UNIVERSIDADE FEDERAL DO PARANÁ

JOÃO MARIO CLEMENTIN DE ANDRADE

AUXÍLIO COMPUTACIONAL NO DIAGNÓSTICO POR IMAGEM

CURITIBA

2022

JOÃO MARIO CLEMENTIN DE ANDRADE

AUXÍLIO COMPUTACIONAL NO DIAGNÓSTICO POR IMAGEM

Dissertação apresentada ao Programa de Pós-Graduação em Medicina Interna e Ciências da Saúde, Setor de Ciências da Saúde, Universidade Federal do Paraná, como parte dos requisitos exigidos para obtenção do título de Mestre em Medicina interna e Ciências da Saúde.

Orientador: Prof. Dr. Dante Luiz Escuissato.
Coorientador: Prof. Dr. Lucas Ferrari de Oliveira
Professor do Departamento de Informática, UFPR.

CURITIBA

2022

A545 Andrade, João Mario Clementin de
Auxílio computacional no diagnóstico por imagem [recurso eletrônico] / João Mario Clementin de Andrade. – Curitiba, 2024.

Dissertação (mestrado) – Programa de Pós-Graduação em Medicina Interna. Setor de Ciências da Saúde. Universidade Federal do Paraná.
Orientador: Prof. Dr. Dante Luiz Escuissato.
Coorientador: Prof. Dr. Lucas Ferrari de Oliveira

1. Embolia pulmonar. 2. Tomografia. 3. Inteligência artifical. 4. Base de dados. I. Escuissato, Dante Luiz. II. Oliveira, Lucas Ferrari de. III. Programa de Pós-Graduação em Medicina Interna. Setor de Ciências da Saúde. Universidade Federal do Paraná. IV. Título.

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIAS DA SAÚDE, BIBLIOTECÁRIA: RAQUEL PINHEIRO COSTA JORDÃO CRB 9/991

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS DA SAÚDE
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MEDICINA INTERNA E CIÊNCIAS DA SAÚDE - 40001016012P1

ATA Nº016

ATA DE SESSÃO PÚBLICA DE DEFESA DE MESTRADO PARA A OBTENÇÃO DO GRAU DE MESTRE EM MEDICINA INTERNA E CIÊNCIAS DA SAÚDE

No dia vinte e oito de julho de dois mil e vinte e dois às 09:00 horas, na sala https://bit.ly/3Ou4ZGi, MICROSOFT TEAMS, foram instaladas as atividades pertinentes ao rito de defesa de dissertação do mestrando **JOÃO MARIO CLEMENTIN DE ANDRADE**, intitulada**: "AUXÍLIO COMPUTACIONAL NO DIAGNÓSTICO POR IMAGEM".**, sob orientação do Prof. Dr. DANTE LUIZ ESCUISSATO. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação MEDICINA INTERNA E CIÊNCIAS DA SAÚDE da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: DANTE LUIZ ESCUISSATO (UNIVERSIDADE FEDERAL DO PARANÁ), LUCAS FERRARI DE OLIVEIRA (PPG INFORMÁTICA - UFPR), BERNARDO CORRÊA DE ALMEIDA TEIXEIRA (UNIVERSIDADE FEDERAL DO PARANÁ). A presidência iniciou os ritos definidos pelo Colegiado do Programa e, após exarados os pareceres dos membros do comitê examinador e da respectiva contra argumentação, ocorreu a leitura do parecer final da banca examinadora, que decidiu pela APROVAÇÃO. Este resultado deverá ser homologado pelo Colegiado do programa, mediante o atendimento de todas as indicações e correções solicitadas pela banca dentro dos prazos regimentais definidos pelo programa. A outorga de título de mestre está condicionada ao atendimento de todos os requisitos e prazos determinados no regimento do Programa de Pós-Graduação. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, DANTE LUIZ ESCUISSATO, lavrei a presente ata, que vai assinada por mim e pelos demais membros da Comissão Examinadora.

Curitiba, 28 de Julho de 2022.

Assinatura Eletrônica
28/07/2022 21:36:40.0
DANTE LUIZ ESCUISSATO
Presidente da Banca Examinadora

Assinatura Eletrônica
28/07/2022 11:54:54.0
LUCAS FERRARI DE OLIVEIRA
Avaliador Externo (PPG INFORMÁTICA - UFPR)

Assinatura Eletrônica
28/07/2022 18:54:05.0
BERNARDO CORRÊA DE ALMEIDA TEIXEIRA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS DA SAÚDE
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MEDICINA INTERNA E CIÊNCIAS DA SAÚDE - 40001016012P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação MEDICINA INTERNA E CIÊNCIAS DA SAÚDE da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **JOÃO MARIO CLEMENTIN DE ANDRADE** intitulada: **"AUXÍLIO COMPUTACIONAL NO DIAGNÓSTICO POR IMAGEM".**, sob orientação do Prof. Dr. DANTE LUIZ ESCUISSATO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 28 de Julho de 2022.

Assinatura Eletrônica
28/07/2022 21:36:40.0
DANTE LUIZ ESCUISSATO
Presidente da Banca Examinadora

Assinatura Eletrônica
28/07/2022 11:54:54.0
LUCAS FERRARI DE OLIVEIRA
Avaliador Externo (PPG INFORMÁTICA - UFPR)

Assinatura Eletrônica
28/07/2022 18:54:05.0
BERNARDO CORRÊA DE ALMEIDA TEIXEIRA
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

RESUMO

Foi desenvolvida uma base de dados com exames de angiotomografia computadorizada de tórax de pacientes com tromboembolismo pulmonar (TEP) agudo, com imagens obtidas dos arquivos de dados digitais (PACS), do Hospital de Clínicas (HC-UFPR) e da clínica Diagnóstico Avançado por Imagem (DAPI). Foram coletados 40 exames, 20 deles realizados com tomógrafo Toshiba Aquilion 64 no HC-UFPR e 20 no tomógrafo GE Revolution na clínica DAPI. Os exames foram preparados para utilização no desenvolvimento de programas de informática voltados ao diagnóstico de TEP, sendo os trombos marcados ao nível do pixel e do corte. Paralelamente foi desenvolvido pelo departamento de informática da UFPR um programa baseado em algoritmos de “machine learning” voltado para a pesquisa de TEP por meio de angiotomografias computadorizadas (TC) de tórax. Este foi desenvolvido e testado utilizando a base de dados criada e outras duas outras bases públicas de exames com TEP.

Palavras-chave: Tromboembolismo pulmonar. Tomografia de tórax, Angiotomografia de tórax. Aprendizado de máquinas. Base de dados. Inteligência artificial. Aprendizado de máquinas.

ABSTRACT

We developed a database of chest computed angiotomography exams from patients with acute pulmonary thromboembolism (PE), with scans from the digital data files of “Hospital de Clínicas- Universidade Federal do Paraná (HC-UFPR) and “Clínica Diagnóstico Avançados por Imagem”(DAPI). We collected 40 exams, 20 performed with Toshiba Aquilion 64 CT scanner at HC-UFPR and 20 with GE Revolution CT scanner at DAPI. The exams were prepared for the development of computer programs aimed at the diagnosis of PE, with thrombi being segmented at pixel and slice level. At the same time, a computer program, based on “machine learning” algorithms, aimed at the research of PE in computed angiotomography of the chest was developed in partenership with the informatics department of UFPR. This program was developed and tested with the new database and two other public available PE databases.

Keywords: Pulmonary embolism. Chest computed tomography. Chest computed angiotomography. Machine learning, Computer aided diagnostics. Database. Artificial intelligence. Deep learning.

LISTA DE FIGURAS

LISTA DE QUADROS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

Art.	-	Artéria
angioTC	-	Angiotomografia computadorizada
BPM	-	Batimentos por minuto
CAD	-	*Computer- aided diagnosis*
DAPI	-	Diagnóstico Avançado por Imagem
DICOM	-	*Digital Imaging and Communications in Medicine*
EUA	-	Estados Unidos da América
HC-UFPR	-	Hospital de Clínicas da Universidade Federal do Paraná
IA	-	Inteligência Artificial
PACS	-	*Picture Archiving and Communication Systems*
ROI	-	*Region Of Interest*
RSNA	-	*Radiological Society of North America*
TAP	-	Tronco da Artéria Pulmonar
TC	-	Tomografia Computadorizada
TEP	-	Tromboembolismo Pulmonar
TVP	-	Tromboembolismo Venoso Profundo
VCI	-	Veia Cava Inferior
VD	-	Ventrículo Direito
VE	-	Ventrículo Esquerdo

SUMÁRIO

1 INTRODUÇÃO

O tromboembolismo pulmonar é uma condição potencialmente grave, de grande incidência e alta morbimortalidade. Dados dos EUA revelam uma incidência hospitalar de 200 a 300 mil casos de TEP por ano (WHITE, 2003). Estudos de autópsia demonstraram que a incidência é similar nos EUA e no Brasil (TERRA-FILHO, 2010). A variabilidade da sintomatologia apresentada pelos pacientes é responsável pela grande dificuldade no reconhecimento desta doença (DALEN, 2016). Mais de 90% das mortes devido ao TEP ocorrem em pacientes que não foram tratados devido à ausência de diagnóstico. Menos de 10% das mortes ocorrem entre os pacientes diagnosticados e tratados, indicando um potencial benefício na redução de mortalidade com a melhora da acurácia diagnóstica (DALEN, 2016).

O diagnóstico é fundamental para o rápido manejo e definição do tratamento do paciente (ROY; MEYER, 2006). Atualmente, a angiotomografia computadorizada (angioTC) de tórax é o método de escolha para pesquisa de TEP (KONSTANTINIDES; TORBICKI; AGNELLI, 2014; REMY-JARDIN; PISTOLESI; GOODMAN, 2007).

A interpretação deste exame exige alto grau de treinamento e diversos fatores podem levar a erros diagnósticos. Muitos destes são conhecidos e podem ser agrupados nas seguintes categorias: fatores relacionados ao paciente, fatores técnicos, fatores anatômicos e patologias (CONRAD et al., 2004). Entre os relacionados ao paciente incluem-se: artefatos de movimento respiratório, ruídos na imagem devido a variações morfológicas e composição corporal, presença de cateter na artéria pulmonar e artefato relacionado ao fluxo sanguíneo. Nos fatores técnicos incluem-se: janelamento inadequado, artefatos devido ao endurecimento de feixe, artefatos relacionados ao algoritmo de reconstrução pulmonar, artefatos de volume parcial e artefatos em escada nas reconstruções em outros planos. Em relação aos fatores anatômicos incluem-se: volumes parciais devido a linfonodos adjacentes a artérias, bifurcações vasculares podendo simular falhas de enchimento e falha na diferenciação de veias e artérias. Entre as patologias temos: rolhas de muco, edema perivascular, aumentos localizados da resistência vascular, trombose in situ, sarcoma primário da artéria pulmonar e embolia tumoral.

Outras importantes causas de diagnósticos errôneos são as falhas do médico na leitura do exame (MAIZLIN et al., 2007). Com os avanços tecnológicos que possibilitaram a aquisição de maior número de cortes e reconstruções multiplanares, o número de imagens a serem analisadas pelo radiologista aumentou de maneira significativa, prolongando o tempo de leitura dos exames. O radiologista deve sistematicamente identificar e distinguir cada artéria e seus ramos das demais estruturas anatômicas, corte a corte, da área central até a periferia pulmonar, na busca de falhas de enchimento que correspondam a trombos intraluminais. A complexidade das imagens e o grande número de cortes e vasos a serem analisados aumentam a chance de ocorrência de erros devido a fadiga do examinador.

Um programa computadorizado de auxílio diagnóstico (CAD) que ajude na redução dos erros diagnósticos é algo desejável e que pode levar a um melhor manejo dos pacientes e reduzir a mortalidade desta patologia. Diversas abordagens de CAD já foram propostas (QUADRO 1), contudo, ainda não se chegou a uma solução definitiva. Com os recentes avanços tecnológicos de hardware e software, surgiram diversas aplicações da inteligência artificial na radiologia (CHOY et al, 2018; REMY-JARDIN et al., 2020). Observa-se uma propensão ao uso deste tipo de aplicação (QUADRO 1), em especial utilizando redes neurais artificiais, no desenvolvimento de programas para o diagnóstico de TEP.

O desempenho deste tipo de programa é altamente dependente da base de dados utilizada para o treinamento. Esta deve ser uma amostra realista dos exames que o programa irá processar em uma aplicação clínica. A acurácia do programa tende a aumentar com a expansão da base de dados disponível para o treinamento (COLAK et al., 2021).

Para utilização no desenvolvimento de programas de aprendizado supervisionado de máquinas, a base de dados deve receber marcações que indiquem a presença do trombo. Dois tipos diferentes de marcação foram utilizados nas três bases de dados públicas disponíveis, a marcação ao nível dos pixels, na qual todos os pixels do trombo são demarcados, e a marcação ao nível do corte e do exame, nas quais as imagens em que os trombos são visualizados recebem uma marcação, sem a delimitação do exato local do trombo. A marcação ao nível do pixel é mais versátil e pode ser convertida na marcação ao nível do corte e do exame, contudo, o oposto não é verdadeiro.

Outra vantagem da anotação ao nível de pixel é permitir que o algoritmo preveja a região exata do TEP e verifique se esta foi a região utilizada pelo programa para gerar o diagnóstico. Isto é especialmente útil para avaliar programas destinados ao diagnóstico de TEP, uma vez que a alta taxa de falsos positivos é um desafio para diversos programas desenvolvidos.

Atualmente existem duas bases de dados públicas de exames de angioTC de tórax com TEP com anotações ao nível dos pixels. Uma delas contendo 91 exames com TEP, realizados em tomógrafos SIEMENS (GONZÁLEZ, 2019) e outra contendo 35 exames, dos quais 33 têm TEP, e dois não têm TEP (falsos positivos) realizados com tomógrafos da PHILIPS e Neusoft Medical System Co (MASOUDI et al., 2018). As informações publicadas destas bases não detalham o processo de seleção dos exames e os critérios de inclusão e exclusão.

Outra base de dados disponível publicamente disponível, foi deenvolvida pela sociedade norte americana de radiologia para uma competição de desenvolvimento de programas de inteligência artificial para a detecção de TEP (COLAK el al., 2021). Os exames foram obtidos de hospitais das universidades de Stanford (Palo Alto, California), Unity Health Toronto (Toronto, Canada), Universidade Federal de São Paulo (São Paulo, Brasil), Alfred Health (Melbourne, Australia), and Koç University (Istanbul, Turquia), totalizando 12.195 pacientes. A marcação dos exames foi realizada ao nível do corte e do exame, sem a marcação ao nível dos pixels.

