UAB
Universitat Autònoma de Barcelona

INTELIGENCIA ARTIFICIAL APLICADA A LA RADIÓMICA EN PATOLOGÍA PULMONAR

SONIA BAEZA MENA

Universitat Autònoma de Barcelona

Programa de Doctorado en Medicina
Departamento de Medicina
Universidad Autónoma de Barcelona

INTELIGENCIA ARTIFICIAL APLICADA A LA RADIÓMICA EN PATOLOGÍA PULMONAR

AUTORA: SONIA BAEZA MENA

TESIS DOCTORAL

Directores: Dr. Antoni Rosell Gratacós
Dra. Debora Gil Resina
Tutora: Dra. Maria Cristina Tural Llacher

Barcelona, 2024

A mis padres, Manuel y Sonia.

Por apoyarme y estar siempre presentes a pesar de la distancia.

“Vive como si fueras a morir mañana;

aprende como si el mundo fuera a durar para siempre”

Mahatma Gandhi

AGRADECIMIENTOS

Comienzo por agradecer a mis directores de tesis Toni Rosell y Debora Gil, pieza clave en todo el proyecto que verán reflejado en las siguientes páginas.

Toni, gracias por ser mi mentor durante estos 5 años, por formarme como investigadora, guiarme como neumóloga, y sobre todo por creer en mí. Hemos formado un buen equipo, no solo en el ámbito de la radiómica, sino también con proyecto el cribado de cáncer de pulmón. Han sido años de intenso trabajo, pero con frutos que ya estamos comenzando a visualizar. Como olvidar las largas tardes de reuniones que hemos tenido junto a Natalia, (quien también ha contribuido mucho en esta tesis) en las que divagábamos en ideas que nos parecían geniales y Natalia no sabía cómo centrarnos en el prepuesto del estudio. Espero que sigamos trabajando juntos más allá de esta tesis, sería un honor.

Debora, si me remonto al verano de 2019, la radiómica era un área totalmente nueva para mí, y yo, una médico que no había hecho más que trabajar con otros médicos, me vi entrando por las puertas del Centro de Visión por Computador sin tener idea, y me encontré un grupo excepcional que me recibió y compartió su conocimiento. Gracias por todo lo que me has enseñado en estos años, eres una investigadora de excelencia y aún tengo mucho que aprender de ti, así que espero que sigamos trabajando en este equipo híbrido que hemos formado entre matemáticos/ingenieros y médicos.

Cristina Tural mi tutora de tesis, gracias por aceptarme como doctoranda y acompañarme en estos años.

Quiero agradecer a todos quienes han colaborado directamente en esta tesis, como Carles y Guillermo, parte del equipo de CVC con quienes he trabajado muchísimo durante estos 5 años. Felipe Andreo y Carmen Centeno, gracias por todo lo que me han enseñado sobre el cáncer de pulmón, pero también como neumóloga desde mis primeros años de residencia. Al equipo de Radiología, en especial a Ignasi Guasch, Isa Nogueira y Carles Gómez,

gracias por la buena disposición y ayuda cada vez que voy a verlos con algún proyecto diferente, por el interés que han puesto en la radiómica y sobre todo en el cribado de cáncer de pulmón. Al equipo de Cirugía Torácica, en especial a Samuel Garcia-Reina, te has convertido en nuestro cirujano de referencia, gracias por tu apoyo. No puedo dejar de agradecer a también a Jose Luis Mate, Maite Salcedo, Jordi Deportós, Gloria Moragas y Cristian Tebé por su colaboración en este proyecto.

Agradezco también a todo el servicio de Neumología, a enfermería, fisioterapia, administrativas, residentes y compañeros neumólogos, son un apoyo fundamental y un grupo de personas maravillosas. En especial quiero dar las gracias a Jorge Abad, por su apoyo que va más allá del transcurso de esta tesis, se remonta a los inicios de mi residencia, cuantos buenos momentos en estos años. Nacho Garcia-Olivé, gracias por ser una especie de “subdirector” de mi tesis, aconsejándome y ayudándome en incontables ocasiones sin esperar nada a cambio. Alicia Marín, gracias por tu ayuda sobre todo al inicio de la tesis, cuando me diste una mano en un momento complejo a nivel laboral. No puedo dejar de mencionar a Juan Ruiz, mi primer jefe, un gran ejemplo como médico y persona, gracias por estar siempre atento a mis logros.

También tengo que agradecer a personas que aunque indirectamente, con su amistad han sido un soporte fundamental durante estos años. Adela, cuanto y que bien hemos trabajado juntas, tu apoyo como amiga/madre ha sido clave, me has ayudado a centrarme en muchos momentos y me haces tener una visión clara del trabajo y también de la vida. Carlos y Filipe, mis queridos viejes, como podría dejar de agradecer a quienes se han convertido en una familia para mí y con quienes he compartido tantos buenos momentos. También quiero agradecer a compañeros que se han convertido durante estos años en muy buenos amigos, como son Paloma, Marta, Agnes, Adri, Salo, Rox, Laurita, Carmen, Zoran, Maca y Toni M, gracias por estar siempre, compartir juntos el día a día, las barbacoas (o asados como a mí me gusta decir) y tantas risas. No puedo olvidar agradecer a la Domi, mi amiga incondicional desde la infancia, no sabes como me alegran tus mensajes y llamadas, me transporta a nuestra adolescencia, parece como si el tiempo no hubiese pasado.

Como no mencionar a Alex, mi marido y compañero de vida y viajes. Cuanta paciencia has tenido conmigo, nos conocimos cuando comenzaba la tesis así que si hemos sobrevivido victoriosos a todo este periodo, podemos sobrepasar cualquier obstáculo. Gracias por cuidarme tanto, entenderme siempre, y hacer que nuestro hogar sea un refugio lleno paz y sobre todo de risas y café. También tengo que agradecer a Nico nuestro perro, mí terapeuta emocional que es capaz de cambiarme el chip después de una guardia o un mal día de trabajo.

Gracias a mi familia en Chile, “los Baeza” y “los Mena”, que están conmigo apoyándome a la distancia no solo ahora, desde siempre. En especial quiero dar las gracias a mis abuelos Margarita y Orlando, dos personas rompedoras para su época, un ejemplo maravilloso y admirable. A mis queridísimas hermanas Natalia y Constanza, gracias por la alegría y cariño de siempre. A Rosi, por cuidarme y hacerme sentir como una hija, tienes una bondad única, soy una afortunada de que estes presente en mi día a día.

Finalizo agradeciendo a quienes he dedicado esta tesis, mis padres Manuel y Sonia. Aunque nunca hemos sido una familia convencional, me han dado una infancia preciosa de la que estoy enormemente agradecida. Más allá de ser un ejemplo de perseverancia, trabajo y optimismo, me han ayudado a trazar mi camino como persona y como médico, y han tenido la grandeza de apoyarme a pesar de que ese camino me lleve a miles de kilómetros de distancia de casa, Viña del Mar. Mamá, eres un ejemplo de entrega y perseverancia, siempre preocupadas de tus “niñitas”, intentando darnos lo mejor en cada momento. Como podría olvidar los desayunos saludables con zumo de fruta natural recién exprimido a las 7 de la mañana, o tu esfuerzo incansable porque nos gustara la natación o el tenis, llegando a apuntarte con nosotras para hacer una actividad juntas. Papá, tu forma alegre y positiva de ver la vida a pesar de la adversidad me ha forjado como persona, sin contar que tu confianza en mí ha hecho que sienta que puedo lograr todo lo que me proponga.

ABREVIATURAS

TEP = Tromboembolismo pulmonar

AngioTC = Angiografía pulmonar por TC

SPECT/TC-ventilación/perfusión = SPECT/TC-V/Q

SPECT/TC-perfusión = SPECT/TC-Q

IA = Inteligencia artificial

IC = intervalos de confianza

VPP= Valor predictivo positivo

VPN = Valor predictivo negativo

AUC = área bajo la curva

TABLA DE CONTENIDOS

RESUMEN

Esta tesis expone los resultados dos estudios en el que se desarrollan modelos diagnósticos basados en radiómica, una novedosa técnica de análisis cuantitativo y de alto rendimiento de la imagen médica, potenciada por inteligencia artificial principalmente Deep Learning, orientados a optimizar el diagnóstico por imagen de TEP y de nódulo pulmonar.

El primer estudio muestra un modelo radiómico inteligente que es capaz de hacer un diagnóstico local de diferentes patrones de imagen (TEP, neumonía y pulmón sano) en pacientes con y sin COVID-19 por medio de SPECT-TC/Q. El segundo estudio muestra un modelo integrativo que fusiona un modelo Deep radiómico con un modelo de datos clínicos para predecir malignidad en nódulos pulmonares incidentales o detectados por cribado de cáncer de pulmón.

Ambos modelos experimentales han demostrado un buen rendimiento diagnóstico en el grupo de pacientes analizados, siendo capaces de dar respuesta a la problemática que se ha propuesto resolver. El uso de nuevas tecnologías como la presentada en esta tesis, nos aproxima a un diagnóstico avanzado por imagen que en un futuro podrá ser trasladado a la clínica diaria para beneficio de los pacientes y la práctica médica.

SUMMARY

This thesis presents the results of two studies in which diagnostic models based on radiomics, a novel quantitative and high-throughput medical imaging analysis technique powered by artificial intelligence, mainly Deep Learning, are developed to optimize imaging diagnosis of pulmonary embolism (PE) and pulmonary nodules.

The first study demonstrates an intelligent radiomic model capable of making a localized diagnosis of different imaging patterns (PE, pneumonia, and healthy lung) in patients with and without COVID-19 through SPECT-CT/Q. The second study presents an integrative model that merges a Deep radiomic model with a clinical data model to predict malignancy in incidental or screening-detected pulmonary nodules.

Both experimental models have shown good diagnostic performance in the group of patients analyzed, being able to address the proposed issues. The use of new technologies, such as those presented in this thesis, brings us closer to advanced imaging diagnosis that can potentially be applied to daily clinical practice for the benefit of patients and medical practice.

1. INTRODUCCIÓN

Las enfermedades respiratorias son una causa importante de morbimortalidad tanto en Europa como a nivel mundial, siendo responsables de 3 de las 10 principales causas de muerte en 2019 de acuerdo con World Health Organization (WHO) (1) . Desde 2020, debido a la pandemia mundial de COVID-19, la relevancia de las enfermedades respiratorias cobra aún más interés, considerando que en solo 24 meses se produjeron > 5.7 millones de muertes debido a esta causa (2) . Por otra parte, el cáncer de pulmón es el cáncer que provoca mayor mortalidad a nivel mundial, ocasionando el 18.7% de todas las muertes por cáncer (3). Según los datos del Instituto Nacional de Estadística (INE), en 2022 ocurrieron 22.712 muertes por cáncer de pulmón en España, aumentando un 1.3% respecto a 2021 (4). Estos datos evidencian la importancia de investigar y optimizar los procesos diagnósticos en las patologías respiratorias, para que los pacientes puedan acceder de manera rápida a un tratamiento que permita la curación y prevenga la mortalidad por esta causa.

1.1. Tromboembolismo pulmonar

El tromboembolismo pulmonar agudo (TEP) es una forma frecuente de tromboembolismo venoso, muchas veces infradiagnosticado, cuya mortalidad puede alcanzar el 5% incluso bajo tratamiento (5). Entre los factores predisponentes se encuentran los traumatismos, cirugía mayor, infarto agudo de miocardio, cáncer, hospitalización por insuficiencia cardiaca o fibrilación auricular, uso de anticonceptivos orales, procesos infecciosos, trombofilias y, enfermedades autoinmunes como el síndrome antifosfolipídico, entre otros (6).

Sus síntomas son inespecíficos, pudiendo dificultar su diagnóstico (7), es por esto que es necesario utilizar escalas de predicción clínicas validadas para estimar la probabilidad pre-test en pacientes en los que se está considerando el diagnóstico de TEP, siendo la más conocida la escala de Wells clasificando clínicamente en probabilidad baja, intermedia y alta de padecer la enfermedad (8). La determinación en sangre del dímero D es una herramienta útil en los casos con probabilidad pre-test intermedia, ya que, es una prueba

de valor predictivo negativo alto para TEP (9). La realización de una prueba de imagen para el diagnóstico se recomienda en pacientes con probabilidad intermedia/alta de TEP, siendo de elección la angiografía pulmonar por TC (AngioTC), aunque si esta no está disponible o hay contraindicación se recomienda realizar SPECT-TC/ventilación-perfusión (SPECT/TC-V/Q) (9), ya que es una técnica validada para el diagnóstico de TEP incluso en presencia de neumonía (10,11).

Para entender el SPECT/TC-V/Q, es importante explicar la gammagrafía planar (2D) de ventilación/perfusión, prueba radiológica no invasiva más antigua para el diagnóstico de TEP (12), que consiste en una inyección intravenosa de partículas de albúmina macroagregadas marcadas con ^{99m} Tc, que bloquean arteriolas pulmonares y permiten la evaluación de la perfusión pulmonar. Cuando hay una oclusión arterial se observan áreas sin captación gammagráfica. Al asociarse a ventilación, se administran también radiofármacos (tipo aerosol o gases inertes) para identificar la hipoventilación como causa de hipoperfusión. En resumen, la firma distintiva para el diagnóstico de TEP es un desajuste ventilación-perfusión en el cual las áreas de la imagen de perfusión muestran un defecto (hipoperfusión) en cuña consistente con la obstrucción y las áreas correspondientes en la imagen de ventilación permanecen normales, consistente con la estructura pulmonar preservada. Es muy útil para excluir TEP en casos con probabilidad media y alta.

El SPECT a diferencia de la gammagrafía planar, proporciona imágenes en 3D debido a la rotación de la gammacámara alrededor del eje corporal, lo que permite una mejor caracterización y visión anatómica, aumenta sensibilidad y especificidad, y disminuye los casos indeterminados (13). SPECT/TC-V/Q, al igual que el PET-TC se basa en la fusión de la imagen anatomo funcional, permitiendo mejorar aún más la localización anatómica, con hallazgos de imagen que pueden confirmar la EP, como coágulos hiperatenuantes o tamaño diferencial de los vasos, pudiendo además demostrar causas alternativas a los síntomas del paciente, como neumonía, edema pulmonar, entre otros (14,15). Algunos estudios sugieren que el TC puede reemplazar de manera confiable la ventilación, especialmente en la población más joven y en aquellos sin enfermedad obstructiva de las vías respiratorias. La figura 1 muestra un SPECT/TC-V/Q que resulta compatible con TEP.

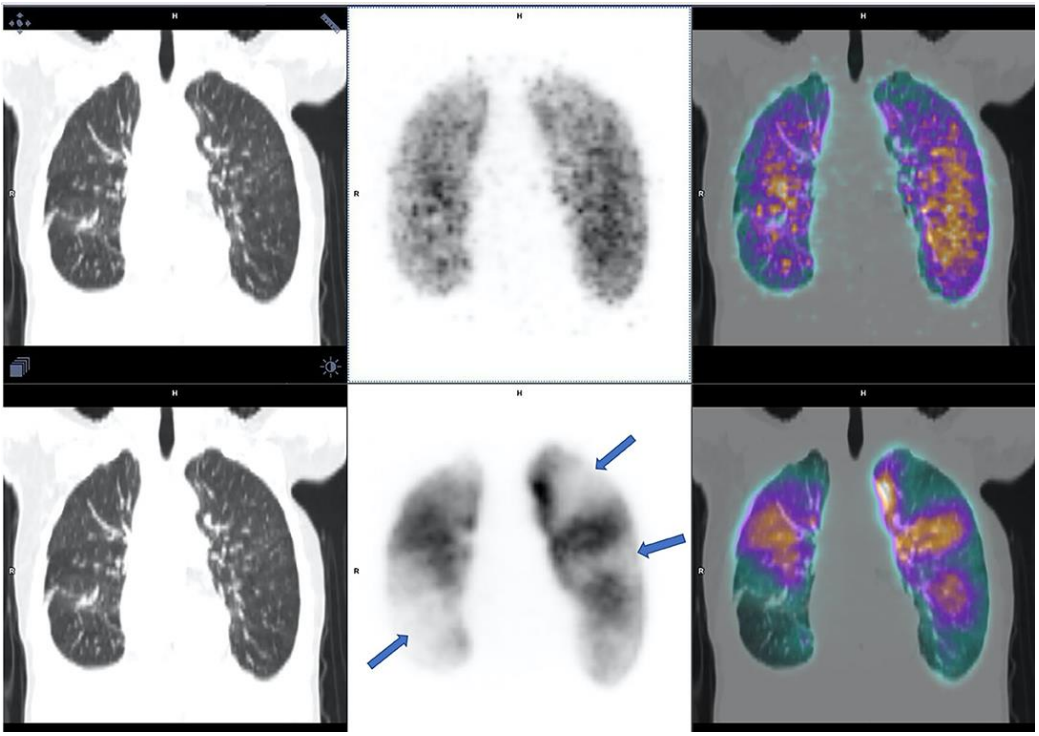


Figura 1: SPECT/TC-V/Q en un paciente con TEP. Le Pennec, Romain et al. (16)

En contexto de la pandemia de COVID-19 en 2020, hubo un aumento significativo de casos de TEP extensos, incluso en pacientes que habían recibido profilaxis farmacológica al ser hospitalizados. Las causas que contribuyen a esta trombosis generalizada son una inflamación excesiva con tormenta de citoquinas, activación plaquetaria, disfunción endotelial y estasis (17). Los eventos tromboembólicos venosos y arteriales se reportaron en un 31–59% de los pacientes hospitalizados con COVID-19, especialmente en casos graves que necesitaban ingreso en cuidados intensivos (18–20). El diagnóstico de embolia pulmonar en esta última situación puede ser desafiante.

Debido a la necesidad de realizar el diagnóstico de TEP en este grupo de pacientes, aumentaron las solicitudes de pruebas para descartarlo, y aunque la AngioTC es la prueba de elección, cuando esta prueba no se encuentra disponible o en pacientes con contraindicaciones para su realización (insuficiencia renal, alergia al contraste yodado, etc.),

el SPECT/TC-Q es una alternativa que tuvo que ser adaptada evitando la ventilación para poder responder a los estándares de seguridad requeridos para evitar contagios (21–25). La figura 2 muestra un SPECT/TC-Q que es compatible con TEP en un paciente con neumonía por COVID-19.

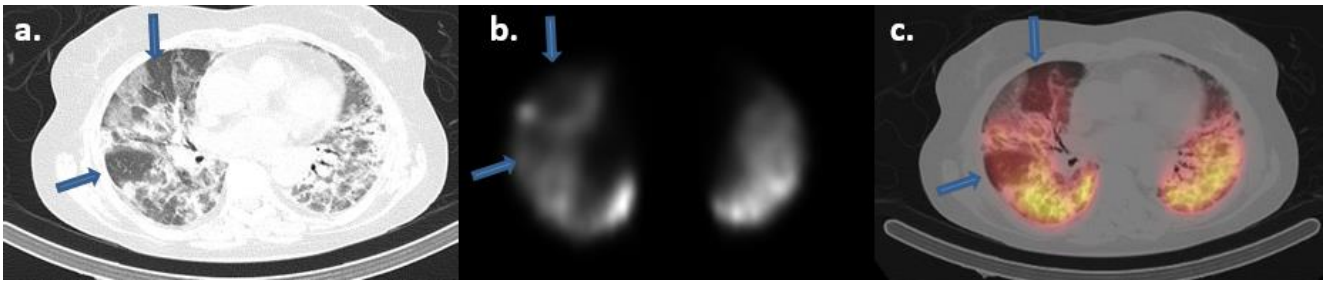


Figura 2: SPECT/TC-Q realizado en paciente con neumonía por COVID-19, que se diagnostica de TEP. Caso realizado en el Hospital Germans Trias i Pujol

El valor de SPECT/TC-Q fue reportado antes de la pandemia de COVID-19, mostrando una sensibilidad y especificidad del 100% y 83%, respectivamente para el diagnóstico de TEP (26). En un metaanálisis publicado recientemente, muestra una sensibilidad y especificidad del 92% para SPECT/TC-Q en 1196 pacientes pertenecientes a 7 estudios (27). Sin embargo, el diagnóstico de TEP puede ser difícil de realizar en áreas extensas de neumonía como ocurrió durante la pandemia de COVID-19. En estos casos, el informe tarda más tiempo en realizarse y necesita ser validado con frecuencia por médicos nucleares experimentados para prevenir errores.

El diagnóstico de COVID-19 y la detección de sus complicaciones mediante imágenes médicas ha despertado un gran interés dentro de la comunidad científica dedicada al desarrollo de inteligencia artificial (IA). El uso de la radiómica y las técnicas de IA para

mejorar el diagnóstico de TEP en COVID-19 a través de SPECT/CT-Q no había sido explorado al momento de iniciar nuestra investigación, siendo un área innovadora y con un amplio campo a explorar.

1.2. Cáncer de pulmón y nódulo pulmonar

El cáncer de pulmón es el cáncer más diagnosticado y la principal causa de muerte por cáncer en todo el mundo (3). Según Globocan 2022, ese año se registraron 2.480.301 casos nuevos de cáncer de pulmón y se contabilizaron 1.817.172 muertes (3). En Cataluña, en 2020 se registraron 4.509 casos nuevos y 3.548 muertes (28). Las cifras de mortalidad están directamente relacionadas con el hecho de que hasta un 70% de los casos se diagnostican en estados avanzados, cuando no es posible un tratamiento curativo.

El tabaquismo es el principal factor de riesgo para el desarrollo de cáncer de pulmón, estando relacionado con hasta un 80% de los casos en población occidental. Sin embargo, hay que tener en cuenta que existen otros factores asociados como el contacto con amianto (29), radón (30), contaminación ambiental (31,32), enfermedades pulmonares crónicas previas como la EPOC y la fibrosis pulmonar (33,34), infecciones previas por tuberculosis (35), y antecedentes familiares de cáncer de pulmón (36), entre otros.

El cáncer de pulmón es una enfermedad que no da síntomas en estadios iniciales, lo que ocasiona que el diagnóstico sea tardío en la mayoría de los casos si es que no se establecen herramientas orientadas a la detección precoz de esta patología. En los últimos años está surgiendo un cambio de tendencia en la detección de cáncer de pulmón debido al aumento del uso de Tc de tórax en diferentes ámbitos, lo que ha llevado a un aumento en el número de nódulos pulmonares detectados incidentalmente. Por otro lado, este cambio también se atribuye a la implementación de programas de cribado que han demostrado un aumento en la detección de cáncer de pulmón en estadio I (37,38). En ambos casos, el resultado es el diagnóstico en etapas tempranas. Es así como el informe anual de la Sociedad Americana del Cáncer de 2022 muestra una disminución en la incidencia de cáncer de pulmón en etapas avanzadas, con un aumento anual simultáneo

del 4.5% en un estadio localizado, principalmente vinculado al inicio del cribado de cáncer de pulmón en Estados Unidos hace una década (39).

El cribado de cáncer de pulmón ha demostrado de forma consistente que es capaz de reducir mortalidad en cáncer de pulmón (37,38,40), sin embargo, es importante tener en cuenta que no todos los nódulos pulmonares detectados resultan ser cáncer de pulmón. En una revisión sistemática (40), el 21% de los Tc de tórax de baja radiación basal detectaron un nódulo pulmonar falso positivo, lo que genera un aumento en la realización de TC de seguimiento e incluso de biopsias innecesarias. La detección de estos nódulos pulmonares falsos positivos puede causar daño potencial a los pacientes y generar costos elevados para los servicios de salud (41–43). De manera similar, ocurre con el aumento de la detección incidental de nódulos pulmonares (fuera del cribado), ya que muchos de estos son benignos y requieren un seguimiento cercano, similar al enfoque adoptado con los nódulos pulmonares identificados a través cribado de cáncer de pulmón (44–47). En este contexto, ya sea porque el nódulo pulmonar se detecte incidentalmente o a través de cribado de cáncer de pulmón, es crucial identificar los casos de alto riesgo y descartar aquellos con bajo riesgo (48,49).

Centrándonos en el nódulo pulmonar, este se define como una opacidad pulmonar bien delimitada con un diámetro de menos de 3 cm. De acuerdo con su radiomorfología, los nódulos pulmonares se clasifican como sólidos o subsólidos, siendo los subsólidos diferenciados adicionalmente en nódulos de vidrio esmerilado puro (GGO) y nódulos mixtos-sólidos (parcialmente-sólido) (50), como se puede observar en la figura 3.

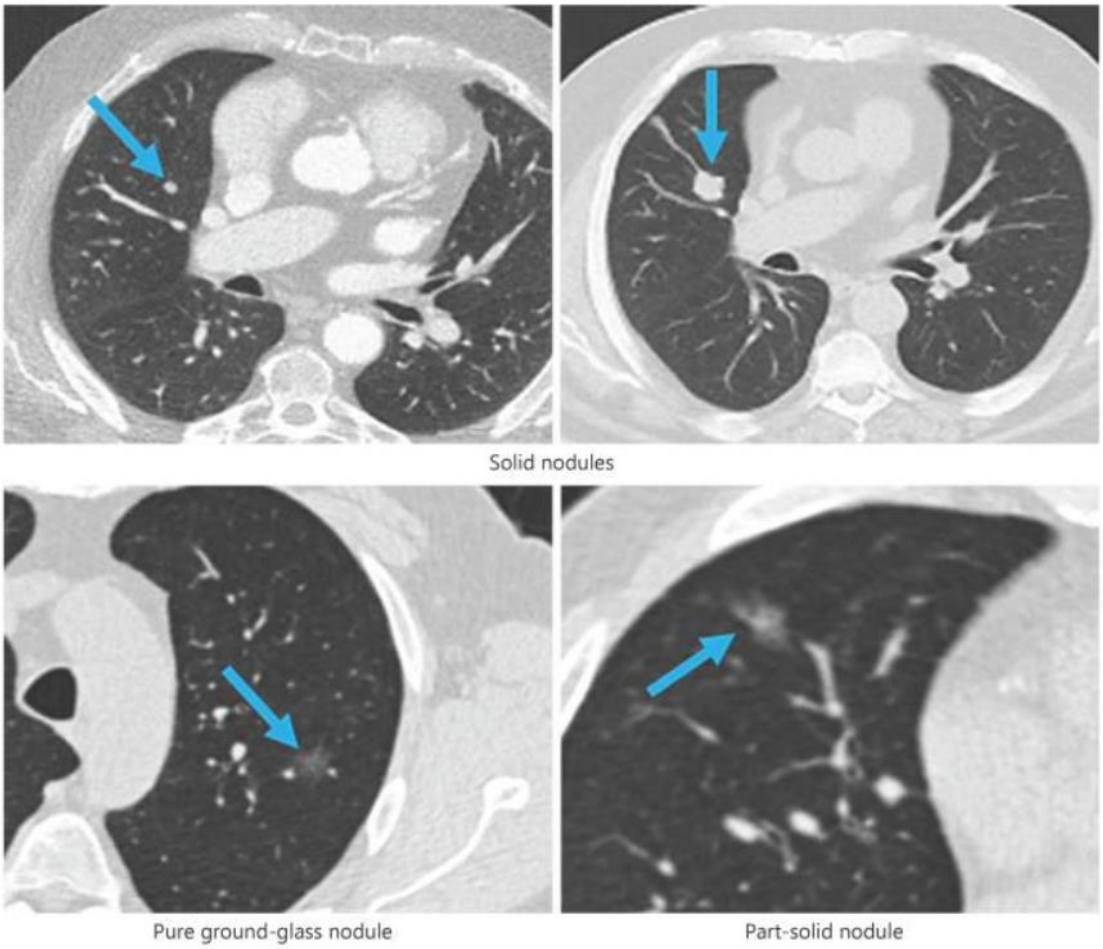


Figura 3: Clasificación de nódulos por radiomorfología según la Guía Fleischner. Schmid-Bindert et al. (51)

Los nódulos pulmonares pueden tener diferentes etiologías (52), como se muestra en la figura 4. Es por esto que lo más relevante es detectar dentro de este amplio grupo a los nódulos pulmonares que pueden ser malignos. Para obtener el diagnóstico histológico, existen diferentes vías de abordaje: biopsia quirúrgica, biopsia endobronquial o transbronquial por vía broncoscópica o biopsia guiada por TC de tórax. El problema es que muchas veces por su tamaño es difícil hacer un diagnóstico no invasivo, debido a esto y para estandarizar su manejo se han desarrollado diferentes guías clínicas (50,53,54) que se basan principalmente en la forma y el tamaño, destacando que en los últimos se han asociado al uso de la volumetría como forma más exacta de medición. Aun así, el manejo

de los nódulos pulmonares es mejorable con el uso de modelos de riesgo que permiten asociar datos individuales de cada paciente y del nódulo pulmonar en cuestión para predecir la posibilidad de malignidad (51). La guía BTS (53) respalda el uso de los modelos de Herder et al. (55) y el modelo de Brock (48) en nódulo pulmonar incidental, siendo este último el más utilizado, aunque hay que considerar que ha sido validado en población de cribado de cáncer de pulmón y no de nódulo pulmonar incidental.

La tecnología también puede ser una herramienta útil para optimizar el manejo y diagnóstico no invasivo de los nódulos pulmonares tanto de cribado como en nódulo pulmonar incidental. En este sentido, la radiómica potenciada por la IA es capaz de aproximarse al diagnóstico, tratamiento y pronóstico en cáncer de pulmón (56–59). Concretamente en nódulo pulmonar existe un amplio campo de desarrollo de la radiómica, tanto para llegar a un diagnóstico de malignidad, como para histología y para la detección de marcadores moleculares y PD-L1 (60–62). Este enfoque del nódulo pulmonar nos conduce a la obtención de resultados no invasivos similares a una biopsia (también llamado biopsia virtual), con lo que se podrían evitar biopsias innecesarias, reducir las complicaciones para los pacientes tanto clínicas como ansiedad por un resultado incierto de un nódulo pulmonar, además de permitir una disminución en los costos sanitarios.

