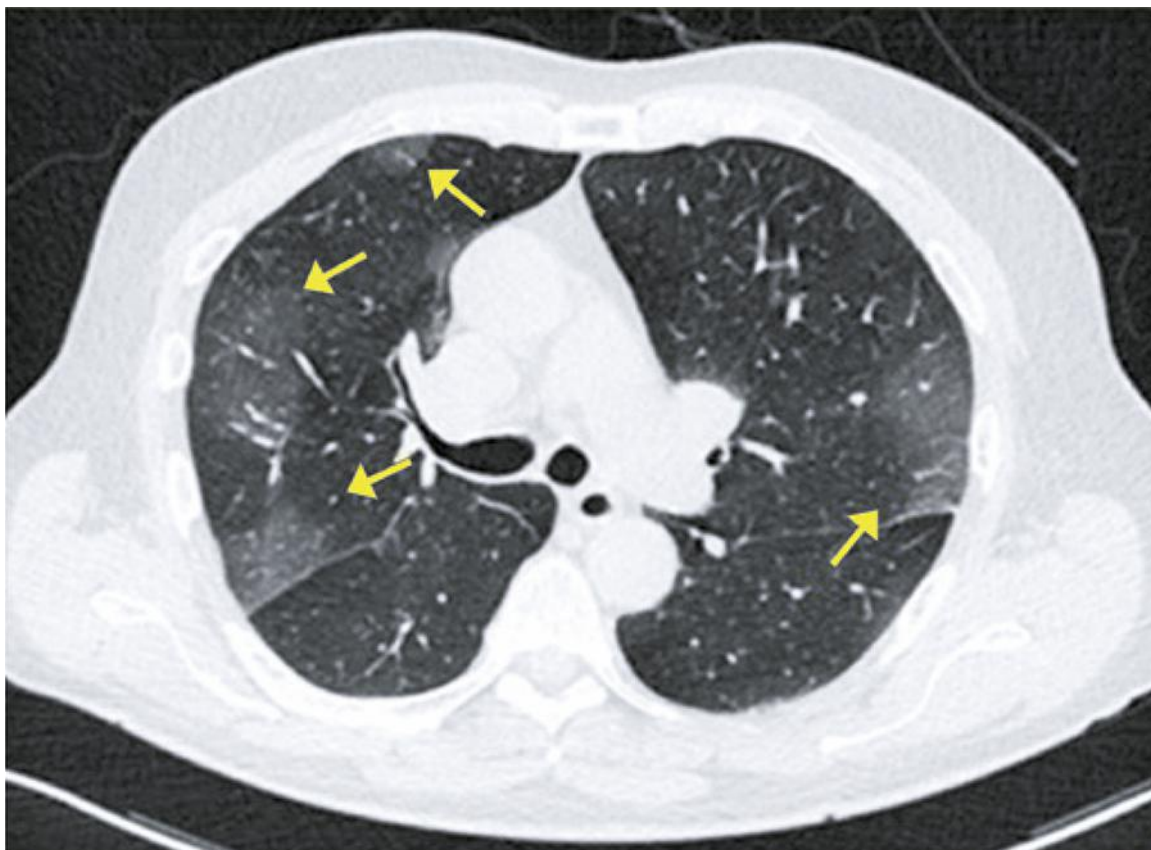


Figura A.1: Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado en los lóbulos superiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]