Recentemente a revista Radiology, da Sociedade Norte América de Radiologia (RSNA), publicou um guia para produção de trabalhos com inteligência artificial (BLUEMKE et al., 2020). Este guia ressalta a importância de se utilizar exames realizados com aparelhos de diferentes fabricantes, devido a variabilidade das imagens.

A obtenção de bases de dados é um desafio para pesquisadores da área, uma vez que é um processo demorado e requer a participação de radiologistas com experiência no reconhecimento desta patologia.

Há uma escassez de bancos de dados públicos com uma amostra realista de exames que o programa processará em uma aplicação clínica real. Na prática médica, existe uma grande variabilidade entre das imagens de angioTC de tórax. Isso ocorre devido a fatores relacionados aos pacientes, como diferentes biótipos e comorbidades que podem obscurecer o TEP, e fatores técnicos, como aquisição de

imagem em tempo inadequado após a infusão do meio de contraste ou com fluxo de infusão inadequado.

Para suprir esta carência, desenvolvemos uma base de dados com uma amostra realista de casos de TEP anotados tanto ao nível de pixel quanto ao nível de imagem, tornando-o compatível com programas desenvolvidos utilizando as duas metodologias. Esta base de dados é composta por exames do arquivo digital (PACS) do HC-UFPR, realizados em um tomógrafo Toshiba Aquilion 64, e da clínica DAPI, realizados em um tomógrafo GE Revolution 128.

A base foi primeiramente utilizada no desenvolvimento de um programa para o diagnóstico de TEP em parceria com o departamento de informática da Universidade Federal do Paraná (Olescki et al., 2022), e em breve será disponibilizada para que outros grupos de pesquisa possam utilizá-la.

1.1 JUSTIFICATIVA

O tromboembolismo pulmonar é uma doença de alta incidência e grande morbimortalidade. O diagnóstico é fundamental para o adequado manejo dos pacientes e diversos fatores podem levar a erros na interpretação das imagens. Um programa que auxilie o diagnóstico pode impactar na mortalidade dos pacientes.

Existe uma grande tendência ao uso de algoritmos de inteligência artificial para esta tarefa. Estes algoritmos têm sua acurácia dependente da base de dados utilizada para o seu treinamento. A baixa disponibilidade de bases públicas com exames anotados para o desenvolvimento de programas é um desafio para pesquisadores da área.

2 OBJETIVOS

2.1 OBJETIVO PRINCIPAL

Desenvolver uma base de dados contendo uma amostra realista de angiotomografias computadorizadas de tórax para utilização no desenvolvimento de programas de aprendizado supervisionado de máquinas.

2.2 OBJETIVO SECUNDÁRIO

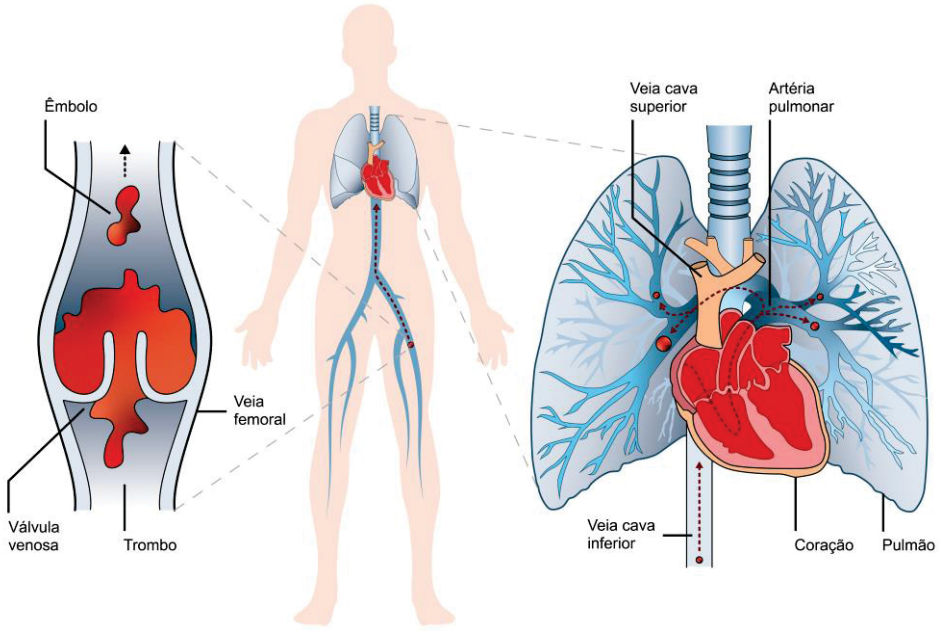
Contribuir com o departamento de informática da UFPR para o desenvolvimento de um programa para o diagnóstico de TEP em angioTC de tórax.

3. REVISÃO DA LITERATURA

3.1 TROMBOEMBOLISMO VENOSO

O tromboembolismo venoso é uma condição caracterizada pela formação inadequada de coágulos na circulação venosa. Este processo é regulado pela cascata da coagulação e pode ser desencadeado por alterações no fluxo sanguíneo (estase), alterações endoteliais e estados de hipercoagulabilidade (HUISMAN et al., 2018). O tromboembolismo venoso engloba a trombose venosa profunda, que ocorre na circulação sistêmica, e o tromboembolismo pulmonar. O tromboembolismo pulmonar ocorre quando um coágulo se desprende de uma veia sistêmica e se desloca para a circulação arterial pulmonar (BECKMAN et al., 2010) (FIGURA 1).

FIGURA 1- MECANISMO DO TEP



FONTE: Adaptado de Huisman (2018).

LEGENDA: Mecanismo do tromboembolismo pulmomar. Após se desprender da circulação venosa profunda o trombo se deslocada em direção ao coração com a circulação venosa e então é transportado para a circulação arterial pulmonar.

No tromboembolismo pulmonar agudo, a mortalidade ocorre devido a disfunção ventricular direita que pode estar ou não associada a arritmias cardíacas, resultando

em colapso vascular, fluxo arterial inadequado para os órgãos e óbito (HUISMAN et al., 2018).

3.1.1 EPIDEMIOLOGIA DO TROMBOEMBOLISMO PULMONAR

O tromboembolismo venoso, incluindo a trombose venosa profunda e o tromboembolismo pulmonar, é a terceira principal causa de mortalidade cardiovascular, sendo menos frequente apenas que infarto agudo do miocárdio e acidente vascular encefálico (RASKOB et al., 2014). A incidência populacional é estimada em 39-115 casos para 100 000 indivíduos por ano. Estima-se que ocorram até 300 000 mortes por ano nos Estados Unidos devido ao TEP (KONSTANTINIDES et al., 2020).

Estudos de autópsia demonstraram que a incidência do TEP é similar nos EUA e no Brasil (TERRA-FILHO, 2010). A variabilidade da sintomatologia apresentada pelos pacientes é responsável pela grande dificuldade no reconhecimento desta doença (DALEN, 2016). Mais de 90% das mortes devido ao TEP ocorrem em pacientes que não foram tratados devido à ausência de diagnóstico. Menos de 10% das mortes ocorrem entre os pacientes diagnosticados e tratados, indicando um potencial benefício na redução de mortalidade com a melhora da acurácia diagnóstica (DALEN, 2016).

3.1.2 DIAGNÓSTICO DO TROMBOEMBOLISMO PULMONAR

A abordagem de pacientes com suspeita de TEP é iniciada com a avaliação da probabilidade pré-teste de apresentar esta patologia. Tradicionalmente esta avaliação é feita com os escores de risco de Genova revisado (TABELA 1) ou de Wells (TABELA 2). Pacientes classificados como de risco baixo e intermediário devem realizar a dosagem do D-dímero, que é um produto da degradação da fibrina (proteína presente nos coágulos). O teste do D-dímero é altamente sensível para o TEP, sendo um resultado negativo útil para exclusão desta condição. Contudo, diversas outras condições podem elevar o D-dímero, o que reduz a sua especificidade e leva a resultados falso positivos. Pacientes com o D-dímero positivo, assim como os classificados inicialmente como de alto risco pelos escores

de Wells ou de Genebra modificado, devem realizar angioTC de tórax (HUISMAN et al., 2018).

TABELA 1 – ESCORE DE GENEBRA MODIFICADO

Itens	**Versão original**	**Versão Simplificada**
TEP ou TVP prévio	3	1
Frequência cardíaca: 75-94 b.p.m.	3	1
Frequência cardíaca: ≥ 95 b.p.m.	5	2
Cirurgia ou fratura no último mês	2	1
Hemoptise	2	1
Câncer ativo	2	1
Edema unilateral da perna	3	1
Dor unilateral da perna ou dor à palpação de trajeto venoso	4	1
Idade >65 anos	1	1
Probabilidade clinica (3 níveis)		
Baixa	0 - 3	0 – 1
Intermediária	4 – 10	2 – 4
Alta	≥11	≥5
Probabilidade clinica (2 níveis)		
TEP – improvável	0 – 5	0 – 2
TEP – provável	≥ 6	≥3

Fonte: O autor (2022).

LENGENDA: Critério do escore de Genebra modificado para a classificação da probabilidade clínica de TEP.

TABELA 2 – ESCORE DE WELLS

Itens	Pontos
TVP ou TEP prévios	1,5
Cirurgia recente ou Imobilização	1,5
Sinais clínicos de TVP	3
Diagnóstico alternativo menos provável que TEP	3
Hemoptise	1
Câncer	1
Probabilidade clinica	
Baixa	0-1
Intermediária	2-6
Alta	≥7

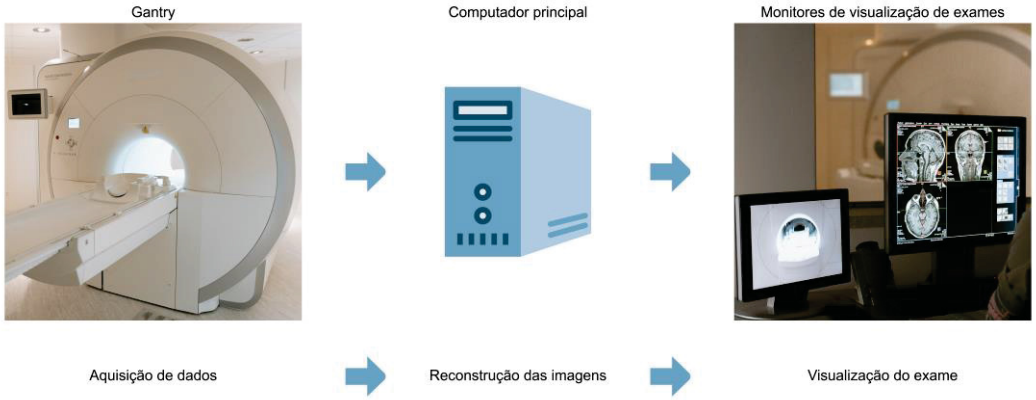
Fonte: O autor (2022).

LENGENDA: Critério do escore de Wells para a classificação da probabilidade clínica de TEP.

3.2 TOMOGRAFIA COMPUTADORIZADA

A tomografia computadorizada é um método de imagem seccional baseado nos diferentes coeficientes de atenuação do feixe de raios X apresentados pelos tecidos da região examinada (HOUNSFIELD, 1973). O tomógrafo é composto pelo *gantry* associado a uma maca móvel, responsável pela aquisição dos dados, um computador, no qual é feita a reconstrução das imagens, e o monitor para visualização das imagens (FIGURA 2).

FIGURA 2 – COMPONENTES DO TOMÓGRAFO



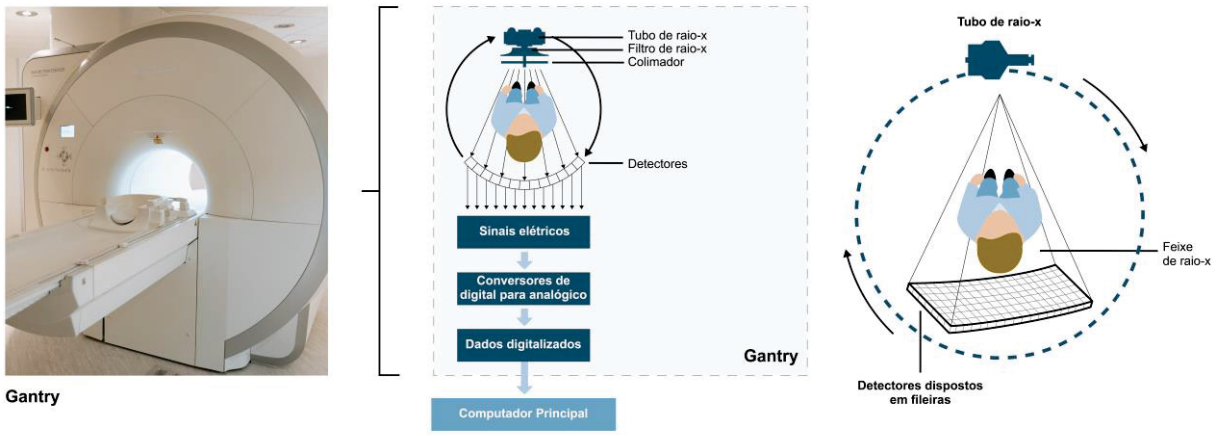
CC

FONTE: Adaptado de Seeram (2018).

LENGENDA: O tomógrafo é composto pelo Gantry, onde está localizado o tubo de raio-X e os detectores de raio-X que irão adquirir os dados referentes a região estudada, o computador principal, que processa estes dados resultando em arquivos de imagem que serão visualizados através de monitores.

O *gantry* possui um tubo de raio X que emite um feixe em leque. Este feixe atravessa o paciente e incide sobre os receptores, que são organizados em diferentes linhas, sendo cada linha denominada de canal (FIGURA 3). O *gantry* realiza movimentos rotacionais, enquanto a maca com o paciente se move pelo seu interior, fazendo a varredura helicoidal da área de interesse.

FIGURA 3 – TOMÓGRAFO MULTISLICE.

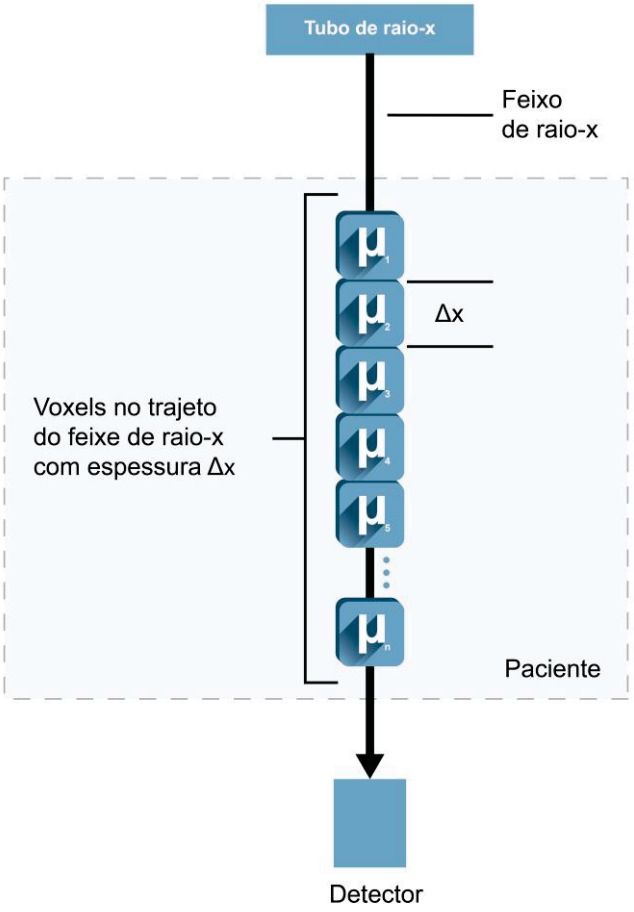


FONTE: Adaptado de Seeram (2018).