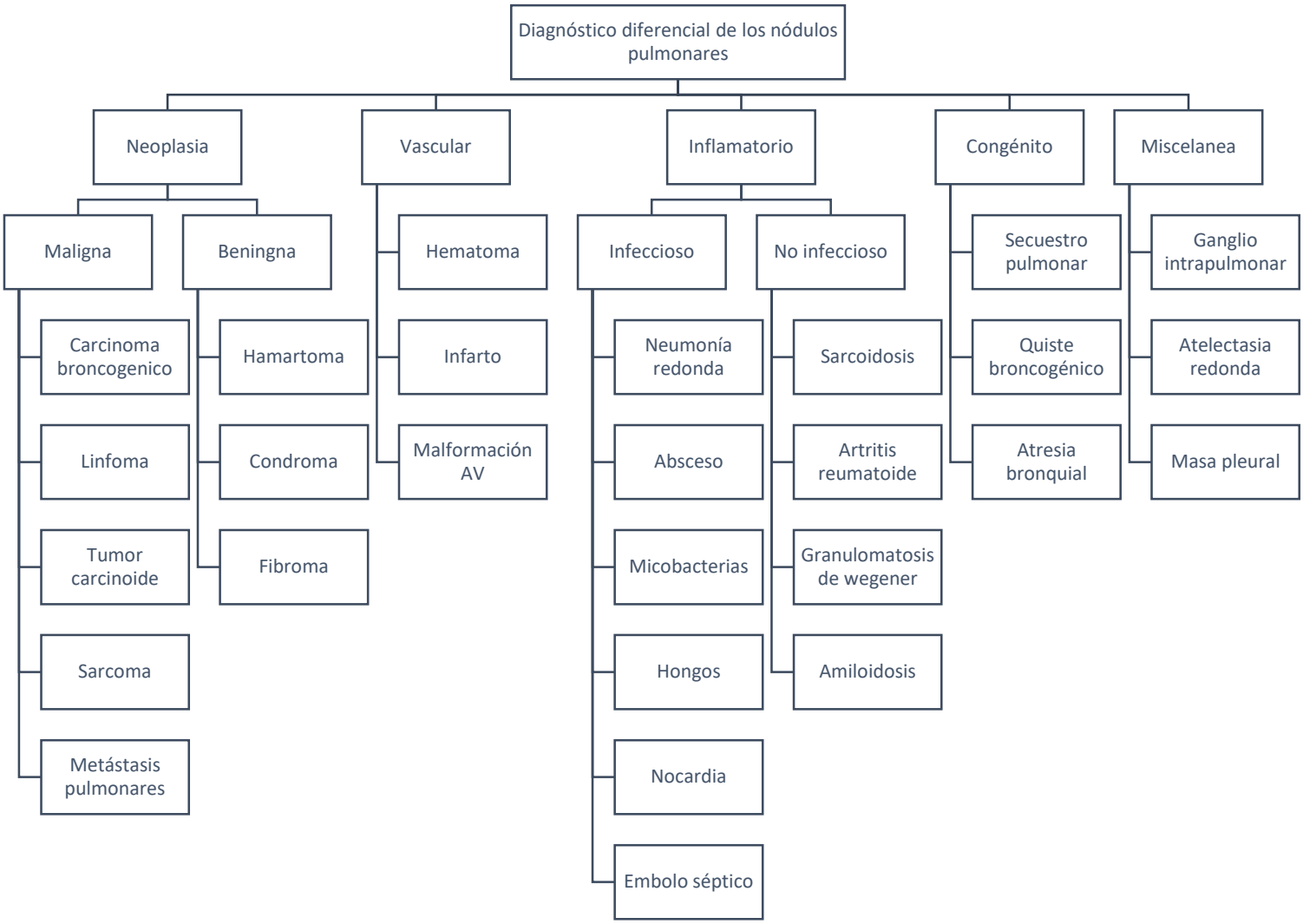


Figura 4: Diagnóstico diferencial de los nódulos pulmonares

1.3. Radiómica e Inteligencia Artificial

En los últimos 12 años ha habido un cambio en la imagen médica debido al desarrollo de la radiómica, una técnica creada para resolver problemáticas en cáncer, aunque actualmente se utiliza para diferentes patologías. Consiste en la extracción de alto rendimiento de características cuantitativas de las imágenes médicas (TC, RM, PET-TC), considerando las imágenes como datos, con el objetivo principal de explotar estas características para predecir un resultado clínico o diagnóstico al extraer el fenotipo de la imagen (63,64). Es capaz de extraer características no visibles al ojo humano, siendo su principal ventaja poder visualizar de manera no invasiva y global la apariencia de un cáncer, lo que le permite capturar la heterogeneidad tumoral macroscópica, a diferencia de una biopsia convencional, con el potencial de informar sobre un diagnóstico, agresividad, respuesta a tratamiento y pronóstico. Por otra parte es capaz de determinar la expresión del genotipo, el microambiente tumoral y la susceptibilidad del cáncer a los tratamientos (65,66), lo que se denomina radiogenómica.

Trabaja con un gran conjunto de datos, debido a que se pueden analizar miles de características para buscar la más representativa de la firma radiómica que estamos creando, a esto se le llama datos minables (en inglés “mineable” or “data mining”). Usar conjuntos de datos suficientemente grandes, puede usarse para descubrir marcadores y patrones previamente desconocidos de evolución de la enfermedad, progresión y respuesta al tratamiento.

La IA con sus métodos de aprendizaje automático se aplica de manera eficiente a la radiómica para construir, entrenar y validar modelos que pueden ayudar en la predicción de enfermedades y resultados del tratamiento, así como en la estratificación de pacientes, a lo cual llamamos medicina de precisión (67). Por otra parte la IA potencia a la radiómica al permitirle combinarse con datos clínicos, de laboratorio, histológicos, genómicos u otros, logrando la creación de modelos integrativos.

La IA se inicia en 1956 con la invención de los robots, y consiste en el uso de una computadora para modelar comportamientos inteligentes con una intervención humana

mínima (68,69). La Asociación Médica Americana definió recientemente el papel de la IA en la atención médica como "inteligencia aumentada", afirmando que la IA se diseñará y utilizará para mejorar la inteligencia humana en lugar de reemplazarla. Una subdisciplina de la IA es el Machine Learning o aprendizaje automático, técnica que abarca todos los enfoques que permiten que las computadoras aprendan a partir de varios tipos de datos. Por otra parte, el Deep Learning o aprendizaje profundo es un subtipo de Machine Learning basado en redes neuronales artificiales multicapa para resolver problemas altamente complejos (56,70). Tanto el Machine Learning como el Deep Learning se han aplicado para potenciar a la radiómica, utilizándose concretamente para modelar características multi-vista con el fin de predecir resultados clínicos.

La forma de trabajo en radiómica consiste en varias fases (Figura 5):

1. Definir una pregunta clínica y seleccionar la cohorte de pacientes adecuada.

2. Optimizar y estandarizar los protocolos de adquisición de la técnica de imagen que analizaremos.

3. Pre-procesamiento de imagen.

4. Segmentación de lesiones en las imágenes, identificando las zonas de interés.

5. Extraer características de imagen, ya sea mediante métodos manuales o de aprendizaje profundo.

6. Reducir la dimensionalidad de los datos generados mediante métodos de selección de características.

7. Entrenar y validar el modelo radiómico.

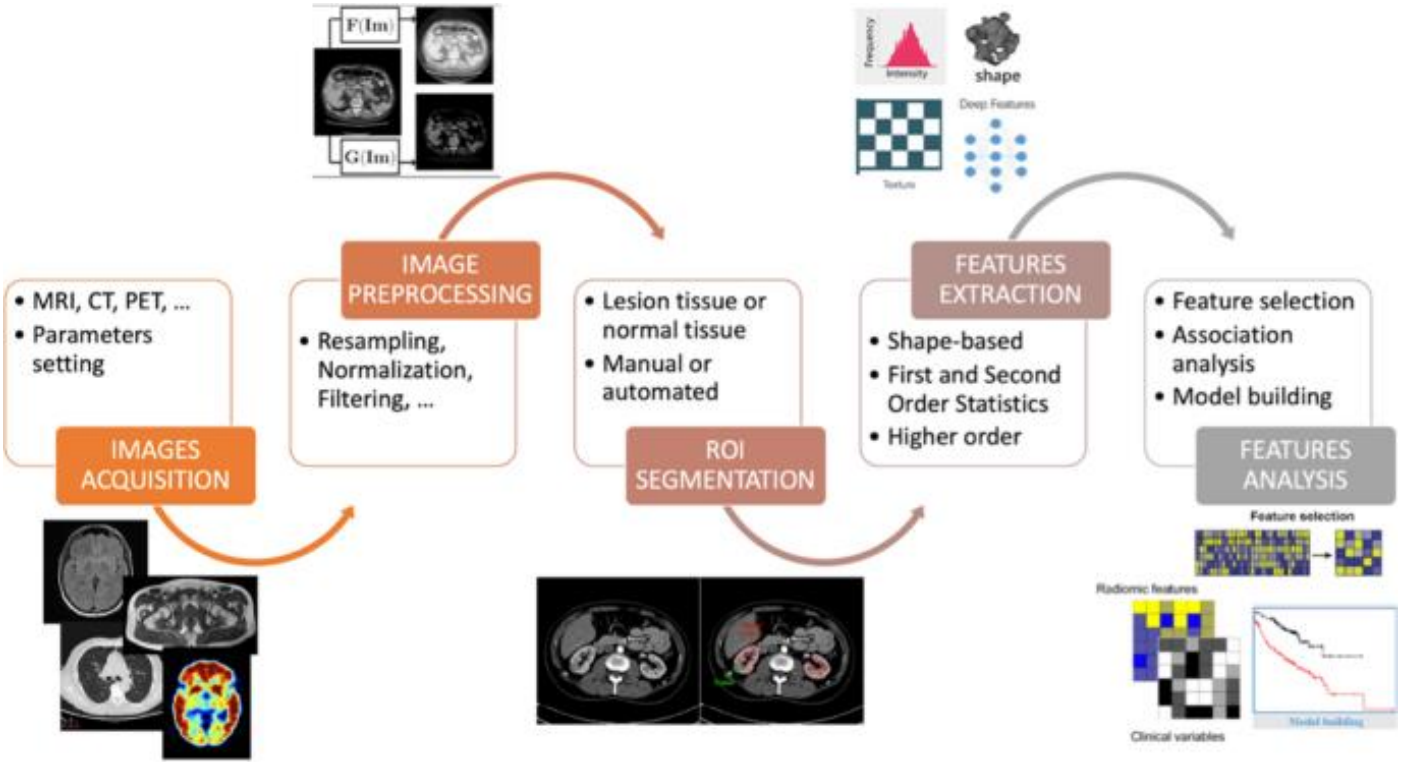


Figura 5: Flujo de trabajo en radiómica (71)

Respecto a las características radiómicas, estas se extraen de regiones de interés (ROI) bidimensionales (2D) o tridimensionales (3D). Estas características son centenares y se agrupan basadas en el tamaño y la forma; descriptores del histograma de intensidad de la imagen; descriptores de las relaciones entre voxels; texturas derivadas; texturas extraídas de imágenes filtradas; y características fractales (61,72). Las características radiómicas con frecuencia muestran altas correlaciones, lo que indica redundancia de datos, lo que significa que algunas características pueden descartarse, agruparse, y/o reemplazarse con una característica representativa, mediante el uso de análisis de componentes principales o análisis discriminante lineal. Se debe tener cuidado en este proceso para evitar el llamado sobreajuste (en inglés “overfitting”), que conduce a estimaciones demasiado optimistas de las precisiones predictivas. Si el número de características es lo suficientemente alto, se pueden detectar correlaciones incluso en datos aleatorios.

Encontrar las características de imagen más relevantes para el problema clínico que queremos desarrollar puede convertirse en un desafío, y es en este punto en que podemos utilizar técnicas de IA . Una capacidad clave de Deep Learning es la explotación de grupos de predictores o características tomadas en conjunto, llamados patrones multivariados. Los métodos de bagging y boosting (métodos de agregación), como los llamados bosques aleatorios (en inglés “random forest”), han permitido el entrenamiento de clasificadores o regresores robustos al integrar la selección de características y el muestreo efectivo en el proceso de entrenamiento para lograr modelos predictivos. Por otra parte, contamos con las redes neuronales convolucionales que han resurgido como modelos potentes de clasificación. Dado un conjunto de datos de entrenamiento suficiente, las redes neuronales convolucionales superan los esquemas de selección de características, ya que en lugar de seleccionar características de conjuntos finitos y predefinidos, construyen características óptimas a partir de los propios datos de imagen. Una limitación es la necesidad de una cantidad excesiva de datos de entrenamiento. Dado que muchos de los enfoques tienen parámetros, las curvas ROC se utilizan para representar la precisión de la predicción a lo largo de un rango de parámetros.

A pesar de las ventajas evidentes de esta tecnología, existen desafíos por resolver. Uno de los obstáculos más importantes para obtener una imagen que se pueda analizar correctamente de forma cuantitativa, en es el hecho de que los TC actuales están diseñados para producir imágenes visualmente atractivas en lugar de producirlas como dispositivos de medición (73). Además existe una capacidad limitada para estandarizar la forma de adquirir datos de imágenes lo que dificulta su validación en cohortes externas. Según una revisión sistemática de Traverso et al (74), la repetibilidad y reproducibilidad de las características radiómicas son sensibles en diversos grados a detalles de procesamiento como la configuración de adquisición de imágenes, el algoritmo de reconstrucción de imágenes, el preprocesamiento digital de imágenes y el software utilizado para extraer características radiómicas. Otro factor por considerar es la necesidad de un número elevado de datos (imágenes) para entrenar y desarrollar un algoritmo, lo cual requiere de estudios multicéntricos con las complicaciones que conlleva compartir imágenes de forma segura para resguardar la confidencialidad de datos del paciente.

Finalmente, a pesar de los desafíos por resolver, contamos con una tecnología potente que es capaz de explotar al máximo los datos de imagen médica de una manera objetiva, lo que nos permite y nos permitirá aún más en un futuro, resolver diferentes problemas médicos como es en el caso del TEP en pacientes con y sin neumonía por COVID-19, y también el diagnóstico del nódulo pulmonar.

2. JUSTIFICACIÓN

La patología respiratoria tiene relevancia clínica por ser clave en la salud global de los pacientes y por su elevada frecuencia. Ha cobrado especial interés tras la pandemia de COVID-19 que ocasionó millones de muertes en todo el mundo por neumonía grave y en muchos casos enfermedad tromboembólica como el TEP. Mucho antes de la pandemia, aunque posiblemente con menos visibilidad global, el cáncer de pulmón ya era una patología altamente relevante debido a su incidencia y sobre todo a su mortalidad, ya que es el cáncer que causa más muertes en todo el mundo.

El diagnóstico de las patologías respiratorias ha avanzado en los últimos años debido en gran medida al desarrollo y evolución de la imagen médica, sin embargo, existe un amplio margen de mejora para aproximarnos en mayor medida a un diagnóstico no invasivo que se asemeje a la información que nos puede dar una biopsia anatomopatológica.

Nuevas tecnologías en el análisis de la imagen médica como es la radiómica, potenciada por técnicas de inteligencia artificial (IA) como el Machine Learning y el Deep Learning, están en auge en la última década debido a que están demostrando utilidad para predecir el diagnóstico, tratamiento y pronóstico sobre todo en patología neoplásica, aunque actualmente también se abren paso en otras patologías.

Durante la pandemia de COVID-19 se observó un aumento de casos de TEP en pacientes que ingresaban por una neumonía por COVID-19, lo que hizo que aumentara la cantidad de pruebas diagnósticas solicitadas para esta patología. En los pacientes en que no era posible realizar un AngioTC por contraindicación o falta de rapidez en la realización de la prueba por sobrecarga asistencia, se debía realizar SPECT/TC-V/Q. Debido al riesgo de contagio con las técnicas de ventilación se recomendaba realizar SPECT/TC-Q con las dificultades que suponía no disponer de la ventilación para la valoración de esta prueba. Es por esto que surgió la idea de desarrollar una firma radiómica potenciada por IA para detectar patrones de imagen de TEP o neumonía en pacientes con COVID-19 basándose únicamente en SPECT/TC-Q. Para ello, planteamos un enfoque combinado que utiliza

características radiómicas extraídas de cada imagen y una red de clasificación entrenada basada en el análisis de una región de interés (análisis local) de las imágenes.

Por otra parte, teniendo en cuenta la importancia de optimizar la detección precoz del cáncer de pulmón como herramienta para disminuir la mortalidad por esta patología, y el potencial que tiene la radiómica en su diagnóstico, surge la idea de ir más allá en diagnóstico de imagen e incorporar datos clínicos relacionados con la enfermedad para poder desarrollar un algoritmo capaz de predecir malignidad en nódulos pulmonares detectados por vía convencional (nódulo pulmonar incidental) o por medio de programas de cribado de cáncer de pulmón.

La finalidad de esta tesis doctoral es demostrar que la radiómica potenciada por IA es capaz de mejorar el diagnóstico por imagen de patologías respiratorias relevantes en la clínica diaria como son el diagnóstico de TEP y cáncer de pulmón, para acercar las nuevas tecnologías a la práctica asistencial.

3. HIPÓTESIS

La radiómica potenciada por la inteligencia artificial, es capaz de mejorar el diagnóstico por imagen de patologías pulmonares relevantes.

4. OBJETIVOS

4.1. Objetivo principal

Desarrollar modelos radiómicos potenciadas por inteligencia artificial, para la creación de algoritmos que optimicen el diagnóstico de TEP en neumonía por COVID-19 y de nódulo pulmonar.

4.2. Objetivos específicos

- Clasificar e identificar diferentes tipos de tejido pulmonar mediante un análisis por regiones de las imágenes de SPECT-TC-Q.
- Desarrollar un modelo predictivo para TEP que sea capaz de clasificar diferentes imágenes de SPECT/TC-Q para poder dar un diagnóstico de cada región de interés en pacientes con y sin neumonía por COVID-19.
- Crear un modelo radiómico asociado a técnicas de Deep Learning para el diagnóstico de malignidad de nódulos pulmonares.
- Detectar las características clínicas relevantes que sean capaces de optimizar los resultados de un modelo radiómico y valorar el aporte de la integración de ambos modelos para el diagnóstico de nódulos pulmonares.

5. COMPENDIO DE PUBLICACIONES

5.1. Artículo I

Título: A novel intelligent radiomic analysis of perfusion SPECT/CT images to optimize pulmonary embolism diagnosis in COVID-19 patients

Autores: Sonia Baeza, Debora Gil, Ignasi Garcia-Olivé, Maite Salcedo-Pujantell, Jordi Deportós, Carles Sanchez, Guillermo Torres, Gloria Moragas y Antoni Rosell.

Revista: EJNMMI Physics

Volumen: 9

Año:2022

Factor de impacto: 4

Cuartil: 1 (Q1)

DOI: https://doi.org/10.1186/s40658-022-00510-x

Citar como: Baeza S, Gil D, Garcia-Olivé I, et al. A novel intelligent radiomic analysis of perfusion SPECT/CT images to optimize pulmonary embolism diagnosis in COVID-19 patients [published correction appears in EJNMMI Phys. 2023 Feb 17;10(1):13. doi: 10.1186/s40658-023-00532-z]. EJNMMI Phys. 2022;9(1):84. Published 2022 Dec 5. doi:10.1186/s40658-022-00510-x

Baeza *et al. EJNMMI Physics* 2022, **9**(1):84
https://doi.org/10.1186/s40658-022-00510-x

EJNMMI Physics

ORIGINAL RESEARCH Open Access

A novel intelligent radiomic analysis of perfusion SPECT/CT images to optimize pulmonary embolism diagnosis in COVID-19 patients

Sonia Baeza[1,2,3*], Debora Gil[4], Ignasi Garcia-Olivé[1,2,3], Maite Salcedo-Pujantell[5], Jordi Deportós[5], Carles Sanchez[4], Guillermo Torres[4], Gloria Moragas[5] and Antoni Rosell[1,2,3]

*Correspondence: smbaeza.germanstrias@gencat.cat

[1] Respiratory Medicine Department, Hospital Universitari Germans Trias I Pujol, Badalona, Barcelona, Spain
[2] Germans Trias I Pujol Research Institute (IGTP), Badalona, Barcelona, Spain
[3] Departament de Medicina, Universitat Autónoma de Barcelona, Barcelona, Spain
[4] Computer Vision Center and Computer Science Department, UAB, Barcelona, Spain
[5] Nuclear Medicine Department, Hospital Universitari Germans Trias I Pujol, Badalona, Barcelona, Spain

Abstract

Background: COVID-19 infection, especially in cases with pneumonia, is associated with a high rate of pulmonary embolism (PE). In patients with contraindications for CT pulmonary angiography (CTPA) or non-diagnostic CTPA, perfusion single-photon emission computed tomography/computed tomography (Q-SPECT/CT) is a diagnostic alternative. The goal of this study is to develop a radiomic diagnostic system to detect PE based only on the analysis of Q-SPECT/CT scans.

Methods: This radiomic diagnostic system is based on a local analysis of Q-SPECT/CT volumes that includes both CT and Q-SPECT values for each volume point. We present a combined approach that uses radiomic features extracted from each scan as input into a fully connected classification neural network that optimizes a weighted cross-entropy loss trained to discriminate between three different types of image patterns (pixel sample level): healthy lungs (control group), PE and pneumonia. Four types of models using different configuration of parameters were tested.

Results: The proposed radiomic diagnostic system was trained on 20 patients (4,927 sets of samples of three types of image patterns) and validated in a group of 39 patients (4,410 sets of samples of three types of image patterns). In the training group, COVID-19 infection corresponded to 45% of the cases and 51.28% in the test group. In the test group, the best model for determining different types of image patterns with PE presented a sensitivity, specificity, positive predictive value and negative predictive value of 75.1%, 98.2%, 88.9% and 95.4%, respectively. The best model for detecting pneumonia presented a sensitivity, specificity, positive predictive value and negative predictive value of 94.1%, 93.6%, 85.2% and 97.6%, respectively. The area under the curve (AUC) was 0.92 for PE and 0.91 for pneumonia. When the results obtained at the pixel sample level are aggregated into regions of interest, the sensitivity of the PE increases to 85%, and all metrics improve for pneumonia.

Conclusion: This radiomic diagnostic system was able to identify the different lung imaging patterns and is a first step toward a comprehensive intelligent radiomic system to optimize the diagnosis of PE by Q-SPECT/CT.

Highlights: Artificial intelligence applied to Q-SPECT/CT is a diagnostic option in patients with contraindications to CTPA or a non-diagnostic test in times of COVID-19.

Keywords: Pulmonary embolism, SPECT, CT, COVID-19, Radiomics, Neural networks

Background

Although CT pulmonary angiography (CTPA) is the most widely used imaging test for diagnosing pulmonary embolism (PE) [1, 2], another test is needed for patients with contraindications (allergy to iodinated contrast, kidney failure) or non-diagnostic on CTPA. Ventilation/perfusion single-photon emission computed tomography/computed tomography SPECT/CT (V/P-SPECT/CT) is a validated technique for PE diagnosis even in the presence of pneumonia [3, 4], but during the COVID-19 pandemic the use of ventilation has been discouraged due to the high risk of aerosol production [2, 5].

Thromboembolic disease has been a major complication in patients with COVID-19 pneumonia. In severe cases of COVID-19, an association with hyperinflammatory syndrome [6, 7] has been reported, as well as coagulation abnormalities and thrombosis [8–10]. Venous and arterial thromboembolic events occur in 31–59% of hospitalized patients with COVID-19 [8, 11], especially if they require critical care [12–16]. Diagnosis of PE in the latter situation can be challenging [17].

Due to the need to adapt the test to the safety standards required by the pandemic and in search of an alternative for patients for whom CTPA is not an option, it is worth considering perfusion single-photon emission computed tomography/computed tomography (Q-SPECT/CT) as a diagnosis alternative [18–21]. The value of Q-SPECT/CT was reported even prior to the COVID-19 pandemic [22], showing a sensitivity and specificity of 100% and 83%, respectively, for the diagnosis of PE with the Q-SPECT/CT technique [23]. Nevertheless, the diagnosis of PE can be difficult in large areas of pneumonia. In these cases, reporting takes longer and frequently needs to be validated by senior nuclear physicians to prevent errors.

The early detection of COVID-19 by medical imaging has aroused great interest within the artificial intelligence community. However, to the best of our knowledge, to date there are no published studies that use artificial intelligence to improve the diagnosis of COVID-19 PE through Q-SPECT/CT.

An artificial intelligence system to support physicians in the diagnosis of PE through Q-SPECT/CT would improve the specificity of PE diagnosis in cases of associated pneumonia and reduce the reading time of the test, and for this reason, we have decided to apply radiomics and artificial intelligence techniques as an innovative alternative that responds to the requirement of an alternative test to the CTPA to rule out PE during the COVID-19 pandemic.

Existing methods addressing the diagnosis of COVID-19 pathologies focus mainly on the detection of COVID-19 pneumonia from an analysis performed through a single modality (either X-ray or CT scans). Most methods are deep learning approaches based on well-known architectures that have proved successful in other fields (like ResNet [24], U-net [25] or EfficientNet [26], among others) that have been adapted and fine-tuned for managing COVID-19 diagnosis.

Regardless of the imaging modality and architecture, the usual approach is to use a classification scheme that provides a single diagnosis for each image/scan [27, 28]. Patients with severe COVID-19 have several lung image patterns at the same time. A main challenge is the accurate localization inside the lungs of pathologies affecting a small percentage of lung tissue (as is the case with PE).

The objective of the present work is to seek the best radiomic signature to detect image patterns of PE or pneumonia in patients with COVID-19 based only on Q-SPECT/CT. For that purpose, we present a combined approach that uses radiomic characteristics extracted from each scan and a classification network trained from scratch based on an analysis of a segmented region of interest (local analysis) of the images.

Material and methods

Patient recruitment

Patients were recruited from Hospital Universitari Germans Trias i Pujol (HUGTiP), a university referral hospital covering an area of 800,000 inhabitants in Barcelona (Spain).

The Q-SPECT/CT protocol is prospective and all COVID-19 patients (positive RT-PCR) were collected prospectively in 2020 (April to September 2020). In order to have a control group of patients not affected by COVID-19, we reviewed patients from a 2018 database (January–December 2018) and selected a retrospective group to ensure that they had not presented with a COVID-19 infection.

The cases were divided into two subsets, one of which was used to perform the algorithm training (training group) and the other to perform the validation (test group).

The study's exclusion criteria were the patient's refusal to participate in the study, and the detection of severe alterations in the patients' lungs caused by other pathologies unrelated to COVID-19 infection such as severe emphysema, bullae or interstitial lung diseases.

This study was performed in accordance with the principles of the Declaration of Helsinki. The research protocol was approved by the regional ethics committee (Ethics Committee for Clinical Research of Germans Trias i Pujol University Hospital), with the reference PI-20–161, and subjects gave their written informed consent.

Demographics

The total number of patients included was 61, from which two cases were discarded due to technical failures in the image acquisition process or to the presence of severe abnormalities in the patients' lungs caused by pathologies other than COVID-19.

Thus, the resulting database contained data from 59 patients, 20 in the training group (4,927 sets of samples of three types of image patterns) and 39 in the test group (4,410 sets of samples of three types of image patterns). In the training group, COVID-19 infection corresponded to 45% of the cases and to 51.28% in the test group. The demographic characteristics of each group are given in Table 1.

Table 1 Demographic data–patients

	Training set	Test set
Patients (*n*)	20	39
Age (average, years, SD)	60	68
Gender female, *n* (%)	12 (60%)	18 (46.15%)
COVID-19, *n* (%)	9 (45%)	20 (51.28%)
Pneumonia (*n*)	4	9
Pneumonia + PE (*n*) †	3	10
PE (*n*)	4	10
Healthy lungs	6	10
Extra categories*	3	0

* Extra categories: areas with abnormal CT and low perfusion; black background (areas with no tissue uptake); body tissue (areas of the body not belonging to the lungs)

† All cases registered with Pneumonia + PE were cases with COVID-19

Image data acquisition

A perfusion single-photon emission computed tomography/computed tomography (Q-SPECT/CT) based on the intravenous administration of 6 mCi (222 MBq) of 99mCt-macroaggregates of human albumin (99mCt-MAA) was acquired for each patient, with the subsequent acquisition of a tomo-scintigraphy (SPECT) and a CT in two hybrid devices indistinctly: a Symbia T2 Gamma camera (brand Siemens, based in Munich, Germany) and a Discovery NM/CT 670 ES Gamma camera (brand General Electric, based in Boston, Massachusetts, US).

The acquisition parameters were the following. The SPECT was obtained with a circular orbit with a 360º arc, 128×128 matrix, zoom 1, 140 keV photopeak, obtaining 90 images of eight seconds per image. The acquired CTs used 120 kV, 50–350 mA, with a slice thickness and interval of 1.25 mm (General Electric) and 3 mm (Siemens). With reconstructions of B41S, B80S, B08 and 1 soft, recon 2 lung, respectively, for CT and three-dimensional reconstruction and the format with attenuation correction for scintigraphy were used, 512×512 matrix.

Ventilation lung scintigraphy was not allowed due to the risk of COVID-19 cross-contamination, and ventilation alterations were therefore determined only by CT scans.

The images were extracted from PACS in DICOM format and were anonymized to guarantee the confidentiality of the data.

All the CT images obtained for the Q-SPECT/TC were reviewed by a pulmonologist, nuclear medicine physician or radiologist with more than seven-year experience, and their opinion of the expert clinicians was considered the gold standard. These experts segmented areas of healthy lung, PE or pneumonia in Q-SPECT/CT images from each patient using in-house software. For each patient scan, an expert annotated 2D regions containing pixels of one of the target tissues (healthy lungs, TEP or pneumonia). In the cases with COVID-19 pneumonia and PE, we selected and noted areas in which only one of the image patterns existed. 2D regions were identified by a single reader. The average number of 2D ROIs annotated for each patient was 25. The number of pixels of each ROI was 473.4738 on average, with a minimum of 48 and a maximum of 2,964.

Radiomic features were computed pixel-wise using the values of neighboring pixels contained in windows centered at each pixel in the 2D region. Each pixel in the 2D region such that a window of side equal to *sze* (in our case 3, 5 pixels) that is contained in the region is a sample to be used for either training or test. The analysis was performed upon these multiple samples and not on the final diagnosis of the patient. Since the number of samples is given by the pixels whose window is contained in the annotated ROIs, we have a different number of samples for each window size. In order to perform the statistical analysis, the data obtained for windows of size 3×3 pixels was resampled to match the samples obtained for the windows of size 5×5 pixels. The training group had 4,927 image pattern samples (2,096 healthy lung patterns, 1,204 pulmonary embolism patterns and 1,627 pneumonia patterns), while the test group had 4,410 (2.442 healthy lung patterns, 597 pulmonary embolism patterns and 1,371 pneumonia patterns) (Fig. 1).

Image processing

Since the final diagnosis requires a combination of information from both the SPECT and CT scans, the first step consists of the registration of the two volumes in order to fuse both image modalities. In order to register volumes, first SPECT volumes were resized to match the same number of CT slices using HOROS, an open-source medical image viewer (Horos Project). Then, we used an affine unimodal [29]

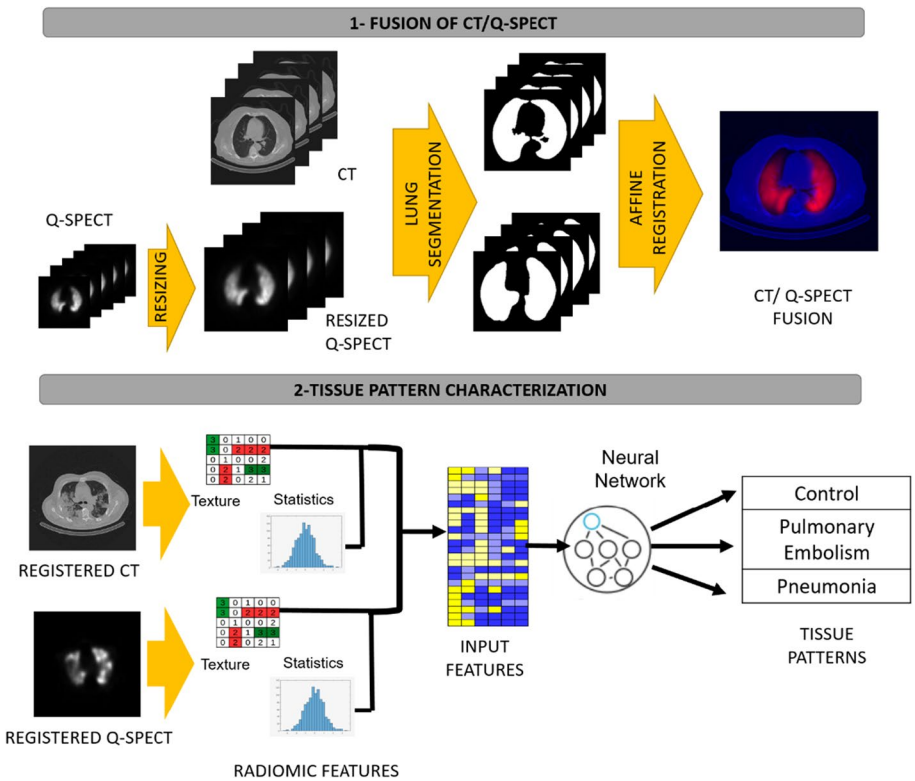


Fig. 1 Illustrative flowchart of the stages of the system pipeline

transformation to register a segmentation of the lungs in CT and SPECT resized volumes. By registering binary masks instead of intensity volumes, we can account for multimodal differences using the mean squared error as a cost function and thus minimize the risk of premature convergence associated with multimodal approaches using mutual information. The computed transformation was applied to intensity SPECT volumes to register them to CT scans.