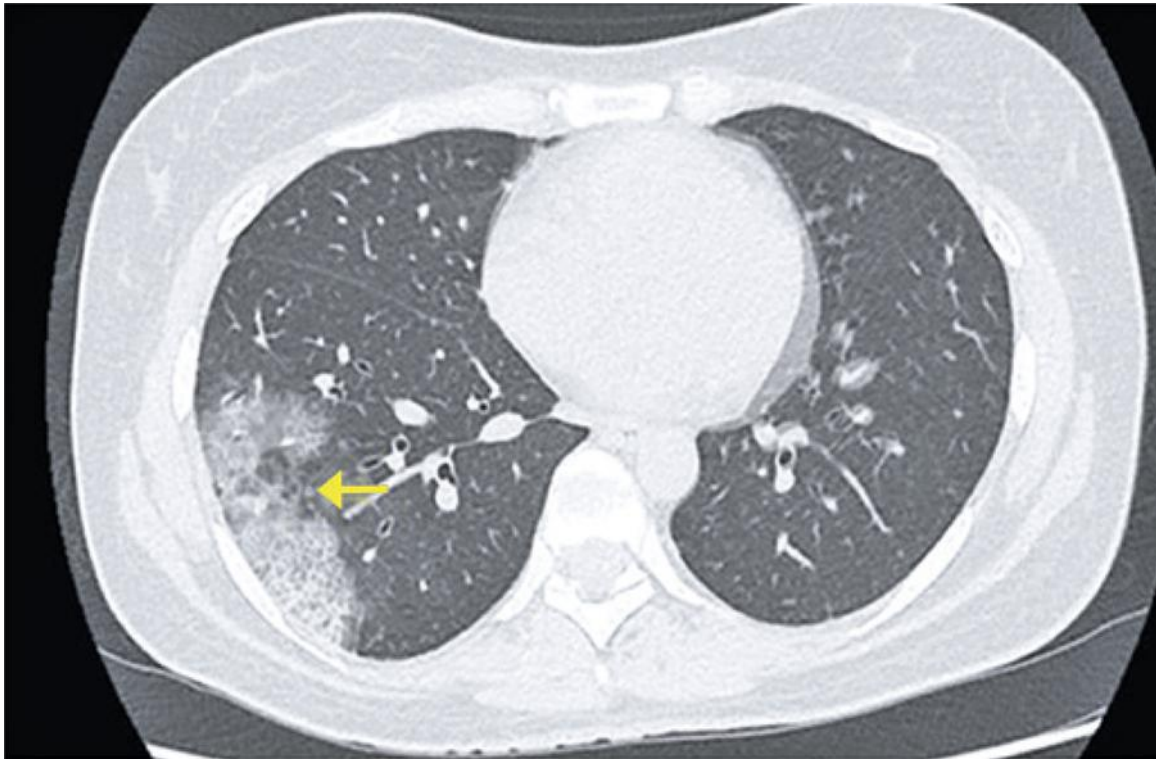


Figura A.2: Tomografía computarizada que muestra patrón de pavimentación loca en la periferia del lóbulo inferior derecho. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]