LENGENDA: O tomógrafo multislice é composto por um tubo de raio-X que emite raios-X que irão passar através da região examinada e serão captados por detectores, organizados em fileiras, que constituem os canais do tomógrafo.

A partir dos dados de atenuação do feixe ao atravessar a região examinada, medida pela diferença entre as intensidades dos feixes emitido e captado pelo receptor, medida em diferentes ângulos, os algoritmos de reconstrução calculam o coeficiente de atenuação (μ) de cada voxel da estrutura estudada (SEERAM, 2018) (FIGURA 4).

FIGURA 4 – FEIXE DE RAIO X



FONTE: Adaptado de Seeram (2018).

LENGENDA: Ilustração do trajeto de um feixe de raio-X na região em estudo. O feixe emitido atravessa a região de estudo, que pode ser dividida em pequenos voxels. Cada voxel é responsável por parte da atenuação do feixe emitido. Oposto ao tubo de raio-X esta posicionado o detector, que irá captar o feixe resultante após ser atenuado na passagem pela região de estudo.

Obtidos os coeficientes de atenuação de cada voxel do tecido estudado, estes são padronizados utilizando o coeficiente de atenuação da água, resultando no número tomográfico, também conhecido como “densidade tomográfica”, conforme a fórmula:

Número tomográfico = (μ tecido - μ água) × K / μ água

μ tecido: coeficiente de atenuação do tecido presente no voxel

μ água: coeficiente de atenuação da água

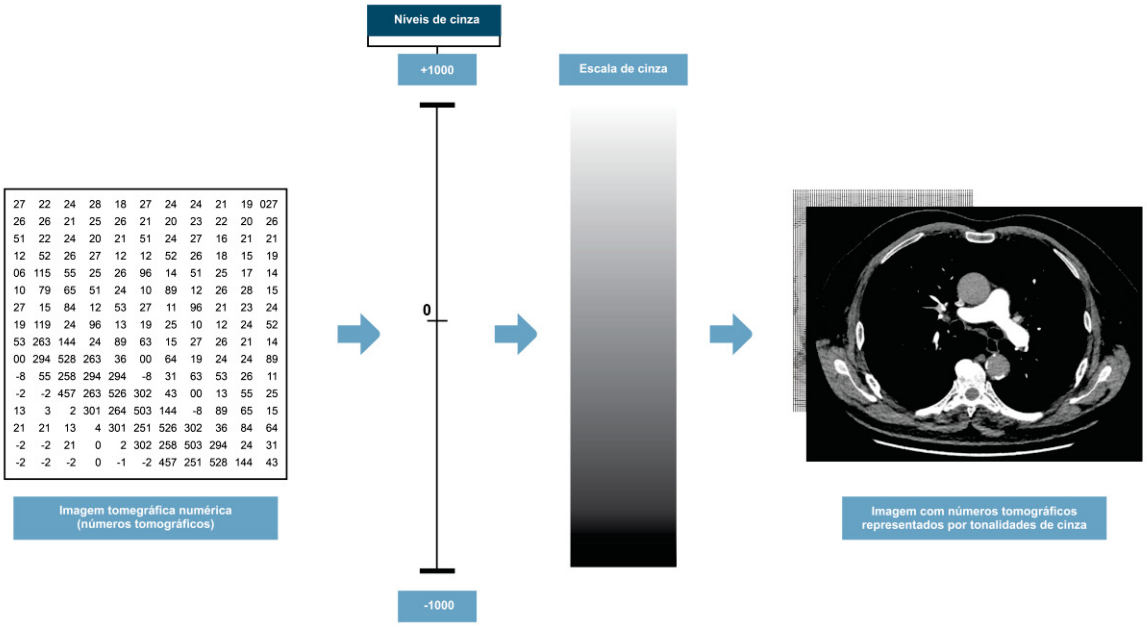
k: escalar variável de acordo com o fabricante do aparelho (geralmente equivalente a 1000).

A unidade de medida do número tomográfico é Hounsfield (H), em homenagem ao inventor da tomografia computadorizada.

3.2.1 JANELAMENTO DE IMAGENS TOMOGRÁFICAS

Para visualização das imagens tomográficas, a matriz numérica contendo os valores em unidades Hounsfield, que representam as densidades tomográficas, devem ser representadas em diferentes tons de cinza (FIGURA 5). A este processo denominamos janelamento. De acordo com as estruturas que serão analisadas na imagem são definidas a amplitude da faixa de números tomográficos a ser representada em tons de cinza (amplitude da janela) e o número central desta faixa numérica (centro da janela). Estruturas que apresentam densidades tomográficas acima do limite superior da janela escolhida são representadas em branco, enquanto as que apresentam densidade tomográfica abaixo do limite inferior da janela escolhida são representadas em preto.

FIGURA 5 – REPRESENTAÇÃO DOS NÚMEROS TOMOGRÁFICOS EM TONS DE CINZA.



FONTE: Adaptado de Seeram (2018).

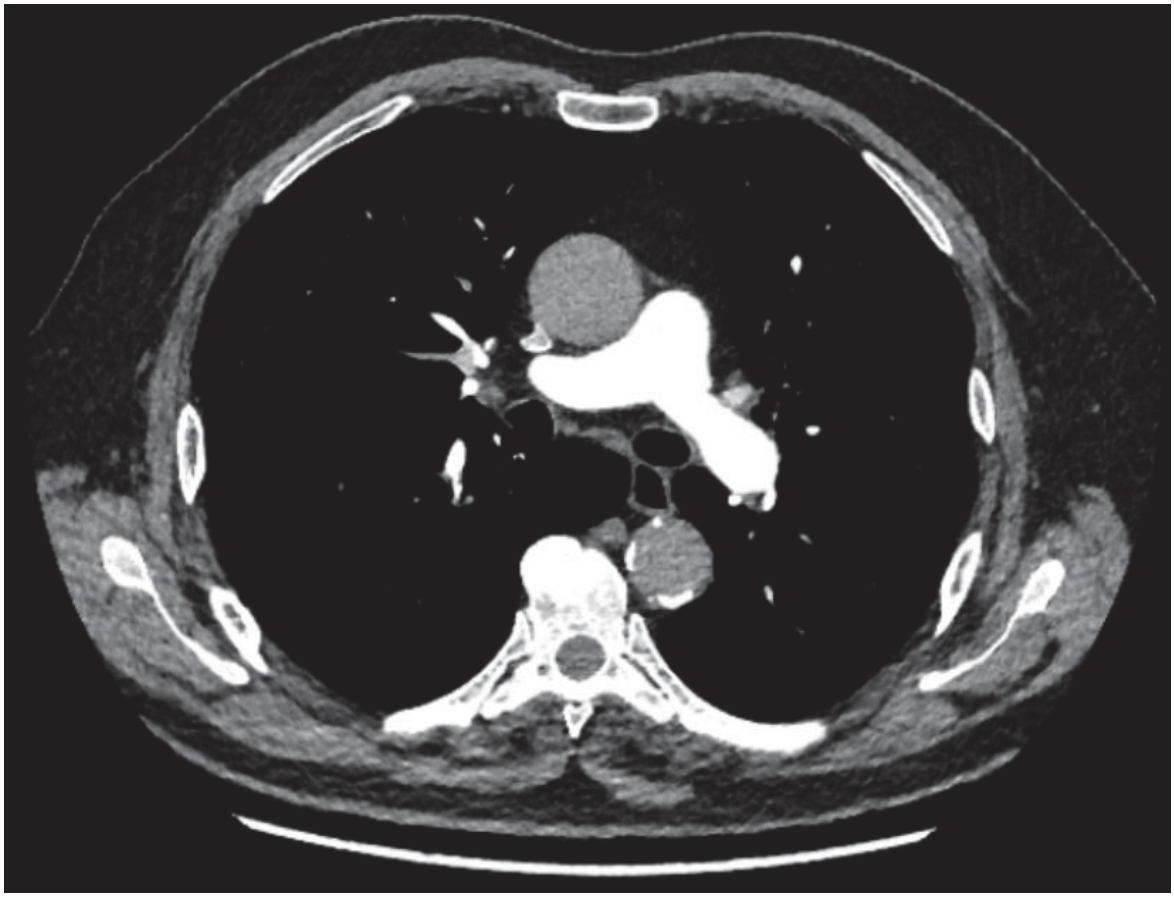
LENGENDA: Após o processamento dos dados obtidos, o arquivo de cada imagem é composto por uma matriz em que cada elemento representa o número tomográfico de cada pixel. Estes números são representados por diferentes tons de cinza nas imagens resultantes.

3.2.2 ANGIOTOMOGRAFIA ARTERIAL PULMONAR

A técnica da angioTC pulmonar tem como objetivo principal a obtenção de imagens com realce pelo meio de contraste da circulação arterial pulmonar (MORITZ et al., 2017) (FIGURA 6). Idealmente, busca-se a obtenção de imagens com as menores doses possíveis de contraste e de radiação. A fim de reduzir os artefatos relacionados a movimentos respiratórios, busca-se obter as imagens com a maior velocidade praticável sem prejudicar a qualidade das imagens. O tempo decorrido entre a injeção do meio de contraste e a opacificação da vascularização arterial pulmonar depende de características intrínsecas do paciente. Atualmente duas técnicas são utilizadas para a obtenção de imagens com o tempo adequado entre a injeção do contraste e a aquisição das imagens. A primeira consiste em monitorar a chegada do contraste através de um ROI posicionado no tronco da artéria pulmonar, sendo a aquisição das imagens iniciada quando o contraste atinge a circulação

pulmonar. A segunda é a injeção prévia de um pequeno bolus de contraste, sendo avaliado o tempo necessário para que o contraste atinja a circulação pulmonar (MORITZ et al., 2017).

FIGURA 6 - ANGIOTC DE TÓRAX REALIZADA COM TÉCNICA ADEQUADA



Fonte: O autor (2020).

LENGENDA: Imagem de angioTC de tórax obtida ao nível da bifurcação da artéria pulmonar com opacificação adequada da circulação arterial pulmonar.

3.3 PROGRAMAS COMPUTADORIZADOS DE AUXÍLIO DIAGNÓSTICO DE TEP

Um programa computadorizado de auxílio diagnóstico (CAD) que ajude na redução dos erros diagnósticos é algo desejável e que pode levar a um melhor manejo dos pacientes. Diversas abordagens de CAD já foram propostas (QUADRO 1), contudo, ainda não se chegou a uma solução definitiva com sensibilidade e especificidade adequadas. Os avanços na informática, que permitiram a criação de programas utilizando técnicas de inteligência artificial, em especial utilizando redes neurais convolucionais com capacidade de análise de um grande volume de informações, levaram ao surgimento de uma nova geração de CAD.

Na revisão da literatura realizada, observa-se forte tendência a utilização de técnicas de inteligência artificial (IA). Foram encontradas 42 publicações de trabalhos com programas para diagnóstico de TEP publicados entre 2002 e 2022. Entre estas publicações, 20 trazem informações que permitem concluir que foram utilizados algoritmos com algum tipo IA. Desde 2015, foram encontradas 17 publicações, todas com programas que utilizam IA.

QUADRO 1- PUBLICAÇÕES COM CAD PARA DIAGNÓSTICO DE TEP

Autores		**Ano de publicação**	**Utiliza IA**
MASATUNI et al.		2002	
PICHON et al.		2004	
ZHOU et al.		2005	
MAIZLIN et al.		2007	
BUHMANN et al.		2007	
SCHOEPF et al.		2007	
LIANG, J., BI, J.		2007	
SEBBE		2007	
ENGELKE et al.		2008	
BOUMA		2008	
DAS et al.		2008	
WALSHAM et al.		2008	
BOUMA et al.		2009	
DINESH et al.		2009	Sim
ZHOU et al.		2009	
WITTENBERG et al.		2010	
DEWAILLY et al.		2010	
LEE et al.		2011	
PARK et al.		2011	Sim
BLACKMON et al.		2011	Sim
WITTENBERG et al.		2011	
WITTENBERG et al.		2012	
KLIGERMAN et al.		2014	
ÖZKAN et al.		2014	
LAHIJI et al.		2014	
TAJBAKHSH et al.		2015	Sim
OZKAN et al.		2017	Sim
TAJBAKSH et al.		2019	Sim
RAJAN et al.		2019	Sim
YANG et al.		2019	Sim
CANO-ESPINOSA et al.		2020	Sim
HUANG et al.		2020	Sim
YU et al.		2020	Sim
HUANG et al.		2020	Sim
LIU et al.		2020	Sim

WEIKERT et al.		2020	Sim
VAINIO et al.		2021	Sim
LI et al.		2021	Sim
LONG, K. et al.		2021	Sim
HUHTANEN et al.		2022	Sim
YADLAPALLI et al.		2022	Sim
OLESCKI et al.		2022	Sim

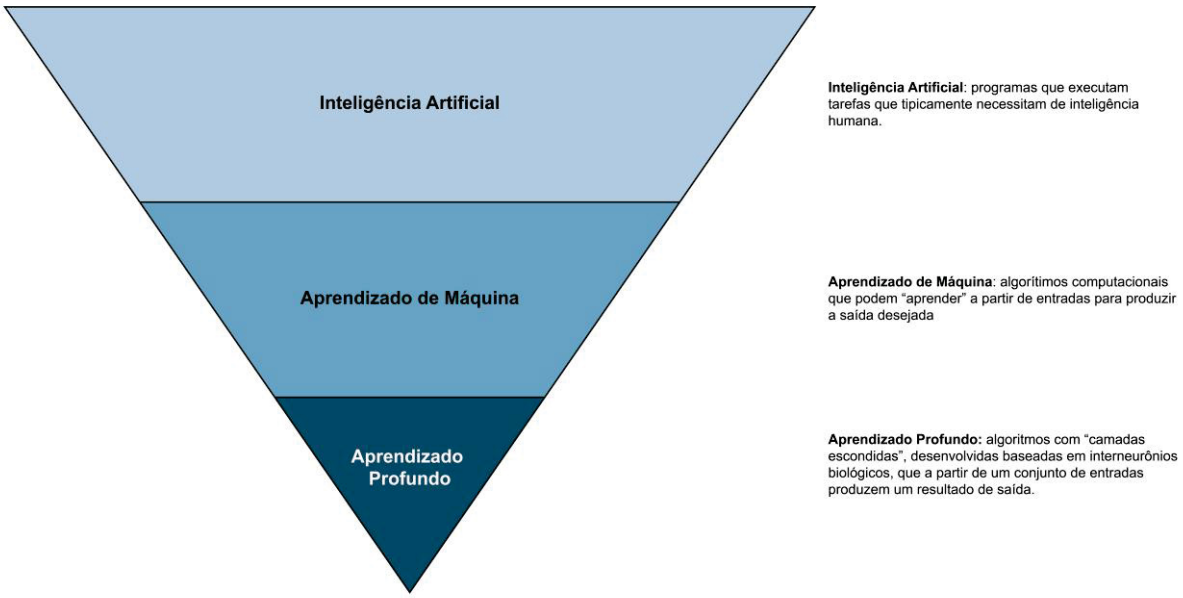
FONTE: O autor (2022).