The segmentation of lungs was computed using thresholding and morphological operations. CT lungs were selected as the largest connected component of the voxels with an intensity between 950 and -300 Hounsfield Units. The thresholded volumes undergo a post-processing step called morphological closing [30] that fills in interior holes of the lung mask and small concave entries at the boundary of the lung mask. The size and shape of the filled mask is given by a binary volume representing a 3D shape called a structuring element. In our case, the structuring element is a sphere of radius 5 (also refer to as size among the image processing community).

For the lung segmentation of the perfusion volumes, a threshold of intensity of 20 was selected. An intensity of 20 is over the SPECT values which reflect the amount of radioactive tracer absorbed by the lung tissue which reflects how blood is flowing within the lungs. Given that background pixels do not have a contrast agent, their baseline value is very low. Thus, voxels with an intensity over 20 were considered to be part of the lung.

The quality and interpretability of segmentations was performed visually according to the quality of the registered volumes.

In order to account for differences in SPECT intensities due to variations in contrast agent concentrations, SPECT scans were normalized using two different approaches. The normalization step was performed after the registration of volumes.

The first approach normalized SPECT values in the range [0, 1], using the maximum and minimum values of the intensities of each scan:

$$\frac{\text{SPECT} - \min(\text{SPECT})}{\max(\text{SPECT}) - \min(\text{SPECT})}$$

SPECT might have spots of high concentration of the agent that are not associated with a higher lung perfusion. Those artifacts associated with the technique itself deviate maximum intensity values and introduce a drop in the intensity of normalized volumes not related to a drop in perfusion. In order to alleviate the impact of this artifact in the predictions of PE, we have also applied a normalization based on the superior quantile of SPECT intensity values:

$$\frac{\text{SPECT} - \min(\text{SPECT})}{\text{percentile}(\text{SPECT}, 99) - \min(\text{SPECT})}$$

For percentile(SPECT, 99), denoting the percentile leaving a right probability of 0.01 in the distribution of SPECT intensity values.

Intelligent radiomics for the detection of pulmonary embolism

The proposed radiomic system analyzes the intensity values of CT and SPECT volumes for each voxel in order to discriminate the three clinically relevant types of image patterns defined by pulmonologist and nuclear medicine experts:

1. Control: healthy—normal CT and SPECT.
2. Pulmonary Embolism (PE): normal CT with a localized perfusion defect in the SPECT.
3. Pneumonia: abnormal CT with normal perfusion or with an atypical perfusion defect that does not meet PE criteria.

In order to avoid overfitting, three extra categories were added:

4. Areas with abnormal CT and low perfusion.
5. Black Background: areas with no tissue uptake.
6. Body Tissue: areas of the body not belonging to the lungs.

Our intelligent radiomic system was applied in three main steps. Firstly, scans were preprocessed as described in the "Image processing" section. Secondly, radiomic features selected according to their reproducibility were extracted from the registered volumes to define a radiomic feature space. Finally, a machine learning method was used to disseminate each value of the feature space between the four types of image patterns.

The radiomic features are a subset of PyRadiomics [31], an open-source python package for the extraction of Radiomic features from medical imaging volumes. PyRadiomics features include shape features, first-order features, and textural features (Gray Level Co-occurrence Matrix (GLCM), Gray Level Size Zone (GLSZM), Gray Level Run Length Matrix (GLRLM) and Gray Level Dependency Matrix (GLDM)) describing several aspects of the lesion. The subset was selected according to reproducibility against different image acquisition conditions and interobserver variability in lesion identification. Reproducibility is based on the correlation of feature values obtained from data collected using different conditions and settings [32]. The selected set of (17) features are given in Table 2.

Radiomic features are computed pixel-wise using the values of pixels in the neighborhood of each pixel. These neighborhoods are squared areas (called windows) of side equal to sze (in our case 3, 5) pixels centered at each pixel of the annotated ROI. That is, for each pixel of coordinates (i0,j0) a window, w(i0,j0), of side sze is given by the pixels of coordinates (i,j) that satisfy:

w(i0,j0): = ROI(i,j), for i Î{i0-sze, i0-sze + 1, …, i0 + sze-1, i0 + sze}, j Î{j0-sze, j0-sze + 1, …, j0 + sze-1, j0 + sze}.

being ROI(i,j) the image area annotated by the expert. Given the formulation of the window used to compute radiomic features, we observe that radiomic features could

Table 2 Features selected according to reproducibility

First order	Texture GLCM	Texture GLDM	Texture LGRM
Entropy	Id	Dependence Non-Uniformity	Gray Level Non-Uniformity Normalized
TotalEnergy	Idm	Normalized	Run Length Non-Uniformity Normalized
Uniformity	JoinEnergy	Dependence Variance	Run Percentage
	MaximumProbability	Large Dependence Emphasis	Short Run Emphasis

only be computed for pixels such that its neighboring window pixels are also inside the ROI. We call this set of pixels samples used for both training and testing.

The radiomic features computed for each sample, using the values of the neighboring pixels contained in a square window of side sze are the input for a fully connected network which combines them in a multi-classification approach. The input for the network is the concatenation of the radiomic features extracted for the two pre-processed scans. The network has two fully connected layers with 128 neurons linked with one Relu layer and an output classification layer with sigmoid activation. To account for unbalancing in the training data, the loss function is the weighted cross-entropy given by:

$$\text{Loss} = \sum_i w_i \text{Loss}_i$$

where Loss_i is the cross-entropy loss for the i-th class and the weights, w_i are given by the inverse of the class frequency, w_i, normalized to sum one:

$$w_i = \frac{\rho_i}{\sum_j \rho_j}$$

for $\rho_i = N\text{Samp}_i/N\text{Samp}$, $N\text{Samp}_i$ and$N$Sampbeing, respectively, the number of samples of the i-th class and the total number of samples.

Experiments

In order to assess the impact of the size of the window used to compute radiomic features, we trained two different model with features extracted using windows of size, sze x sze$=3\times3$ and sze x sze$=5\times5$, to compute the pyradiomics texture feature in the neighborhood of each sample. To analyze the effect of the normalization of SPECT scans, a different model was trained with data normalized using the maximum and percentile criteria for each window size. A total number of four models were trained using the configurations reported in Table 3. From now on, models will be referred to using the labels given in Table 3.

Two different experiments were carried out:

1. *Training and Selection of Models*. A leave-1-out validation on a training set of samples from patients to select the best model for the detection of PE and pneumonia.

Table 3 Configuration of the different radiomic models

	SPECT Normalization	
Window size	**Max**	**Percentile**
3 × 3	ModelMx3: SPECT normalization based on maximum values with 3 × 3 windows	ModelPrct3: SPECT normalization based on upper percentile values with 3 × 3 windows
5 × 5	ModelMx5: SPECT normalization based on maximum values with 5 × 5 windows	ModelPrct5: SPECT normalization based on upper percentile values with 5 × 5 windows

2. *Testing and Assessment of Models Reproducibility*. Validation of the best models on an independent set of samples from test patients to assess the reproducibility of results

Statistical analysis

All calculations were conducted in version 15 of STATA. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), area under the curve (AUC) and diagnostic accuracy were calculated using the *dt* command. The equality of ROC areas was tested using the *rocgold* command. This command independently tests the equality of the ROC area of each of several test modalities against a "gold standard" ROC curve. For each comparison, *rocgold* reports the raw and the Bonferroni-adjusted significance probability. A p-value of 0.05 was considered significant. The computation of the power and effect size of the tests was conducted using G*Power 3.1.9.2.

Results

Training and selection of models

Table 4 provides a statistical summary of the AUC score for the training set used for the selection of models. For each radiomic model (rows) and diagnoses (columns), we report the AUC, its 95% confidence intervals (CI) and p-values for the comparison of AUC. The *P*-value was adjusted with the Bonferroni correction. The blank cells with * correspond to the best AUC used to make the comparison.

For the diagnosis of PE, models normalized using the upper percentile performed significantly better (with AUC > 0.9) than those normalized using maximum values, with ModelPrct3 performing significantly better than ModelPrct5. For the diagnosis of pneumonia, models using the upper percentile normalization also performed significantly better (with AUC > 0.9). In this case, although the analysis detected that ModelPrct3 AUC was significantly higher than ModelPrct5 AUC, the analysis of CIs

Table 4 AUC statistics for the leave-1-out training for model selection

	Pulmonary Embolism			Pneumonia			Healthy Lung		
Radiomic model	**AUC**	**95% CI**	***P*-value***	**AUC**	**95% CI**	***P*-value***	**AUC**	**95% CI**	***P*-value**
ModelPrct3	0.919	0.909–0.929	*	0.922	0.914–0.931	*	0.971	0.966–0.977	0.09
ModelPrct5	0.902	0.892–0.913	0.028	0.905	0.896–0.914	< 0.001	0.978	0.973–0.983	*
ModelMx3	0.835	0.822–0.848	< 0.001	0.825	0.814–0.837	< 0.001	0.928	0.920–0.937	< 0.001
ModelMx5	0.823	0.810–0.836	< 0.001	0.818	0.806–0.830	< 0.001	0.935	0.928–0.943	< 0.001

Table 5 Sensitivity, specificity, PPV and NPV statistics for the leave-1-out training of the best AUC model

	Pulmonary embolism	Pneumonia	Healthy lung
Sensitivity	89.6% (95% CI 87.7–91.2)	92.5% (95% CI 91.1–93.7)	98.1% (95% CI 97.3–98.7)
Specificity	95.8% (95% CI 92.5–96.4)	94.2% (95% CI 93.4–94.9)	97.5% (95% CI 97.0–98.0)
Positive predictive value	87.0% (95% CI 85.0–88.8)	88.0% (95% CI 86.3–89.4)	95.2% (95% CI 94.1–96.1)
Negative predictive value	96.7% (95% CI 96.1–97.2)	96.5% (95% CI 95.8–97.1)	99.0% (95% CI 98.6–99.3)

suggests that this difference is not very large. In fact, according to the effect size associated with the sample size of the training set, this difference is at most 0.05. For both PE and pneumonia diagnosis, the AUC of ModelPrct3 was above 0.91, which confers percentile-based models a high diagnostic value. Regarding healthy lung, models normalized using the upper percentile performed better (AUC > 0.97), with no significant differences between them.

Table 5 reports the sensitivity, specificity, positive predictive value, and negative predictive value for ModelPrct3, and we also report the average values and 95% confidence intervals (CI) of each score. Sensitivity for PE, pneumonia and healthy lung were 90% or higher with a PPV of over 87%. Specificity and NPV were over 94% and 96%, respectively (Figs. 2, 3, 4).

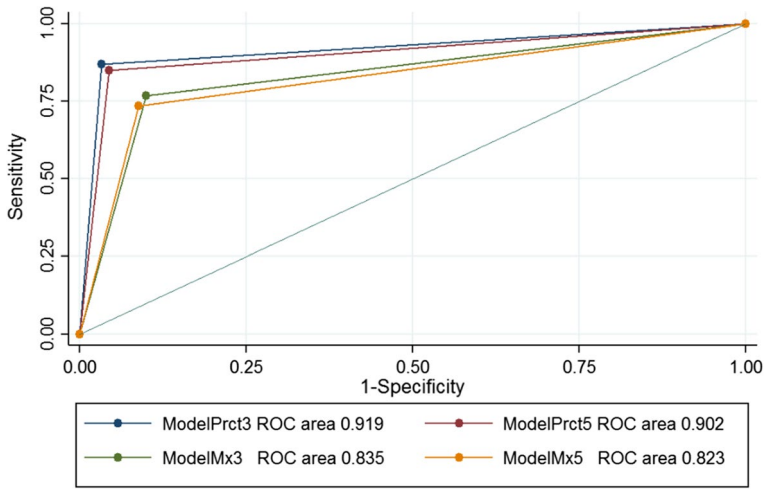


Fig. 2 Pulmonary embolism ROC curve in the training set

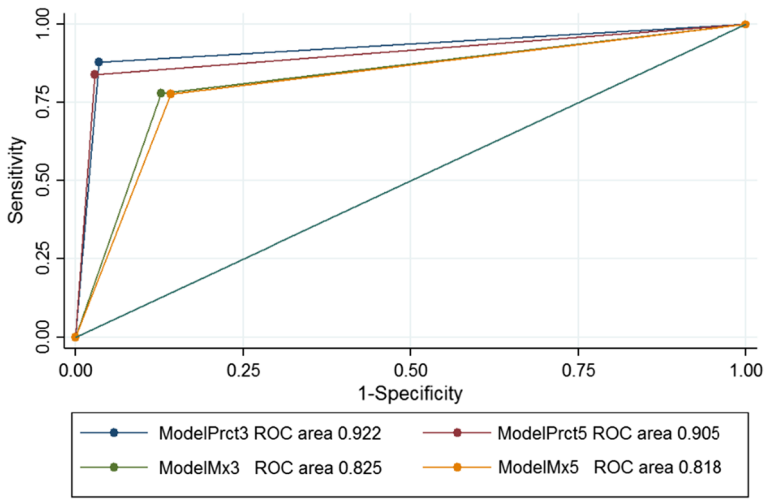


Fig. 3 Pneumonia ROC curve in the training set

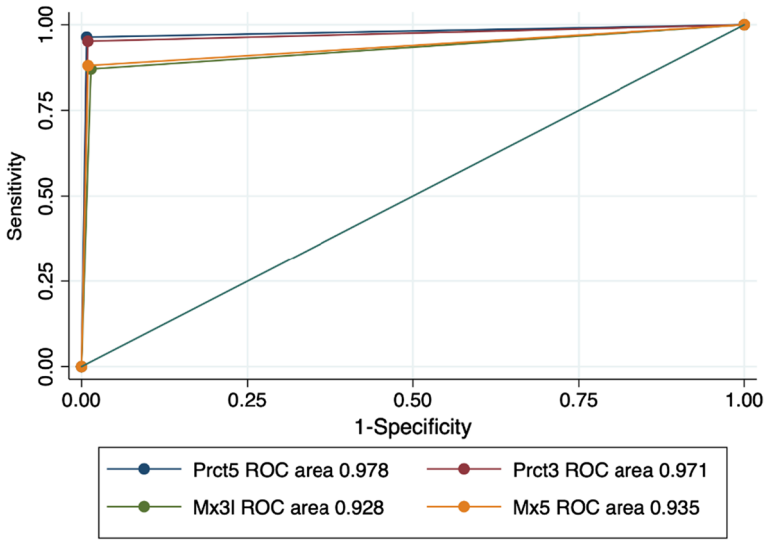


Fig. 4 Healthy lung ROC curve in the training set

Test group

The AUC for ModelPrct3 in the test group is 0.922 (95% CI 0.090–0.935) for PE, 0.906 (95% CI 0.896–0.916) for pneumonia and 0.908 (95% CI 0.900–0.917) for healthy lung, which proves that the model predictions are reproducible.

Table 6 shows the sensitivity, specificity, positive predictive value and negative predictive value for ModelPrct3, and we also report the average values and 95% confidence intervals (CI) of each score. Values for pneumonia diagnosis are almost the same as for the training set. In the case of healthy lung, although there is a decrease in the metrics in relation to training, the values are still around 90%. Values for PE diagnosis are also comparable to those of the training set, except for sensitivity, which is lower (75%), while specificity reaches 98% (Figs. 5, 6, 7).

To further verify the diagnostic value of the proposed method, we aggregated the results obtained at pixel sample level to obtain a diagnosis for each annotated region (corresponding to a specific pulmonary lesion or region of interest). The final diagnosis is given by the most frequent image pattern on the diagnostic target: PE/noPE, Pneumonia/No Pneumonia and Healthy/not-Healthy.

Table 7 reports the sensitivity, specificity, positive predictive value, and negative predictive value for ModelPrct3, and we also report the average values and 95% confidence intervals (CI) of each score. The metrics increase for all image patterns. In particular, values for PE sensitivity increase to 85%, while keeping a high specificity very close to 90%. For the detection of healthy lung, we have a score of 100% for all metrics.

Table 6 Sensitivity, specificity, PPV and NPV statistics for the test set of ModelPrct3

	Pulmonary embolism	Pneumonia	Healthy lung
Sensitivity	75.1% (95% CI 71.8–78.2)	93.3% (95% CI 91.8–94.6)	92.2% (95% CI 91.0–93.2)
Specificity	98.2% (95% CI 97.7–98.6)	93.0% (95% CI 92.1–93.9)	89.3% (95% CI 87.9–90.6)
Positive predictive value	88.9% (95% CI 86.2–91.2)	83.9% (95% CI 81.8–85.7)	91.3% (95% CI:90.1–92.3)
Negative predictive value	95.4% (95% CI 94.7–96.0)	97.3% (95% CI 92.3–93.8)	90.4% (95% CI89.0–91.6)

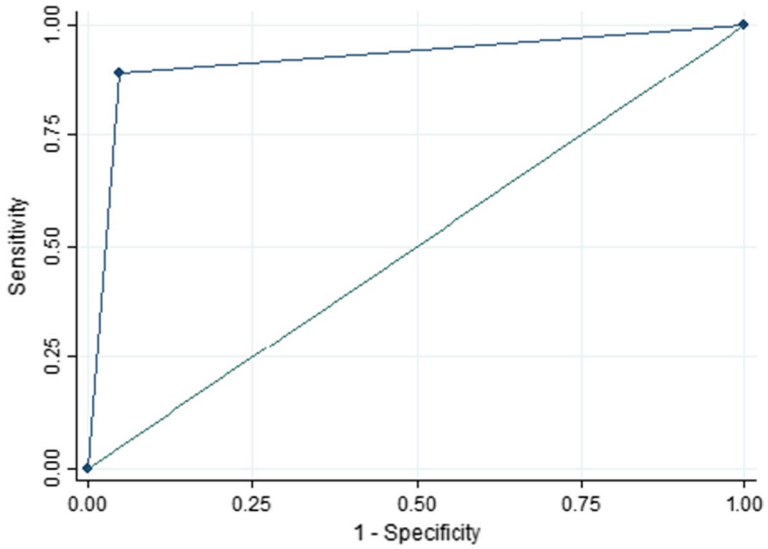


Fig. 5 Pulmonary embolism ROC curve in the test set

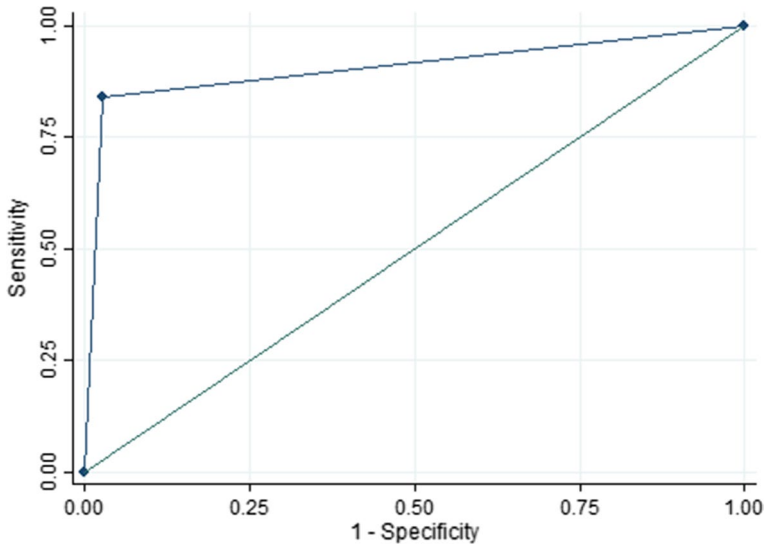


Fig. 6 Pneumonia ROC curve in the test set

Discussion

The diagnosis of pulmonary embolism in COVID-19 patients is challenging. In this sense, radiomics is a tool that could help the physician in the diagnosis of PE and pneumonia, since it is able to detect different image patterns obtained by Q-SPECT/CT, without the need to perform the ventilation technique.

To the best of the authors' knowledge, there are no published studies that apply artificial intelligence to ventilation/perfusion SPECT/CT or Q-SPECT/CT. This is an experimental study and is a first step toward a complete intelligent radiological system capable of diagnosing PE and pneumonia image patterns only by Q-SPECT/CT.

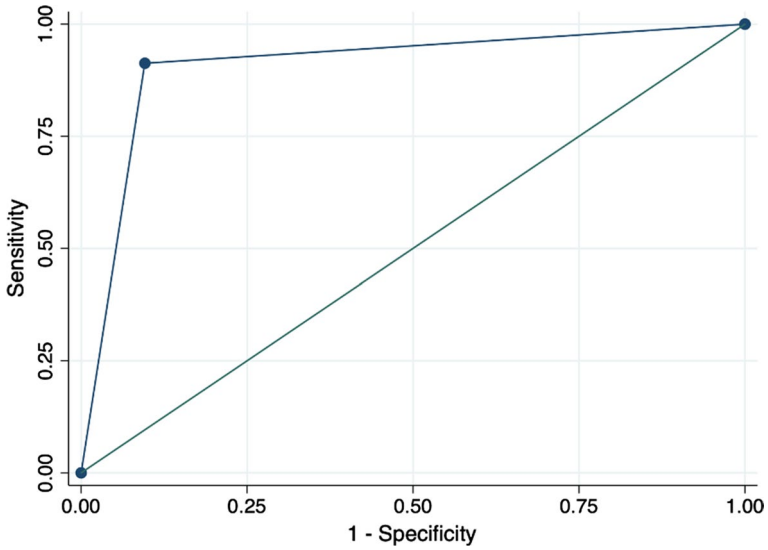


Fig. 7 Healthy lung ROC curve in the test set

Table 7 Sensitivity, specificity, PPV and NPV statistics for diagnosis at lesion level

	Pulmonary embolism	Pneumonia	Healthy lung
Sensitivity	85.0% (95% CI 84.1–85.8)	95.0% (95% CI 93.9–95.6)	100% (95% CI 100–100)
Specificity	89.5% (95% CI 88.6–90.3)	100% (95% CI 99.1–100)	100% (95% CI 100–100)
Positive predictive value	89.5% (95% CI 88.6–90.3)	95.0% (95% CI 93.9–95.6)	100% (95% CI 100–100)
Negative predictive value	85.0% (95% CI 84.2–85.8)	100% (95% CI 99.1–100)	100% (95% CI 100–100)

Our algorithm uses images obtained by Q-SPECT/CT and have proved to be useful to detect both PE and pneumonia image patterns. The results obtained for a training set of 20 patients using leave-1-out sampling show that the size of the window used to compute local radiomic features is not a critical parameter. In contrast, the normalization of the PE needed to compensate different concentrations of the contrast agent has a significant impact on the performance of the models, with a normalization based on percentiles being particularly recommended. The results obtained for an independent set of 39 test patients show the reproducibility of results and a high diagnosis capability and rule out for both PE and pneumonia. The sensitivity and AUC for the detection of pneumonia (93.3% recall and AUC = 0.906) is comparable to deep learning approaches using a larger number of training cases including COVNet [24] (90% recall and 0.96 AUC) and the early work of Zhang et al. [33] (88% recall and 0.92 AUC). Regarding PE, in spite of an AUC = 0.92 in tests, the sensitivity dropped to 75%. Nevertheless, with a positive predictive value of over 88%, we consider that this does not invalidate the model for PE. Moreover, the specificity and negative predictive values reach 98% and 95%, respectively, which indicates that it is an adequate technique to rule out this pathology, which is clinically useful. When the results obtained at the pixel sample level are aggregated into regions of interest, the sensitivity of the PE increases significantly to 85% but the specificity is lower (89.5%). Nevertheless, all metrics improve for both pneumonia and

healthy lung. These results suggest that aggregating the results into regions of interest provides the clinician with a better tool for diagnosing PE.

In relation to the working approach, an artificial intelligence diagnosis based on the analysis of whole images/scans might only detect the main pathology while ignoring secondary ones which are also clinically relevant. For this reason, the local analysis approach adopted by this study is important. Another issue is the clinical interpretability of results. Although interpretability can be improved with the use of a heatmap (such as gradient-weighted class activation mapping [34]), deep learning approaches are still difficult to interpret and lack the ability to accurately locate the injured tissue. Another recently identified concern [35] is the sensitivity of models to the quality and quantity of the cases used for training and testing, which can lead to overestimating the results. Deep learning methods require a large number of annotated images for training. This is not a major issue in most fields of application, but, unfortunately, there is a limited availability of images in the case of COVID-19 patients. Moreover, it is suspected that its protocol of acquisition may introduce bias in models [36]. In particular, it has been reported [37] that the high-performance of machine learning methods could be attributed mainly to the presence of image patterns (such as corner labels or instrumentation), device acquisition parameters or population factors (such as sex or age). If these characteristics are specific to some of the classes (groups of patients), models may learn to recognize these biases in the data set, rather than focusing on the pathologies they are trying to detect. This bias, of course, limits the generalization and reproducibility of results when tested on data sets with a different origin from the ones used in training and testing. We believe that adopting a local approach that analyzes small regions rather than the entire scan could minimize the need for a large number of annotated cases, as well as reducing the impact of image bias.

This study has some limitations. First, the local identification of image patterns must be added for each case (similar to [24]) to produce a multiple clinical diagnosis and to be able to produce a global diagnosis. Second, although the training group and the test groups are different, the models must be tested in cases originating from other hospitals to fully validate the generalizability and clinical applicability of the models. Third, we have excluded cases with emphysema and pulmonary fibrosis due to the difficulties involved in interpreting Q-SPECT/CT without ventilation, and for future studies we will include this type of cases in order to be able to make a useful algorithm in real life. Fourth, despite the fact that there are many samples analyzed, the number of patients is small, so for future studies we must have a larger number of patients that allows us to consolidate our algorithm. Finally, a model based on 2D radiomics, though similar to a histological analysis of the lesion, might not reflect 3D aspects of the lesion, like the spatial distribution of tissue patterns and volumetric measures of the extent of each pathology. However, given that the goal is to detect the presence of either TEP or pneumonia, this is not a main limitation for this particular study. For the computation of the extension and volume of the diseased lung tissue a 3D analysis should be carried out. This could be done by either accumulating the presented 2D radiomics or training a 3D radiomic model using 3D ROIs.

The main strength of this study is the novelty of using artificial intelligence to focus on PE and COVID-19 pneumonia diagnosed with Q-SPECT/CT images, which may

provide a diagnostic alternative for patients for whom CT pulmonary angiography (CTPA) is not capable of diagnosing PE or for when the technique is contraindicated. It is important to mention that it also provides a safer alternative to V-Q/SPECT-CT by avoiding aerosol contamination that occurs in ventilation, preventing the spread of diseases such as COVID-19. Although our system does not provide a global diagnosis of the patient, it is capable of classifying different image patterns and identifying areas affected by PE, as well as providing reliable information on the areas affected by pneumonia and healthy lung areas.

Our next step is to analyze the impact of our method on the final diagnosis of patients and to apply and adjust this algorithm to the current practice of V/Q-SPECT for PE not associated with COVID-19 and to include all types of respiratory diseases (emphysema, pulmonary fibrosis).

Our ultimate goal is to build a software for clinical use that can provide us with a diagnosis of PE and pneumonia only with an "intelligent Q-SPECT/CT" to be validated with our standard Ventilation/Perfusion-SPECT/CT. If this is achieved, ventilation could even be avoided, saving time and healthcare costs. This would mean a paradigm shift in the use of SPECT/CT for the diagnosis of these pathologies in times of COVID-19 or future epidemics, not only to minimize the risk of contagion by aerosols, but also to avoid unnecessary examinations and reduce test times.

Conclusion

In summary, a combined approach based on artificial intelligence and radiomics can detect areas of pneumonia and areas of pulmonary embolism using a limited amount of annotated data. The ability to detect different image patterns in patients with COVID-19 encourages us to continue working to achieve a global diagnosis through a local analysis that allows us to develop a software for clinical use based on SPECT/CT.

Abbreviations

PE	Pulmonary embolism
CTPA	CT pulmonary angiography
Q-SPECT/CT	Perfusion single-photon emission computed tomography/computed tomography
V/P-SPECT/CT	Ventilation/Perfusion single-photon emission computed tomography/computed tomography SPECT/CT
PPV	Positive predictive value
NPV	Negative predictive value

Acknowledgements
Not applicable

Author contributions
S B involved in conceptualization, methodology, data curation, formal analysis, validation, writing—original draft and visualization. M S involved in methodology and writing—review and editing. J D involved in methodology and writing—review and editing. I G O involved in formal analysis, writing—original draft and writing—review and editing [1]. C S involved in software, formal analysis, writing—review and editing, G T involved in software, formal analysis and writing—Review and Editing. G M involved in conceptualization, methodology and writing—review and editing. A R involved in conceptualization, writing—review and editing and supervision. D G involved in conceptualization, software, formal analysis, validation, writing—original Draft, writing—review and editing and supervision. S B also participated in review and editing. All authors read and approved by the final manuscript.

Funding
This work was supported by grants PID2021-126776OB-C21, RTI2018-095209-B-C21 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe"; Agència de Gestió d'Ajuts Universitaris i de Recerca grant numbers 2017-SGR-1624 and CERCA Programme / Generalitat de Catalunya; Barcelona Respitatory Network (BRN)- Fundació Ramon Pla i Armengol; "Clinical project" grant, Fundació Acadèmia Ciències Mèdiques de Catalunya i de

Balears; "Talents" grant 2020, Fundació La Pedrera and Hospital Universitari Germans Trias i Pujol; Lung Ambition Alliance Grant.

Availability of data and materials
Anonymized data are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study was performed in accordance with the principles of the Declaration of Helsinki. The research protocol was approved by the regional ethics committee (Ethics Committee for Clinical Research of Germans Trias i Pujol University Hospital), with the reference PI-20–161 and subjects have given their informed consent.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no relevant competing interests.