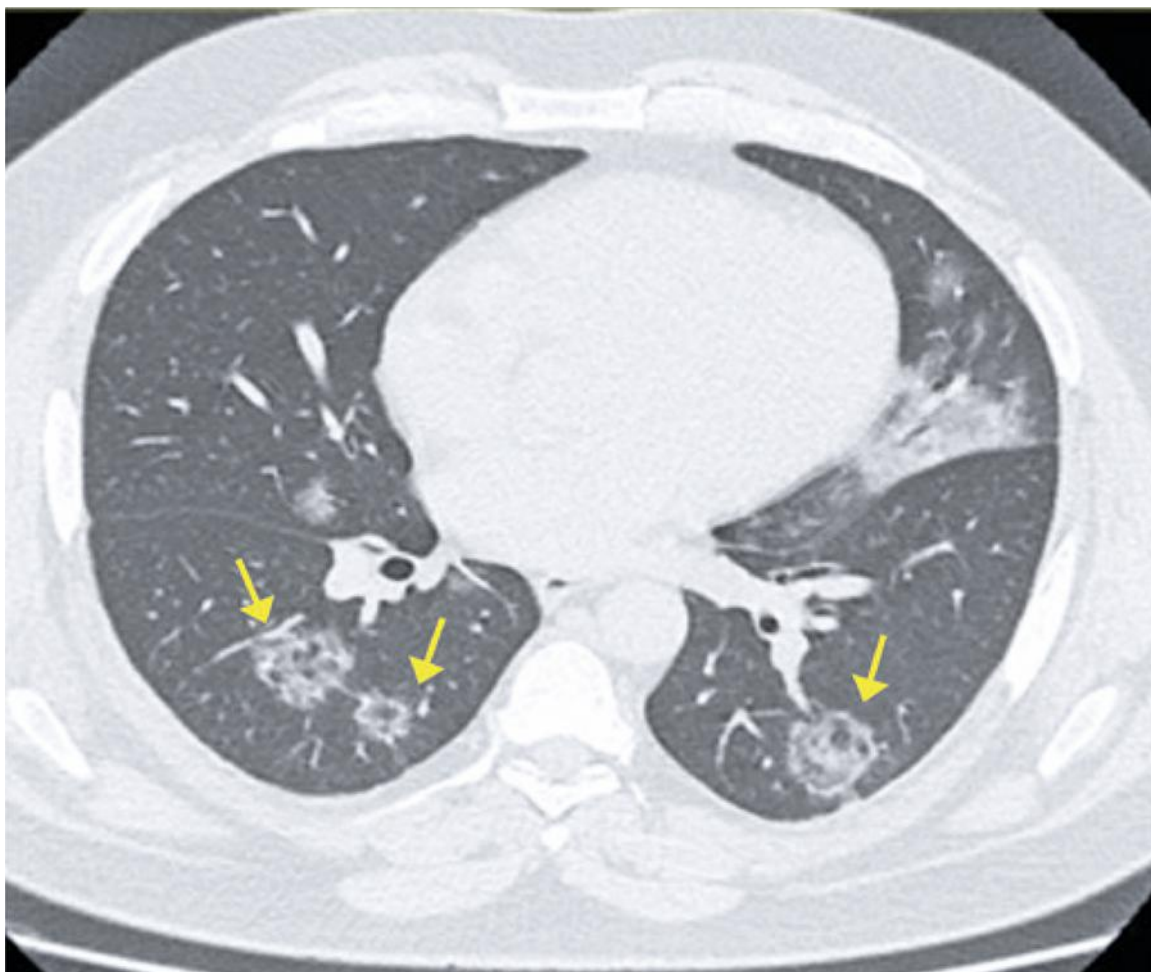


Figura A.3: Tomografía computarizada que muestra signo de halo invertido en los lóbulos inferiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]