LENGENDA: Publicações de trabalhos com programas computacionais para o diagnóstico de TEP com o respectivo ano de publicação sendo sinalizado quais trabalhos apresentam algoritmos que utilizam IA.

3.4 INTELIGÊNCIA ARTIFICIAL

Inteligência artificial é uma subárea da ciência da computação que se dedica a criar programas para realizar tarefas que habitualmente requerem inteligência humana (CHARTRAND et al., 2017). O aprendizado de máquinas, por sua vez, é uma subárea da inteligência artificial que dá aos computadores a habilidade de aprender sem que sejam explicitamente programados (YAO et al., 2020) (FIGURA 7). Este resultado é atingido por métodos que permitem a descoberta de padrões em um conjunto de dados e a utilização deste padrão descoberto para prever novos dados ou tomar decisões sob incerteza (MURPHY, 2012). Diferentemente da programação tradicional, em que o programador define explicitamente as regras que serão executadas para chegar ao resultado pretendido, no aprendizado de máquinas as regras são abstraídas a partir dos dados apresentados.

FIGURA 7 – CLASSIFICAÇÕES DE DIFERENTES ARQUITETURAS EM INTELIGÊNCIA ARTIFICIAL



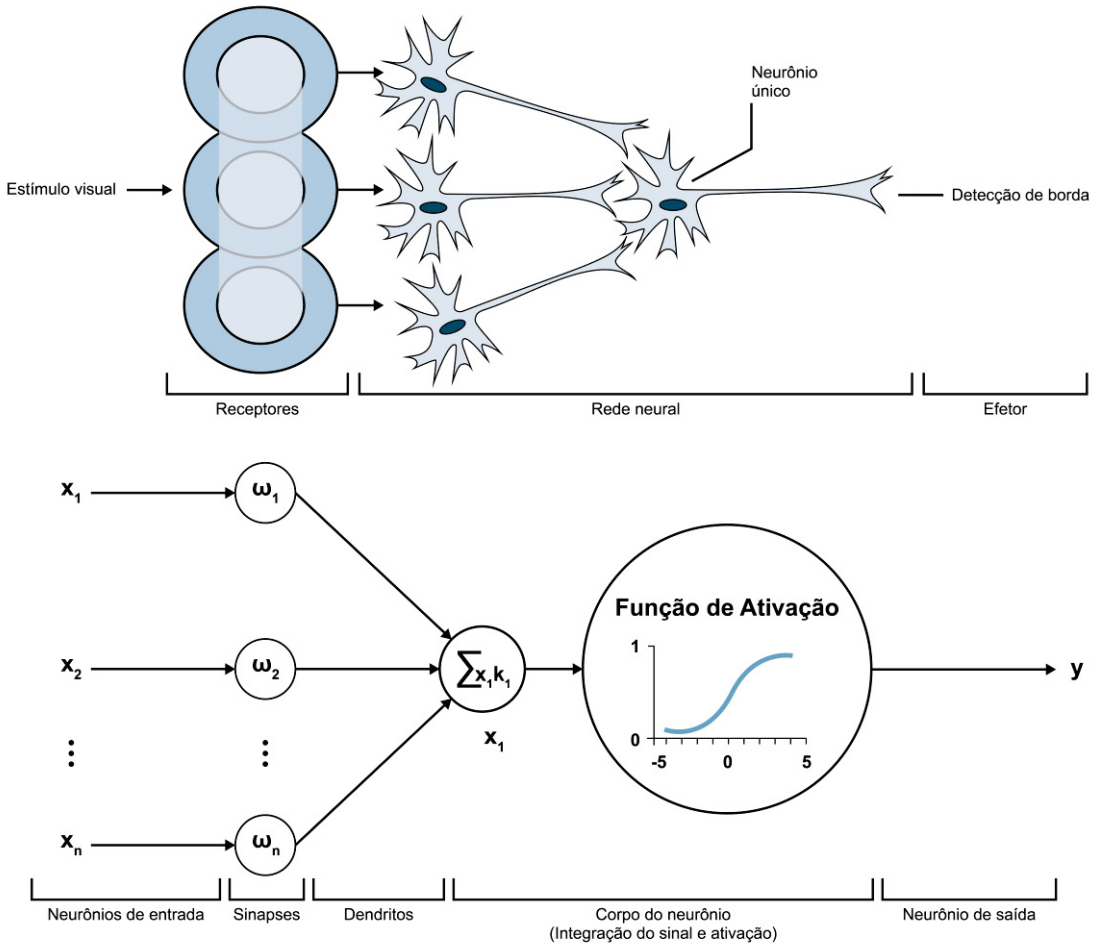
FONTE: Adaptado de YAO et al (2020).

LENGENDA: O aprendizado profundo é uma subárea do aprendizado de máquina, que por sua vez é uma subárea da inteligência artificial.

Podemos distinguir duas classes dentro do aprendizado de máquinas, o aprendizado supervisionado e o aprendizado não supervisionado (LEE et al., 2017). No aprendizado supervisionado, os dados são marcados e busca-se que o programa extraia padrões destes dados que possibilitem previsões quando fornecidos novos dados. Já no aprendizado não supervisionado, os dados são apresentados sem qualquer marcação e o programa busca extrair algum padrão intrínseco aos dados.

Dentre os diferentes algoritmos utilizados no aprendizado de máquinas, as redes neurais profundas têm ganho destaque. No campo da visão computacional as redes neurais convolucionais tem obtido resultados promissores. As redes neurais artificiais imitam o funcionamento do cérebro humano (YAO et al, 2020) (FIGURA 8). Elas são formadas por camadas neuronais conectadas (FIGURA 9), sendo que as conexões apresentam diferentes pesos. Os pesos são ajustados na etapa de treinamento da rede. Ao longo do treinamento da rede neural os pesos são reajustados, buscando reduzir a diferença entre o resultado obtido e o resultado esperado. A necessidade de ajuste dos pesos das interconexões da rede neural torna seu desempenho dependente do tamanho e da qualidade da base de dados disponível para o treinamento.

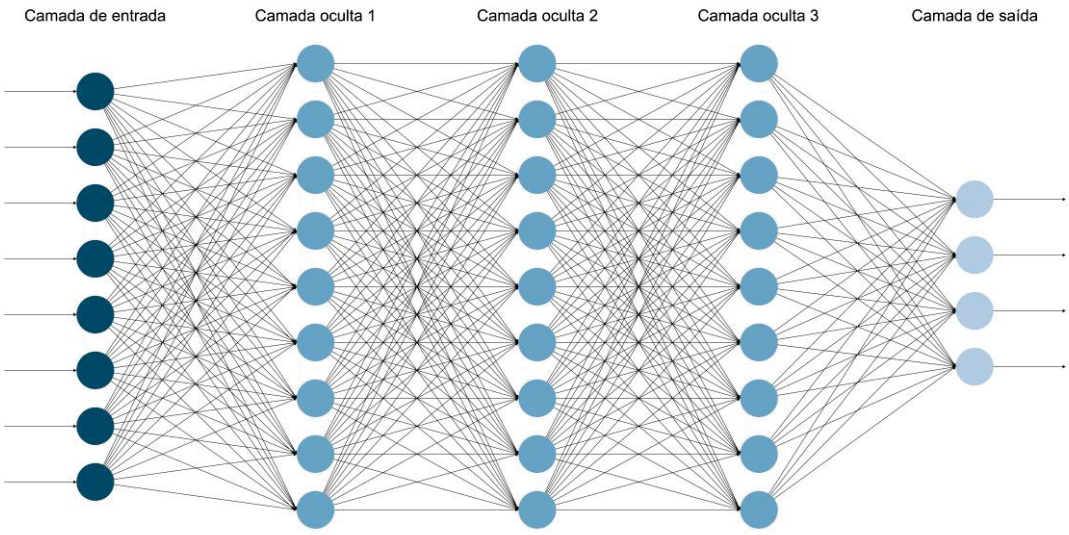
FIGURA 8 – ANALOGIA ENTRE NEURÔNIOS E NEURÔNIOS ARTIFICAIS



FONTE: Adaptado de CHARTRAND et al (2017).

LENGENDA: Analogia entre neurônio e suas sinapses e um neurônio artificial e sua rede. Assim como o neurônio recebe estímulo por seus dendritos e produz uma resposta que será enviada por seu axônio para um próximo neurônio, cada neurônio artificial é composto por um modelo de classificação simples, com uma saída a partir da soma de informações com diferentes pesos e de uma função de ativação. Múltiplas destas unidades associadas constituem uma rede neural artificial. Os pesos da rede são treinados através dos algoritmos de aprendizagem em que sinais de entrada e resultado esperado são apresentados de forma pareada, de forma análoga ao cérebro humano, que aprende a partir de estímulos externos.

FIGURA 9 – REDE NEURAL COM 3 CAMADAS PROFUNDAS



FONTE: Adaptado LEE et al (2018).

LENGENDA: Exemplo de rede neural profunda com três camadas ocultas completamente conectadas. Na camada de entrada são inseridos os dados a serem processados. Na camada de saída obtêm-se o resultado do processamento das informações. Neste exemplo temos três camadas profundas completamente conectadas com as demais camadas, ou seja, cada neurônio artificial recebe informações e envia informações para todos os neurônios artificiais da camada anterior e seguinte, respectivamente.

3.5 MARCAÇÃO DE IMAGENS

Existem diferentes formas de realizar a marcação das bases de dados para a utilização em programas de aprendizado supervisionado de máquinas. A marcação é utilizada para ressaltar as características a serem reconhecidas pelo programa (BOESCHIMAGE, 2022) que deverá reconhecê-las em imagens não marcadas.

A escolha da forma de marcação depende de diversos fatores, como o tipo de dado a ser marcado, a complexidade da tarefa que o programa irá executar e o resultado esperado do programa que utilizará a base de dados no seu desenvolvimento.

Quanto ao grau de automatização, a marcação da base de dados pode ser realizada de forma manual, semiautomatizada ou automatizada (ZHANG, 2006).

O tipo de anotação utilizado depende, entre outros fatores, da tarefa a ser executada pelo programa. Programas que realizam a classificação de imagens, que consiste em verificar a presença de determinado elemento na imagem, demandam a marcação ao nível da imagem quanto a presença deste elemento. No caso específico do TEP, uma base nas quais os cortes contendo trombo são marcados, pode ser utilizada para o desenvolvimento de um programa que classifique novas imagens quanto a presença desta patologia, sem identificar o local onde o trombo é visualizado (BOESCHIMAGE, 2022).

A segmentação semântica, também conhecida como segmentação ao nível do pixel, consiste na atribuição de uma classe para cada pixel da imagem (POKHREL, 2020). É mais complexa e permite definir a localização, o volume e o número de objetos de determinada classe nas imagens.

No caso de imagens de angioTC de tórax, exitem duas bases de dados públicas disponíveis marcadas utilizando esta técnica. Os pixels que representam o trombo são marcados como pertencentes a classe “trombo presente” e os demais pixels como pertencentes a classe “trombo ausente”, representados respectivamente pelos números zero e 1 em uma matriz de dimensões semelhantes ao número de pixels na imagem. As bases de dados com este tipo de marcação permitem desenvolver programas que tenham como objetivo delimitar o trombo em exames novos.

3.6 BASES DE DADOS PÚBLICAS DE TROMBOEMBOLISMO PULMONAR ANOTADAS AO NÍVEL DO PIXEL

Os exames tabulados são necessários para treinar e testar programas computacionais de auxílio diagnóstico. Especialmente no caso dos programas com algoritmos de aprendizado de máquina, a acurácia diagnóstica tende a subir com o aumento do volume de imagens tabuladas disponíveis para o treinamento (PAIVA; PREVEDELLO, 2017).

Por ser um processo lento e dependente de médicos radiologistas com experiência no reconhecimento desta patologia, a obtenção de um grande banco de dados é um dos obstáculos encontrados por pesquisadores da área. Até o momento, existem apenas duas bases de dados públicas de tomografias de tórax com

tromboembolia pulmonar tabuladas ao nível do pixel, uma com 91 (GONZÁLEZ, G. et al, 2020) e outra com 35 exames (MASOUDI, M. et al., 2018).

Ambas as bases consistem em dois arquivos por paciente, um contendo o exame no formato digital (DICOM) e outro com o gabarito no qual os trombos estão marcados. Quando abertos em conjunto o gabarito se sobrepõe ao trombo permitindo a identificação destes.

A maior base disponível é composta por 91 exames de pacientes com tromboembolismo pulmonar realizados em seis hospitais de Madrid (GONZÁLEZ. et al, 2020). Todos os exames foram feitos em tomógrafos SIEMENS, modelo Somaton Sensation 40, com o tamanho dos pixels variando de 0,58 mm a 0,85 mm e espessura dos cortes entre 0,75 mm e 1,5 mm. Os exames foram tabulados de forma semiautomática pelo método STAPLE (GROOT; BIRLUTIU; HESKES, 2011). As informações publicadas desta base não divulgam os critérios de inclusão e exclusão, apenas salientam que exames com outras patologias pulmonares não foram excluídos.

A segunda base disponível é composta por 35 exames, sendo 33 com tromboembolismo pulmonar, também marcados com um método semiautomatizado (MASOUDI et al., 2018). Os exames foram obtidos com o tomógrafo da PHILIPS, utilizando 120 kVp, colimação de 0.75 × 16 mm, com velocidade de rotação do *gantry* 0.75 segundos e *pitch* de 1,2. Assim como na primeira base de dados, o processo de seleção dos exames e os critérios de inclusão e exclusão não foram divulgados. Um destaque feito pelos autores desta base é a seleção de exames com êmbolos em vasos pulmonares periféricos, que é um ponto crítico para detecção dos programas de auxílio diagnóstico de TEP.