Received: 4 March 2022 Accepted: 7 November 2022
Published: 5 December 2022

References

1. Ooi MW, Rajai A, Patel R, Gerova N, Godhamgaonkar V. Liong SY Pulmonary thromboembolic disease in COVID-19 patients on CT pulmonary angiography–Prevalence, pattern of disease and relationship to D-dimer. Eur J Radiol. 2020;132:109336.
2. Zhai Z, Li C, Chen Y, Gerotziafas G, Zhang Z, Wan J, Liu P, Elalamy I, Wang C. Prevention and treatment of venous thromboembolism associated with coronavirus disease 2019 infection: a consensus statement before guidelines. Throm Haemost. 2020;120(06):937–48.
3. Bajc M, Schümichen C, Grüning T, Lindqvist A, Le Roux PY, Alatri A, Bauer RW, Dilic M, Neilly B, Verberne HJ, Delgado Bolton RC. EANM guideline for ventilation/perfusion single-photon emission computed tomography (SPECT) for diagnosis of pulmonary embolism and beyond. Eur J Nucl Med Mol Imaging. 2019;46(12):2429–51.
4. Bhatia KD, Ambati C, Dhaliwal R, Paschkewitz R, Hsu E, Ho B, Young A, Emmett L. SPECT-CT/VQ versus CTPA for diagnosing pulmonary embolus and other lung pathology: pre-existing lung disease should not be a contraindication. J Med Imaging Radiation Oncol. 2016;60(4):492–7.
5. Zuckier LS, Moadel RM, Haramati LB, Freeman LM. Diagnostic evaluation of pulmonary embolism during the COVID-19 pandemic. J Nucl Med. 2020;61(5):630–1.
6. Gustine JN, Jones D. Review immunopathology of hyperinflammation in COVID-19. Am J Pathol. 2021. https://doi.org/10.1016/j.ajpath.2020.08.009.
7. McGonagle D, Sharif K, O'Regan A, Bridgewood C. The role of cytokines including interleukin-6 in COVID-19 induced pneumonia and macrophage activation syndrome-like disease. Autoimmun Rev. 2020;19(6):102537.
8. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. Lancet Haematol. 2020;7(6):e438–40.
9. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, Vanstapel A, Werlein C, Stark H, Tzankov A, Li WW. Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19. New England J Med. 2020;383(2):120–8.
10. Jenner WJ, Kanji R, Mirsadraee S, Gue YX, Price S, Prasad S, Gorog DA. Thrombotic complications in 2928 patients with COVID-19 treated in intensive care: a systematic review. J Thrombos Thrombol. 2021;51(3):595–607.
11. Middeldorp S, Coppens M, van Haaps TF, Foppen M, Vlaar AP, Müller MC, Bouman CC, Beenen LF, Kootte RS, Heijmans J, Smits LP. Incidence of venous thromboembolism in hospitalized patients with COVID-19. J Thrombos Haemos. 2020;18(8):1995–2002.
12. Klok FA, Kruip MJ, Van der Meer NJ, Arbous MS, Gommers DA, Kant KM, Kaptein FH, van Paassen J, Stals MA, Huisman MV, Endeman H. Incidence of thrombotic complications in critically ill ICU patients with COVID-19. Thrombos Res. 2020;191:145–7.
13. Poyiadji N, Cormier P, Patel PY, Hadied MO, Bhargava P, Khanna K, Nadig J, Keimig T, Spizarny D, Reeser N, Klochko C. Acute pulmonary embolism and COVID-19. Radiology. 2020;297(3):E335.
14. Suh YJ, Hong H, Ohana M, Bompard F, Revel MP, Valle C, Gervaise A, Poissy J, Susen S, Hékimian G, Artifoni M. Pulmonary embolism and deep vein thrombosis in COVID-19: a systematic review and meta-analysis. Radiology. 2021;298(2):70.
15. Ippolito D, Giandola T, Maino C, Pecorelli A, Capodaglio C, Ragusi M, Porta M, Gandola D, Masetto A, Drago S, Allegranza P. Acute pulmonary embolism in hospitalized patients with SARS-CoV-2-related pneumonia: multicentric experience from Italian endemic area. La Radiol Med. 2021;126(5):669–78.
16. Grillet F, Behr J, Calame P, Aubry S, Delabrousse E. Acute pulmonary embolism associated with COVID-19 pneumonia detected by pulmonary CT angiography. Radiology. 2020;296:186.
17. Stals M, Kaptein F, Kroft L, Klok F, Huisman M. Challenges in the diagnostic approach of suspected pulmonary embolism in COVID-19 patients. Postgrad Med. 2021;133:36–41.
18. Das JP, Yeh R, Schöder H. Clinical utility of perfusion (Q)-single-photon emission computed tomography (SPECT)/CT for diagnosing pulmonary embolus (PE) in COVID-19 patients with a moderate to high pre-test probability of PE. Eur J Nucl Med Molecul Imaging. 2021;48(3):794–9.

19. Burger IA, Niemann T, Patriki D, Fontana F, Beer JH. Is there a role for lung perfusion [99mTc]-MAA SPECT/CT to rule out pulmonary embolism in COVID-19 patients with contraindications for iodine contrast? Eur J Nucl Med Molecul Imaging. 2020;47(9):2062–3.
20. Lu Y, Macapinlac HA. Perfusion SPECT/CT to diagnose pulmonary embolism during COVID-19 pandemic. Eur J Nucl Med Molecul Imaging. 2020;47(9):2064–5.
21. Vöö S, Dizdarevic S. Single photon emission computed tomography-computed tomography lung perfusion imaging during the COVID-19 pandemic: does nuclear medicine need to reconsider its guidelines? Nuclear Med Commun. 2020;41:991–3.
22. Lu Y, Lorenzoni A, Fox JJ, Rademaker J, Vander Els N, Grewal RK, Strauss HW, Schöder H. Noncontrast perfusion single-photon emission CT/CT scanning: a new test for the expedited, high-accuracy diagnosis of acute pulmonary embolism. Chest. 2014;145(5):1079–88.
23. Mazurek A, Dziuk M, Witkowska-Patena E, Piszczek S, Gizewska A. The utility of hybrid SPECT/CT Lung perfusion scintigraphy in pulmonary embolism diagnosis. Respiration. 2015;90(5):393–401.
24. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. Radiology. 2020;296(2):E65-71.
25. Wang M, Xia C, Huang L, Xu S, Qin C, Liu J, Cao Y, Yu P, Zhu T, Zhu H, Wu C. Deep learning-based triage and analysis of lesion burden for COVID-19: a retrospective study with external validation. Lancet Dig Health. 2020;2(10):e506–15.
26. Bai HX, Wang R, Xiong Z, Hsieh B, Chang K, Halsey K, Tran TM, Choi JW, Wang DC, Shi LB, Mei J. AI augmentation of radiologist performance in distinguishing COVID-19 from pneumonia of other etiology on chest CT. Radiology. 2020;296(3):E156.
27. Summers RM. Artificial intelligence of COVID-19 imaging: a hammer in search of a nail. Radiology. 2021;298:E169–71.
28. shah FM, Joy SKS, Ahmed F, Humaira M, Ami AS, Paul S, Jim MARK, Hossain T, Ahmed S A comprehensive survey of COVID-19 detection using medical images. SN Computer Science 2(6): 1-22 https://doi.org/10.31224/OSF.IO/9FDYP
29. Wells WM, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. Med Image Anal. 1996;1:35–51.
30. Jean S (1982) Image analysis and mathematical morphology ISBN 0–12–637240–3
31. Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RG, Fillion-Robin JC, Pieper S, Aerts HJ. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77(21):e104–7.
32. Ligero M, Torres G, Sanchez C, Diaz-Chito K, Perez R, Gil D (2019) Selection of radiomics features based on their reproducibility. Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS 403–408.
33. Zhang R, Tie X, Qi Z, Bevins NB, Zhang C, Griner D, Song TK, Nadig JD, Schiebler ML, Garrett JW, Li K. Diagnosis of coronavirus disease 2019 pneumonia by using chest radiography: value of artificial intelligence. Radiology. 2021;298(2):E88.
34. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2019) Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vision 128(2) 128:336–359.
35. López-Cabrera JD, Orozco-Morales R, Portal-Diaz JA, Lovelle-Enríquez O, Pérez-Díaz M. Current limitations to identify COVID-19 using artificial intelligence with chest X-ray imaging. Health Technol. 2021;11(2):411–24.
36. Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. AI Soc. 2020;35(3):761–5.
37. Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-ray images. Inform Fusion. 2021;1(76):1–7.
38. Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. Comput Math Methods Med. 2017. https://doi.org/10.1155/2017/3762651.

Publisher's Note

5.2. Artículo II

Título: Radiomics and clinical data for the diagnosis of incidental pulmonary nodules and lung cancer screening: Radiolung integrative predictive model

Autores: Sonia Baeza, Debora Gil, Carles Sanchez, Guillermo Torres, João Carmezim, Cristian Tebé, Ignasi Guasch, Isabel Nogueira, Samuel García-Reina, Carlos Martínez-Barenys, Jose Luis Mate, Felipe Andreo y Antoni Rosell.

Revista: Archivos de Bronconeumología

Volumen: por publicar. Disponible online: 14 de junio de2024

Año: 2024

Factor de impacto 2024: 8

Cuartil: 1 (Q1)

DOI: https://doi.org/10.1016/j.arbres.2024.05.027

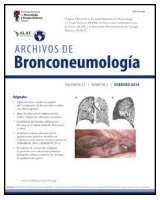
Citar como: Baeza S, Gil D, Sanchez C, et al. Radiomics and Clinical Data for the Diagnosis of Incidental Pulmonary Nodules and Lung Cancer Screening: Radiolung Integrative Predictive Model. Arch Bronconeumol. Published online May 31, 2024. doi:10.1016/j.arbres.2024.05.027

Archivos de Bronconeumología xxx (xxxx) xxx–xxx

Sociedad Española de Neumología y Cirugía Torácica SEPAR

ARCHIVOS DE
Bronconeumología

www.archbronconeumol.org



Original Article

Radiomics and Clinical Data for the Diagnosis of Incidental Pulmonary Nodules and Lung Cancer Screening: Radiolung Integrative Predictive Model

Sonia Baeza[a,b,c,*], Debora Gil[d], Carles Sanchez[d], Guillermo Torres[d], João Carmezim[b,e], Cristian Tebé[b,e], Ignasi Guasch[f], Isabel Nogueira[f], Samuel García-Reina[g,h], Carlos Martínez-Barenys[g,h], Jose Luis Mate[i], Felipe Andreo[a,b,c], Antoni Rosell[a,b,c]

[a] *Respiratory Medicine Department, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain*
[b] *Germans Trias i Pujol Research Institute (IGTP), Badalona, Barcelona, Spain*
[c] *Departament de Medicina, Universitat Autònoma de Barcelona, Barcelona, Spain*
[d] *Computer Vision Center and Computer Science Department, UAB, Barcelona, Spain*
[e] *Biostatistics Support and Research Unit, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain*
[f] *Radiodiagnostic Department, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain*
[g] *Thoracic Surgery Department, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain*
[h] *Departament de Cirugia, Universitat Autònoma de Barcelona, Barcelona, Spain*
[i] *Pathology Department, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain*

ARTICLE INFO

Article history:
Received 16 March 2024
Accepted 21 May 2024
Available online xxx

Keywords:
Pulmonary nodules
Lung cancer screening
Radiomics
Predictive model

ABSTRACT

Introduction: Early diagnosis of lung cancer (LC) is crucial to improve survival rates. Radiomics models hold promise for enhancing LC diagnosis. This study assesses the impact of integrating a clinical and a radiomic model based on deep learning to predict the malignancy of pulmonary nodules (PN).
Methodology: Prospective cross-sectional study of 97 PNs from 93 patients. Clinical data included epidemiological risk factors and pulmonary function tests. The region of interest of each chest CT containing the PN was analysed. The radiomic model employed a pre-trained convolutional network to extract visual features. From these features, 500 with a positive standard deviation were chosen as inputs for an optimised neural network. The clinical model was estimated by a logistic regression model using clinical data. The malignancy probability from the clinical model was used as the best estimate of the pre-test probability of disease to update the malignancy probability of the radiomic model using a nomogram for Bayes' theorem.
Results: The radiomic model had a positive predictive value (PPV) of 86%, an accuracy of 79% and an AUC of 0.67. The clinical model identified DLCO, obstruction index and smoking status as the most consistent clinical predictors associated with outcome. Integrating the clinical features into the deep-learning radiomic model achieves a PPV of 94%, an accuracy of 76% and an AUC of 0.80.
Conclusions: Incorporating clinical data into a deep-learning radiomic model improved PN malignancy assessment, boosting predictive performance. This study supports the potential of combined image-based and clinical features to improve LC diagnosis.

Abbreviations: AI, artificial intelligence; AUC, area under the curve; LC, lung cancer; LCS, lung cancer screening; LDCT, low-dose computed tomography; PN, pulmonary nodules; BMI, body mass index; COPD, chronic obstructive pulmonary disease; FVC, forced vital capacity; FEV_1, forced expiratory volume in one second; DLCO, diffusing capacity for carbon monoxide; PPV, positive predictive value; NPV, negative predictive value.

* Corresponding author.
E-mail address: smbaeza.germanstrias@gencat.cat (S. Baeza).

https://doi.org/10.1016/j.arbres.2024.05.027
0300-2896/

Please cite this article as: S. Baeza, D. Gil, C. Sanchez et al., Radiomics and Clinical Data for the Diagnosis of Incidental Pulmonary Nodules and Lung Cancer Screening: Radiolung Integrative Predictive Model, Archivos de Bronconeumología, https://doi.org/10.1016/j.arbres.2024.05.027

Introduction

Lung cancer (LC) remains a major health problem, causing the highest mortality in Europe and worldwide.[1] These data are explained by more than 70% of cases diagnosed at an advanced stage with a low survival; therefore, early diagnosis is key because early-stage treatment improves LC prognosis.[2] Recently, a shift has been occurring due to the increasingly expanded use of computed tomography across various settings, leading to a rise in the number of incidentally detected pulmonary nodules (PN). On the other hand, this shift is also attributed to the implementation of lung cancer screening (LCS). In both cases, the outcome is early-stage diagnosis. That's how the 2022 American Cancer Society annual report reveals a decline in the incidence of advanced-stage cancer, with a simultaneous annual increase of 4.5% in a localised stage primarily linked to the initiation of LCS in the US a decade ago.[3]

Despite the known advantages in reducing mortality due to LCS,[4–6] not all PN detected are LC. In a systematic review, 21% of trial screens yielded a false positive result (range 1–42%) in baseline LDCT,[6] requiring follow-up images or even biopsy. Thus, these false positives can cause potential harm to patients, and they can generate high costs for health services.[7–9] Similarly, they occur with the increased incidental finding of PN, as many of these are benign and require close follow-up, which is similar to the approach taken with PN identified through LCS.[10–13] In this context, whether the PN is detected incidentally or through LCS, it is crucial to identify high-risk cases and rule out those with low risk.[14,15]

Fortunately, applying artificial intelligence (AI) techniques to radiomics might make it possible to discriminate between benign and malignant PN.[16] Radiomics is an emerging area based on the high-throughput mining of quantitative image features from medical images that allows data to be applied in clinical decision-making to improve diagnostics and prognostics. It is mainly applied to cancer research.[17,18] Image analysis with a structured step workflow[19] is used to extract complementary features from multiple views (the shape, intensity and texture of features, among other things) which encapsulate different aspects of the lesion. Machine learning is a subfield of AI that encompasses all approaches and allows computers to learn from various types of data. In contrast, deep learning is part of machine learning based on multi-layered artificial neural networks to solve highly complex problems.[20,21] This has been applied in radiomics and is used to model such multi-view features to predict clinical outcomes.

The usefulness of radiomics in the early diagnosis of LC is promising.[22,23] Although many of these studies yield favourable outcomes, most are based solely on imaging data. Incorporating clinical data related to LC can enhance these findings by improving their application in daily clinical practice, yet there are still few studies with this approach. Among these, Zhang et al., using retrospective data, developed a radiomic and deep learning model to predict the malignancy of PN, achieving an AUC of 0.819 (95% CI: 0.76–0.88).[24] Employing a multi-omics approach could represent a significant advancement in applying AI models to the clinical prognosis of LC patients.[25]

The Radiolung project aims to design an algorithm based on a radiomic signature that, associated with clinical data, can accurately discriminate between LC and benign tumours.

Methods

Design and recruitment

Prospective, cross-sectional, comparative, and experimental study investigating the radiomic signatures of malignant and benign resected PNs, along with clinical risk factors. Recruitment took place from December 2019 to September 2023 at a tertiary care hospital. The inclusion criteria were patients aged 35–85 years-old with a clearly identifiable PN detected incidentally by chest CT or in LCS that qualified for surgery according to a multidisciplinary tumour board. The exclusion criteria were a slice thickness greater than 2.5 mm in the chest CT, a PN larger than 3 cm in diameter, or lung metastasis. The patients underwent lobectomy, segmentectomy, or atypical resection according to thoracic surgery criteria.

In each case, clinical and demographic were selected. Preoperative results of pulmonary function tests were obtained. Pre-surgery chest CT images, PN characteristics, and the pathological results of the lung tissue obtained during surgery were collected.

This study was performed in accordance with the principles of the Declaration of Helsinki. The research protocol was approved by the regional ethics committee (reference PI-19-169). All the patients gave their written informed consent.

Image data acquisition

Patients underwent multislice chest CT scans using regular radiation or low-dose radiation according to the protocols. Most of these scans were performed without intravenous contrast. Images were acquired with a GE Revolution scanner (General Electric Healthcare, Milwaukee, WI, USA), a SOMATOM Drive (Siemens Healthineers AG, Forchheim, Germany) and a Philips Incisive scanner (Philips Healthcare, Best, the Netherlands) equipped with 128 mm × 0.625 mm, 128 mm × 0.6 mm and 64 mm × 0.64 mm detector collimation, respectively. They all used automatic tube voltage and automatic tube current modulation. Chest CT image reconstructions were performed on a 512 × 512 matrix using a high spatial frequency algorithm and thin slice thickness applying the lung window setting (WW: 1600 and WL: −600) for the lung series. All chest CTs were interpreted by thoracic radiologists with over 10 years of experience.

The images were extracted from the picture archiving and communication system (PACS) in the Digital Imaging and Communications in Medicine (DICOM) format and were anonymized to guarantee patient confidentiality. Subsequently, the anonymized DICOMs were uploaded to a website hosted at the Autonomous University of Barcelona (UAB) for processing by the Computer Vision Center (CVC) to construct the radiomic model.

Data analysis

The study analysis followed a three-step strategy. In the first step, a radiomic prediction model was fitted to estimate a PN's malignancy probability based on the chest CT image. In the second step, a clinical prediction model was fitted to estimate a patient's malignancy probability based on their clinical profile. In the third step, the malignancy probability predicted by the estimated clinical model was used as the best estimate of the pretest probability of disease to update the radiomic model's node malignancy probability using a nomogram for Bayes' theorem.[26]

Deep radiomic model

An AI system designed to diagnose PN based on radiomic analysis of chest CT scans has the three main steps sketched in Fig. 1 (pipeline). These steps have the following goals:

1. Nodule detection. The first step is to identify the position and volumetric region (volume of interest, VOI) in the CT scan that contain the lesion of interest.
2. Nodule representation. Features are extracted by computing or using intensity values from the volume (3D) or slices (2D) of

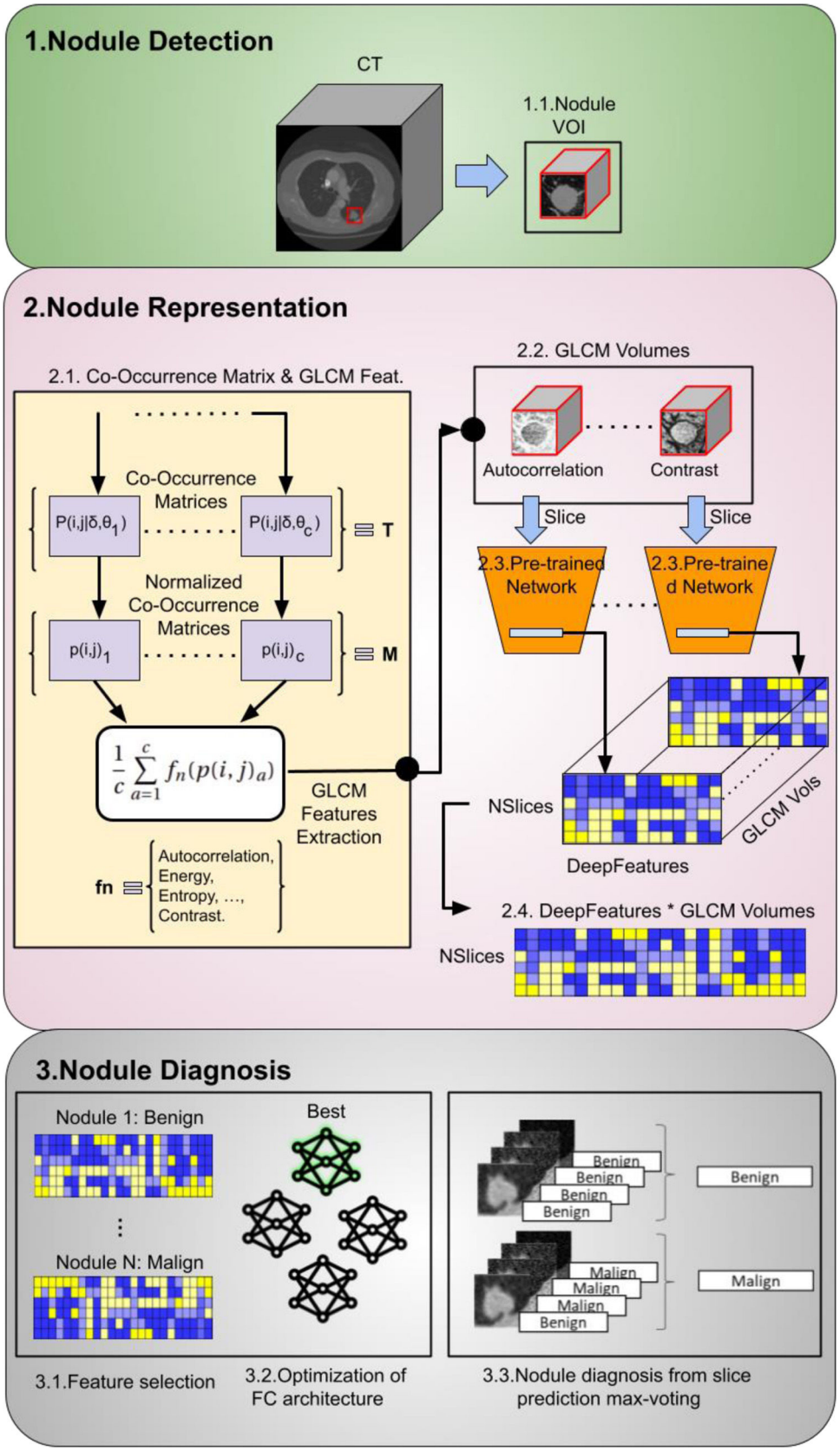


Fig. 1. Overview of the steps involved in the creation of the deep-radiomic model.

the VOI, thereby characterising the visual appearance of nodules. These features define a crucial representation space for the malignancy characterisation of the nodules. The representation space of nodules is given by visual features describing the content of textural volumes extracted from the intensity VOI. The textural volumes are given by 3D GLCM (Grey Level Co-occurrence Matrix) textural descriptors derived from co-occurrence matrices.[27] The visual features describing the content of textural volumes are extracted using a pre-trained convolutional neural network and concatenated for each slice. Finally, most discriminant features were the input to a fully connected neural network for the classification of benign and malign slices.

3. Nodule diagnosis. In this step, a fully connected neural network is trained to determine the values of the nodule representation space that best discriminate malignancy. Furthermore, throughout the nodule representation and diagnosis process, various methods are used, with the potential to optimise their hyperparameters.

The probabilities of the deep radiomic model were calculated using a nodule *k*-fold ($K = 10$) validation scheme in order to mitigate overfitting of the deep learning approach.

For more implementation details of deep-radiomic model see supplementary material.

Clinical model

A descriptive analysis of potential diagnostic factors for the diagnosis of PN malignancy was performed. The set of diagnostic factors included age, sex, educational level, body mass index (BMI), smoking status, living in an area with air pollution, family history of cancer, personal history of cancer, chronic obstructive pulmonary disease (COPD), and spirometric profile like forced vital capacity (FVC), forced expiratory volume in one second (FEV_1), diffusing capacity for carbon monoxide (DLCO) and FEV_1/FVC or obstruction index.

To proceed with the selection of variables according to the Akaike information criterion (AIC criteria), the initial dataset was bootstraped with repetition 2000 times,[28] and a logistic regression model was fit in each sample. Malignancy diagnosis (yes/no) was used as the outcome. Variables that were retained in more than 70% of the models were candidates for the final model. The nonlinear relationship between age and the log odd of the outcome was assessed with no relevant results. The final set of variables included in the model was then approved by a pulmonologist with more than 10 years of experience. Internal validation of the resulting model was performed on the whole cohort and was based on discrimination, calibration, and bootstrap validation.[29]

Discrimination was assessed by estimating the area under the receiver operating characteristic (ROC) curve (AUC). Calibration was assessed by the Brier score and graphically comparing the observed versus expected probabilities of malignancy diagnoses by deciles of predicted risk. Due to the imbalance in the malignancy distribution a precision recall curve was estimated and used to complement the model performance analysis. Bootstrap validation was performed to account for model overfitting and correct for optimism the model performance on the development data. Due to a lack of data, external validation was not performed.

Integrative model – Bayes update

The Bayes theorem underlies the Fagan nomogram.[26] This method allowed us to update the probability that a patient had a condition of interest given the probability that the subject had the condition before the test was performed and the likelihood ratio of the test. In our study, we used the probability of the clinical model as the probability that the subject had the condition before testing the PN, and the likelihood ratio of the test was based on the deep radiomic model. The resulting probability, using the Fagan nomogram, estimates the probability that the PN is malignant based on the patient's clinical probability and the deep radiomic model test result. Therefore, given a positive result in the radiomic model, the final probability of lung nodule malignancy is the malignancy clinical probability of the patient multiplied by the positive likelihood ratio of the deep radiomic model. Given a negative result in the radiomic model, the final probability of malignancy is the malignancy clinical probability of the patient multiplied by the negative likelihood ratio of the deep radiomic model.

All analyses were conducted using Python and R software version 4.1.0.[30]

Results

Demographic and clinical data

The demographic and clinical characteristics of the patients are presented in Table 1. The mean diameter of the PNs was 17.98 mm (18.60 mm in malignant, 15.76 mm in benign) with a median of 18.00 mm. 73 were malignant, and 20 were benign. In terms of the histological type of malignant PN, 57 were adenocarcinoma, and 15 were squamous cell carcinoma. In benign PN, the majority corresponded to fibrosis/inflammation processes (80%), with less frequency attributed to infectious causes such as aspergillosis and tuberculosis. In both benign and malignant nodules, over 50% were located in the upper lobes. The flowchart of patients and PN is in Fig. 2.

Deep radiomic model

This model includes data from 90 PN that met all the technical requirements for the analysis, 69 (77%) of which were malignant. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were 87% (95% CI: 0.77–0.94), 52% (95% CI: 0.30–0.74), 86% (95% CI: 0.75–0.93) and 55% (95% CI: 0.32–0.77), respectively, with a brier score 0.16 (95% CI: 0.1–0.23), accuracy of 79% (95% CI: 0.59–0.79) and AUC of 0.67 (95% CI: 0.51–0.83) (Fig. 3).

Clinical model

The model was carried out using the information of the 90 patients with complete cases (Fig. 4), 72 (80%) of which presented a malignant diagnosis. Based on the AIC and considering only variables selected more than 50% of the estimated models in the bootstrap samples, the optimal clinical model is displayed in Table 2.

In this clinical model, the sensitivity, specificity, PPV, and NPV were 86% (95% CI: 0.76–0.93), 61% (95% CI: 0.36–0.83), 90% (95% CI: 0.80–0.96) and 52% (95% CI: 0.30–0.74), respectively, with a brier score 0.14 (95% CI: 0.10–0.19), accuracy of 81% (95% CI: 0.71–0.89) and AUC of 0.71 (95% CI: 0.55–0.87)

Integrative model

The integrative model includes data from 89 PN that met all requirements for the radiomic and clinical model, 69 (78%) of which presented a malignant diagnosis. Integrating selected clinical features into a radiomic deep learning model exhibits a sensitivity, specificity, PPV and NPV of 74% (95% CI: 0.62–0.84), 85% (95% CI: 0.62–0.97), 94% (95% CI: 0.85–0.99) and 49% (95% CI: 0.31–0.66), respectively. The brier score is 0.14 (95% CI: 0.09–0.19), the accuracy is 76% (95% CI: 0.66–0.85), and the AUC is 0.80 (95% CI:

Table 1
Demographic and clinical characteristics of the patients.

	Malignant *N* = 73	Benign *N* = 20	Total *N* = 93
Age			
Mean (SD)	69.22 (8.5)	65.91 (10.6)	68.50 (9.0)
Sex, *n* (%)			
Woman	23 (31.5%)	10 (50.0%)	33 (35.0%)
BMI			
Mean (SD)	27.41 (3.8)	27.18 (6.6)	27.36 (4.5)
Education, *n* (%)			
College studies	4 (5.5%)	1 (5.0%)	5 (5.4%)
Post-high school training	31 (42.5%)	12 (60.0%)	43 (46.2%)
High school or less	38 (52.0%)	7 (35.0%)	45 (48.4%)
Air pollution, *n* (%)			
Yes	17 (23.3%)	3 (15.0%)	20 (22.0%)
Smoking, *n* (%)			
Never	8 (11.0%)	4 (20.0%)	12 (12.9%)
Current	28 (38.4%)	8 (40.0%)	36 (38.7%)
Former	37 (50.6%)	8 (40.0%)	45 (48.4%)
Pack year – index			
Mean (SD)	39.1 (26.2)	27.6 (23.6)	42.1 (23.4)
Family history of cancer, *n* (%)			
Yes	29 (39.7%)	8 (40.0%)	37 (39.8%)
Lung cancer	13 (44.8%)	2 (25.0%)	15 (40.5%)
Others	16 (55.2%)	6 (75.0%)	22 (59.5%)
Personal history of cancer, *n* (%)			
Yes	21 (28.8%)	5 (25.0%)	26 (28.0%)
COPD, *n* (%)			
Yes	24 (32.9%)	5 (25.0%)	29 (31.2%)
FVC (%)			
Mean (SD)	92.2 (18.4)	95.0 (17.0)	92.8 (18.0)
FEV_1 (%)			
Mean (SD)	84.4 (19.2)	88.5 (21.5)	85.3 (19.7)
Index (%)			
Mean (SD)	72.7 (10.1)	77.5 (15.5)	73.7 (11.6)
DLCO (%)			
Mean (SD)	76.6 (16.3)	83.2 (27.6)	77.9 (19.1)

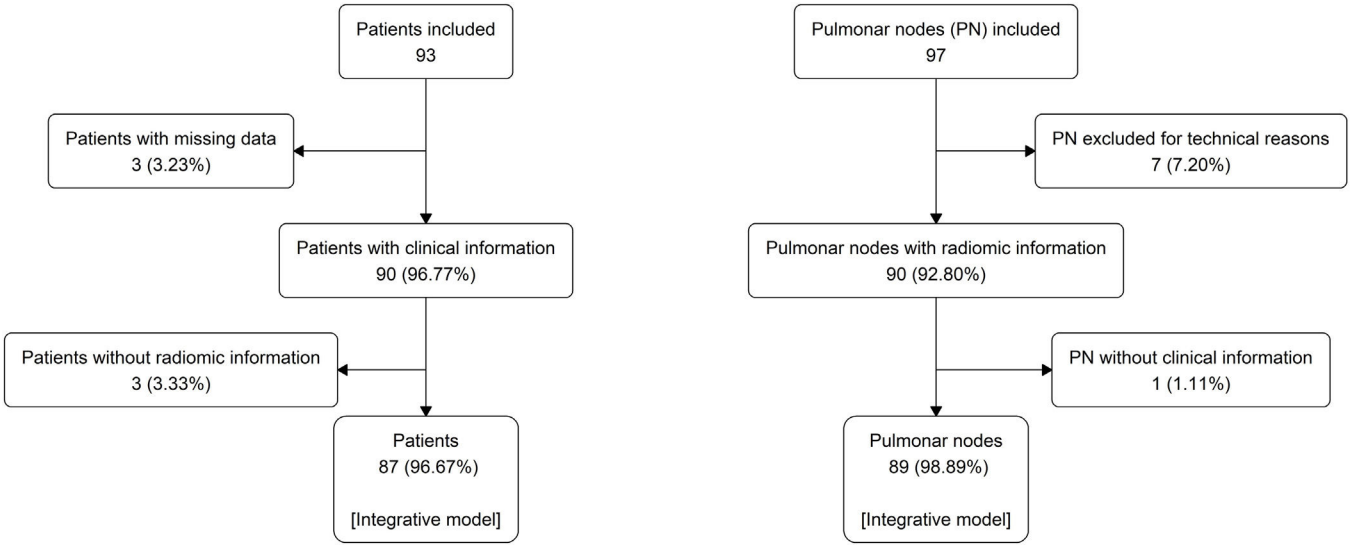


Fig. 2. Flowchart of patients and PN.

0.67–0.92) (Fig. 5). Fig. 6 plots the final predicted PN malignancy by risk decile against the observed incidence of PN malignancy in each decile. The convergence of the two curves indicates good model calibration. The precision-recall curve (supplemental fig. 5) rises sharply and plateaus at high precision with precision-recall AUC of 0.91. This suggests that the estimated model can achieve high precision (true positive rate) without sacrificing recall (sensitivity).