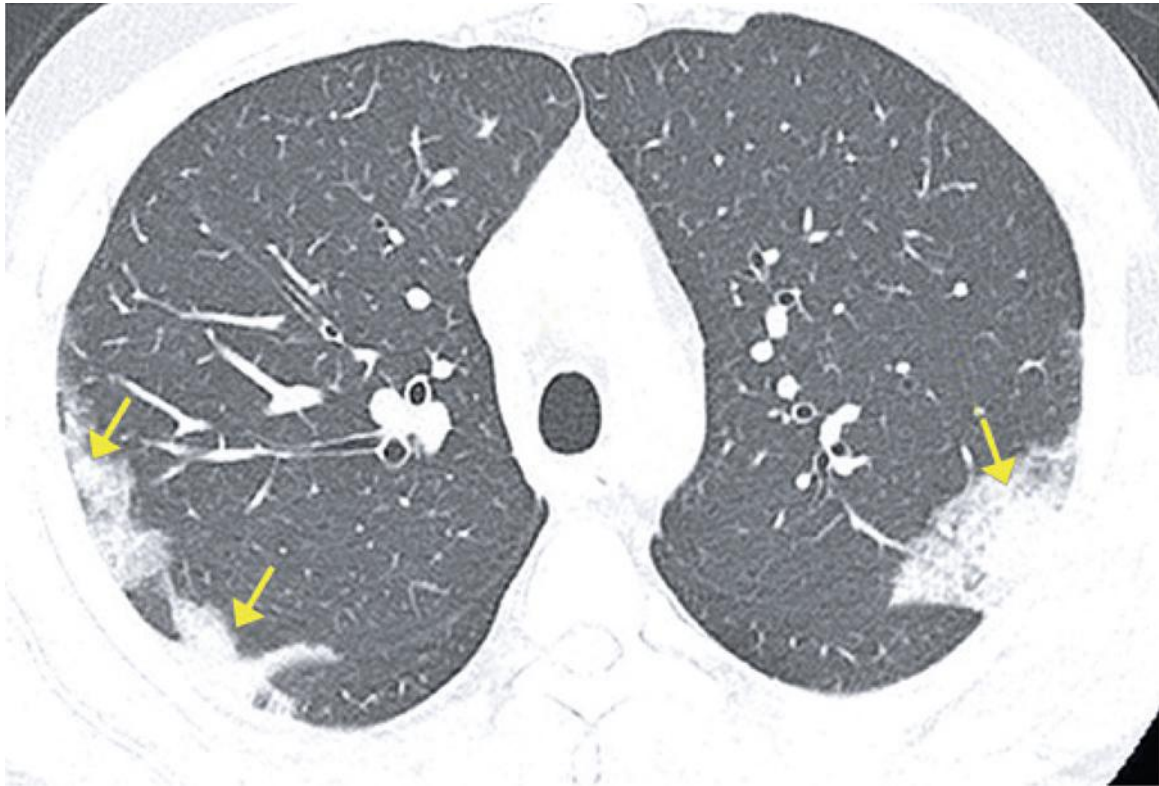


Figura A.4: Tomografía computarizada que muestra consolidaciones periféricas en los lóbulos superiores. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]



Figura A.5: Tomografía computarizada que muestra opacidades periféricas en vidrio esmerilado, consolidaciones posteriores con broncogramas aéreos. Imagen tomada de *COVID-19 Findings Identified in Chest Computed Tomography: a Pictorial Essay* [1]

A.2. Descripción del hardware

Memory	125GiB
CPU	AMD Ryzen Threadripper 1950X 16-Core Processor 3533MHz 64 bits
GPU	GeForce RTX 2080 11016MiB CUDA Version: 11.2 Driver Version: 460.91.03

Tabla A.1: Descripción del hardware utilizado para desarrollar modelos de redes neuronales.

A.3. Grad-CAM

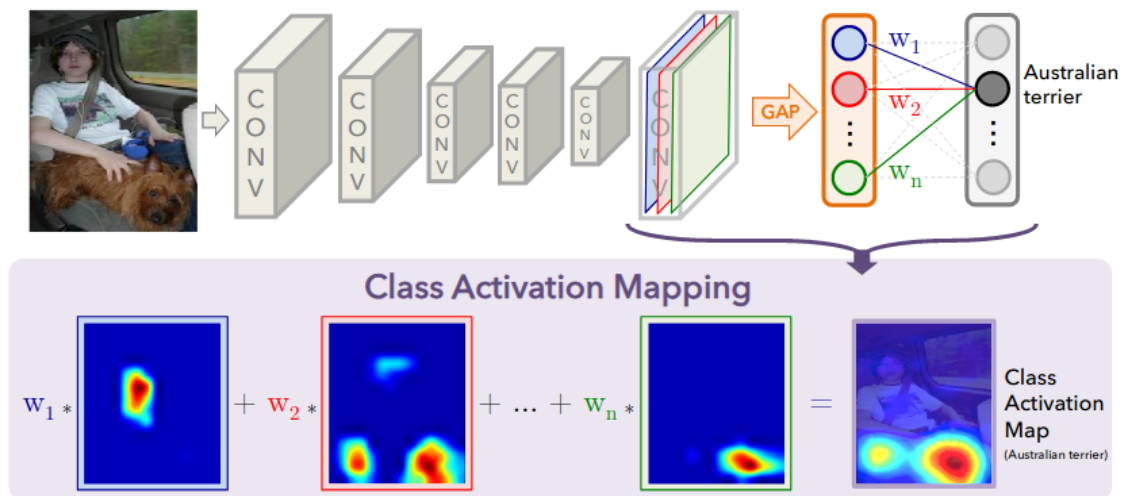


Figura A.6: Mapeo de la activación de la clase: el puntaje de clase predicho se asigna a la capa convolucional anterior para generar los mapas de activación de clase (CAMs). Imagen tomada de *Learning Deep Features for Discriminative Localization* [5].