4 METODOLOGIA

4.1 TOMÓGRAFOS

As imagens do HC-UFPR foram obtidas com o tomógrafo da Toshiba Aquilion de 64 canais (FIGURA 10). Foi utilizado protocolo com 120 KVp, espessura de corte de 1,0 mm, *gap* de 8 mm, tempo de rotação do tubo de 0,5 s e *pitch de 1,485* e modulação da dose.

Os exames da clínica DAPI foram realizados no tomógrafo da GE Revolution 512 cortes (FIGURA 10). Foi utilizado protocolo com 120 KVp, espessura de corte de 0,625 mm, *gap* de 0,625 mm, tempo de rotação do tubo de 0,5 segundos, *pitch* de 0,9 e modulação de dose.

FIGURA 10 – TOMÓGRAFOS

TOSHIBA AQUILION 64 GE REVOLUTION

FONTE: Adaptado de LEITÃO (2020).

LENGENDA: Fotos dos dois equipamentos utilizados para realização dos exames. Do lado esquerdo o Toshiba Aquilion 64 do HC-UFPR. Do lado esquerdo o GE Revolution da clínica DAPI.

4.2 PLANO PARA RECRUTAMENTO DOS PARTICIPANTES DA PESQUISA

Os exames foram obtidos por busca retrospectiva nos arquivos digitais (PACS) do HC-UFPR e da clínica DAPI. Coletamos 20 exames com TEP agudo realizados no HC-UFPR no período de 05/11/2018 a 25/02/2019 e 20 exames com TEP agudo

realizados na DAPI no período 09/11/2018 a 20/11/2019. O período de coleta dos exames foi definido de forma arbitrária para que se obtivessem 20 exames realizados com cada aparelho, sendo respeitado rigorosamente os critérios de inclusão e exclusão, sem qualquer critério omisso de seleção dos exames. Todos os exames incluídos foram revisados por um residente de radiologia do terceiro ano e por um radiologista com 32 anos de experiência em radiologia torácica.

4.2.1 CRITÉRIOS DE INCLUSÃO

Exames de angioTC de tórax de pacientes com TEP agudo.

Disponibilidade do exame no arquivo digital do HC – UFPR e do DAPI.

4.2.2 CRITÉRIOS DE EXCLUSÃO

Exames com artefatos que prejudiquem a análise das imagens a ponto de impedir a sua interpretação.

Exames que tenham tido parte ou todo o arquivo corrompido.

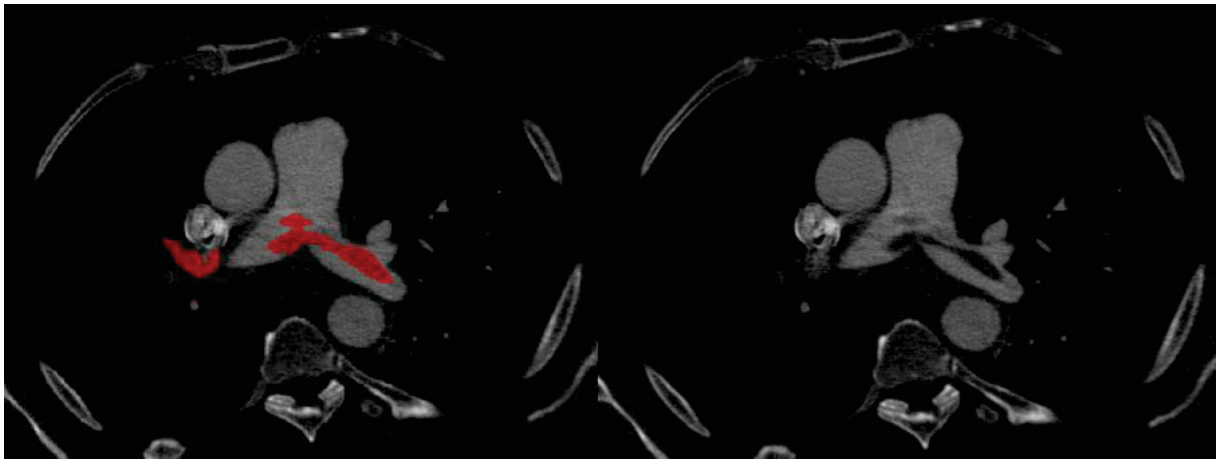
Exames repetidos do mesmo paciente. Neste caso apenas o primeiro exame foi incluído.

4.3 MARCAÇÃO DE EXAMES

Após a revisão e confirmação da presença de TEP agudo por um radiologista torácico com 32 anos de experiência, os exames foram marcados de forma manual com o programa ITK-SNAP (YUSHKEVICH, 2006), sendo obtido um segundo arquivo para cada exame contendo uma máscara na qual os pixels correspondentes ao trombo estão marcados (FIGURA 11). Estas marcações foram realizadas por médicos residentes do terceiro ano de radiologia e revisadas por um radiologista habilitado com 1 ano de experiência.

Baseado na máscara obtida, a segmentação ao nível da imagem foi gerada tabulando as imagens que continham o trombo.

FIGURA 11 – EXAME COM E SEM A MARCAÇÃO DO TROMBO



FONTE: O autor (2020).

LENGENDA: Ilustração contendo duas imagens tomográficas. Do lado esquerdo o exame sem a marcação. Do lado direito o trombo marcado em vermelho.

Adicionalmente, todos os exames foram classificados quanto a localização dos trombos na vascularização pulmonar, sendo considerado a localização do trombo mais central. Também foram verificados os sinais tomográficos possivelmente relacionados a sobrecarga de câmaras cardíacas direitas e hipertensão pulmonar. Foram mensurados o diâmetro do tronco da artéria pulmonar (TAP), os diâmetros transversos dos ventrículos direito (VD) e esquerdo (VE), medidos entre as superfícies endocárdicas nos maiores eixos perpendiculares ao eixo longitudinal, sendo calculado a relação entre VD e VE. Os exames também foram classificados quanto a presença de refluxo de contraste para a veia cava inferior (VCI), sendo considerado como presente quando visualizado contraste na VCI ao nível das veias hepáticas, quanto a retificação e inversão da concavidade do septo interventricular.

5 RESULTADOS

5.1 BASE DE DADOS DESENVOLVIDA

Artigo publicado na revista “nature scientific data” (novo qualis A1).

scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Pixel-level annotated dataset of computed tomography angiography images of acute pulmonary embolism

João Mario Clementin de Andrade[1]✉, Gabriel Olescki[2], Dante Luiz Escuissato[1], Lucas Ferrari Oliveira[2], Ana Carolina Nicolleti Basso[1] & Gabriel Lucca Salvador[1]

Pulmonary embolism has a high incidence and mortality, especially if undiagnosed. The examination of choice for diagnosing the disease is computed tomography pulmonary angiography. As many factors can lead to misinterpretations and diagnostic errors, different groups are utilizing deep learning methods to help improve this process. The diagnostic accuracy of these methods tends to increase by augmenting the training dataset. Deep learning methods can potentially benefit from the use of images acquired with devices from different vendors. To the best of our knowledge, we have developed the first public dataset annotated at the pixel and image levels and the first pixel-level annotated dataset to contain examinations performed with equipment from Toshiba and GE. This dataset includes 40 examinations, half performed with each piece of equipment, representing samples from two medical services. We also included measurements related to the cardiac and circulatory consequences of pulmonary embolism. We encourage the use of this dataset to develop, evaluate and compare the performance of new AI algorithms designed to diagnose PE.

Background & Summary

Pulmonary embolism (PE) has a high incidence and mortality. It occurs when a blood clot, most commonly from the deep venous system, moves into the pulmonary arterial circulation[1]. Up to 300,000 deaths per year are estimated to occur in the United States due to PE[2]. Less than 10% of deaths occur among diagnosed and treated patients, indicating a potential reduction in mortality by improving the diagnostic accuracy for the disease[3].

Computed tomography pulmonary angiography (CTPA) is the examination of choice for evaluating patients with PE[4,5]. After intravenous infusion of iodinated contrast medium, CT is performed when there is optimal opacification of the pulmonary arterial circulation, and the thrombus is identified as an intraluminal filling defect.

CTPA image interpretation is a complex task: radiologists must carefully search for contrast filling defects in the entire pulmonary arterial vasculature across a large number of images. Technical problems, patient-related factors, anatomical issues and the presence of other pathologies[6] can lead to misdiagnosis.

Computer-aided diagnosis (CAD) programs aimed at reducing these errors can reduce mortality. Several approaches have already been proposed[7–48]; however, a definitive solution has not yet been reached. More recently, there has been increased interest in the creation of artificial intelligence (AI) techniques, especially using artificial neural networks (ANNs), for addressing this problem.

The diagnostic performance of these techniques is highly dependent on the dataset used for training them, as it must contain examinations as diverse as those in real applications. The diagnostic accuracy of these methods tends to increase by augmenting the training dataset[49].

Obtaining reliable datasets is a considerable obstacle encountered by researchers, as it is time-consuming, requires radiologists with experience to recognize PE and depends on medical center cooperation. Furthermore, to be suitable for supervised learning applications, the dataset must be annotated. Two different annotation approaches have been used in the three public datasets available: a pixel-level annotation, in which all pixels of

[1]Department of Radiology and Image Diagnosis, Hospital de Clínicas, Federal University of Paraná, Curitiba, Brazil. [2]Department of Informatics, Federal University of Paraná, Curitiba, Brazil. ✉e-mail: joaoclementin@gmail.com

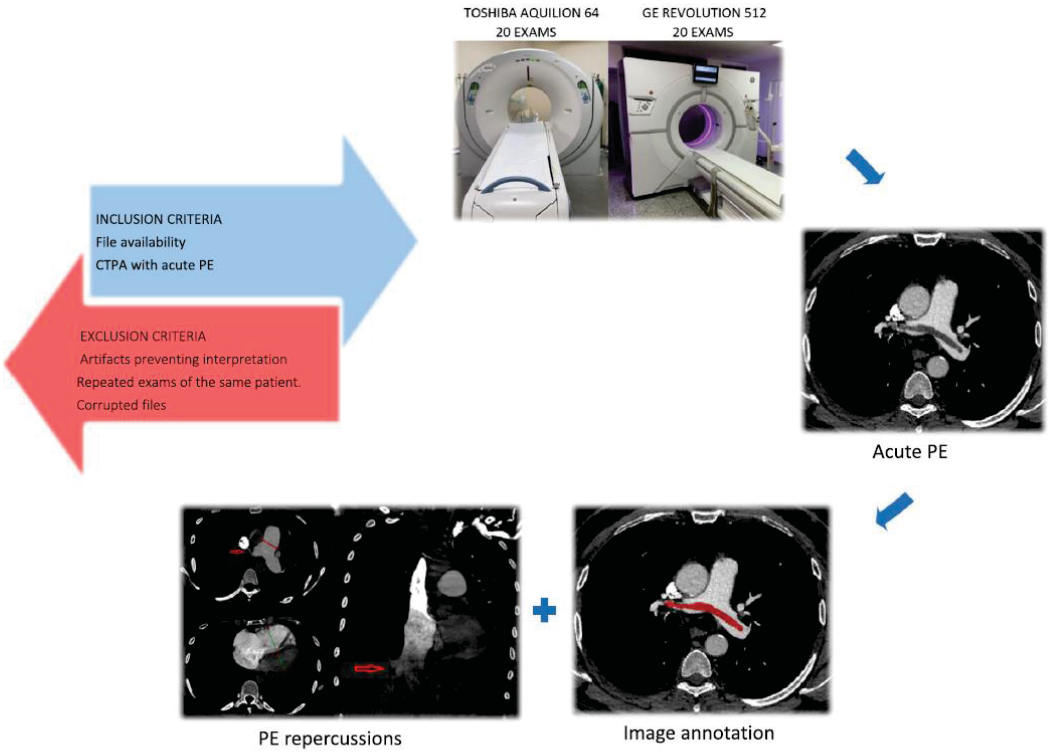


Fig. 1 Steps for producing the dataset. First step: definition of inclusion and exclusion criteria. Second and third steps: selection of 20 examinations performed with each machine on patients with acute PE. Fourth step: image annotation at the pixel and image levels. Fifth step: evaluation of features related to right heart strain and pulmonary artery hypertension.

the thrombus are demarcated as a ground truth, and annotations at the image and study levels[50], in which images with a visible PE receive a label, but the thrombus itself is not demarcated. The first approach is more versatile and can be converted into an annotation at the image level, but the opposite is not true.

The pixel-level annotation may help the algorithm to predict the exact region in which the embolus is located and to verify if this was the region used by the algorithm to generate the diagnostic output. It is especially useful for evaluating CADs designed to diagnose PE, as a high false-positive rate is a notable challenge for many of these algorithms.

To date, there are only two public datasets containing CTPA examinations with pulmonary emboli annotated at the pixel level. The first one contains 91 examinations of patients with PE obtained using SIEMENS CT scanners[51], and the second contains 35 examinations, 33 of which were conducted in patients with PE, obtained with CT scanners from PHILIPS and Neusoft Medical System Co[52]. The information published from these datasets does not disclose the examination selection process or the inclusion and exclusion criteria.

A recently published guide for research on AI[53] highlights the importance of using datasets containing images from devices from multiple vendors due to the variability inherent in such images.

There is a shortage of public datasets with a representative sample of examinations annotated at the pixel level that the AI algorithm can process in a real clinical setting. In medical practice, there is great variability among CTPA images, which occurs due to patient-related factors, such as different biotypes and comorbidities that can obscure PE, and technical factors, such as delays in image acquisition following infusion of the contrast media or with inadequate infusion flow, that lead to suboptimal image quality.

We developed a dataset with a sample of cases of acute PE annotated at the pixel and image levels, making it suitable for algorithms developed using both approaches (Fig. 1). Our dataset contains 40 examinations performed with multidetector scanners, half from a Toshiba CT and the other half from a GE CT[54].

The dataset was primarily used in conjunction with the two public datasets of CTPA examinations with pulmonary emboli annotated at the pixel level[51,52] to develop a program for the diagnosis of PE[48]. A method capable of finding and segmenting PEs in CT images using deep learning was subsequently developed.

We encourage the use of this dataset to develop, evaluate and compare the performance of AI algorithms designed to diagnose PE.