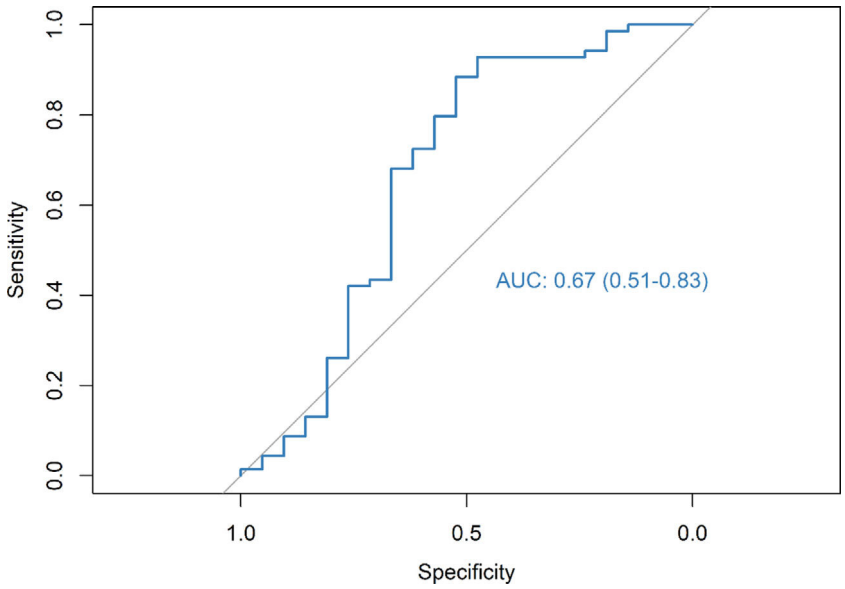


Fig. 3. ROC curve – deep radiomic model.

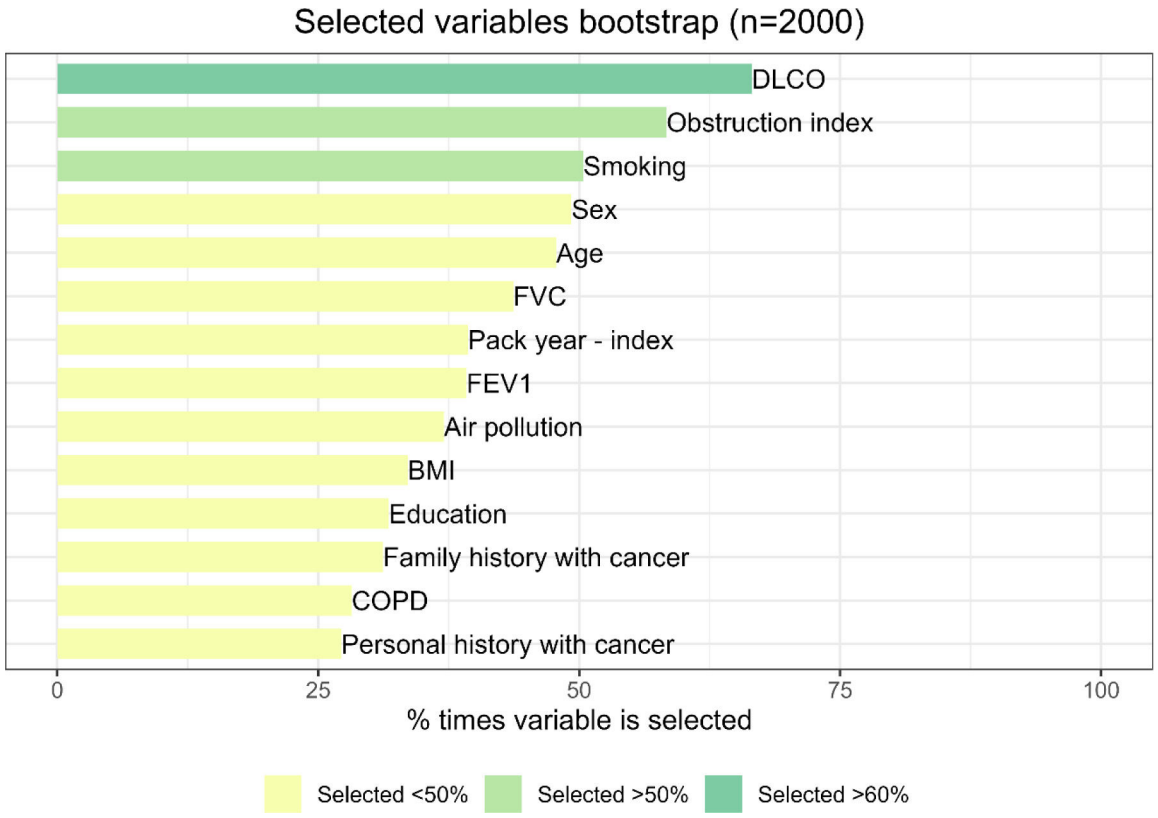


Fig. 4. Variables selected from the 2000 bootstrap model.

Discussion

Radiomics has revolutionised medical imaging, and there are currently many working groups dedicated to this research in different fields.[18,31] Other AI techniques, such as deep learning, allow us to associate these radiomic characteristics with other data, such as clinical information. Regarding this integrative approach, few working groups are dedicated to LC and PN diagnosis using radiomics and clinical data together, and the retrospective study by Lui et al.[32] stands out. They conducted an analysis utilising 20 radiomic features, in addition to age, gender, and PN location. Successfully, they developed a predictive model that achieved an AUC of 0.81 (95% CI: 0.75–0.87). Similar results were published in 2023 by Lin et al.[33] With a similar approach, our study yields promising results, as the integration of selected clinical features with radiomic image data based on deep learning techniques can formulate a predictive model that significantly improve results compared to relying solely on image analysis, achieving a positive predictive value of 94% (95% CI: 0.85–0.99) and an AUC of 0.80 (95% CI: 0.67–0.92). It is worth noting that the accuracy showed a slight but

Table 2
Optimal clinical model on 90 patients.

	Optimal clinical model			
Predictors	OR[a]	SE[a]	95% CI[a]	*p*-Value
Index (per 10 units)	0.67	0.25	0.41–1.08	0.11
Smoking				
Never	–	–	–	
Current	1.64	0.79	0.33–7.61	0.5
Former	3.37	0.78	0.69–15.8	0.12
DLCO (per 10 units)	0.82	0.15	0.61–1.10	0.2

[a] OR = odds ratio, SE = standard error, CI = confidence interval.

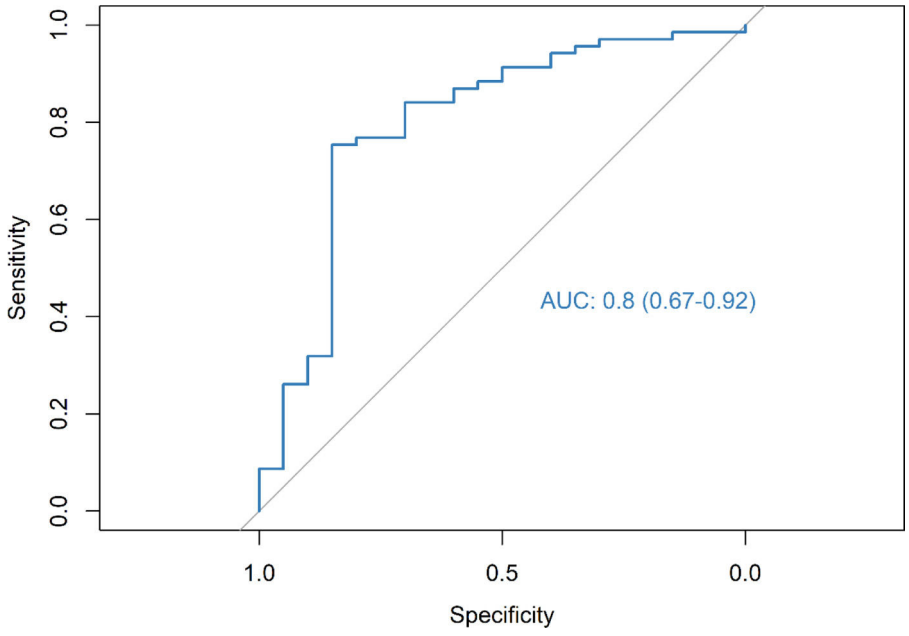


Fig. 5. ROC curve – integrative model.

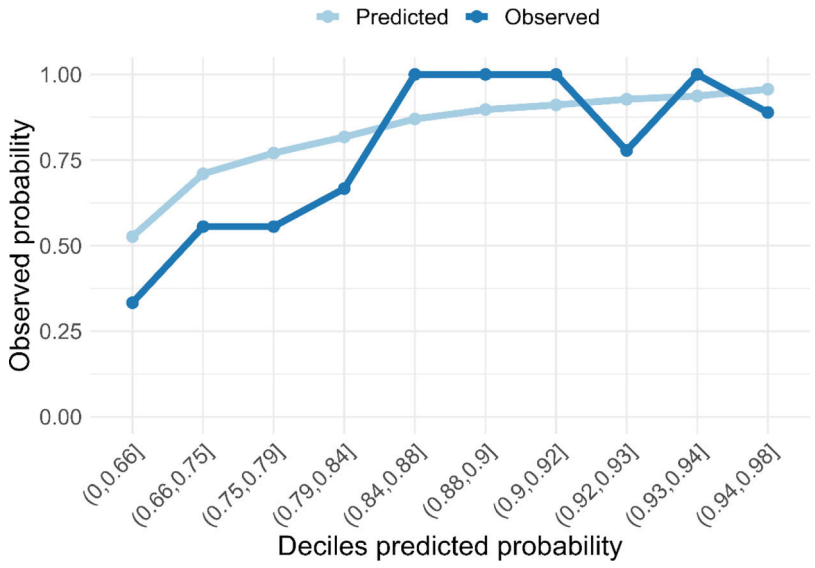


Fig. 6. Model predicted PN malignancy against the observed incidence of PN malignancy by risk decile.

non-significant decrease in the integrative model, which could be related to the imbalance in the number of benign and malignant cases in our cohort.

Other studies such as Marmon et al.[34] and Kammer et al.[35] have developed models based on blood biomarkers, clinical data, and radiomics with favourable results. However, it is important to note that these studies use blood biomarkers that are not specific to LC, which may add an additional cost without necessarily improving performance compared to models using only clinical and radiomic data. Furthermore, technically, they use classic machine learning hand-crafted features for radiomic descriptors without incorporating information from surrounding lung tissue. While these studies are relevant, our research stands out for its innovative deep learning approach and its ability to integrate multiple data sources.

Regarding validated predictive models based on clinical data, there are those designed for incidental PN such as the Mayo model,[36] while others focus on LCS such as the Brock model.[37] Please see the comparative table in the supplementary material. We have applied the Mayo model in our cohort and it exhibits similar performance in terms of PPV of 88% (95% CI: 0.75–0.95), but worst accuracy 63% (95% CI: 0.52–0.74), and AUC of 0.54 (95% CI: 0.36–0.72) (see supplemental material).

A strength of this study is that it employs prospective patient data for individuals who have undergone PN surgery, allowing for the collection of clinical data, chest CT images and histology. Another point to highlight is the novelty of this exploratory study, as it tests various clinical characteristics to identify the most representative ones for association with a deep radiomic model. This emphasizes the importance of integrating tools from daily medical practice with new technologies to enable early cancer diagnosis. Finally, we have compared our model with a conventional and validated clinical model such as the Mayo model, allowing us to demonstrate its validity in our cohort.

As for the limitations of this study, the most significant one is the need to expand the number of cases to improve both the clinical and radiomic models. Therefore, we are considering making the next study multicentric in the upcoming year to validate our exploratory analysis. Another limitation is the low number of benign cases, as these are real-life cases involving PN that have undergone surgery, which are inherently more diagnostically challenging. On the flip side, this situation contributes to the model training to detect suspicious malignancy cases of PN that ultimately prove benign.

Radiomics and AI models focused on LC are evolving across different phases, encompassing aspects such as PN management, diagnosis, treatment, and relapse of LC.[38,39] Integrative models with clinical data will be crucial to optimise their effectiveness. These advancements will positively impact patients, particularly in managing PN, because these models will aid in more accurately predicting the likelihood of malignancy, thereby avoiding unnecessary follow-ups, biopsies or surgeries. This situation not only reduces patient anxiety but also minimises potential harm, decreases waiting lists and reduces healthcare system costs.

In conclusion, we firmly believe that this integrative model, combining clinical data and radiomics through deep learning, can aid in diagnosing and managing PN. Although external validation is necessary in a subsequent phase, in the near future it may be directly useful in LCS programmes, thus optimising the early detection of LC and bringing technological advances closer to clinical practice.

Funding

This project is supported by the Ministerio de Economía, Industria y Competitividad, Gobierno de España grant number PID 2021-126776OB-C21; Barcelona Respiratory Network (BRN) Fundació Ramon Pla i Armengol; "Clinical project" grant, Fundació Acadèmia Ciències Mèdiques de Catalunya i de Balears; "Talents" grant 2020, Fundació La Pedrera and Hospital Universitari Germans Trias i Pujol; Lung Ambition Alliance Grant; and the JMC Legacy Research Fund of Germans Trias i Pujol University Hospital.

Financial/non-financial disclosures

The authors declare not to have any conflicts of interest that may be considered to influence directly or indirectly the content of the manuscript.

Authors' contributions

Sonia Baeza: conceptualization, methodology, data curation, formal analysis, validation, writing-original draft, writing-review and editing, visualisation. Debora Gil: conceptualization, methodology, software, formal analysis, validation, writing-review and editing, supervision. Carles Sanchez: methodology, software, data curation, formal analysis, validation, writing-review and editing. Guillermo Torres: methodology, software, data curation, formal analysis, validation, writing-review and editing. Cristian Tebé: validation, formal analysis, writing-review and editing. João Carmezim: validation, formal analysis, writing-review and editing. Ignasi Guasch: methodology, writing-review and editing. Isabel Nogueira: methodology, writing-review and editing. Samuel Garcia-Reina: methodology, writing-review and editing. Carlos Matínez-Barenys: methodology, writing-review and editing. Jose Luis Mate: methodology, writing-review and editing. Felipe Andreo: conceptualization, writing-review and editing. Antoni Rosell: conceptualization, methodology, formal analysis, writing-review and editing, supervision.

Artificial intelligence involvement

The text is original, and no artificial intelligence techniques have been used to write its content.

Acknowledgements

We thank Adela González, our nurse navigator, and Laia Ruiz, a UAB medical student, for their contribution and support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.arbres.2024.05.027.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209–49, http://dx.doi.org/10.3322/caac.21660.
2. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM Classification for lung cancer. J Thorac Oncol. 2016;11:39–51, http://dx.doi.org/10.1016/j.jtho.2015.09.009.
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72:7–33, http://dx.doi.org/10.3322/caac.21708.
4. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365:395–409, http://dx.doi.org/10.1056/NEJMoa1102873.
5. De Koning HJ, Van der Aalst CM, De Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced lung-cancer mortality with volume ct screening in a randomized trial. N Engl J Med. 2020;382:503–13, http://dx.doi.org/10.1056/NEJMoa1911793.

6. Bonney A, Malouf R, Marchal C, Manners D, Fong KM, Marshall HM, et al. Impact of low-dose computed tomography (LDCT) screening on lung cancer-related mortality. Cochrane Database Syst Rev. 2022;8:CD013829, http://dx.doi.org/10.1002/14651858.CD013829.pub2.
7. Mulshine JL, D'Amico TA. Issues with implementing a high-quality lung cancer screening program. CA Cancer J Clin. 2014;64:351–63, http://dx.doi.org/10.3322/caac.21239.
8. Mazzone PJ. Obstacles to and solutions for a successful lung cancer screening program. Semin Respir Crit Care Med. 2016;37:659–69, http://dx.doi.org/10.1055/s-0036-1592114.
9. Wang GX, Baggett TP, Pandharipande PV, Park ER, Fintelmann FJ, Percac-Lima S, et al. Barriers to lung cancer screening engagement from the patient and provider perspective. Radiology. 2019;290:278–87, http://dx.doi.org/10.1148/radiol.2018180212.
10. Gould MK, Tang T, Liu ILA, Lee J, Zheng C, Danforth KN, et al. Recent trends in the identification of incidental pulmonary nodules. Am J Respir Crit Care Med. 2015;192:1208–14, http://dx.doi.org/10.1164/rccm.201505-0990OC.
11. Larici AR, Farchione A, Franchi P, Ciliberto M, Cicchetti G, Calandriello L, et al. Lung nodules: size still matters. Eur Respir Rev. 2017;26:170025, http://dx.doi.org/10.1183/16000617.0025-2017.
12. Mazzone PJ, Lam L. Evaluating the patient with a pulmonary nodule: a review. JAMA. 2022;327:264–73, http://dx.doi.org/10.1001/jama.2021.24287.
13. Karius P, Lembcke A, Sokolowski FC, Gandara IDP, Rodríguez A, Hamm B, et al. Extracardiac findings on coronary computed tomography angiography in patients without significant coronary artery disease. Eur Radiol. 2019;29:1714–23, http://dx.doi.org/10.1007/s00330-018-5688-4.
14. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med. 2013;369:910–9, http://dx.doi.org/10.1056/NEJMoa1214726.
15. Simon M, Zukotynski K, Naeger DM. Pulmonary nodules as incidental findings. CMAJ. 2018;190:E167, http://dx.doi.org/10.1503/cmaj.171223.
16. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in lung cancer screening. Transl Lung Cancer Res. 2021;10:1186–99, http://dx.doi.org/10.21037/tlcr-20-708.
17. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48:441–6, http://dx.doi.org/10.1016/j.ejca.2011.11.036.
18. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14:749–62, http://dx.doi.org/10.1038/nrclinonc.2017.141.
19. Aerts HJWL. The potential of radiomic-based phenotyping in precision medicine. JAMA Oncol. 2016;2:1636, http://dx.doi.org/10.1001/jamaoncol.2016.2631.
20. Samuel AL. Some studies in machine learning using the game of checkers. IBM J Res Dev. 1959;3:210–29.
21. Lee J-G, Jun S, Cho Y-W, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: general overview. Korean J Radiol. 2017;18:570–84, http://dx.doi.org/10.3348/kjr.2017.18.4.570.
22. Chen B, Yang L, Zhang R, Luo W, Li W. Radiomics: an overview in lung cancer management—a narrative review. Ann Transl Med. 2020;8:1191, http://dx.doi.org/10.21037/atm-20-4589.
23. Wu Y-J, Wu F-Z, Yang S-C, Tang E-K, Liang C-H. Radiomics in early lung cancer diagnosis: from diagnosis to clinical decision support and education. Diagnostics. 2022;12:1064, http://dx.doi.org/10.3390/diagnostics12051064.
24. Zhang R, Wei Y, Shi F, Ren J, Zhou Q, Li W, et al. The diagnostic and prognostic value of radiomics and deep learning technologies for patients with solid pulmonary nodules in chest CT images. BMC Cancer. 2022;22:1118, http://dx.doi.org/10.1186/s12885-022-10224-z.
25. Li J, Wu J, Zhao Z, Zhang Q, Shao J, Wang C, et al. Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: a narrative review. J Thorac Dis. 2021;13:7021–33, http://dx.doi.org/10.21037/jtd-21-864.
26. Fagan TJ. Nomogram for Bayes's theorem. N Engl J Med. 1975;293:257, http://dx.doi.org/10.1056/NEJM197507312930513.
27. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973;SMC-3:610–21.
28. Austin PC, Tu JV. Bootstrap methods for developing predictive models. Am Stat. 2004;58:131–7, http://dx.doi.org/10.1198/0003130043277.
29. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594, http://dx.doi.org/10.1136/bmj.g7594.
30. The Comprehensive R Archive Network, 2024. https://Cran.r-ProjectOrg [accessed 01.3.24].
31. Frix A-N, Cousin F, Refaee T, Bottari F, Vaidyanathan A, Desir C, et al. Radiomics in lung diseases imaging: state-of-the-art for clinicians. J Pers Med. 2021;11:602, http://dx.doi.org/10.3390/jpm11070602.
32. Liu A, Wang Z, Yang Y, Wang J, Dai X, Wang L, et al. Preoperative diagnosis of malignant pulmonary nodules in lung cancer screening with a radiomics nomogram. Cancer Commun. 2020;40:16–24, http://dx.doi.org/10.1002/cac2.12002.
33. Lin R, Zheng Y, Lv F, Fu B, Li W, Liang Z, et al. A combined non-enhanced CT radiomics and clinical variable machine learning model for differentiating benign and malignant sub-centimeter pulmonary solid nodules. Med Phys. 2023;50:2835–43, http://dx.doi.org/10.1002/mp.16316.
34. Marmor HN, Jackson L, Gawel S, Kammer M, Massion PP, Grogan EL, et al. Improving malignancy risk prediction of indeterminate pulmonary nodules with imaging features and biomarkers. Clin Chim Acta. 2022;534:106–14, http://dx.doi.org/10.1016/j.cca.2022.07.010.
35. Kammer MN, Lakhani DA, Balar AB, Antic SL, Kussrow AK, Webster RL, et al. Integrated biomarkers for the management of indeterminate pulmonary nodules. Am J Respir Crit Care Med. 2021;204:1306–16, http://dx.doi.org/10.1164/rccm.202012-4438OC.
36. Swensen SJ. The probability of malignancy in solitary pulmonary nodules. Arch Intern Med. 1997;157:849, http://dx.doi.org/10.1001/archinte.1997.00440290031002.
37. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med. 2013;369:910–9, http://dx.doi.org/10.1056/NEJMoa1214726.
38. Pan F, Feng L, Liu B, Hu Y, Wang Q. Application of radiomics in diagnosis and treatment of lung cancer. Front Pharmacol. 2023:14, http://dx.doi.org/10.3389/fphar.2023.1295511.
39. Khawaja A, Bartholmai BJ, Rajagopalan S, Karwoski RA, Varghese C, Maldonado F, et al. Do we need to see to believe? Radiomics for lung nodule classification and lung cancer risk stratification. J Thorac Dis. 2020;12:3303–16, http://dx.doi.org/10.21037/jtd.2020.03.105.

6. RESUMEN GLOBAL DE LOS RESULTADOS

El primer estudio presentado muestra los resultados de un modelo radiómico inteligente que busca la mejor firma radiómica para detectar patrones de imagen de TEP o neumonía en pacientes con COVID-19 basándose únicamente en SPECT/TC-Q.

El número total de pacientes reclutados fue de 61, de los cuales dos casos fueron descartados debido a fallos técnicos en el proceso de adquisición de imágenes, y a la presencia de anomalías pulmonares graves causadas por patologías distintas a la COVID-19. El modelo fue entrenado con 20 pacientes (4,927 muestras en total, correspondiendo 2.096 a patrones de pulmón sano, 1.204 a patrones de TEP y 1.627 a patrones de neumonía) y validado en un grupo de 39 pacientes (4,410 muestras es total, correspondiendo 2.442 a patrones de pulmón sano, 597 a patrones de embolia pulmonar y 1.371 a patrones de neumonía). En el grupo de training, la infección por COVID-19 correspondió al 45% de los casos y al 51.28% en el grupo de test.

Se entrenaron modelos con diferentes ventanas (3x3 y 5x5) y normalización. Para el diagnóstico de TEP, los modelos normalizados que utilizaron el percentil superior tuvieron un rendimiento significativamente mejor comparado con aquellos normalizados utilizando los valores máximos, siendo de estos el modelo ModelPrct3 (ventana 3x3) el que tuvo un rendimiento significativamente superior.

En el grupo de training, tanto para el diagnóstico de TEP como para el de neumonía del ModelPrct3, el área bajo la curva (AUC) fue superior a 0.91, lo que confiere a los modelos basados en percentiles un alto valor diagnóstico. La sensibilidad para TEP, neumonía y pulmón sano en ModelPrct3 fue >90%, con un valor predictivo positivo (VPP) del >87% en todos los casos. La especificidad y el valor predictivo negativo (VPN) fueron superiores al 94% y al 96%, respectivamente.

En relación con el grupo de validación (test), el ModelPrct3 presentó un AUC de 0.92 para TEP, 0.91 para neumonía y 0.91 para pulmón sano, lo que demuestra que las predicciones del modelo son reproducibles. Además, para la detección de TEP presentó una sensibilidad, especificidad, VPP y VPN del 75.1%, 98.2%, 88.9% y 95.4%, respectivamente. Para

detectar zonas de neumonía fue observada una sensibilidad, especificidad, VPP y VPN del 94.1%, 93.6%, 85.2% y 97.6%, respectivamente.

Cuando los resultados obtenidos a nivel de muestra de píxeles se agregan en regiones de interés, y el diagnóstico final se da por el patrón de imagen más frecuente en la zona evaluada, la sensibilidad para TEP aumenta al 85%, manteniendo una especificidad cercana al 90%. Además, todas las métricas mejoran en los casos de neumonía.

En el segundo estudio presentado, evalúa los resultados diagnósticos de un modelo radiómico basado en Deep Learning integrado a un modelo clínico, para poder detectar malignidad en nódulos pulmonares.

Se evaluó un total de 93 pacientes, con 97 nódulos pulmonares. El diámetro promedio de los nódulos pulmonares fue de 17.98 mm (18.60 mm en malignos, 15.76 mm en benignos) con una mediana de 18.00mm. Del total de nódulos pulmonares, se contabilizaron 73 malignos y 20 benignos. En cuanto al tipo histológico de los nódulos pulmonares malignos, se contabilizaron 57 adenocarcinomas y 15 carcinomas de células escamosas. En los nódulos pulmonares benignos, la mayoría correspondía a procesos de fibrosis/inflamación (80%), con menor frecuencia atribuida a causas infecciosas como aspergilosis y tuberculosis. En ambos tipos de nódulos, benignos y malignos, más del 50% se ubicaron en los lóbulos superiores.

El modelo radiómico basado en Deep Learning fue entrenado con 90 nódulos pulmonares que cumplieron con todos los requisitos técnicos para el análisis, de los cuales 69 (77%) eran malignos. La sensibilidad, especificidad, valor predictivo positivo (VPP) y valor predictivo negativo (VPN) fueron del 87% (IC 95%: 0.77-0.94), 52% (IC 95%: 0.30-0.74), 86% (IC 95%: 0.75-0.93) y 55% (IC 95%: 0.32-0.77), respectivamente, con una puntuación de Brier de 0.16 (IC 95%: 0.1-0.23), una precisión del 79% (IC 95%: 0.59-0.79) y un AUC de 0.67 (IC 95%: 0.51-0.83).

El modelo clínico fue llevado a cabo utilizando la información de 90 pacientes, de los cuales 72 (80%) presentaron un diagnóstico de malignidad. El modelo clínico óptimo seleccionado considera datos de tabaquismo, obstrucción bronquial (FEV1/FVC) y DLCO. La sensibilidad, especificidad, VPP y VPN fueron del 86% (IC 95%: 0.76-0.93), 61% (IC 95%: 0.36-0.83), 90% (IC 95%: 0.80-0.96) y 52% (IC 95%: 0.30-0.74), respectivamente, con una puntuación de Brier de 0.14 (IC 95%: 0.10-0.19), una precisión del 81% (IC 95%: 0.71-0.89) y un AUC de 0.71 (IC 95%: 0.55-0.87).

El modelo integrativo de datos clínicos y radiómicos incluye datos de 89 nódulos pulmonares que cumplían con los requisitos para ambos modelos, de los cuales 69 (78%) presentaron un diagnóstico maligno. La sensibilidad, especificidad, VPP y VPN de 74% (IC 95%: 0.62-0.84), 85% (IC 95%: 0.62-0.97), 94% (IC 95%: 0.85-0.99) y 49% (IC 95%: 0.31-0.66), respectivamente. La puntuación de Brier score es de 0.14 (IC 95%: 0.09-0.19), la precisión es del 76% (IC 95%: 0.66-0.85) y el AUC es de 0.80 (IC 95%: 0.67-0.92).

7. RESUMEN GLOBAL DE LA DISCUSIÓN

En relación con los dos estudios anteriormente presentados, podemos afirmar que la radiómica potenciada por la inteligencia artificial es capaz de ayudarnos a optimizar el diagnóstico por imagen de patologías respiratorias relevantes como son el TEP y el cáncer de pulmón en estadios iniciales. La radiómica es un campo de investigación en franco desarrollo, que ha revolucionado la imagen médica en los últimos años a través de un análisis cuantitativo de alta dimensionalidad que nos aporta información valiosa sobre la biología de diferentes lesiones/patologías, que no nos entregan los medios convencionales de interpretación de la imagen médica, y que además tiene la ventaja de ser potenciada por técnicas de inteligencia artificial como el Deep Learning sacando el máximo provecho de ella. Su uso inicialmente fue enfocado hacia diferentes tipos de cáncer, pero como se explica en esta tesis, también tiene utilidad en patología no neoplásica, lo que lo convierte en una herramienta útil para los clínicos.

Respecto al primer estudio, hasta el momento de su publicación no existían otros estudios que aplicaran la radiómica e inteligencia artificial a SPECT/TC-V/Q o Q-SPECT/TC-Q. Aunque este estudio es experimental, representa el primer paso hacia un sistema radiómico inteligente, capaz de diagnosticar patrones de imagen de TEP y neumonía únicamente mediante SPECT/TC-Q. Respecto al modelo con mejor rendimiento en este estudio, el tamaño de las ventanas (grupo de pixeles) utilizada para calcular las características radiómicas no fue un parámetro crítico, pero si lo fue la normalización para compensar las diferencias en la administración de contraste, siendo especialmente útil una normalización basada en percentiles. En cuanto a los resultados, para el diagnóstico de patrones de TEP, se obtuvo un AUC de 0.92. Es destacable que la especificidad y VPN alcanzan el 98% y el 95%, respectivamente, lo que indica que es una técnica adecuada para descartar TEP. En el caso del diagnóstico de neumonía el AUC fue de 0.91, comparable con los resultados presentados por otros grupos que han trabado en neumonía por COVID-19 (86,87). Por otra parte, cuando los resultados obtenidos a nivel de ventanas (conjunto de píxeles) se agregan en regiones de interés mejoran las métricas y por lo tanto el diagnóstico. Es así como este modelo predictivo de TEP por medio del análisis radiómico de SPECT-TC puede

ser una alternativa innovadora que responda a la necesidad de una prueba alternativa al AngioTC para descartar TEP.

El segundo estudio muestra los resultados de un modelo predictivo que integra datos radiómicos y clínicos para detectar malignidad en nódulos pulmonares. Nuestro modelo radiómico es potenciado por Deep Learning, utilizando una red neuronal totalmente conectada para seleccionar las características radiómicas más relevantes para discriminar malignidad. Otro punto por destacar es su capacidad exploratoria al probar varias características clínicas para identificar las más representativas para su asociación con el modelo radiómico. Aunque no hay muchos estudios que utilicen esta fusión de datos para optimizar este diagnóstico por medio de la imagen, destacan unos pocos estudios de referencia como el de Lui et al. (88) y Lin et al (89), los que logran un AUC de 0.80 y 0.81 respectivamente, siendo similar a nuestros resultados (AUC 0.8). Por otra parte, el VVP de nuestro estudio que es del 94%, lo que hace que sea una buena prueba para detectar los nódulos pulmonares malignos. Al comparar nuestro modelo radiómico integrativo con un modelo de referencia para la detección de malignidad de nódulos pulmonares, como es el modelo de Mayo (90), vemos que al aplicarlo en nuestra cohorte el modelo de Mayo presenta un AUC del 0.54 y VVP del 88%, siendo un rendimiento menor que el alcanzado por nuestro modelo.

Como fortalezas de ambos estudios y por lo tanto de esta tesis, es innovación que aporta al utilizar nuevas tecnologías como la radiómica y técnicas de inteligencia artificial como Deep Learning, para resolver problemas de la clínica diaria, acercando así a la medicina a una nueva era de digitalización que nos llevará a mejorar los diagnósticos y optimizar tiempos y resultados. En particular, el primer estudio cuenta la ventaja de ser el primer estudio que utilizó radiómica e inteligencia artificial para optimizar el diagnóstico en SPECT-TC, presentando originalidad tanto en la problemática que busca resolver, como en el modelo desarrollado. En relación con el segundo estudio, aunque hay varios grupos trabajando en radiómica para el diagnóstico de nódulo pulmonar, pocos de ellos integran datos clínicos para optimizar los resultados radiómicos, lo que supone un aporte a la evidencia actual en este campo, sobre todo en una era en la que el diagnóstico precoz de

cáncer de pulmón es fundamental, más aún si podemos hacer un diagnóstico certero por imagen, asemejándose a una biopsia virtual.

Por otra parte, es destacable la objetividad que aporta la radiómica al realizar un análisis cuantitativo de la imagen, las que son valoradas como datos. El análisis actual de las imágenes médicas presenta en muchos casos subjetividad en su interpretación debido a variaciones inter observador, sobre todo cuando se trata de casos de nódulos pulmonares de pequeño tamaño valorados por TC de tórax o en el caso del TEP detectado por SPECT-TC/Q cuando existen otras patologías asociadas que pueden causar confusión como es el caso de la neumonía por COVID-19.

Como debilidad nos encontramos problemáticas que afectan a la radiómica de manera conocida como es la falta de un amplio número de datos. Es por esto que en ambos estudios, sobre todo al ser realizados en un único centro, tuvieron dificultad en la obtención de un amplio número imágenes de calidad asociados a datos clínicos detallados y completos. Esto ha hecho que en el segundo estudio no contáramos con suficientes datos para realizar un grupo de validación, sin embargo, esto nos ha motivado a ampliar la cohorte de forma multicéntrica para poder mejorar nuestro modelo y a la vez poder validarlo en una cohorte externa, necesario para poder llegar a implementarlo como un modelo predictivo a disposición de la práctica clínica.

El uso de radiómica potenciada por técnicas de inteligencia artificial en medicina, además de las ventajas clínicas que ya han sido explicadas, supondrá una mejora en economía de la salud. En el caso de un sistémico radiómico inteligente para el diagnóstico de TEP por medio de SPECT-TC-Q aplicado a la práctica clínica, podría ahorrar los recursos que supone asociar esta prueba a la técnica de ventilación (SPECT/TC-V/Q) y además reducir el tiempo de lectura de la prueba, suponiendo todo ello un ahorro en recursos asistenciales. En el caso del diagnóstico de malignidad de los nódulos pulmonares por medio de análisis radiómico de la imagen, puede disminuir los costes sanitarios al evitar biopsias innecesarias, cirugías en pacientes con patología benigna y en un futuro incluso lograr un diagnóstico histológico y molecular completo en los casos en que un nódulo pulmonar corresponda a un cáncer de pulmón.

Otra ventaja de las técnicas de análisis radiómico inteligente aplicadas a la imagen médica, tiene que ver con el beneficio en la atención que pueden recibir los pacientes, al poder darle una información más detallada y certera de su patología con solo una prueba de imagen no invasiva. Esto cobra especial interés en el diagnóstico de nódulos pulmonares en una era en que se fomenta la detección precoz y se detectan más lesiones de pequeño tamaño que suponen un desafío para los clínicos, teniendo muchas veces que someter a los pacientes a seguimientos periódicos y/o diferentes tipos de pruebas como biopsias por técnicas broncoscópicas o guiadas por TC, que no están exentas de complicaciones y que pueden producir malestar y ansiedad a los pacientes. Esto es mejorable si logramos dar un diagnóstico fiable con una primera imagen desde nuestra consulta médica, por lo que debemos trabajar por el desarrollo de este tipo de tecnología, que cambiará positivamente nuestra forma de trabajo y la calidad asistencial.

El futuro de la radiómica va enfocada a potenciar el desarrollo de la medicina de precisión en imagen médica, lo que nos permitirá individualizar en cada caso y avanzar hacia un diagnóstico completo de la patología. En el caso del cáncer, gracias a la radiómica se podrá ir más allá de un diagnóstico de benigno/maligno y de la histología, para adentrarnos a marcadores de agresividad, de respuesta a tratamientos como PD-L1 y en un futuro de mutaciones del tumor que puedan ser tributarios de tratamientos dirigidos, lo que se denomina radiogenómica. En el caso de la patología no neoplásica a pesar de que la radiómica tiene menos tiempo de desarrollo, ya que la radiómica inicialmente surgió para aplicarse en cáncer, existe un amplio campo por trabajar, no solo en el diagnóstico de TEP en el caso de las enfermedades respiratorias, sino también en diversas enfermedades como patología pulmonar intersticial y la tuberculosis. La revolución que está ocurriendo actualmente en la imagen médica debido a la radiómica y la inteligencia artificial, hará que en un futuro cercano cambie nuestra forma de trabajo en medicina debido a que esta tecnología será una herramienta fundamental en la práctica clínica, de la que probablemente podremos disponer tan fácilmente como por medio de un software desde nuestro sitio de trabajo. Por todo esto es relevante la investigación y el aprendizaje en esta área, siendo una excelente forma de aportar en el desarrollo de la medicina del futuro.

8. CONCLUSIONES

1. Con un sistema radiómico inteligente es posible clasificar e identificar correctamente diferentes patrones pulmonares mediante un análisis por regiones de las imágenes de SPECT/TC-Q.

2. El modelo predictivo desarrollado para el análisis de SPECT/TC-Q logra detectar con un buen rendimiento diagnóstico regiones de la imagen que corresponden a TEP, neumonía por COVID-19 y pulmón sano.

3. Un modelo basado en características radiómicas potenciado por Deep Learning es capaz de predecir malignidad en nódulos pulmonares.

4. Al incorporar datos clínicos y radiómicos en un modelo predictivo integrativo, se alcanza un buen rendimiento diagnóstico, capaz de detectar correctamente la mayoría de los nódulos pulmonares malignos tanto incidentales como los detectados en cribado de cáncer de pulmón.

9. LINEAS FUTURAS

Las nuevas tecnologías son el futuro en medicina, pero su desarrollo depende de que las problemáticas clínicas puedan llegar a los científicos que crean estas herramientas, lo que hace fundamental la existencia de grupos de investigación multidisciplinares.

El grupo de investigación híbrido que hemos conformado con investigadores clínicos del Hospital Universitario Germans Trias i Pujol y matemáticos e ingenieros expertos en imagen médica del Centro de Visión por Computador de la Universidad Autónoma de Barcelona, nos ha permitido desarrollar una línea de trabajo innovadora y apasionante, que va más allá de los estudios presentados en esta tesis doctoral. Tanto por los buenos resultados como por el potencial que creemos que tiene la investigación en imagen médica, para nosotros es fundamental consolidar esta línea de trabajo.

La principal línea de futuro que potenciaremos en los próximos años está enfocada en el diagnóstico del nódulo pulmonar a diferentes niveles:

1. Validar y retroalimentar el modelo integrativo con datos clínicos y radiómicos para el diagnóstico de malignidad de los nódulos pulmonares que se ha presentado en esta tesis doctoral.
2. En el nódulo pulmonar maligno, desarrollar un modelo que nos permita por medio del análisis de imagen predecir el tipo histológico.
3. Explorar la presencia de marcadores orientados a tratamiento, como son marcadores de PD-L1 y de mutaciones accionables como EGFR, ALK, ROS1, entre otros.

Con todo esto, nos hemos propuesto lograr en un futuro el desarrollo de un software que pueda darnos un diagnóstico avanzado del nódulo pulmonar por medio de imagen de TC de tórax, siendo aplicable en la práctica clínica tanto por radiólogos, como por neumólogos que realizan una consulta de nódulo pulmonar o de cribado de cáncer de pulmón. De esta manera creemos firmemente que podremos ayudar a acercar las nuevas tecnologías a la práctica clínica asistencial.

10. BIBLIOGRAFÍA

1. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death. 2020. The top 10 causes of death.
2. World Health Organization. https://covid19.who.int. 2024. Coronovirus (COVID-19) Dashbaord, World Health Organization.
3. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024 Apr 4;
4. Instituto nacional de estadística (INE). https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176780&menu=ultiDatos&idp=1254735573175 . 2023. Defunciones según la causa de muerte. .
5. RIETE Registry. https://rieteregistry.com/graphics-interactives/dead-30-days. 2024. RIETE Registry. Death within 30 days.
6. Konstantinides S V, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). Eur Heart J. 2020 Jan 21;41(4):543–603.
7. Moser KM, Fedullo PF, LitteJohn JK, Crawford R. Frequent asymptomatic pulmonary embolism in patients with deep venous thrombosis. JAMA. 1994 Jan 19;271(3):223–5.
8. Wells P, Anderson D, Rodger M, Ginsberg J, Kearon C, Gent M, et al. Derivation of a Simple Clinical Model to Categorize Patients Probability of Pulmonary Embolism: Increasing the Models Utility with the SimpliRED D-dimer. Thromb Haemost. 2000 Dec 14;83(03):416–20.
9. Raja AS, Greenberg JO, Qaseem A, Denberg TD, Fitterman N, Schuur JD. Evaluation of Patients With Suspected Acute Pulmonary Embolism: Best Practice Advice From the Clinical Guidelines Committee of the American College of Physicians. Ann Intern Med. 2015 Nov 3;163(9):701–11.
10. Bhatia KD, Ambati C, Dhaliwal R, Paschkewitz R, Hsu E, Ho B, et al. SPECT-CT/VQ versus CTPA for diagnosing pulmonary embolus and other lung pathology: Pre-existing lung disease should not be a contraindication. J Med Imaging Radiat Oncol. 2016 Aug 11;60(4):492–7.
11. Bajc M, Schümichen C, Grüning T, Lindqvist A, Le Roux PY, Alatri A, et al. EANM guideline for ventilation/perfusion single-photon emission computed tomography (SPECT) for diagnosis of pulmonary embolism and beyond. Eur J Nucl Med Mol Imaging. 2019 Nov 13;46(12):2429–51.
12. Quinn JL, Whitley JE, Hudspeth AS, Prichard RW. Early Clinical Applications of Lung Scintiscanning. Radiology. 1964 Feb;82(2):315–7.

13. M B, C S, T G, A L, PY LR, A A, et al. EANM guideline for ventilation/perfusion single-photon emission computed tomography (SPECT) for diagnosis of pulmonary embolism and beyond. Eur J Nucl Med Mol Imaging [Internet]. 2019 Nov 1 [cited 2021 Jul 19];46(12):2429–51. Available from: https://pubmed.ncbi.nlm.nih.gov/31410539/

14. Mortensen J, Gutte H. SPECT/CT and pulmonary embolism. Eur J Nucl Med Mol Imaging. 2014 May 9;41(S1):81–90.

15. Derenoncourt PR, Felder GJ, Royal HD, Bhalla S, Lang JA, Matesan MC, et al. Ventilation-Perfusion Scan: A Primer for Practicing Radiologists. RadioGraphics. 2021 Nov;41(7):2047–70.

16. Le Pennec R, Iravani A, Woon B, Dissaux B, Gest B, Le Floch PY, et al. Gallium-68 Ventilation/Perfusion PET-CT and CT Pulmonary Angiography for Pulmonary Embolism Diagnosis: An Interobserver Agreement Study. Front Med (Lausanne). 2021 Feb 16;7.

17. Joseph Loscalzo AFDKSHDLJLJ. Harrison's Principles of Internal Medicine. . 29th ed. Vol. 279. 2022.

18. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. Lancet Haematol. 2020 Jun;7(6):e438–40.

19. Middeldorp S, Coppens M, van Haaps TF, Foppen M, Vlaar AP, Müller MCA, et al. Incidence of venous thromboembolism in hospitalized patients with COVID-19. Journal of Thrombosis and Haemostasis. 2020 Aug;18(8):1995–2002.

20. Suh YJ, Hong H, Ohana M, Bompard F, Revel MP, Valle C, et al. Pulmonary Embolism and Deep Vein Thrombosis in COVID-19: A Systematic Review and Meta-Analysis. Radiology. 2021 Feb;298(2):E70–80.

21. JP D, R Y, H S. Clinical utility of perfusion (Q)-single-photon emission computed tomography (SPECT)/CT for diagnosing pulmonary embolus (PE) in COVID-19 patients with a moderate to high pre-test probability of PE. Eur J Nucl Med Mol Imaging [Internet]. 2021 Mar 1 [cited 2021 Jul 19];48(3):794–9. Available from: https://pubmed.ncbi.nlm.nih.gov/32959115/

22. IA B, T N, D P, F F, JH B. Is there a role for lung perfusion [99m Tc]-MAA SPECT/CT to rule out pulmonary embolism in COVID-19 patients with contraindications for iodine contrast? Eur J Nucl Med Mol Imaging [Internet]. 2020 Aug 1 [cited 2021 Jul 19];47(9):2062–3. Available from: https://pubmed.ncbi.nlm.nih.gov/32350559/

23. Y L, HA M. Perfusion SPECT/CT to diagnose pulmonary embolism during COVID-19 pandemic. Eur J Nucl Med Mol Imaging [Internet]. 2020 Aug 1 [cited 2021 Jul 19];47(9):2064–5. Available from: https://pubmed.ncbi.nlm.nih.gov/32383092/

24. S V, S D. Single photon emission computed tomography lung perfusion imaging during the COVID-19 pandemic: does nuclear medicine need to reconsider its guidelines? Nucl Med Commun [Internet]. 2020 [cited 2021 Jul 19];41(9):991–3. Available from: https://pubmed.ncbi.nlm.nih.gov/32796489/

25. Cengiz TB, Abdelrahman A, Rohren SA, Doucette J, Ghesani M. The diagnostic accuracy of perfusion-only scan in the diagnosis of pulmonary embolism in the era of COVID-19: A single-center study of 434 patients. Ann Thorac Med. 2023;18(4):199–205.

26. A M, M D, E WP, S P, A G. The Utility of Hybrid SPECT/CT Lung Perfusion Scintigraphy in Pulmonary Embolism Diagnosis. Respiration [Internet]. 2015 Nov 1 [cited 2021 Jul 19];90(5):393–401. Available from: https://pubmed.ncbi.nlm.nih.gov/26418469/

27. Iftikhar IH, Iftikhar NH, Naeem M, BaHammam A. SPECT Ventilation/Perfusion Imaging for Acute Pulmonary Embolism: Meta-analysis of Diagnostic Test Accuracy. Acad Radiol. 2024 Feb;31(2):706–17.

28. Asociación Española contra el Cáncer (AEEC). https://www.contraelcancer.es/sites/default/files/2021-03/Informe%20Impacto%20C%C3%A1ncer%20Catalu%C3%B1a_2.pdf. 2021. Impacto del Cáncer en Cataluña 2020.

29. Curiel-García T, Rey-Brandariz J, Varela-Lema L, Ruano-Ravina A, Candal-Pedreira C, Mourino N, et al. Asbestos exposure and small cell lung cancer: Systematic review and meta-analysis. J Occup Environ Hyg. 2023 Oct 3;20(10):427–38.

30. Rodríguez-Martínez Á, Ruano-Ravina A, Torres-Durán M, Provencio M, Parente-Lamelas I, Vidal-García I, et al. Residential Radon and Small Cell Lung Cancer. Final Results of the Small Cell Study. Arch Bronconeumol. 2022 Jul;58(7):542–6.

31. Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). Lancet Oncol. 2013 Aug;14(9):813–22.

32. Hamra GB, Guha N, Cohen A, Laden F, Raaschou-Nielsen O, Samet JM, et al. Outdoor Particulate Matter Exposure and Lung Cancer: A Systematic Review and Meta-Analysis. Environ Health Perspect. 2014 Sep;122(9):906–11.

33. de Torres JP, Marín JM, Casanova C, Cote C, Carrizo S, Cordoba-Lanus E, et al. Lung Cancer in Patients with Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med. 2011 Oct 15;184(8):913–9.

34. Ballester B, Milara J, Cortijo J. Idiopathic Pulmonary Fibrosis and Lung Cancer: Mechanisms and Molecular Targets. Int J Mol Sci. 2019 Jan 30;20(3):593.

35. Qin Y, Chen Y, Chen J, Xu K, Xu F, Shi J. The relationship between previous pulmonary tuberculosis and risk of lung cancer in the future. Infect Agent Cancer. 2022 Dec 7;17(1):20.

36. Cannon-Albright LA, Carr SR, Akerley W. Population-Based Relative Risks for Lung Cancer Based on Complete Family History of Lung Cancer. Journal of Thoracic Oncology. 2019 Jul;14(7):1184–91.

37. Koning HJ de, Aalst CM van der, Jong PA de, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. https://doi.org/101056/NEJMoa1911793 [Internet]. 2020 Jan 29 [cited 2021 Oct 4];382(6):503–13. Available from: https://www.nejm.org/doi/10.1056/NEJMoa1911793

38. DR A, AM A, CD B, WC B, JD C, RM F, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med [Internet]. 2011 Aug 4 [cited 2021 Dec 29];365(5):395–409. Available from: https://pubmed.ncbi.nlm.nih.gov/21714641/

39. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin [Internet]. 2022 Jan [cited 2023 Apr 10];72(1):7–33. Available from: https://pubmed.ncbi.nlm.nih.gov/35020204/

40. Bonney A, Malouf R, Marchal C, Manners D, Fong KM, Marshall HM, et al. Impact of low-dose computed tomography (LDCT) screening on lung cancer-related mortality. Cochrane Database of Systematic Reviews. 2022 Aug 3;2022(8).

41. Mulshine JL, D'Amico TA. Issues with implementing a high-quality lung cancer screening program. CA Cancer J Clin [Internet]. 2014 Sep 10 [cited 2023 Dec 6];64(5):351–63. Available from: https://pubmed.ncbi.nlm.nih.gov/24976072/

42. Mazzone PJ. Obstacles to and Solutions for a Successful Lung Cancer Screening Program. Semin Respir Crit Care Med [Internet]. 2016 Oct 1 [cited 2023 Dec 6];37(5):659–69. Available from: https://pubmed.ncbi.nlm.nih.gov/27732988/

43. Wang GX, Baggett TP, Pandharipande P V., Park ER, Fintelmann FJ, Percac-Lima S, et al. Barriers to Lung Cancer Screening Engagement from the Patient and Provider Perspective. Radiology [Internet]. 2019 Feb 1 [cited 2023 Dec 6];290(2):278–87. Available from: https://pubmed.ncbi.nlm.nih.gov/30620258/

44. Gould MK, Tang T, Liu ILA, Lee J, Zheng C, Danforth KN, et al. Recent Trends in the Identification of Incidental Pulmonary Nodules. Am J Respir Crit Care Med [Internet]. 2015 Nov 15 [cited 2023 Dec 6];192(10):1208–14. Available from: https://pubmed.ncbi.nlm.nih.gov/26214244/

45. Larici AR, Farchione A, Franchi P, Ciliberto M, Cicchetti G, Calandriello L, et al. Lung nodules: size still matters. Eur Respir Rev [Internet]. 2017 Dec 31 [cited 2023 Dec 6];26(146). Available from: https://pubmed.ncbi.nlm.nih.gov/29263171/

46. Mazzone PJ, Lam L. Evaluating the Patient With a Pulmonary Nodule: A Review. JAMA [Internet]. 2022 Jan 18 [cited 2023 Dec 6];327(3):264–73. Available from: https://pubmed.ncbi.nlm.nih.gov/35040882/

47. Karius P, Lembcke A, Sokolowski FC, Gandara IDP, Rodríguez A, Hamm B, et al. Extracardiac findings on coronary computed tomography angiography in patients without significant coronary artery disease. Eur Radiol [Internet]. 2019 Apr 1 [cited 2023 Dec 6];29(4):1714–23. Available from: https://pubmed.ncbi.nlm.nih.gov/30255246/

48. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med [Internet]. 2013 Sep 5 [cited 2023 Dec 6];369(10):910–9. Available from: https://pubmed.ncbi.nlm.nih.gov/24004118/

49. Simon M, Zukotynski K, Naeger DM. Pulmonary nodules as incidental findings. CMAJ : Canadian Medical Association Journal [Internet]. 2018 Feb 2 [cited 2023 Dec 6];190(6):E167. Available from: /pmc/articles/PMC5809217/

50. Bueno J, Landeras L, Chung JH. Updated Fleischner Society Guidelines for Managing Incidental Pulmonary Nodules: Common Questions and Challenging Scenarios. Radiographics [Internet]. 2018 Sep 1 [cited 2023 Apr 16];38(5):1337–50. Available from: https://pubmed.ncbi.nlm.nih.gov/30207935/

51. Schmid-Bindert G, Vogel-Claussen J, Gütz S, Fink J, Hoffmann H, Eichhorn ME, et al. Incidental Pulmonary Nodules – What Do We Know in 2022. Respiration. 2022;101(11):1024–34.

52. Loverdos K, Fotiadis A, Kontogianni C, Iliopoulou M, Gaga M. Lung nodules: A comprehensive review on current approach and management. Ann Thorac Med. 2019;14(4):226.

53. Callister MEJ, Baldwin DR, Akram AR, Barnard S, Cane P, Draffan J, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. Thorax [Internet]. 2015 Aug 1 [cited 2023 Apr 16];70 Suppl 2:ii1–54. Available from: https://pubmed.ncbi.nlm.nih.gov/26082159/

54. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. Radiology. 2017 Jul;284(1):228–43.

55. Herder GJ, van Tinteren H, Golding RP, Kostense PJ, Comans EF, Smit EF, et al. Clinical Prediction Model To Characterize Pulmonary Nodules. Chest. 2005 Oct;128(4):2490–6.

56. Chen B, Yang L, Zhang R, Luo W, Li W. Radiomics: an overview in lung cancer management—a narrative review. Ann Transl Med. 2020 Sep;8(18):1191–1191.

57. Chetan MR, Gleeson F V. Radiomics in predicting treatment response in non-small-cell lung cancer: current status, challenges and future perspectives. Eur Radiol. 2021 Feb 1;31(2):1049–58.

58. Pan F, Feng L, Liu B, Hu Y, Wang Q. Application of radiomics in diagnosis and treatment of lung cancer. Front Pharmacol. 2023 Nov 1;14.

59. Wen Q, Yang Z, Dai H, Feng A, Li Q, Solaiyappan M, et al. Radiomics Study for Predicting the Expression of PD-L1 and Tumor Mutation Burden in Non-Small Cell Lung Cancer Based on CT Images and Clinicopathological Features. 2021;11:1. Available from: www.frontiersin.org

60. Wu YJ, Wu FZ, Yang SC, Tang EK, Liang CH. Radiomics in Early Lung Cancer Diagnosis: From Diagnosis to Clinical Decision Support and Education. Diagnostics. 2022 Apr 24;12(5):1064.

61. Binczyk F, Prazuch W, Bozek P, Polanska J. Radiomics and artificial intelligence in lung cancer screening. Transl Lung Cancer Res [Internet]. 2021 Feb 1 [cited 2023 Dec 6];10(2):1186. Available from: /pmc/articles/PMC7947422/

62. A K, BJ B, S R, RA K, C V, F M, et al. Do we need to see to believe?-radiomics for lung nodule classification and lung cancer risk stratification. J Thorac Dis [Internet]. 2020 Jun 1 [cited 2021 Jul 19];12(6):3303–16. Available from: https://pubmed.ncbi.nlm.nih.gov/32642254/

63. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Stiphout RGPM van, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer [Internet]. 2012 Mar [cited 2021 Oct 1];48(4):441. Available from: /pmc/articles/PMC4533986/

64. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016 Feb;278(2):563–77.

65. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clin Radiol. 2017 Jan;72(1):3–10.

66. Zanfardino M, Franzese M, Pane K, Cavaliere C, Monti S, Esposito G, et al. Bringing radiomics into a multi-omics framework for a comprehensive genotype–phenotype characterization of oncological diseases. J Transl Med. 2019 Dec 7;17(1):337.

67. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017 Dec 4;14(12):749–62.

68. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM J Res Dev. 1959 Jul;3(3):210–29.

69. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism. 2017 Apr;69:S36–40.

70. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. Korean J Radiol. 2017;18(4):570.

71. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics. Radiol Med. 2021 Oct 2;126(10):1296–311.

72. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. Journal of Nuclear Medicine. 2020 Apr;61(4):488–95.

73. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp. 2018 Dec 14;2(1):36.

74. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. International Journal of Radiation Oncology*Biology*Physics. 2018 Nov;102(4):1143–58.

75. Wells WM, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. Med Image Anal. 1996 Mar;1(1):35–51.

76. Jean Serra. Image Analysis and Mathematical Morphology ISBN 0-12-637240-3 . 1982.

77. JJM van G, A F, C P, A H, N A, V N, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res [Internet]. 2017 Nov 1 [cited 2021 Jul 19];77(21):e104–7. Available from: https://pubmed.ncbi.nlm.nih.gov/29092951/

78. Ligero M, Torres G, Sanchez C, Diaz-Chito K, Perez R, Gil D. Selection of Radiomics Features based on their Reproducibility. Annu Int Conf IEEE Eng Med Biol Soc [Internet]. 2019 Jul 1 [cited 2021 Dec 29];2019:403–8. Available from: https://pubmed.ncbi.nlm.nih.gov/31945924/

79. Nomogram for Bayes's Theorem. New England Journal of Medicine. 1975 Jul 31;293(5):257–257.

80. Beig N, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, et al. Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. Radiology. 2019 Mar;290(3):783–92.

81. Calheiros JLL, de Amorim LBV, de Lima LL, de Lima Filho AF, Ferreira Júnior JR, de Oliveira MC. The Effects of Perinodular Features on Solid Lung Nodule Classification. J Digit Imaging. 2021 Aug 31;34(4):798–810.

82. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern. 1973 Nov;SMC-3(6):610–21.

83. Tomassini S, Falcionelli N, Sernani P, Burattini L, Dragoni AF. Lung nodule diagnosis and cancer histology classification from computed tomography data by convolutional neural networks: A survey. Comput Biol Med. 2022 Jul;146:105691.

84. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE; 2018. p. 4510–20.

85. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A Next-generation Hyperparameter Optimization Framework. 2019 Jul 25;

86. L L, L Q, Z X, Y Y, X W, B K, et al. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. Radiology [Internet]. 2020 Aug 1 [cited 2021 Jul 19];296(2):E65–71. Available from: https://pubmed.ncbi.nlm.nih.gov/32191588/

87. R Z, X T, Z Q, NB B, C Z, D G, et al. Diagnosis of Coronavirus Disease 2019 Pneumonia by Using Chest Radiography: Value of Artificial Intelligence. Radiology [Internet]. 2021 Feb 1 [cited 2021 Jul 19];298(2):E88–97. Available from: https://pubmed.ncbi.nlm.nih.gov/32969761/

88. Liu A, Wang Z, Yang Y, Wang J, Dai X, Wang L, et al. Preoperative diagnosis of malignant pulmonary nodules in lung cancer screening with a radiomics nomogram. Cancer Commun. 2020 Jan 3;40(1):16–24.

89. Lin R, Zheng Y, Lv F, Fu B, Li W, Liang Z, et al. A combined non-enhanced CT radiomics and clinical variable machine learning model for differentiating benign and malignant sub-centimeter pulmonary solid nodules. Med Phys. 2023 May 2;50(5):2835–43.

90. Swensen SJ. The Probability of Malignancy in Solitary Pulmonary Nodules. Arch Intern Med. 1997 Apr 28;157(8):849.

11. ANEXOS

11.1. Premios

- 1er premio por la mejor comunicación en el congreso XLI Diada pneumològica – SOCAP (Societat Catalana de Pneumologia) 2024: “Millora de la predicció de la malignitat del nòdul pulmonar mitjançant la integració clínica en un model radiòmic de deep learning”

- 1er premio por la mejor comunicación en el congreso XXXIX Diada pneumològica – SOCAP (Societat Catalana de Pneumologia) 2022: “Aplicació de la radiòmica en el diagnòstic del nòdul pulmonar. Resultats preliminars del projecte Radiolung”

11.2. Abstract presentados en congresos internacionales

11.2.1. Artificial intelligence to optimize pulmonary embolism diagnosis during Covid-19 pandemic by perfusion SPECT/CT, a pilot study.

S. Baeza Mena; R. Domingo; M. Salcedo-Pujantell; et al; D. Gil. Artificial Intelligence to Optimize Pulmonary Embolism Diagnosis During Covid-19 pandemic by Perfusion SPECT/CT, a Pilot Study. American Journal of Respiratory and Critical Care Medicine 2021;203:A3726.
DOI: https://doi.org/10.1164/ajrccm-conference.2021.203.1_MeetingAbstracts.A3726

Artificial Intelligence to Optimize Pulmonary Embolism Diagnosis During Covid-19 Pandemic by Perfusion SPECT/CT, a Pilot Study

S. Baeza Mena[1], R. Domingo[2], M. Salcedo-Pujantell[3], G. Moragas[3], J. Deportós[3], I. Garcia-Olivé[1], C. Sanchez[2], D. Gil[2], A. Rosell[1]; [1]Respiratory Medicine, Hospital Universitari Germans Trias i Pujol, Badalona, Spain, [2]Image Processing and Artificial Intelligence, Computer Vision Center-UAB, Bellaterra, Spain, [3]Nuclear Medicine, Hospital Universitari Germans Trias i Pujol, Badalona, Spain.

INTRODUCTION: COVID-19 pneumonia is associated with a high rate of pulmonary embolism (PE) and its diagnosis can be changeling. In patients with contraindications for CT pulmonary angiography (CTPA) or nondiagnostic on CTPA, perfusion single photon emission computed tomography/computed tomography (Q-SPECT/CT) is a diagnosis option in patients with COVID-19 disease. Our aim was to develop an artificial intelligence (AI) model based on Q-SPECT/CT images of patients during the COVID-19 pandemic which is able to classify lung lesions to optimize PE diagnosis. METHODS: Single center study with a prospective observational branch with patients who tested positive for COVID-19 and underwent a perfusion SPECT-CT study for diagnosis of PE, from April to September 2020. The second branch is retrospective, with patients in pre-pandemic period who underwent Q-SPECT/CT for diagnosis of PE from January to December 2018. For each patient, a Q-SPECT/CT based on intravenous administration of 6 mCi (222MBq) of 99mTC-macroaggregates of human-albumin (99mTC-MAA) was performed, with the subsequent tomo-scintigraphy (SPECT) and a CT. The entire image pre-processing (volume exploration, segmentation analysis, registration analysis) was conducted with MATLAB. The final diagnosis for each patient and the different tissue type segments were analyzed and validated by two senior nuclear physicians. Our Intelligent Radiomic System for the identification of patients with PE (with or without pneumonia) is based on a local analysis of SPECT-CT volumes that considers both CT and SPECT values for each volume point. Volumes are first co-registered and their intensity is normalized in [0, 1] to account for differences due to acquisition. A support vector machine (SVM) model was trained to discriminate among 4 different types of tissue: no pneumonia without PE (control group), no pneumonia with PE, pneumonia without PE and pneumonia with PE. We followed a k-fold (with k=30) scheme for statistical analysis of results. RESULTS: We collected 133 patients, 63 in prospective branch (26 with PE, 22 without PE, 15 indeterminate-PE) and 70 in retrospective branch (31 healthy/control, 39 PE). Concerning the local analysis, for all cases we obtain an average sensitivity and positive predictive value over 92%. CONCLUSION: This study represents a first step towards a complete intelligent radiomic system to optimize the diagnosis of PE by Q-SPECT/CT. The capability to detect alterations in perfusion for COVID-19 pneumonia encourages developing a tool in the cloud for clinical use and further investigates if it can also predict long term complications.

	No Pneumonia		Pneumonia	
	No PE	PE	No PE	PE
Recall (Sensitivity)	0.99±0.0007	0.98±0.002	0.99±0.0005	0.93±0.017
Precision (Positive predictive value)	0.98±0.007	0.95±0.014	0.99±0.0002	0.93±0.017

Table 1. Recall and Precision (mean +/- SD) of each tissue type.