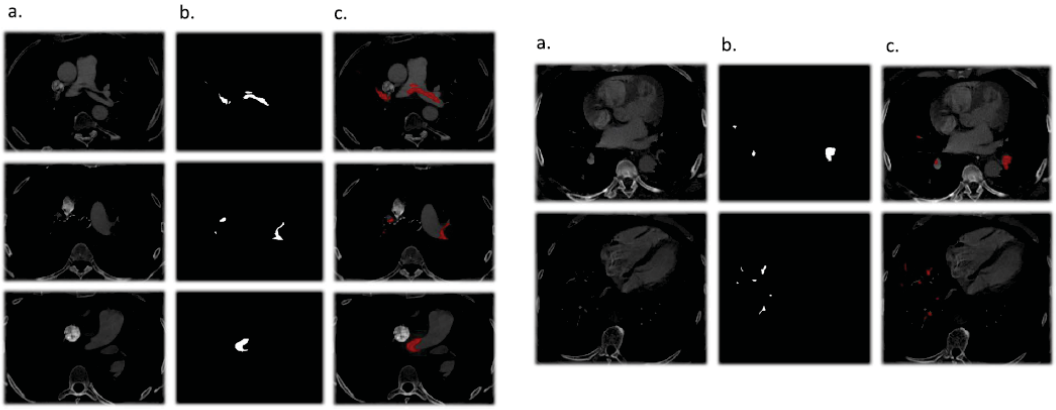


Fig. 2 Pixel level annotation. Column (**a**) shows CTPA images of pulmonary emboli in different anatomical locations. In column (**b**), pixel-level annotations show all pixels of the embolus of the corresponding CTPA image in column (**a**) in white. In column (**c**), the corresponding images of columns (**a**) and (**b**) are superimposed, and the thrombus is shown in red.

Methods

The study was approved by the Research Ethics Committee of Hospital de Clínicas-Federal University of Paraná. Given the retrospective nature of the study, only retrospective access to anonymized scan files was necessary, and so the need for informed consent was waived by the ethics committee. Aiming to preserve the identity of the participants, the examinations were deidentified by deleting patients' personal information, such as names, date of birth and identification numbers from the CT scans. Fields corresponding to time or numbers were replaced by "000000.00". Fields corresponding to dates were replaced by "00010101". Written Fields were replaced by "Anonymized" or removed.

Imaging. Twenty CTPA scans were performed in a public university hospital with a 64-channel Toshiba Aquilion CT scanner, with a tube voltage of 120 KVp, slice thickness of 1.0 mm, gantry rotation time of 0.5 seconds, beam pitch of 1.485 and dose modulation protocol.

The other twenty scans were performed in a private imaging practice with a GE Revolution 512 CT scanner, with a tube voltage of 120 KVp, slice thickness of 0.625 mm, slice interval slice thickness of 0.625 mm, gantry rotation time of 0.5 seconds, beam pitch of 0.9 and dose modulation protocol.

Image segmentation. All CTPA scans were used to make a diagnosis of acute PE by a staff radiologist. The diagnosis and location of the acute pulmonary embolism were confirmed by a thoracic radiologist with 32 years of experience. After this, a third-year resident generated the pixel-level ground-truth masks, which were revised by a certified radiologist.

The examinations were segmented using manual mode in ITK-SNAP[55], generating the ground-truth mask in which the thrombus pixels are demarcated (Fig. 2). Based on this mask, image-level segmentation was performed, labeling slices containing the thrombi.

CTPA features related to right heart strain and pulmonary artery hypertension. Obstruction of the pulmonary vasculature due to PE can increase vascular resistance, leading to an increase in pulmonary arterial pressure and right cardiac strain. Indirect signs such as pulmonary artery dilatation, right ventricular enlargement (increase in the right ventricle-to-left ventricle diameter ratio), inferior vena cava (IVC) contrast reflux, and abnormal positioning of the interventricular septum (flattening or even paradoxically bowing toward the left ventricle), can be observed on CTPA scans (Fig. 3).

In all CTPA scans, we evaluated the largest artery involved, inferior vena cava reflux, interventricular septum flattening or paradoxical bowing, pulmonary artery trunk diameter (PAD), transverse diameters of the right ventricle (RV) and left ventricle (LV)—measured between the endocardial surfaces in the largest place perpendicular to the longitudinal axis—and right ventricle-to-left ventricle diameter ratio (Tables 1, 2).

Data Records

All data records described in this paper are available on a Figshare repository[54]. This repository contains three folders. The first contains CTPA images in Digital Imaging and Communications in Medicine (DICOM) format. The second contains the ground-truth pixel-level segmentation of the location of the pulmonary embolus in Neuroimaging Informatics Technology Initiative (NiFTI) format. The segmentations consist of a three-dimensional matrix in which each element corresponds to a voxel of the CT scan: the elements corresponding to the embolus have a value of "1", and the others have a value of "0". The third is in comma separate values (CSV) format, in which each element corresponds to a slice of the CT scan. The first element represents

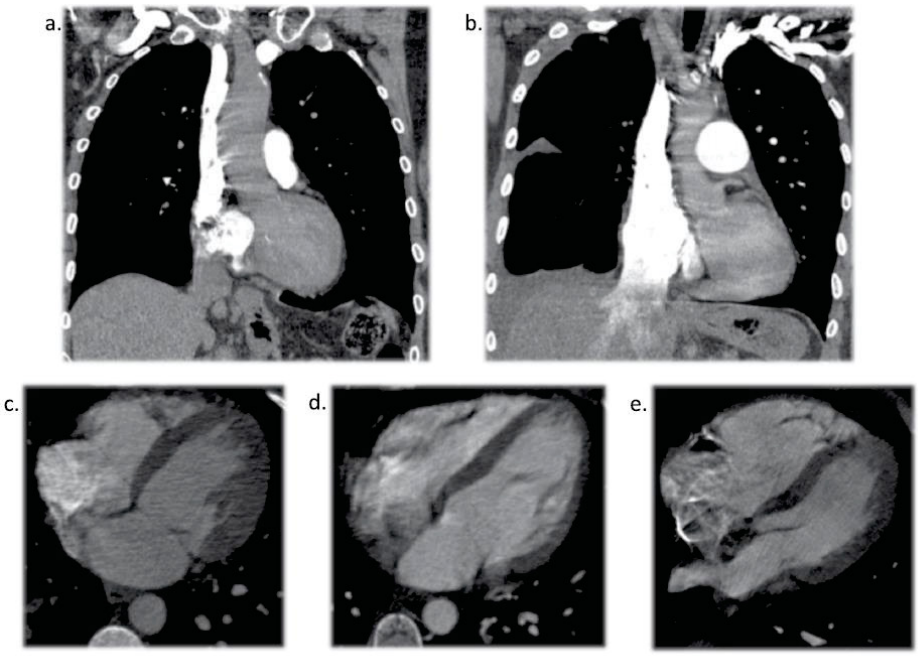


Fig. 3 CTPA features related to right heart strain and pulmonary artery hypertension. (**a**) Exam TS04 without IVC reflux. (**b**) Exam TS19 with IVC reflux. (**c**) Exam TS19 showing the interventricular septum in its normal position. (**d**) Exam TS02 showing a flattened interventricular septum. (**e**) Exam TS10 showing paradoxical interventricular septal bowing.

Case	Sex	Age	PAD (mm)	RV (mm)	LV (mm)	RV/LV	IVC Reflux	IV Septum	Largest affected vessel
1	F	63	24	43	57	0.75	Present	Normal	Lobar
2	F	11	20	37	40	0.92	Absent	Flattened	Trunk bifurcation
3	F	77	29	45	35	1.29	Absent	Flattened	Main pulmonary artery
4	F	38	34	52	29	1.79	Present	Paradoxically bowed	Main pulmonary artery
5	F	68	29	41	43	0.95	Absent	Flattened	Trunk bifurcation
6	F	54	28	39	54	0.72	Present	Flattened	Main pulmonary artery
7	F	58	35	33	35	0.94	Present	Normal	Lobar
8	F	64	31	37	48	0.77	Absent	Flattened	Segmental
9	M	29	28	23	31	0.74	Absent	Flattened	Segmental
10	M	68	26	52	35	1.49	Present	Paradoxically bowed	Main pulmonary artery
11	F	81	27	29	42	0.69	Present	Normal	Lobar
12	F	84	31	52	41	1.27	Present	Normal	Trunk bifurcation
13	F	41	24	32	38	0.84	Absent	Normal	Segmental
14	F	48	34	43	51	0.84	Present	Normal	Segmental
15	F	45	20	40	45	0.89	Absent	Normal	Main pulmonary artery
16	F	75	31	33	46	0.72	Present	Normal	Subsegmental
17	F	59	31	42	36	1.17	Present	Flattened	Trunk bifurcation
18	F	42	24	42	45	0.93	Absent	Normal	Segmental
19	F	26	25	35	50	0.70	Absent	Normal	Lobar
20	F	50	25	34	52	0.66	Absent	Normal	Main pulmonary artery

Table 1. Data from exams obtained with the Toshiba Aquilion 64. F: female. M: male. PAD: pulmonary artery diameter. RV: right ventricle diameter. LV: left ventricle diameter. RV/LV: right ventricle-to-left ventricle diameter ratio. IVC Reflux: inferior vena cava reflux of contrast media. IV Septum: interventricular septum position.

Case	Sex	Age	PAD (mm)	RV (mm)	LV (mm)	RV/LV	IVC Reflux	IV Septum	Largest affected vessel
1	M	59	27	52	57	0.91	Absent	Normal	Main Artery (unilateral)
2	F	72	32	60	40	1.50	Absent	Flattened	Main Artery (unilateral)
3	M	71	36	38	32	1.19	Present	Normal	Lobar
4	M	62	28	33	34	0.97	Absent	Flattened	Trunk bifurcation
5	M	73	30	32	42	0.76	Absent	Normal	Segmental
6	M	82	40	47	30	1.57	Present	Flattened	Trunk bifurcation
7	F	21	32	39	47	0.83	Absent	Normal	Segmental
8	F	82	29	32	42	0.76	Absent	Normal	Segmental
9	M	50	28	22	40	0.55	Present	Normal	Lobar
10	F	57	26	55	38	1.45	Absent	Paradoxically bowed	Subsegmental
11	M	31	27	25	51	0.49	Absent	Normal	Segmental
12	F	36	22	30	35	0.86	Absent	Normal	Lobar
13	F	86	26	35	36	0.97	Present	Normal	Subsegmental
14	F	57	38	50	48	1.04	Present	Flattened	Main Artery (unilateral)
15	F	32	27	44	42	1.05	Absent	Flattened	Segmental
16	M	38	30	46	50	0.92	Absent	Normal	Lobar
17	F	81	31	50	50	1.00	Present	Normal	Lobar
18	M	54	28	44	50	0.88	Absent	Normal	Lobar
19	M	41	29	40	49	0.82	Absent	Normal	Trunk bifurcation
20	M	54	29	38	64	0.59	Present	Normal	Segmental

Table 2. Data from exams obtained with the GE Revolution 512. F: female. M: male. PAD: pulmonary artery diameter. RV: right ventricle diameter. LV: left ventricle diameter. RV/LV: right ventricle-to-left ventricle diameter ratio. IVC Reflux: inferior vena cava reflux of contrast media. IV Septum: interventricular septum position.

the first slice, and the last element represents the lowest slice. Slices in which the embolus can be visualized are represented by the number "1", and the others by the number "0".

The files corresponding to patients scanned with the GE equipment are named 01GE to 20GE, and the files corresponding to patients scanned with the Toshiba equipment are named 01TS to 20TS, according to the number of each patient in Tables 1, 2, which are also available on the repository in Excel binary file (XLS) format.

Technical Validation

Exam selection. The CTPA scans were selected retrospectively by a search of the digital files from a public university hospital and a private imaging practice.

An arbitrary starting date was defined for each medical facility. From the defined starting date, all CTPA scans were sequentially reviewed until we reached 20 examinations with PE from each device that fit the inclusion and exclusion criteria.

For the Toshiba device, the examinations were performed from November 5, 2018 to February 5, 2019. For the GE device, the examinations were performed from November 9, 2018 to September 20, 2019.

Inclusion criteria. Chest CT scans performed using the PE protocol (CTPA) for diagnosing acute PE. File availability in the Picture Archive and Communication System (PACS) from each medical center.

Exclusion criteria. Artifacts that prevented the radiologist from visually interpreting the images. Examinations with fully or partially corrupted files. Follow-up examinations (only the first CTPA was used).

There was no restriction on age, patient status (inpatients or outpatients) or any other inclusion or exclusion criteria different from those mentioned.

Code availability

All code for loading and normalization of the dataset is available in GitHub (https://github.com/glescki/dicom_image_parser).

For parsing the data, the PyDicom library is recommended, and the loading of the labels can be performed with a parser available in GitHub. Each DICOM file should be loaded separately and then joined within a data structure.

For normalization, it is recommended that the spacing in the z-axis of all slices be changed to 1.