This abstract is funded by: None

11.2.2. Radiomics to increase the effectiveness of lung cancer screening programs. Radiolung preliminary results

A Rosell; S Baeza; S García-Reina; JL Mate; I Guasch, et al, D Gil. EP01.05-001 Radiomics to Increase the Effectiveness of Lung Cancer Screening Programs. Radiolung Preliminary Results. JTO. Sep 2022. 17 (9), S182. DOI: https://doi.org/10.1016/j.jtho.2022.07.302

stimulate further research and model development to enhance future and existing LC screening programs. **Keywords:** Lung Cancer Screening, Personalization, Risk scores

EP01.05 EARLY DETECTION AND SCREENING - PULMONARY NODULE

EP01.05-001
Radiomics to Increase the Effectiveness of Lung Cancer Screening Programs. Radiolung Preliminary Results

A. Rosell,[1] **S. Baeza**,[1] **S. Garcia-Reina**,[1] **J.L. Mate**,[1] **I. Guasch**,[1] **I. Nogueira**,[1] **G. Torres**,[2] **C. Sanchez-Ramos**,[2] **D. Gil**[2] [1]*Hospital Germans Trias, Badalona/ES,* [2]*Centre de Visió Per Computador (UAB), Bellaterra/ES*

Introduction: Main lung cancer screening clinical trials demonstrated a 20-25% mortality reduction but rendered 19.7-27.3% false-positive nodules in the chest CT. Radiomics can provide discriminatory capacity beyond what is perceived by the naked eye. **Methods:** GoalEstablish a radiomic signature of pulmonary nodules (PN) to distinguish malignancy from benignity.PatientsProspective observational study with PNs studied and resected according to usual clinical practice.Method CT images are sent to the Computer Vision Center, and the nodules segmented. Gray Level Co-occurrence Matrix radiomic based texture features that significatively correlate with malignancy are extracted. These variables are used to train a neural network with an architecture optimized according to the diagnosis of the nodule. The model has been trained with 8 benign PNs and 43 malignant PNs from our hospital, and subsequently validated with 6 benign and 21 malignant PNs from our hospital and from a public database. **Results:** Results51 PNs of 22.68mm (range 3-45mm), 32 men, mean age 69 years, 41 smokers or ex-smokers were analyzed. The pathological results were: 32 (62.7%) ADK, 7 (13.7%) SCC, 3 (5.8%) NSCLC, 1 (1.9%) carcinoid; 2 (3.9%) nonspecific necrosis, 6 (11.7%) benign tumors. The diagnostic accuracy of our hybrid system is 96.30%, 100% sensitivity, specificity of 83.3%. **Conclusions:** ConclusionsIn our sample, the application of a hybrid radiomic system achieves high diagnostic accuracy(96.3%) to detect malignant nodules on chest CT. External validation in a lung cancer screening program is needed. Funded by: ACMCiB, BRN, Fundació Ramon Pla, Lung Ambition Alliance **Keywords:** RADIOMICS, PULMONARY NODULE, SCREENING

EP01.05-002
Role of the Lung Cancer Screening MDT in the Manchester Lung Health Check Programme

A. Ghoshal, **P. Bradley**, **A. Whales**, **A. Alonso**, **P. Crosbie**, **R. Booton**, **H. Balata** *Manchester University NHS Foundation Trust, Manchester/GB*

Introduction: Low-dose CT (LDCT) screening for lung cancer has been shown to reduce lung cancer mortality and its implementation is currently being examined and considered across multiple countries. In the United Kingdom (UK), the Targeted Lung Health Checks (TLHC) program has initiated small-scale screening programs across several regions of the country. A significant challenge to implementation is the increased demand for quality-assured radiology reporting, as stipulated by the TLHC service protocols, in order to maximize screening efficacy and reduce potential harm to participants from false-positive and false-negative results. In this analysis, we describe outcomes from the first six months of case discussions at the Manchester Lung Health Checks screening Multidisciplinary Team meeting(MDT). **Methods:** The expanded Manchester Lung Health Check screening program was initiated in 2019, having been successfully piloted between 2016-2019. Ever-smokers, aged 55-80, are invited to a free community-based LHC. Those at higher risk of lung cancer ($PLCO_{m2012NoRACE} \geq 1.5\%$) are invited to annual LDCT screening. Scans are initially analyzed through a computer-aided detection software before being formally reported by an external thoracic radiology reporting company. LDCT results are categorized as either 'Negative', 'Indeterminate' (requiring close surveillance or clinical review), or 'Positive' (requiring assessment for malignancy) as per TLHC reporting protocols. In Manchester, all Indeterminate and Positive screening results are triaged and then discussed in a weekly lung cancer screening MDT, attended by the program's responsible clinicians, responsible radiologists, and specialist nurses, before confirming final screening outcomes for individual participants. When applicable, previous radiology imaging is reviewed as a comparator. This analysis describes initial screening LDCT and subsequent MDT outcomes for all cases discussed at the Manchester screening MDT between August 2021, when the MDT was first established, and February 2022. **Results:** A total of 358 cases were reviewed at the screening MDT in the analyzed time period, including 58 positive and 300 indeterminate results. Of the 58 positive results discussed, 53.4% (n=31) were downgraded from requiring immediate investigation for malignancy (positive) to requiring surveillance imaging only (indeterminate). Seven of the 31 downgraded cases have already had their surveillance imaging performed from which no cancers have been diagnosed to date. Of the 300 indeterminate results, 25.3% (n=76) were downgraded from requiring a 3-month surveillance scan (indeterminate) to an interval 2-year LDCT scan (negative). However, 3% (n=9) cases were upgraded from requiring a short-term surveillance scan (indeterminate) to a positive outcome and investigated immediately for malignancy, four of which have been subsequently diagnosed with lung cancer. Overall, initial screening LDCT outcomes were altered in a third (32.4%; n=116) of cases discussed at the screening MDT. **Conclusions:** A lung cancer screening MDT, where all indeterminate and positive screening outcomes are reviewed, plays an important role in providing prospective quality assurance of initial radiology reporting to maximize screening efficacy, reduce unnecessary harm and optimise the use of limited resources. A limitation of our analysis is that long-term clinical outcomes are not yet available for participants discussed at the MDT. **Keywords:** Lung Cancer Screening, Screening MDT, Targeted Lung Health Check

EP01.05-003
Changes of Repeated Lung Cancer Screening Results and Affected Factors; Analysis of Korean National Lung Cancer Screening Program

E. Kang, **N.-Y. Lee**, **Y. Kim** *National Cancer Center, Gyeonggi-do/KR*

Introduction: Korean national lung cancer screening program (KNLCS) from 2019 provides screening test to high-risk smokers (30 pack-year or more) between the ages of 54-74 with low-dose computed tomography (LDCT) every two years. According to the results of lung cancer screening modeling studies, repeated screening is more beneficial than a single screening. This study aimed to investigate changes of repeated screening compared to initial screening, and to analyze affected factors to the changes. **Methods:** This study analyzed lung cancer screening results in national cancer center data base which includes about 50% of KNLCS. In 2019, 50,666 examinees participated to lung cancer screening, and in 2021, 57,578 participated. A total of 12,802 subjects were screened in both 2019 and 2021. Based on Lung RADS, categories 1 to 2 represent negative screening results, and categories 3 to 4 define as positive results. In this study, frequency analysis and multiple logistic analysis was performed for determination of factors associated with changes in LDCT results to positive among the examinees whose results were negative in initial screening. In logistic model, predictors included age, sex, past medical history (cancer, pulmonary disease),

11.2.3. Enhanced lung nodule malignancy prediction through clinical integration in a deep-learning radiomic model

A. Rosell; S. Baeza; G. Torres; C. Sanchez-Ramos; C. Tebé; et al; D. Gil. 132P Enhanced lung nodule malignancy prediction through clinical integration in a deep-learning radiomic model. ESMO open. March 2024. 9 (3), 102720 DOI: https://doi.org/10.1016/j.esmoop.2024.102720

Table: 131P Regression analysis of PDL1xEGFR co-expression with pSUVmax (ref: both negative)

	pSUVmax	OR	p-value
Positive EGFR- Positive PDL1	<= 10.665 vs >20.51	1.714	0.409
Positive EGFR- Negative PDL1	<= 10.665 vs >20.51	15.714	0.017*
Negative EGFR- Positive PDL1	<= 10.665 vs >20.51	1.107	0.853

Conclusions: Molecular spectrum encompassing oncogenic drivers and immune checkpoints in NSCLCs resulted in variegated metabolic features. Metabolic parameters on FDG PET can provide valuable information about tumor molecular profile and can potentially act as effective non-invasive imaging biomarker.

Legal entity responsible for the study: The authors.

Funding: Has not received any funding.

Disclosure: All authors have declared no conflicts of interest.

https://doi.org/10.1016/j.esmoop.2024.102719

132P Enhanced lung nodule malignancy prediction through clinical integration in a deep-learning radiomic model

A. Rosell[1], S. Baeza[1], G. Torres[2], I. Garcia-Olivé[1], I. Nogueira[1], F. Andreo[1], A. Hernandez-Gallego[1], J.L. Mate[1], S. Peñafiel[1], C. Martinez-Barenys[1], J. Carmezin[3], C. Sanchez-Ramos[2], C. Tebé[3], D. Gil[2]

[1]Hospital Universitari Germans Trias i Pujol, Barcellona, Spain; [2]UAB - Universitat Autonoma de Barcelona, Cerdanyola Del Vallès, Spain; [3]IGTP - The Institute for Health Science Research Germans Trias i Pujol, Barcelona, Spain

Background: Radiomics, a quantitative imaging analysis technique powered by artificial intelligence, holds promise for lung cancer diagnosis. However, conventional radiomic diagnosis models are often based only on image-based features, potentially limiting their accuracy. This study assesses the impact of integrating clinical information into a deep-learning model for determining the malignancy of pulmonary nodules.

Methods: A cross-sectional study of 87 resected nodules from 90 patients was conducted. Clinical data included demographic variables, epidemiological risk factors, and pulmonary function tests. The region of interest of each Chest CT containing the nodule was extracted and analyzed. The radiomic model used the MobileNetV2 extractor followed by a selection of features having positive standard deviation. This filter reduced the 1280 MobileNetV2 features to 333 and the input to an optimized, fully connected network architecture. This architecture had a hidden layer with 150 neurons, dropout with a probability of 0.8, sigmoid activation function, and batch normalization. The clinical model was a logistic regression to predict benignity using 2000 samples generated by bootstrapping to select the best predictors from the clinical variables.

Results: The nodules had a mean diameter of 17.9 mm, histology: 59% adeno, 15% squamous, 2% other malignant, 1% carcinoid, and 23% benign. In the study sample with a malignancy prevalence of 77%, the radiomic model based on visual features exhibited an accuracy of 62% (95% CI: 0.52-0.73) and an AUC of 0.59 (0.49-0.73). The clinical model identified DLCO, tobacco pack-year index, and smoking status as the most consistent clinical predictors associated with the outcome. Integrating the clinical features into the radiomic model resulted in significantly improved accuracy, achieving an accuracy of 79% (95% CI: 69-87%) and an AUC of 0.75 (95% CI: 56-90), increasing 27% in both cases.

Conclusions: Integrating clinical information into a deep learning radiomic model for lung nodule malignancy assessment enhanced the model's accuracy and predictive performance by 27%. This study supports the potential of combined image-based and clinical features to improve lung cancer diagnosis.

Legal entity responsible for the study: Research Institute Germans Trias (IGTP).

Funding: Acadèmia de Ciències Mèdiques de Catalunya i Balears | Barcelona Respiratory Network (BRN)- Fundació Ramon Pla i Armengol Lung Ambition Alliance | Hosptial Germans Trias i Pujol (Llegat JMC) Ministerio de Economía, Industria y Competitividad, Gobierno de España grant number PID 2021-126776OB-C21 Agència de Gestió d'Ajuts Universitaris i de Recerca grant numbers 2021 SGR 01623 and CERCA Programme / Generalitat de Catalunya.

Disclosure: All authors have declared no conflicts of interest.

https://doi.org/10.1016/j.esmoop.2024.102720

133P Development of CT image-based atelectasis simulation and non-invasive lung nodule localization system

J. Hwang[1], I. Hwang[2], C. Kim[2], S. Ham[2]

[1]Thoracic and Cardiovascular Surgery, Korea University Ansan Hospital, Ansan, Republic of Korea; [2]Korea University Ansan Hospital, Ansan, Republic of Korea

Background: This study creates an atelectasis model by applying a force in the direction of gravity according to the patient's surgical posture based on a three-dimensional model for effective tracking of the location of the pulmonary nodule during video-assisted thoracic surgery(VATS).

Methods: After obtaining Contrast-enhanced computed tomography(CT) images from 12 patients with pulmonary nodules. The images are converted into 3d models. Each vector point of the 3d model is relocated in the direction of gravity according to the direction of surgery. The shape of the deformed lung and the nodule's location were compared with the surgical video to evaluate the simulation similarity. To measure the accuracy of the simulation, images from surgical videos were obtained. After selecting a similar location based on the location of the sternum, an image of the same location was obtained in the simulation.

Results:

Table: 133P

Patient Number	Age	Gender	Location	Number of matching	DSC	JSC	HD	Pearson coefficient r
1	53	M	RUL	12	88.10	80.63	8.86	0.78
2	44	F	LUL	27	93.26	92.92	2.15	0.88
3	57	M	LUL	13	85.48	80.78	8.25	0.76
4	84	F	LUL	9	87.78	86.89	7.62	0.79
5	72	M	RUL	15	93.88	92.94	2.85	0.89
6	57	M	LLL	30	93.76	92.95	3.47	0.8
7	76	F	LLL	16	88.06	86.82	9.92	0.71
8	64	F	RUL	13	93.88	92.97	2.11	0.84
9	46	F	LUL	13	92.82	90.68	3.85	0.88
10	50	M	LUL	20	85.23	84.39	8.24	0.83
11	68	F	RUL	14	93.25	92.84	2.87	0.89
12	53	F	RLL	17	85.01	84.93	10.03	0.72

Conclusions: Through this retrospective study, we developed a system that can predict the position of nodules in thoracoscopic surgery. Also, this research presented a methodology for implementing and verifying the accuracy of the atelectasis simulation developed.

Legal entity responsible for the study: The authors.

Funding: National Research Foundation of Korea.

Disclosure: All authors have declared no conflicts of interest.

https://doi.org/10.1016/j.esmoop.2024.102721

134P Predictors of success and complications in electromagnetic navigation bronchoscopy of peripheral lung lesions: A retrospective cohort study

D. Lang[1], M. Negele[2], D. Sima[2], V. Rambousek[1], A. Horner[1], R. Schimbäck[1], B. Kaiser[1], B. Lamprecht[1]

[1]Department of Pulmonology, Johannes Kepler University Linz, Kepler University Hospital, Linz, Austria; [2]Johannes Kepler University Linz, Faculty of Medicine, Linz, Austria

Background: We evaluated diagnostic performance and complication rate of electromagnetic navigation bronchoscopy (ENB) as well as predictors of outcomes in a single-center cohort.

Methods: All patients having undergone ENB using a superDimension system (Medtronic, Minneapolis, Minnesota) at Kepler University Hospital Linz, Austria, between 2014 and 2022 were retrospectively enregistered. ENB was performed under general anaesthesia, standardized post-interventional follow-up included a chest X-ray after three hours. Severe complications were defined as pneumothorax or bronchial bleeding with need for intervention, procedure-related death, or ICU admission. Success was defined as reaching the lesion to <1cm of the predefined target and no need for further diagnostic intervention.

Results: A total of 238 separate ENB procedures in 231 patients (mean age 66 years, 58% male) were evaluated. The procedure was successful in 124 (52%), the target lesion was reached but information obtained was insufficient in 98 (41%), and the target lesion could not be reached in 16 (7%) patients. Success was significantly associated with smoking history <30 pack-years (OR 2.75, p=0.0013), lesion location in the middle (OR 3.68, p=0.002) or upper (OR 2.79, p=0.01) versus the lower lung third in the frontal plane, and with a positive bronchus sign (OR 2.95, p<0.001). Pneumothorax occurred in 32 (13.5%) patients, bronchial bleeding in 5 (2.1%), overall

11.3. Otras publicaciones relevantes

11.3.1. An Intelligent Radiomic Approach for Lung Cancer Screening

Título: An Intelligent Radiomic Approach for Lung Cancer Screening

Autores: Guillermo Torres, Sonia Baeza, Carles Sanchez, Ignasi Guasch, Antoni Rosell, Debora Gil.

Revista: Applied Sciences.

Volumen:12

Año: 2022

Factor de impacto: 2.7

Cuartil: 2 (Q2)

Citar como: Torres G, Baeza S, Sanchez C, Guasch I, Rosell A, Gil D. An Intelligent Radiomic Approach for Lung Cancer Screening. Applied Sciences. 2022; 12(3):1568. https://doi.org/10.3390/app12031568

applied sciences

MDPI

Article

An Intelligent Radiomic Approach for Lung Cancer Screening

Guillermo Torres [1,*], Sonia Baeza [2,3,4], Carles Sanchez [1], Ignasi Guasch [2,3], Antoni Rosell [2,3,4,5] and Debora Gil [1,6]

1 Computer Vision Center (CVC), Computer Science Department, Universitat Autònoma de Barcelona (UAB), 08193 Barcelona, Spain; csanchez@cvc.uab.cat (C.S.); debora@cvc.uab.cat (D.G.)
2 Respiratory Medicine Department, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain; smbaeza.germanstrias@gencat.cat (S.B.); iguasch.germanstrias@gencat.cat (I.G.); arosellg.germanstrias@gencat.cat (A.R.)
3 Germans Trias i Pujol Research Institute (IGTP), 08916 Badalona, Spain
4 Medicine Deptartment, Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain
5 Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), 28029 Madrid, Spain
6 Serra Hunter Fellow, Escola d'Enginyeria, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain
* Correspondence:gtorres@cvc.uab.cat

Abstract: The efficiency of lung cancer screening for reducing mortality is hindered by the high rate of false positives. Artificial intelligence applied to radiomics could help to early discard benign cases from the analysis of CT scans. The available amount of data and the fact that benign cases are a minority, constitutes a main challenge for the successful use of state of the art methods (like deep learning), which can be biased, over-fitted and lack of clinical reproducibility. We present an hybrid approach combining the potential of radiomic features to characterize nodules in CT scans and the generalization of the feed forward networks. In order to obtain maximal reproducibility with minimal training data, we propose an embedding of nodules based on the statistical significance of radiomic features for malignancy detection. This representation space of lesions is the input to a feed forward network, which architecture and hyperparameters are optimized using own-defined metrics of the diagnostic power of the whole system. Results of the best model on an independent set of patients achieve 100% of sensitivity and 83% of specificity (AUC = 0.94) for malignancy detection.

Keywords: lung cancer; early diagnosis; screening; neural networks; image embedding; architecture optimization

Citation: Torres, G.; Baeza, S.; Sanchez, C.; Guasch, I.; Rosell, A.; Gil, D. An Intelligent Radiomic Approach for Lung Cancer Screening. *Appl. Sci.* **2022**, *12*, 1568. https://doi.org/10.3390/app12031568

Academic Editor: Vera Yuk Ying Chung

Received: 20 December 2021
Accepted: 25 January 2022
Published: 31 January 2022

1. Introduction

Lung cancer is both the most frequently diagnosed cancer and cause of cancer death [1]. The National Lung Screening Trial (NLST) (Study performed in United States) [2] and Dutch-Belgian Randomized Lung Cancer Screening Trial (NELSON) [3] have shown that lung cancer screening (LCS) with computed tomography of low dose (CTLD) reduces mortality by 20–25%. However, the average of false positive rate of the radiological diagnosis obtained by visual inspection of scans was 23% of the nodules detected. This inaccuracy meant long follow-up of patients with repetitive CT or performing an invasive procedure like a biopsy or surgery, which accounted to be futile in 73% of the cases. A reduction of false positives would increase the efficiency of screening for early detection of lung cancer.

The largest screening program in Europe, the NELSON study, introduced volumetry of the nodules in consecutive CT, which meant a significant reduction of the average of false positive rate to 13%. This suggests that the application of radiomics [4] (a recent discipline that extracts a large number of image features correlating to treatment outcome), could represent a critical shift in the reduction of the false positive rate and an improvement of early diagnosis of lung cancer.

In this sense, in a pilot study [2] the authors retrospectively extracted 150 quantitative image features and performed a random forest classification, which finally obtained a

significantly better predictive value than volumetry alone (AUC = 0.87 vs. 0.74). More recently, Peikert et al. [5] built a radiomic classifier based upon eight quantitative radiologic features selected by the least absolute shrinkage and selection operator (LASSO) method from 726 indeterminate nodules of the LCST. These eight features include variables capturing location, size, shape descriptors and texture analysis. In this retrospective study, the optimism-corrected AUC for these eight features was 0.939 with a sensitivity and specificity of, respectively, 90% and 85%.

An alternative to classic radiomics is the use of machine learning methods that extract image features using well known methods such as, Gabor, Local Binary Patterns (LBP), or SIFT descriptor to represent a nodule. Then machine learning technics (e.g., Support Vector Machine (SVM) and Random Forest) are used to define a classification of nodules in this representation space according to their diagnosis [6,7]. This methods achieve better diagnostic power than radiomic methods with AUC equal to 0.97, sensitivity equal to 96% with 95% of specificity for [6].

Recently, deep CNN (CNN stands for Convolutional Neural Network) have achieved great success in various computer vision tasks, such as image classification, segmentation, and enhancement. Researchers are therefore inspired to classify nodules by using CNNs. Existing works based on CNNs can be classified according to some key points: input data (2D or 3D), level of the input image (volume, slice, nodule), CNN architecture, resources needed and obtained performance.

The early work of Shen et al. proposed to use a multi-crop CNN [8] to make the model robust to scales of nodules while keeping 2D input images. Results showed an overall accuracy (including malign and benign cases) of 87%. However, the authors did not report sensitivity for malignancy detection and specificity for discarding benign nodules and, thus, its true clinical value is uncertain.

Since nodules are 3D structures, recent works have addressed the problem using 3D CNNs. Yan et al. [9] explored 3D CNNs for pulmonary nodule classification in comparison to a slice-level 2D CNN and a nodule-level 2D CNN analysis. The 3D approach was the best performer with a 87% of overall accuracy and similar specificity and sensitivity at the cost of a significantly higher demand of computational resources and annotated data. Zhu et al. [10] used 3D deep dual path networks (DPNs) a 3D Faster Regions with Convolutional Neural Net (R-CNN) designed for nodule detection with 3D dual path blocks and a U-net-like encoder-decoder structure to effectively learn nodule features. Despite the complex architecture used, this approach could only achieve a 81% of sensitivity and specificity was not reported. Jiang et al. [11] sequentially deployed a contextual attention module and a spatial attention module to 3D DPN to increase the representation ability. A main novelty of this work is that it ensembles different model variants to improve the prediction robustness. Results show an increase of sensitivity to 90% while keeping a specificity similar to [9]. A main concern is the huge amount of parameters that require extensive data and computational resources for training. In an attempt to minimize computational and data costs, the very recent [12] uses automatic Neural Architecture Search (NAS) technique [13] to design optimal 3D network architectures including attention modules. Results on a subset of the LIDC-IDRI database [14] show a specificity of 95% at the cost of a drop in sensitivity to 85%.

A main challenge in the application of deep learning to biomedical problems is the limited amount of good quality data with annotations, which is a must for training new models with complex architectures. Besides, in the case of benign nodule screening, this is aggravated with the fact that the problem is highly unbalanced with benign cases being the minority class. Under such experimental settings, models are often over-fitted [15] results are non-reproducible [15,16] and most times [9–12] do not outperform conventional machine learning approaches [6]. Another pitfall, especially for deep methods is models should also be easily interpreted from a clinical point of view to allow the analysis of the clinical factors that have an impact on the clinical decision [17].

The output of a classic CNN are features that have no meaning from radiological point of view. In this way, introducing classic radiomic features in the models will be helpful for radiologist in the interpretability of the results by means of most correlated features to malignancy of tumours. It is worth to mention that radiomic features can describe tumour heterogeneity [18], which is a parameter related to malignancy and well known from radiologist. To address the above challenges, we propose to embed 2D slices into a low dimensional radiomic space defined by the classic radiomic features that significantly correlate to malignancy. These features are the input to a fully connected network with an architecture optimized in order to ensure maximum clinical outcome. To do so, novel specific criteria and metrics measuring diagnostic power are presented. Models were optimized in a set of 51 nodules coming from an own collected data base. Results on an independent set of patients from the same data base and LIDC-IDRI database show that our approach outperforms deep approaches with only requiring 290 parameters (in contrast to the thousands required by deep methods).

2. Materials and Methods

This work complies with the fundamental ethical principles of research (Declaration of Helsinki—Fortaleza/Brazil, 2013). In each case, informed consent is requested, and both the images with clinical data are treated anonymously, safeguarding the confidentiality of the patient. This study was approved by the ethics committee of the HUGTiP prior to the start of recruitment (CEIC H. Germans Trias i Pujol: PI-19-169).

2.1. Dataset Description

The patients were recruited at the Germans Trias i Pujol University Hospital (HUGTiP), Barcelona, Spain, from which images and clinical/demographic data were collected between December 2019 and September 2020. The 60 recruited patients have CT-chest and pulmonary nodule (PN) tributary of surgery, and meet the following inclusion and exclusion criteria. The inclusion criteria includes: have a single PN, diameter from 8 to 30 mm, final diagnosis of non-small cell lung carcinoma and non-malignant tumor. And the exclusion criteria includes: have been previously diagnosed with lung cancer, diagnosis of uncured extra-pulmonary cancer (except non-melanoma skin cancer), pregnancy, have received chemotherapy or cytotoxic drugs in the last 6 months and decline to sign the consent. The PNs have been classified in every case by means of a biopsy.

Scans were acquired with GE Medical Systems and Philips CT scan. For both devices, acquisition parameters in all cases were 120 kv, 100–350 mA (dose modulation range), soft tissue reconstructions, high frequency algorithms and 512 $\times$ 512 matrix. These parameters are the gold standard used to ensure enough scan resolution and quality to radiologically evaluate malignancy [19,20]. Table 1 report the acquisition setting for each manufacturer, as well as the number of benign and malign nodules.

Table 1. Details of the acquisition parameters by scanner manufacturer.

Description\Manufacturer	GE Medical Systems	Philips
Model Name	LightSpeed VCT BrightSpeed Optima CT540 Discovery ST	GeminiGXL 16 Brilliance 16 TruFlight Select
Convoluton Kernel	SOFT STANDARD LUNG	B YA YB YC
Pixel XY size	0.56–0.87	0.36–0.72
Slice Thickness	0.63–1.25	1–2
Benign Nodules	3	6
Malignant Nodules	21	30

A respiratory medicine physician with seven years of experience annotated the Region of Interest (ROI) of each nodule with 3D-Slicer (version 4.11.20200930), which is a free, open source and multi-platform software package widely used for medical, biomedical,

and related imaging research. The physician was asked to define ROIs fitting the minimal nodule space as possible. The coordinates of the bounding box defining the ROI were stored in csv format for its further use in the method pre-processing step described in Section 2.3. A ROI per patient was annotated, since the cases conforming our database have one nodule per patient. The Table 2 shows detailed information about our database. We report demographic information, as well as, the minimum, maximum and every number of slices for each nodule type and sex.

Table 2. Details of our database.

	Description	Male	Female	Total
Demographic population	Patients	36	24	60
	Age avg $\pm$ std	70.67 $\pm$ 6.87	63.96 $\pm$ 12.35	67.98 $\pm$ 9.92
	Benign PNs	5	4	9
	Malign PNs	31	20	51
Nodule characterization	Benign Slices min/max/avg	6/111/48	28/39/32	6/111/41
	Malign Slices min/max/avg	8/152/45	12/82/45	8/152/43

2.2. Methodology Description

Our methodology aggregates, for each nodule, the classification at slice level to obtain a prediction of nodule malignancy. The classification of 2D slices bases on radiomic 2D textural features computed on a mask of the nodule to implicitly account for (2D) shape. Texture descriptors are the input to a feedforward neural network with optimized architecture.

Figure 1 shows a general overview of our workflow, which consists in 3 main phases: extraction of nodules from CT scans, embedding of nodules into a space representing malignancy and nodule diagnosis with optimized network architecture. In the extraction step, the nodule is segmented in the ROI volumes using Otsu thresholding and morphological operations. In the embedding phase, PyRadiomics [21] GLCM descriptors are computed in 2D slices of masked volumes and a *t*-test is used to select those features that significantly correlate to malignancy. In the diagnosis phase, the selected features are the input to an optimized feedforward network trained and the most frequent classification among each nodule 2D slices determines the final diagnosis. The architecture and hyperparameters of the diagnostic networks are optimized according to an own-defined metrics measuring the clinical performance of the system.

2.3. Nodule Extraction

In the preprocessing phase we used the anonymized CT-chest DICOMs and the annotated nodule ROIs. A nodule ROI always includes the intranodule region (inside nodule region), but depending on the nodule shape, the perinodular region (around nodule region) is included in greater or lesser extent. Since, in [22,23], it is reported that the importance of using perinodular region in the classification of benign and malignant nodules, the size of ROIs was enlarged 15% of its original size. The volume ROI extracted using the coordinates of the annotated bounding box is the input to the whole workflow. This is the only manual annotation required.

In order to segment the nodule, we applied Otsu thresholding to the ROI volume. Since the segmentation of peripheral nodules can include non-pulmonary tissue, the binarized volumes were masked with a segmentation of lungs. The final nodule segmentation was the largest connected component of the masked volumes. The segmentation of lungs was computed using thresholding and morphological operations [24]. Specifically, CT lungs were selected as the larger connected component of the voxels with intensity between 950 to -300 Hounsfield Units, followed by a closing with a structuring element of size 5.

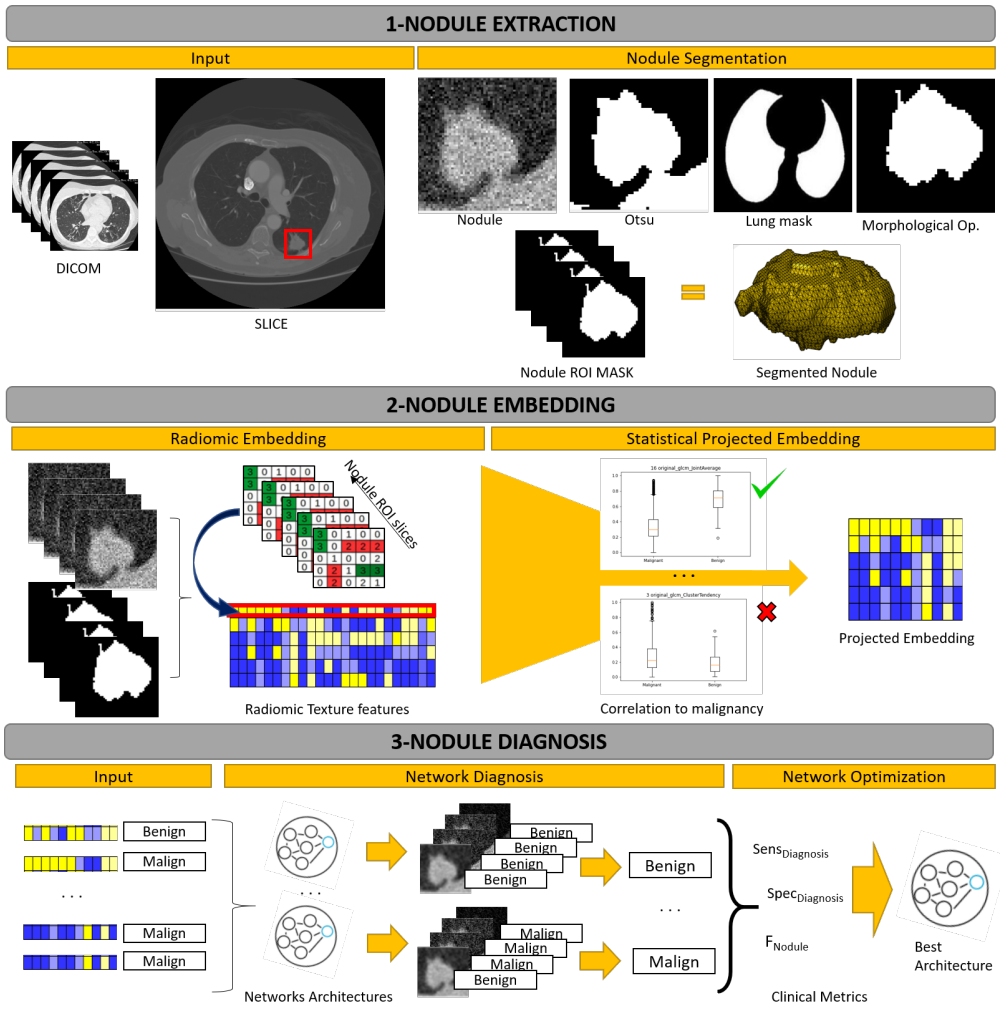


Figure 1. Overall workflow of the proposed work.

2.4. Nodule Embedding

To extract radiomics features we used PyRadiomics [21] (version 3.01), an open-source python package for the extraction of radiomic features from medical imaging volumes. PyRadiomics features include shape features, first order features, and textural features (Gray Level Co-ocurrence Matrix (GLCM), Gray Level Size Zone (GLSZM), Gray Level Run Length Matrix (GLRLM) and Gray Level Dependency Matrix (GLDM)) describing several aspects of the lesion. In this study, we used GLCM texture features [25] by its proven efficiency for cancer diagnosis in a wide range of medical imaging modalities [26–30].

GLCM textural features are probabilistic descriptors computed from a gray level co-ocurrence matrix. This matrix encodes textural patterns in a neighbourhood of each pixel based on different contrasts. To do so, intensity gray values are first discretized using the histogram of the original volume intensity. The width of the histogram bins sets the granularity that GLCM features describe, since neighbouring pixels with a difference in gray level below such parameter are filtered. Bin width, namely Δ, is given by:

$$\Delta = \frac{max(Pixel\_value) - min(Pixel\_value)}{Nbins} \quad (1)$$

for $Pixel\_value$ the intensities of the volume and $Nbins$ the number of histogram bins.

In [29,30], the authors showed the importance of, both, intensity ranges and number of bins. It is reported that a fixed bin count between 30 and 130 bins has good reproducibility and performance. In order to allow for different ranges of intensity in ROIs, while still keeping the texture features informative and comparable inter lesion [26], we first normalize

the CT Hounsfield Units (HU) to a common intensity range. Hounsfield Units are related to intensity values with the following linear transformation:

$$HU = Pixel\_value * Slope + Intercept \quad (2)$$

for $Slope$ and $Intercept$ are acquisition parameters of the CT scan.

HU were mapped to a common intensity range $[0, MaxIntensity]$ using the following formula:

$$Pixel\_value = \frac{(HU - Intercept)/Slope}{\max(HU) - \min(HU)} * MaxIntensity \quad (3)$$

$$= \frac{(HU - Intercept)/Slope}{4000} * MaxIntensity \quad (4)$$

where $\max(HU) = 2976$ was computed as the maximum of the training set volumes HU and $\min(HU) = -1024$ is the HU value for air. The intensity range $MaxIntensity$ and number of bins, $Nbins \in [30, 130]$, are hyper-parameters that were set using grid-search. The optimal values were $MaxIntensity = 24$ and $Nbins = 128$.

The GLCM features (extracted on 2D slices) are selected according to their correlation to lesion malignancy. It is expected that the most discriminant features have values significantly different for slices containing malign and benign nodules. Such differences are detected, for each GLCM feature, with a Student t-test comparing average values for malign and benign slices. In order to account for unbalancing between malignant and benign cases, a k-fold subsampling of malignant slices was performed and max-voting aggregation of significance was used to select the most discriminative features. Sample size was large enough to guarantee normality.

The concatenation of the GLCM features with a p-value under 0.01 are selected to be the input to the classification network. Given that the performance of neural networks is not bias in case of correlated features (unlike other classifiers like logistic regression), no further selection to discard correlations is needed. The most statistically significant 19 GLCM textural features that were selected are shown in Table 3.

2.5. Nodule Diagnosis

The extracted radiomic features are used to feed a feedforward neural network that makes a slice by slice classification. We have defined 4 feedforward neural network architectures composed by a sequence of linear layers with ReLU activation function between them.

Table 4 shows the architectural configurations and the equations to obtain the amount of trainable parameters of each configuration. Each layer of the architecture is described by a tuple $(N^i_{input}, N^i_{output})$ where N^i_{input} represents the number of inputs and N^i_{output} represents the number of outputs of the i-th layer. For the first layer, N^1_{input} is the dimensionality, denoted by N_i, of the input features. For the hidden layers, $N^i_{input} = N^{i-1}_{output}$ and, for the last layer, $N^{last}_{output} = 2$ is the network's output for a binary classification problem with classes equal to benign and malignant (label 1). In our case, N^i_{output} is a function of the number of outputs of the first layer, $N^1_{output} = N_h$. The last column in Table 4 also reports the number of trainable parameters. The number of trainable parameters for a layer is the number of inputs multiplied the number of outputs plus the number of neuron's bias (which is equal to the number of the outputs). Thus, for the i-th layer, the number of parameters is equal to $N^i_{input} * N^i_{output} + N^i_{output}$. Its accumulation is the total amount of trainable parameters of the network that is shown in the last column.

Table 3. GLCM Features Selected with the *t*-test. The first column lists the 24 GLCM textural features. The second one indicates the 19 GLCM features selected by the *t*-student test with a ✓.

GLCM Textural Features	*t*-Test Selection
Autocorrelation	✓
Cluster Prominence	✓
Cluster Shade	✓
Cluster Tendency	✓
Contrast	×
Correlation	✓
Difference Average	×
Difference Entropy	✓
Difference Variance	×
Inverse Difference	✓
Inverse Difference Moment	✓
Inverse Difference Moment Normalized	×
Informational Measure of Correlation 1	✓
Informational Measure of Correlation 2	✓
Inverse Difference Normalized	×
Inverse Variance	✓
Joint Average	✓
Joint Energy	✓
Joint Entropy	✓
Maximum Probability	✓
Maximal Correlation Coefficient	✓
Sum Average	✓
Sum Entropy	✓
Sum Squares	✓

Table 4. Neural network architectures.

Num.	Architecture	# Trainable Parameters
1	$[(N_i, N_h), (N_h, N_h), (N_h, 2)]$	$N_h(N_i + N_h + 4) + 2$
2	$[(N_i, N_h), (N_h, N_h), (N_h, \lfloor \frac{N_h}{2} \rfloor), (\lfloor \frac{N_h}{2} \rfloor, 2)]$	$N_h(N_i + N_h + \lfloor \frac{N_h}{2} \rfloor + 2)$ $+3\lfloor \frac{N_h}{2} \rfloor + 2$
3	$[(N_i, N_h), (N_h, N_h), (N_h, N_h), (N_h, \lfloor \frac{N_h}{2} \rfloor), (\lfloor \frac{N_h}{2} \rfloor, 2)]$	$N_h(N_i + \frac{5}{2}N_h + \frac{7}{2}) + 2$
4	$[(N_i, N_h - 1), (N_h - 1, N_h - 2), (N_h - 2, N_h - 3),$ $(N_h - 3, N_h - 4), (N_h - 4, N_h - 5), (N_h - 5, 2)]$	$(N_h - 1)(N_i + (N_h - 2))+$ $(N_h - 3)(N_h - 6)+$ $(N_h - 5)(N_h - 1)+$ $4N_h - 8$

In order to account for unbalancing in training data, the loss function is a weighted cross entropy given by:

$$\text{loss} = \frac{\sum_{i=1}^{N} weight[class[i]] loss(i, class[i])}{\sum_{i=1}^{N} weight[class[i]]} \tag{5}$$

where $loss(i, class[i])$ is the cross-entropy loss for the i-th class computed from the classifier prediction x and the true class as:

$$\text{loss}(x, class) = -\log\left(\frac{\exp(x[class])}{\sum_{j=1}^{N} \exp(x[j])}\right) \tag{6}$$

and the weight $weight[class[i]]$ is given by the inverse of the class frequency.

The final diagnosis of a nodule is computed from the classification of all the slices of the ROI by an *aggregation operation*. In our case, the aggregation is given by a max voting

of the classifications of 2D slices, so that the most frequent slice classification yields the nodule final diagnosis. That is, if half of the 2D slices are classified as malign, the diagnosis is malign, benign otherwise. In case of tie, a malignant diagnosis is given.

2.6. Network Optimization

In order to select the optimal architectures for malignancy diagnosis, we have defined a *criteria for selection of best models* that uses 3 metrics assessing the diagnostic capability of the system at nodule and slice levels. Taking the malignant nodules as the positive cases, the metrics used to validate the clinical diagnostic accuracy are:

1. **Diagnostic Sensitivity.** This measures the percentage of correctly diagnosed malign nodules:
$$Sens_{Diagnosis} := 100 \frac{TP_{Nodules}}{TP_{Nodules} + FN_{Nodules}} \tag{7}$$
for $TP_{Nodules}$, $FN_{Nodules}$ denoting, respectively, true positives and false negatives for malignancy detection at nodule level.
2. **Diagnostic Specificity.** This measures the percentage of correctly diagnosed benign nodules:
$$Spec_{Diagnosis} := 100 \frac{TN_{Nodules}}{TN_{Nodules} + FP_{Nodules}} \tag{8}$$
for $TN_{Nodules}$, $FP_{Nodules}$ denoting, respectively, true negatives and false positives for benign detection at nodule level.
3. **Slice Diagnostic Index.** This index is an adaptation of the well-known F1-score to measure the percentage of correctly diagnosed slices:
$$F_{Nodule} := \frac{Sens_{Nodule} Spec_{Nodule}}{Sens_{Nodule} + Spec_{Nodule}} \tag{9}$$
being $Sens_{Nodule}$ and $Spec_{Nodule}$, the average sensitivity and specificity for the classification of 2D slices at nodule level. Sensitivity is given by:
$$Sens_{Nodule} := \frac{1}{N_{Malign}} \sum_i \frac{TP^i_{Slice}}{NSlice^i} \tag{10}$$
for N_{Malign} the number of malign nodules, TP^i_{Slice} the true positives for the i-th malign nodule and $NSlice^i$ its number of slices. Specificity is given by:
$$Spec_{Nodule} := \frac{1}{N_{Benign}} \sum_i \frac{TN^i_{Slice}}{NSlice^i} \tag{11}$$
for N_{Benign} the number of benign nodules, TN^i_{Slice} the true positives for the i-th benign nodule and $NSlice^i$ its number of slices.
The F_{Nodule} score measures the trade-off between benign and malign accuracy at nodule level.

Our *criteria for selection of best models* is applied in the next cascade sequence: (1) select the models with highest Diagnostic Sensitivity. If there is only one model, then it is selected. Otherwise, (2) select the models with the highest Diagnostic Specificity. If there is only one model, then it is selected. Otherwise, (3) select the model with highest Slice Diagnostic Index.

3. Results

Two different experiments were carried out:

1. **Model Optimization.** A training and selection of models, which consists in a leave-one-out validation on a training set of patients to select the best model for the benign and malignant classification. In order to assess the benefits of our embedding (labelled *t*-test), models were also trained using all 24 GLCM features (labelled None) and the selection based on reproducibility (labelled Reproducibility) reported in [31] excluding the shape class (see Table 5).
2. **Model Verification.** A testing and assessment of models reproducibility, which is a validation of the best model on an independent set of test patients to assess the reproducibility of results. To assess the advantages of the proposed strategy, the best model selected in the first experiment was compared to state of the art methods.

Table 5. Features from the study [31].

Class	Feature
Fist Order	Entropy TotalEnergy Uniformity
GLCM	Inverse Difference Inverse Difference Moment Joint Energy Joint Entropy Maximum Probability
GLDM	Dependence Non Uniformity Normalized Dependence Variance Large Dependence Emphasis
GLRLM	Run Length Non Uniformity Normalized Run Percentage Short Run Emphasis

3.1. Model Optimization

For this experiment, 51 (85%) patients of the dataset described in Section 2.1 were randomly selected for the optimization of models. This training set had 8 benign and 43 malignant nodules.

The search space for optimizing network architectures given in Table 4 together with their hyperparameters was the following: (1) $N_h \in [6, 7, \ldots, 16, 17]$; (2) optimizer: SGD, Adam, RMSprop; (3) learning rate: 0.01, 0.001, 0.0001; (4) weight initialization: Normal, Xavier, Kaiming, Orthogonal; (5) epochs: 500, 1000, 1500. For each embedding (None, $N_i = 24$, Reproducibility, $N_i = 14$ and *t*-test, $N_i = 19$), we use a grid search to optimize networks and the best configuration was selected according to the criteria given in Section 2.6. Table 6 shows the selected architectures and Table 7 shows the best hyperparameters.

Table 8 reports the diagnostic metrics defined in Section 2.6 with top performance highlighted in boldface. For all architectures, the proposed embedding (corresponding to Models 3, 6, 9 and 12) is the one that achieves better metrics with 100% of diagnostic sensitivity and specificity. Among them, the one with highest Slice Diagnostic Index is Model3, which is the one with the simplest architecture.

Table 6. Architecture of best models.

Model	Radiomic Embedding	Arch. Num.	Arch. Setting N_i, N_h	Architecture	# Param.
Model 1	None	1	24, 6	[(24,6),(6,6),(6,2)]	206
Model 2	Reproducibility	1	14, 8	[(14,8),(8,8),(8,2)]	210
Model 3	*t*-test	1	19, 9	[(19,9),(9,9),(9,2)]	290
Model 4	None	2	24, 9	[(24,9),(9,9),(9,4),(9,2)]	365
Model 5	Reproducibility	2	14, 9	[(14,9),(9,9),(9,4),(9,2)]	275
Model 6	*t*-test	2	19, 9	[(19,9),(9,9),(9,4),(9,2)]	320
Model 7	None	3	24, 8	[(24,8),(8,8),(8,8),(8,4),(4,2)]	382
Model 8	Reproducibility	3	14, 9	[(14,9),(9,9,(9,9),(9,4),(4,2)]	362
Model 9	*t*-test	3	19, 9	[(19,9),(9,9),(9,9),(9,4),(4,2)]	407
Model 10	None	4	24, 8	[(24,8),(8,7)(7,6)(6,5),(5,4),(4,2)]	305
Model 11	Reproducibility	4	14, 14	[(14,14),(14,13),(13,12),(12,11), (11,10),(10,2)]	745
Model 12	*t*-test	4	19, 8	[(19,8),(8,7),(7,6),(6,5),(5,4),(4,2)]	270

Table 7. Hyperparameters of best models.

Model	Radiomic Embedding	Weight Init.	Optimizer	Learning Rate	Epochs
Model 1	None	Kaiming	RMSProp	0.001	1500
Model 2	Reproducibility	Orthogonal	Adam	0.001	1500
Model 3	*t*-test	Xavier	SGD	0.001	1500
Model 4	None	Orthogonal	Adam	0.001	1000
Model 5	Reproducibility	Xavier	Adam	0.01	1000
Model 6	*t*-test	Xavier	Adam	0.001	1000
Model 7	None	Orthogonal	Adam	0.001	1000
Model 8	Reproducibility	Orthogonal	Adam	0.001	1000
Model 9	*t*-test	Kaiming	Adam	0.001	1000
Model 10	None	Kaiming	Adam	0.001	1000
Model 11	Reproducibility	Xavier	Adam	0.001	1000
Model 12	*t*-test	Orthogonal	Adam	0.001	1000

Table 8. Diagnosis scores of best models.

Model	$Sens_{Diagnosis}$	$Spec_{Diagnosis}$	F_{Nodule}
Model 1	100	100	0.856
Model 2	93.02	75	0.683
Model 3	**100**	**100**	**0.903**
Model 4	100	87.5	0.846
Model 5	97.67	37.5	0.595
Model 6	**100**	**100**	**0.839**
Model 7	100	100	0.804
Model 8	100	37.5	0.619
Model 9	**100**	**100**	**0.834**
Model 10	100	87.5	0.840
Model 11	100	37.5	0.617
Model 12	**100**	**100**	**0.831**

3.2. Model Verification

In order to statistically evaluate the reproducibility our system, we have conformed an independent set of test patients from our database and the LIDC-IDRI public database. From our database we used one benign and eight malignant nodules. Regarding the LIDC-IDRI database, since it was not collected to evaluate malignancy, scans are, in general, of a too low quality to assess malignancy. Following [19,20], we selected cases fulfilling the minimum acquisition requirements that allow radiological assessment of malignancy, which are slice thickness $<= 2.5$, resolution $<= 0.71$, except in case thickness is $<= 1.5$, that resolution can be $<= 0.86$ and only taking in consideration those nodules that have been diagnosed through a biopsy as benign or malign. After this filtering, 18 cases with diagnosis (five benign and 13 malign) were selected. In this way, the independent set of test patients is conformed by a total amount of 27 nodules with 6 benign and 21 malign.

We have compared Model3 with state of the art methods which include the three type of approaches: radiomics [5], machine learning [6] and deep CNN [8–12]. In order to compare to the results reported for each of them, we have computed the following metrics from true positive, TP, true negative, TN, false negative, FN, and false positive, FP diagnosis at nodule level:

$$\text{Sensitivity} = 100 \cdot \frac{TP}{TP + FN} \tag{12}$$

Sensitivity measures the percentage of correctly diagnosed malignant nodules:

$$\text{Specificity} = 100 \cdot \frac{TN}{TN + FP} \tag{13}$$

Specificity measures the percentage of benign nodules correctly identified:

$$\text{Accuracy} = 100 \cdot \frac{TP + TN}{Number\ of\ Nodules} \tag{14}$$

for *Number of Nodules* denoting the total amount of nodules. The accuracy measures the percentage of correctly diagnosed nodules (both malign and benign nodules) among the total number of nodules in the dataset:

$$\text{F1 Score} = 100 \cdot \frac{2 \cdot Prec \cdot Rec}{Prec + Rec} \tag{15}$$

for Rec, $Prec$ denoting, respectively, the precision and recall at diagnosis level:

$$Rec = 100 \cdot \frac{TP}{TP + FN} \qquad Prec = 100 \cdot \frac{TP}{TP + FP} \tag{16}$$

The metric (15) measures the trade-off between recall and precision, and in general, a higher F1-score means a better performance. We also computed the receiver operating characteristic (ROC) curves and the area under the curve (AUC).

Table 9 shows the metrics for state of the art methods grouped according to the type of approach (Radiomics, Machine Learning, Deep CNN) and our method. As in Table 8, best performances are in boldface. Table 9 reports the metrics obtained by Model3 in our test set together with the results reported by each state of art method using their datasets. We also report the number of parameters of each method as indicator of its complexity and computational and data cost for training. Our method outperforms in Accuracy, Sensitivity and F1 Score. In computer-aided diagnose, sensitivity is significant because correctly finding out patients with malignant nodules is crucial. Besides, the highest F1 Score implies that our method achieves the best trade-off between precision and recall. Our method has a splendid compromise between the performance of the system and the number of trainable parameters. A remarkable point compared to Deep CNN approaches, is that, our method needs strongly less samples to train the model, which is a must in medical imaging.

Table 9. Results of our method compared to the state of the art with malignant nodules as positive cases.

Approaches	Accuracy	Sensitivity	Specificity	F1 Score	AUC	Param. (M)
Radiomics						
Peikert et al. [5]	–	90.40	85.50	–	0.939	**<0.29**
Machine Learning						
Zhang et al. [6]	96.09	96.84	**95.34**	–	**0.979**	**<0.29**
Deep CNN						
Multicrop [8]	87.14	77.00	93.00	–	0.930	–
Nodule-level 2D [9]	87.30	88.50	86.00	87.23	0.937	–
Vanilla 3D [9]	87.40	89.40	85.20	87.25	0.947	–
DeepLung [10]	90.44	81.42	–	–	–	141.57
AE-DPN [11]	90.24	92.04	88.94	90.45	0.933	678.69
NASLung [12]	90.77	85.37	95.04	89.04	–	16.84
Hybrid						
model3 (Our)	**96.30**	**100**	83.33	**97.67**	0.940	0.29

4. Discussion

Intelligent artificial methods applied to medical imaging have to face two key drawbacks. The available small amount of labelled data and the obligation that methods must ensure good rates avoiding false positives. In order to overcome with these two main challenges, we have proposed an hybrid method that combines an embedded radiomic texture features to characterize nodules and an optimized feedforward network for nodule diagnosis. The nodule embedding step is based on selecting those radiomic features that significantly correlate to malignancy ensuring reproducibility with minimal training data. The fully connected network architecture and hyperparameters are optimized using own-defined metrics of the diagnostic power to ensure maximum clinical outcome.

Results demonstrate the power of the two main contributions. Table 8 results demonstrate the power of using *t*-test analysis for statistical significance and nodule embedding, as best results are achieved on those models (3, 6, 9, 12). Table 9 confirms that the whole hybrid strategy outperforms in Accuracy (96.30), Sensitivity (100) and F1 Score (97.67) the state of the art methods. Notice that in computer-aided diagnose, sensitivity is significant because correctly finding out patients with malignant nodules is crucial. A remarkable outcome is that our approach outperforms deep approaches with only requiring 290 parameters (in contrast to the thousands required by deep methods). This has a direct impact with the small training data needed.

Our work could be improved in the following aspects. Nodule embedding bases on a combination of simple *t*-tests as a first approach to a statistical selection of features, which disregards correlations across GLCM features. Future work will explore the use of other statistics (like regression models) taking into account multiple comparisons across features. In addition, it is planed to increase the database and introduce our own-defined metrics as the loss function of the fully connected network to guide training to ensure maximum clinical outcome.

Author Contributions: Conceptualization, A.R. and D.G.; Data curation, G.T., S.B., C.S., I.G. and D.G.; Formal analysis, D.G.; Funding acquisition, A.R. and D.G.; Investigation, G.T., S.B., C.S., I.G., A.R. and D.G.; Methodology, G.T., S.B., C.S., I.G., A.R. and D.G.; Project administration, A.R. and D.G.; Resources, S.B., I.G., A.R. and D.G.; Software, G.T., C.S. and D.G.; Supervision, C.S., A.R. and D.G.; Validation, G.T., C.S. and D.G.; Visualization, G.T., C.S. and D.G.; Writing—original draft, G.T., S.B., C.S., I.G. and D.G.; Writing—review & editing, G.T., C.S., A.R. and D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported by the Ministerio de Ciencia e Innovación (MCI), Agencia Estatal de Investigación (AEI) and Fondo Europeo de Desarrollo Regional (FEDER), RTI2018-095209-B-C21 (MCI/AEI/FEDER, UE), Generalitat de Catalunya, 2017-SGR-1624 and CERCA-Programme. Debora Gil is supported by Serra Hunter Fellow. Barcelona Respiratory Network (BRN), Acadèmia de Ciències Mèdiques de Catalunya i Balears, i Fundació Ramon Pla i Armengol.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki—Fortaleza/Brazil, 2013, and approved by the Institutional Review Board of Hospital Universitari Germans Trias i Pujol (protocol code PI-19-169 and date 6 September 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The public lung cancer datasets that support the findings of this study are openly available in the RADIOLUNG repository, http://iam.cvc.uab.es/portfolio/radiolung-database, and in the LIDC-IDRI repository, https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI.

Acknowledgments: Debora Gil would like to dedicate this work to her mother Esther Resina Enfedaque, and father, Manel Gil Doria.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area Under the Curve
CNN	Convolutional Neural Network
CT	Computed Tomography
GLCM	Gray Level Co-ocurrence Matrix
HU	Hounsfield Units
PN	Pulmonary Nodule
ROI	Region of Interest

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [CrossRef] [PubMed]
2. National Lung Screening Trial Research Team. The national lung screening trial: Overview and study design. *Radiology* **2011**, *258*, 243–253. [CrossRef] [PubMed]
3. Zhao, Y.R.; Xie, X.; de Koning, H.J.; Mali, W.P.; Vliegenthart, R.; Oudkerk, M. NELSON lung cancer screening study. *Cancer Imaging* **2011**, *11*, S79. [CrossRef] [PubMed]
4. Xie, H.; Yang, D.; Sun, N.; Chen, Z.; Zhang, Y. Automated pulmonary nodule detection in CT images using deep convolutional neural networks. *Pattern Recognit.* **2019**, *85*, 109–119. [CrossRef]
5. de Koning, H.; van der Aalst, C.; de Jong, P. Screening met een thoracale lage-dosis-CT-scan vermindert de sterfte na 10 jaar door longkanker bij mannelijke actieve of ex-rokers. *N. Engl. J. Med.* **2020**, *382*, 503–13. [CrossRef]
6. Zhang, F.; Song, Y.; Cai, W.; Lee, M.Z.; Zhou, Y.; Huang, H.; Shan, S.; Fulham, M.J.; Feng, D.D. Lung nodule classification with multilevel patch-based context analysis. *IEEE Trans. Biomed. Eng.* **2013**, *61*, 1155–1166. [CrossRef]
7. Lee, S.L.A.; Kouzani, A.Z.; Hu, E.J. Random forest based lung nodule classification aided by clustering. *Comput. Med. Imaging Graph.* **2010**, *34*, 535–542. [CrossRef]
8. Shen, W.; Zhou, M.; Yang, F.; Yu, D.; Dong, D.; Yang, C.; Zang, Y.; Tian, J. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognit.* **2017**, *61*, 663–673. [CrossRef]
9. Yan, X.; Pang, J.; Qi, H.; Zhu, Y.; Bai, C.; Geng, X.; Liu, M.; Terzopoulos, D.; Ding, X. Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strategies. In *Asian Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 91–101.
10. Zhu, W.; Liu, C.; Fan, W.; Xie, X. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 673–681.
11. Jiang, H.; Gao, F.; Xu, X.; Huang, F.; Zhu, S. Attentive and ensemble 3D dual path networks for pulmonary nodules classification. *Neurocomputing* **2020**, *398*, 422–430. [CrossRef]
12. Jiang, H.; Shen, F.; Gao, F.; Han, W. Learning efficient, explainable and discriminative representations for pulmonary nodules classification. *Pattern Recognit.* **2021**, *113*, 107825. [CrossRef]

13. Elsken, T.; Metzen, J.H.; Hutter, F. Neural architecture search: A survey. *J. Mach. Learn. Res.* **2019**, *20*, 1997–2017.
14. Armato, S.G., III; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* **2011**, *38*, 915–931. [PubMed]
15. Degenhardt, F.; Seifert, S.; Szymczak, S. Evaluation of variable selection methods for random forests and omics data sets. *Briefings Bioinform.* **2019**, *20*, 492–503. [CrossRef]
16. Cohen, J.P.; Luck, M.; Honari, S. Distribution matching losses can hallucinate features in medical image translation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 529–536.
17. Maicas, G.; Bradley, A.P.; Nascimento, J.C.; Reid, I.; Carneiro, G. Training medical image analysis systems like radiologists. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 546–554.
18. Bernatowicz, K.; Grussu, F.; Ligero, M.; Garcia, A.; Delgado, E.; Perez-Lopez, R. Robust imaging habitat computation using voxel-wise radiomics features. *Sci. Rep.* **2021**, *11*, 20133. [CrossRef]
19. Kim, Y.J.; Lee, H.J.; Kim, K.G.; Lee, S.H. The effect of CT scan parameters on the measurement of CT radiomic features: A lung nodule phantom study. *Comput. Math. Methods Med.* **2019**, *2019*, 8790694. [CrossRef]
20. Xu, Y.; Lu, L.; Sun, S.H.; Lian, W.; Yang, H.; Schwartz, L.H.; Yang, Z.H.; Zhao, B. Effect of CT image acquisition parameters on diagnostic performance of radiomics in predicting malignancy of pulmonary nodules of different sizes. *Eur. Radiol.* **2021**. [CrossRef] [PubMed]
21. van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [CrossRef]
22. Beig, N.; Khorrami, M.; Alilou, M.; Prasanna, P.; Braman, N.; Orooji, M.; Rakshit, S.; Bera, K.; Rajiah, P.; Ginsberg, J.; et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. *Radiology* **2019**, *290*, 783–792. [CrossRef]
23. Calheiros, J.L.L.; de Amorim, L.B.V.; de Lima, L.L.; de Lima Filho, A.F.; Júnior, J.R.F.; de Oliveira, M.C. The Effects of Perinodular Features on Solid Lung Nodule Classification. *J. Digit. Imaging* **2021**, *34*, 798–810. [CrossRef]
24. Gil, D.; Sanchez, C.; Borras, A.; Diez-Ferrer, M.; Rosell, A. Segmentation of distal airways using structural analysis. *PLoS ONE* **2019**, *14*, e0226006. [CrossRef]
25. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
26. Tixier, F.; Le Rest, C.C.; Hatt, M.; Albarghach, N.; Pradier, O.; Metges, J.P.; Corcos, L.; Visvikis, D. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J. Nucl. Med.* **2011**, *52*, 369–378. [CrossRef] [PubMed]
27. Huang, C.L.; Lian, M.J.; Wu, Y.H.; Chen, W.M.; Chiu, W.T. Identification of Human Ovarian Adenocarcinoma Cells with Cisplatin-Resistance by Feature Extraction of Gray Level Co-Occurrence Matrix Using Optical Images. *Diagnostics* **2020**, *10*, 389. [CrossRef] [PubMed]
28. Leijenaar, R.T.; Nalbantov, G.; Carvalho, S.; Van Elmpt, W.J.; Troost, E.G.; Boellaard, R.; Aerts, H.J.; Gillies, R.J.; Lambin, P. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci. Rep.* **2015**, *5*, 11075. [CrossRef] [PubMed]
29. Pomeroy, M.; Lu, H.; Pickhardt, P.J.; Liang, Z. Histogram-based adaptive gray level scaling for texture feature classification of colorectal polyps. In *Medical Imaging 2018: Computer-Aided Diagnosis*; International Society for Optics and Photonics: Houston, TX, USA, 2018; Volume 10575, p. 105752A.
30. Tan, J.; Gao, Y.; Liang, Z.; Cao, W.; Pomeroy, M.J.; Huo, Y.; Li, L.; Barish, M.A.; Abbasi, A.F.; Pickhardt, P.J. 3D-GLCM CNN: A 3-dimensional gray-level Co-occurrence matrix-based CNN model for polyp classification via CT colonography. *IEEE Trans. Med. Imaging* **2019**, *39*, 2013–2024. [CrossRef] [PubMed]
31. Ligero, M.; Torres, G.; Sanchez, C.; Diaz-Chito, K.; Perez, R.; Gil, D. Selection of radiomics features based on their reproducibility. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 403–408.

11.4. Ayudas al grupo de investigación

Por el proyecto Radiolung, nuestro grupo de investigación ha recibido las siguientes becas competitivas:

- Beca Llegat – Hospital Universitario Germans Trias i Pujol. Año 2022
- Beca Lung Ambition Alliance. Año 2021
- Beca MINECO (Ministerio de Economía, Industria y Competitividad, Gobierno de España). Nº: PID 2021-126776OB-C21. Año 2021
- Beca Barcelona Respiratory Network (BRN) – Fundació Ramon Pla Armengol. Año 2020
- Beca a proyecto clínico - Fundació Acadèmia Ciències Mèdiques de Catalunya i de Balears. Año 2020
- Beca “Talents” - Fundació La Pedrera-Hospital Universitario Germans Trias i Pujol. Año 2020

También enfocado en el diagnóstico del cáncer de pulmón, nuestro grupo ha recibido en abril de 2024 la beca FUCAP Dr. Alvar Agustí por el proyecto CatRisk.