Received: 20 June 2022; Accepted: 11 July 2023;
Published online: 04 August 2023

References

1. Beckman, M. G., Hoopeer, W. C., Critchley, S. E. & Ortel, T. L. Venous thromboembolism: a public health concern. *Am J Prev Med.* **4**, 495–501, https://doi.org/10.1016/j.amepre.2009.12.017 (2010).
2. Konstantinides, S. V. *et al.* 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): the task force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). *Eur. Heart Journal.* **41**, 543–603, https://doi.org/10.1093/eurheartj/ehz405 (2020).
3. Dalen, J. E. Pulmonary embolism: what have we learned since Virchow? *Chest.* **122**, 1440–1456, https://doi.org/10.1378/chest.122.4.1440 (2002).
4. Konstantinides, S. V. *et al.* Task force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). ESC guidelines on the diagnosis and management of acute pulmonary embolism. *Eur. Heart J.* **35**, 3033–3069, https://doi.org/10.1093/eurheartj/ehu283 (2014).
5. Remy-Jardin, M. *et al.* Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner society. *Radiology.* **245**, 315–329, https://doi.org/10.1148/radiol.2452070397 (2007).
6. Wittram, C. *et al.* CT angiography of pulmonary embolism: diagnostic criteria and causes of misdiagnosis. *RadioGraphics.* **24**, 219–1238, https://doi.org/10.1148/rg.245045008 (2004).
7. Masutani, Y., Macmahon, H. & Doi, K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. *IEEE Trans. Med. Imaging.* **21**, 1517–1523, https://doi.org/10.1109/TMI.2002.806586 (2002).
8. Pichon, E., Novak, C. L., Kiraly, A. P. & Naidich, D. P. A novel method for pulmonary emboli visualization from high-resolution CT images. *Proceedings of the SPIE.* **5367**, 161–170, https://doi.org/10.1117/12.532892 (2004).
9. Zhou, C. *et al.* Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. *Acad. Radiol.* **12**, 782–792, https://doi.org/10.1016/j.acra.2005.01.014 (2005).
10. Maizlin, Z. V., Vos, P. M., Godoy, M. C. & Cooperberg, P. L. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. *J. Thorac. Imaging.* **22**, 324–329, https://doi.org/10.1097/RTI.0b013e31815b89ca (2007).
11. Buhmann, S. *et al.* Clinical evaluation of a computer-aided diagnosis (CAD) prototype for the detection of pulmonary embolism. *Acad. Radiol.* **14**, 651–658, https://doi.org/10.1016/j.acra.2007.02.007 (2007).
12. Schoepf, U. J. *et al.* Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. *J. Thorac. Imaging.* **22**, 319–323, https://doi.org/10.1097/RTI.0b013e31815842a9 (2007).
13. Liang, J. & Bi, J. Computer aided detection of pulmonary embolism with tobogganing and multiple instance classification in CT pulmonary angiography. *Inf. Process. Med. Imaging.* **20**, 630–641, https://doi.org/10.1007/978-3-540-73273-0_52 (2007).
14. Sebbe, R. *Computer-aided Diagnosis of Pulmonary Embolism in Opacified CT Images*, dissertation, Faculte Polytechnique de Mons. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.3563&rep=rep1&type=pdf (2006).
15. Engelke, C., Schmidt, S., Bakai, A., Auer, F. & Marten, K. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. *Eur. Radiol.* **18**, 298–307, https://doi.org/10.1007/s00330-007-0770-3 (2008).
16. Bouma, H. *Vessel-Diameter Quantification and Embolus Detection in CTA Images*, Phd Thesis, Technische Universiteit Eindhoven. https://research.tue.nl/nl/publications/vessel-diameter-quantification-and-embolus-detection-in-cta-image (2008).
17. Das, M. *et al.* Computer-aided detection of pulmonary embolism: influence on radiologists' detection performance with respect to vessel segments. *Eur. Radiol.* **18**, 1350–1355, https://doi.org/10.1007/s00330-008-0889-x (2008).
18. Walsham, A. C. *et al.* The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. *J. Comput. Assist. Tomogr.* **32**, 913–918, https://doi.org/10.1097/RCT.0b013e31815b3ed0 (2008).
19. Bouma, H., Sonnemans, J. J., Vilanova, A. & Gerritsen, F. A. Automatic detection of pulmonary embolism in CTA images. *IEEE Trans. Med.Imaging.* **28**, 1223–1230, https://doi.org/10.1109/TMI.2009.2013618 (2009).
20. Dinesh, M. S., *et al* Adaptive contrast-based computer aided detection for pulmonary embolism. *Proc. SPIE. 7260, Medical Imaging 2009: Computer-Aided Diagnosis*. https://doi.org/10.1117/12.812223 (2009)
21. Zhou, C. *et al.* Computer-aided detection of pulmonary embolism in computed tomographic pulmonary angiography (CTPA): performance evaluation with independent data sets. *Med. Phys.* **36**, 3385–3396, https://doi.org/10.1118/1.3157102 (2009).
22. Wittenberg, R., Peters, J. F., Sonnemans, J. J., Prokop, M. & Schaefer-Prokop, C. M. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an on-call setting. *Eur. Radiol.* **20**, 801–806, https://doi.org/10.1007/s00330-009-1628-7 (2010).
23. Dewailly, M. *et al.* Computer-aided detection of acute pulmonary embolism with 64-slice multidetector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. *J. Comput. Assist. Tomogr.* **34**, 23–30, https://doi.org/10.1097/RCT.0b013e3181b2e383 (2010).
24. Lee, C. W. *et al.* Evaluation of computer aided detection and dual energy software in detection of peripheral pulmonary embolism on dual-energy pulmonary CT angiography. *Eur. Radiol.* **21**, 54–62, https://doi.org/10.1007/s00330-010-1903-7 (2011).
25. Park, S. C., Chapman, B. E. & Bin, Z. B. A multistage approach to improve performance of computer-aided detection of pulmonary embolisms depicted on CT images: preliminary investigation. *IEEE Trans Biomed Eng.* **58**, 1519–1527, https://doi.org/10.1109/TBME.2010.2063702 (2011).
26. Blackmon, K. N. *et al.* Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? *Eur Radiol.* **21**, 1214–1223, https://doi.org/10.1007/s00330-010-2050-x (2011).
27. Wittenberg, R. *et al.* Impact of image quality on the performance of computer-aided detection of pulmonary embolism. *Am. J. Roentgenol.* **196**, 95–101, https://doi.org/10.2214/AJR.09.4165 (2011).
28. Wittenberg, R. *et al.* Acute pulmonary embolism: effect of a computer-assisted detection prototype on diagnosis—an observer study. *Radiology.* **262**, 305–313, https://doi.org/10.1148/radiol.11110372 (2012).
29. Kligerman, S. J., Lahiji, K., Galvin, J. R., Stokum, C. & White, C. S. Missed pulmonary emboli on CT angiography: assessment with pulmonary embolism–computer-aided detection. *Am. J. Roentgenol.* **202**, 65–73, https://doi.org/10.2214/AJR.13.11049 (2014).
30. Özkan, H., Osman, O., Şahin, S. & Boz, A. F. A novel method for pulmonary embolism detection in CTA images. *Comput. Methods Programs Biomed.* **113**, 757–766, https://doi.org/10.1016/j.cmpb.2013.12.014 (2014).
31. Lahiji, K., Kligerman, S., Jeudy, J. & White, C. Improved accuracy of pulmonary embolism computer-aided detection using iterative reconstruction compared with filtered back projection. *Am. J. Roentgenol.* **203**, 763–771, https://doi.org/10.2214/AJR.13.11838 (2014).
32. Tajbakhsh, N., Gotway, M. B. & Liang, J. Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science.* **9350**, 62–69, https://doi.org/10.1007/978-3-319-24571-3_8 (2015).
33. Ozkan, H., Tulum, G., Osman, O. & Sahin, S. Automatic detection of pulmonary embolism in cta images using machine learning. *Elektron. Elektrotech.* **23**, 63–67, https://doi.org/10.5755/j01.eie.23.1.17585 (2017).
34. Tajbakhsh, N., Shin, J. Y., Gotway, M. B. & Liang, J. Computer-aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation. *Med. Image Anal.* **58**, 101541, https://doi.org/10.1016/j.media.2019.101541 (2019).
35. Rajan, D., Beymer, D., Abedin, S. & Dehghan, E. Pi-PE: A pipeline for pulmonary embolism detection using sparsely annotated 3D CT images. *Proceedings of the Machine Learning for Health NeurIPS Workshop*, **116**, 220–232, http://proceedings.mlr.press/v116/rajan20a/rajan20a.pdf (2020).

36. Yang, X. *et al.* A two-stage convolutional neural network for pulmonary embolism detection from CTPA images. *IEEE Access.* **7**, 84849–84857, https://doi.org/10.1109/ACCESS.2019.2925210 (2019).
37. Cano-espinosa, C., Cazorla, M. & González, G. Computer aided detection of pulmonary embolism using multi-slice multi-axial segmentation. *Appl. Sci.* **10**, 2945, https://doi.org/10.3390/app10082945 (2020).
38. Huang, S., Pareek, A., Zamanian, R., Banerjee, I. & Lungren, M. P. Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection. *Sci. Rep.* **10**, 22147, https://doi.org/10.1038/s41598-020-78888-w (2020).
39. Yu, C. -Y., Cheng, Y. -C. & Kuo, C. Early pulmonary embolism detection from computed tomography pulmonary angiography using convolutional neural networks. 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), https://doi.org/10.1109/ICIEVicIVPR48672.2020.9306659 (2020).
40. Huang, S. *et al.* PENet—a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. *npj Digit. Med.* **3**, 61, https://doi.org/10.1038/s41746-020-0266-y (2020).
41. Liu, W. *et al.* Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. *Eur. Radiol.* **30**, 3567–3575, https://doi.org/10.1007/s00330-020-06699-8 (2020).
42. Weikert, T. *et al.* Automated detection of pulmonary embolism in CT pulmonary angiograms using an AI-powered algorithm. *Eur. Radiol.* **30**, 6545–6553, https://doi.org/10.1007/s00330-020-06998-0 (2020).
43. Vainio, T., Mäkelä, T., Savolainen, S. & kangasniemi, M. Performance of a 3D convolutional neural network in the detection of hypoperfusion at CT pulmonary angiography in patients with chronic pulmonary embolism: a feasibility study. *Eur. Radiol. Exp.* **5**, 45, https://doi.org/10.1186/s41747-021-00235-z (2021).
44. Li, X., Wang, X., Yang, X., Lin, Y. & Huang, Z. Preliminary study on artificial intelligence diagnosis of pulmonary embolism based on computer in-depth study. *Ann. of transl. med.* **9**, 838, https://doi.org/10.21037/atm-21-975 (2021).
45. Long, K. *et al.* Probability-based mask R-CNN for pulmonary embolism detection. *Neurocomputing.* **422**, 345–353, https://doi.org/10.1016/j.neucom.2020.10.022 (2021).
46. Huhtanen, H. *et al.* Automated detection of pulmonary embolism from CT-angiograms using deep learning. *BMC Med. Imaging.* **22**, 43, https://doi.org/10.1186/s12880-022-00763-z (2022).
47. Yadlapalli, P. *et al. Segmentation of pulmonary embolism using deep learning. 2022 International Conference for Advancement in Technology (ICONAT)* https://doi.org/10.1109/ICONAT53423.2022.9726048 (2022).
48. Olescki, G., Andrade, J. M. C., Escuissato, D. L. & Oliveira, L. F. A two-step workflow for pulmonary embolism detection using deep learning and feature extraction. *Comput. Methods in Biomech. Biomed. Eng.: Imaging Vis.* https://doi.org/10.1080/21681163.2022.2060866 (2022).
49. Paiva, O. A. & Prevedello, L. M. The potential impact of artificial intelligence in radiology. *Radiol. Bras.* **50**, 5–6, https://doi.org/10.1590/0100-3984.2017.50.5e1 (2017).
50. Colak, E. *et al.* The RSNA pulmonary embolism CT dataset. *Radiology: Art. Int.* **3**, 2, https://doi.org/10.1148/ryai.2021200254 (2021).
51. González, G., et al. Computer aided detection for pulmonary embolism challenge (CAD-PE). Preprint at https://arxiv.org/abs/2003.13440 (2019).
52. Masoudi, M. *et al.* A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. *Sci. Data.* **5**, 180180, https://doi.org/10.1038/sdata.2018.180 (2018).
53. Bluemke, D. A. *et al.* Assessing radiology research on artificial intelligence: a brief guide for authors, reviewers, and readers—from the *Radiology* editorial board. *Radiology.* **294**, 487–489, https://doi.org/10.1148/radiol.2019192515 (2020).
54. Andrade, JMC. *et al.* Annotated dataset of acute pulmonary embolism computed-tomography images - READ PE CT, *figshare*, https://doi.org/10.6084/m9.figshare.c.6033134.v1 (2023).
55. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage.* **3**, 1116–1128, https://doi.org/10.1016/j.neuroimage.2006.01.015 (2006).

Acknowledgements

We are grateful to the Hospital de Clínicas of the Federal University of Paraná and to the *Clínica Diagnóstico Avançado por Imagem (DAPI)* for allowing the conduction of this project. We are also grateful to the postgraduate programs in internal medicine and health sciences and informatics from the Federal University of Paraná that made this project possible.

Author contributions

João Mario Clementin de Andrade. Project planning. Patient selection following the preestablished inclusion and exclusion criteria. Assistance in the definition of the segmentation method. Pixel-level segmentation. Review of all ground-truth masks. Writing of the manuscript. Gabriel Olescki. Project planning. Definition of the segmentation method. Image-level segmentation. Writing of the manuscript. Dante Luiz Escuissato. Project planning. Assistance in the selection of patients following the preestablished inclusion and exclusion criteria. Review/confirmation of the acute pulmonary embolus location of all patients. Orientation/revision of the manuscript. Lucas Ferrari Oliveira. Project planning. Definition of the segmentation method. Review of this manuscript. Ana Carolina Nicolleti Basso. Pixel level segmentation. Review of the manuscript. Gabriel Lucca Salvador. Pixel level segmentation. Review of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.M.C.d.A.

5.2 PROGRAMA DESENVOLVIDO EM PARCERIA COM O DEPARTAMENTO DE INFORMÁTICA DA UFPR

O programa de informática desenvolvido utilizando a nova base de dados criada e as duas bases de dados públicas com exames marcados ao nível dos pixels foi o projeto de mestrado de Gabriel Olescki, pelo programa de pós-graduação do departamento de Informática da UFPR, com a orientação do Prof. Dr. Lucas Ferrari de Oliveira. O resultado do trabalho foi publicado na revista “Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization”, novo qualis A4 (OLESCKI et al, 2022).

Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tciv20

A two step workflow for pulmonary embolism detection using deep learning and feature extraction

G. Olescki, João M.C. Clementin de Andrade, Dante L. Escuissato & Lucas F. Oliveira

To cite this article: G. Olescki, João M.C. Clementin de Andrade, Dante L. Escuissato & Lucas F. Oliveira (2022): A two step workflow for pulmonary embolism detection using deep learning and feature extraction, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, DOI: 10.1080/21681163.2022.2060866

To link to this article: https://doi.org/10.1080/21681163.2022.2060866

Published online: 14 Apr 2022.

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING: IMAGING & VISUALIZATION
https://doi.org/10.1080/21681163.2022.2060866

A two step workflow for pulmonary embolism detection using deep learning and feature extraction

G. Olescki[a], João M.C. Clementin de Andrade[b], Dante L. Escuissato[b] and Lucas F. Oliveira[a]

[a]Department of Informatics, Federal University of Paraná, Curitiba, Brazil; [b]Department of Internal Medicine, Federal University of Paraná, Curitiba, Brazil

ABSTRACT

Pulmonary embolism is among the leading causes of death all over the world. In order to achieve a fast diagnosis and response from the medical team, a CT exam is used as a means to detect embolisms. Over the years, different methods of computer-aided diagnosis systems (CADs) were implemented to facilitate the analysis of a CT exam for pulmonary embolism detection, and due to the high amount of data produced in these exams, methods that use deep learning are also growing as of late. This paper proposes a pipeline to detect pulmonary embolisms from a CT exam. It uses a U-net network to detect embolisms candidates and classifies them between false positives and true positives using machine learning algorithms. The method achieved a dice score of 0.81 and an IoU of 0.79.

ARTICLE HISTORY
Received 1 October 2021
Accepted 28 March 2022

KEYWORDS
Pulmonary Embolism; Computer Tomography; CNN; U-net network; Feature extraction; Deep Learning

1. Introduction

Pulmonary embolism (PE) is a clinical condition where the patient has a thrombus (clot that moves from elsewhere in the body) in pulmonary vessels, and this thrombus can harshly reduce or even interrupt the blood flow of the artery, which can be fatal (Huisman et al. 2018). It may cause up to 300,000 deaths per year in the US (Konstantinides et al. 2019). The mortality rate (30%) can be reduced to, as low as 2% with an early diagnosis, showing that a fast and accurate diagnosis is critical to saving those lives (Jha et al. 2013; Sadigh et al. 2011).

The modality of choice for diagnosis of the PE is CT exam due to its fast acquisition and reliability using angiography (CTA) (Goldhaber and Bounameaux 2012; Kligerman et al. 2018). This exam produces a 3-dimensional view of the pulmonary trunk, containing hundreds of images (slices) that are carefully analysed by a radiologist, making the process highly time-consuming. Besides that, the interpretation of CTA exams requires a high degree of training due to the high amount of noise or other artifacts that can be confused for a PE. Unlike other pulmonary diseases that have been observed with CT, such as lung nodules or emphysema, PEs are irregularly shaped and can appear in different regions of the lung (Rajan et al. 2020). All of those factors contribute to a high chance of diagnostic error, the difference in the correct diagnosis rate among radiologists that work overnight and in the daytime is 13% (Huang et al. 2020).

Because of this error rate in diagnosis, a computed-aided diagnosis system (CADs) becomes valuable for PE detection. Due to the high number of data produced per examination, systems based on learning data are already able to find the best success rate in the diagnosis of PE (Tajbakhsh et al. 2019).

The objective of this work has been to propose a method capable of finding segmented PEs in CT images using deep learning and classifying them between false positives and true positives. The first step segmented PE candidates, using a network based on U-net (Ronneberger et al. 2015). In the second stage, another network classifies PE candidates between PE and non-PE, reducing the number of false positives.

The proposed method has successfully identified PEs inside CT exams achieving a dice score of 0.81. The model was validated using public data sets and mixing them for testing, one data set used remains private, but it will be publicly available soon. This work also validates the model by analysing the results of each data set separately, observing the model capacity for different PE sizes.

2. Related works

This section reviews some of the literature for the classification and segmentation of pulmonary embolisms (Table 1), for the most part, the data used in the papers were taken from CTA exams or CT exams.

Existing CAD systems follow mainly the same steps to perform PE detection: (1) lung and pulmonary artery segmentation, (2) creation of a set of PE candidates, (3) extraction characteristics of PE candidates and (4) removal of false positives from this set (Tajbakhsh et al. 2015), and the last step was performed using supervised learning techniques.

In the work of Bouma et al. (2009), the authors proposed a segmentation using thresholding of the grey channels and Hessian matrix to generate PE candidates, which classified with decision trees. This work reported a recall of 58% and a rate of 15 false positives per scan (FP/s).

CONTACT Lucas F. Oliveira lferrari@inf.ufpr.br Department of Informatics, Federal University of Paraná, R. Evaristo F. Ferreira da Costa, 383-391 - Jardim das Américas, Curitiba - PRCuritiba - PR, 82590-300, Brazil

Table 1. Summary of each related work reviewed.

Name	Author (Year)	Data set	Results
Automatic detection of pulmonary embolism in CTA images	Bouma et al. (2009)	57 Exams, 38 for training and 19 for testing	Sensitivities of 58%, 63% and 73% at 4, 4.9 and 15 FP/p, respectively
Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks	Tajbakhsh et al. (2015)	Two data sets, a private one with 121 exams and a public one with 20 exams	In the private database, an 83.4% sensitivity with 2.0 FP/p; in the public database a sensitivity of 34.6% at 2 FP/p at a 0 mm error threshold
A two-stage convolutional neural network for pulmonary embolism detection from CTPA images	Yang et al. (2019)	Two data sets, a public with 20 exams and a second data set combining 99 private exams with 30 from another public data set	In the public data set, 75.4% sensitivity and 2 FP/p at 0 mm, 2 mm and 5 mm error threshold; in the private data set, sensitivities were of 76.3%, 78.9% and 84.2% all with 2 FP/p at 0 mm, 2 mm and 5 mm error
PENet – a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging	Huang et al. (2020)	1797 exams, 1461 used in training, 167 for validation and 169 for testing and a second test data set with 200 exams	AUROC 0.84 in the test set and 0.85 in the second test set
Computer aided detection of pulmonary embolism using multi-slice multi-axial segmentation	Cano-Espinosa et al. (2020)	80 exams, 60 for training and 20 for testing	Sensitivities of 55%, 61% and 68% with FP/p of 1, 1 and 0.95 at 0, 2 and 5 mm error threshold, respectively
Probability-based mask *R*-CNN for pulmonary embolism detection	Long et al. (2021)	35 exams from the FUMPE data set	Average precision (AP), AP50 and AP75 of 41.87%, 81.55% and 41.43%, respectively, and a dice score of 0.75

In Tajbakhsh et al. (2015), the PE candidates are found using the segmentation algorithm tobogganing to find candidates and classify them with a CNN. The work used a representation of the PE candidate in 2.5 D using eigenvectors summarising components found by the segmentation algorithm. The method used two databases: one private and one public. The methodology reached 83% recall and 2 FP/s and 34% recall and 2 FP/s in private and public databases, respectively.

The work (Yang et al. 2019) was the first to use 3D U-Net convolutions to find PE candidates and classify them with a ResNet-18 network. The method achieved a recall of 75.4% and 2 FP/s in a public database and a recall of 76.3% with 2 FP/s in the private one.

In Rajan et al. (2020), the authors used a U-net with 2D convolutions using the neighbourhood of the slices as the different input channels. Using an input image with nine channels, each channel being the neighbouring slices from a piece of the CT volume, the work obtained an area under the ROC curve (AUC-ROC) of 0.94.

In Huang et al. (2020), the authors proposed a novel network, nicknamed PENet, based on DenseNets and U-nets that classifies whether an exam contains PE or not. The network also used transfer learning during training using the pre-trained weights with the base Kinects-600, and obtained an AUROC of 0.85.

In Cano-Espinosa et al. (2020), three variations of the U-net network were used. One using 2D convolutions. Another one that also uses 2D convolutions, but with 3D input images, and the last one using 3D convolutions. Of all the models, the one that had better results was the 3D network. The raw output value in the most central voxel of each PE defined if it were a true positive or not, the method got a recall of 55% and 1 FP/s, for training and testing the use of the database (González et al. 2020).

In the work by Long et al. (2021), the authors proposed a deep learning model based on the instance segmentation neural network *R*-CNN (Girshick et al. 2014). The proposed network, the P Mask *R*-CNN, aims to improve small object detection by enlarging the featured maps and anchor selection from the Region Proposal Network (RPN) module of the Mask *R*-CNN. The data used to test the model originated from the FUMPE data set (Masoudi et al. 2018), containing 35 patients, split into 29 patients for training, 2 for validation and 4 for testing, then applying rotation and flipping data augmentation, to prevent over-fitting. The model obtained an average precision (AP), AP50, AP75 and dice score of 41.87%, 81.55%, 41.43% and 0.75, respectively. It had the best results compared to other state-of-the-art models, such as the Mask *R*-CNN.

3. Materials and methods

This section explains the data sets used in this work to validate the proposed method, the algorithms and techniques used during the preprocessing of data, the segmentation of PE candidates and their classification into a true positive or a false positive.

3.1. Data

The data used to train and evaluate the proposed model come from three different data sets, two public and one private. All databases have PE semantic masks.

The first data set available (named 'FUMPE') comes from Masoudi et al. (2018), which is composed of 35 CTAs, with two exams not having any PE. A radiologist with 5 years of experience in chest tomography exams produced the segmentation masks, which then were reviewed by another radiologist with 18 years of experience. The data set has 8,792 slices and 3,438 PE regions of interest, with the height and width of the voxels ranging between 0.52 and 0.78 mm. Most PEs (67%) were found in peripheral arteries. In addition to the markings, the authors have also provided metadata, such as the

Table 2. Count of the number of PEs present in each volume interval (in mm^3) for each data set used.

Volume intervals (mm^3)	FUMPE	CAD-PE	Private
$0.1 \leq n < 0.5$	0	74	35
$0.5 \leq n < 1$	1	117	23
$1 \leq n < 2$	0	93	26
$2 \leq n < 9$	3	96	69
$9 \leq n < 38$	8	68	51
$38 \leq n < 168$	13	157	71
$168 \leq n < 740$	31	155	74
$740 \leq n < 3,270$	32	80	25
$3,270 \leq n < 14,420$	19	54	29
$14,420 \leq n \leq 63,650$	11	27	5

proportion of the size of the right ventricle and left ventricle of the heart, whether there was reflux in the inferior vena cava, whether the interventricular septum is straight, the pulmonary artery diameter and the value of Qanadli score, which is calculated based on the position of the thrombus within the lung.

The second database called 'CAD-PE' (González et al. 2020) is composed of 91 CT angiography exams and a total of 24,624 slices with voxel sizes ranging between 0.52 and 0.92 mm. The segmentation of the first 40 exams was done by a board of radiologists with more than 10 years of experience, while the other 51 exams were segmented by a single board member. All markings were performed with the aid of a semi-automated method, in which the markings performed by the radiologists were of the area of interest (i.e. a rectangle in the region of the PE), with this area of interest a semi-supervised method was used to generate the semantic segmentation masks, each segmentation performed was then averaged to produce the output mask. Many related works utilise data from the CAD-PE challenge, with a new recent wave of papers for PE detection since the republication of the data set with new labelled data.

The third database is private, consisting of 40 exams, 20 produced by the GE Revolution 512 tomographer, 20 by the Toshiba Aquilion 64, totalling 14,463 slices. The sizes of the voxels range from 0.57 and 1.0 mm. The exams were evaluated by a thoracic radiologist with 32 years of experience regarding the presence of PTE and its location. After this review, a physician in the third year of the radiology residency marked the exams producing the masks. No semi-automated method was used to create the segmentation masks. The database was created in partnership with the *Hospital de Clnicas (HC)* at *Universidade do Paraná* (UFPR), the database was approved by the ethics committee of the HC with CAAE: 07724919.0.0000.0096. Table 2 summarises the three databases, showing the PE count in intervals based on sizes in mm^3.

3.2. Proposed method

The technique proposed in this work is composed of three main steps: the data are preprocessed; then a sliding window of five channels from the CT scan is fed to a U-net, which locates and performs segmentation of PE candidates. In the third and final step, the PE candidate has its features extracted and classified as an embolism (true positive) or not embolism (false positive) (Figure 1).

3.2.1. Preprocessing

To facilitate the training of the proposed method, the same preprocessing methods were applied in all three data sets. Firstly, the width of the slice was changed to 1 mm for all patients to standardise the inputs. Secondly, the greyscale values of the images were normalised using a Hounsfield unit (HU) scale windowing, with a window length (WL) of 350 HU and a window centre (WC) of 150 HU, transforming the values between $[-200, 500]$. Thirdly, the values were normalised between 0 and 1 to improve the learning and convergence rate of the network. For each image, a central window of size 364×364 was also cut, removing regions around the rib cage from the images.

A data augmentation step was performed on the training base, with the aim of increasing the number of images for better network training, horizontal and vertical flips with a probability of 50% and Gaussian noise with 20% probability were applied for each image.

To feed the input to the U-net network, the total image volume was split into 5-slice overlapping windows, with each slice being a channel of the image. The central channel (i.e. channel 3) was the aim for the network segmentation.

3.2.2. U-net model

As stated previously, a U-net model was used for the preliminary results obtained. Besides being based on Ronnenberg's original U-net, it does not follow the same amount of layers and hyperparameters. The model is divided into two parts: an encoder and a decoder. The encoder part is built using basic blocks. Each basic block is composed of two convolutions, each followed by batch normalisation and a ReLU activation. The whole encoder is composed of five basic blocks with a number of channels of each convolution of 64, 64, 128, 256 and 512. Subsequently, five decoder blocks are applied and each decoder block is composed of two de-convolution layers, where each is followed by ReLU only and the number of channels of the convolutions is the same as in the encoder part, but decreasing. The input of each decoder block is the output of the last block concatenated with the output of the corresponding encoder block (i.e. the one with the same amount of channels). At the end of the last decoder block, a final 2D convolution is applied with a single channel as output and the sigmoid function is applied. This is the final output of the network, having the same shape as the input image segmentation mask.

Unlike the original U-net network, the proposed network receives as input an image of size $364 \times 364 \times 5$ and generates an output 364×364. The five input channels are the five neighbouring slices in the CT volume, so the 2D convolutions use dimensional data as the channels.

3.2.3. Candidates' extraction

To find PE candidates, the output values of the network were rounded to 0 (values undefined.5) and 1 (values undefined.5), then a connected component algorithm (CCL) was applied, and each component found by the CCL was considered a PE candidate. For an initial reduction of false positives, a closing operation was performed with a kernel of $5 \times 3 \times 3$, aiming to remove small particles and noise, candidates found in only one slice of the exam were also removed.

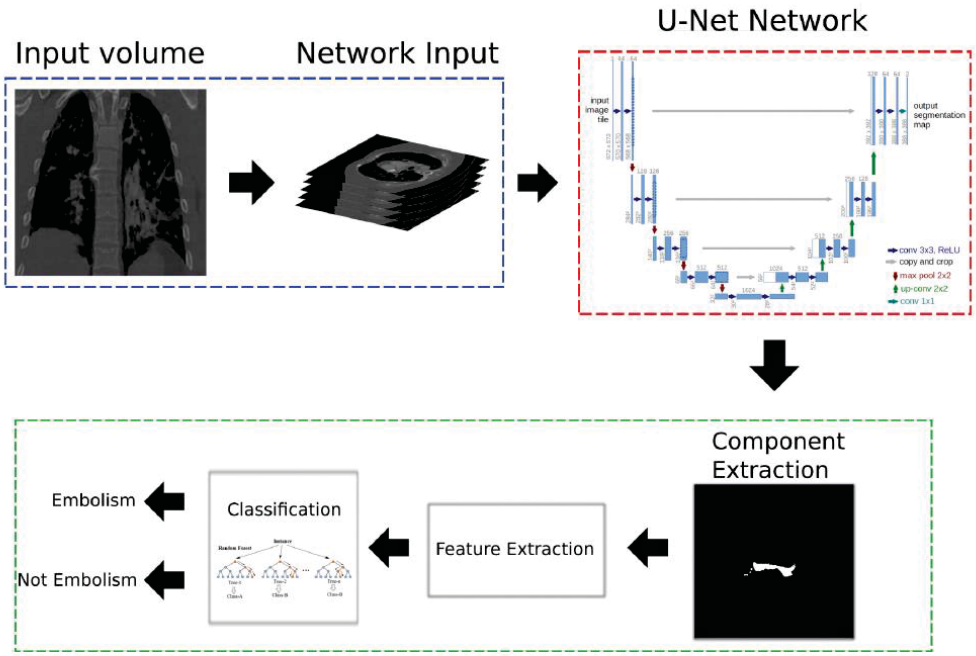


Figure 1. Step-by-step representation of the proposed method. In blue, the steps of preprocessing and slice extraction for the input of the network. In red, the representation of the U-net network adapted from Ronneberger et al. (2015). In green, the steps extraction of PE candidates, feature extraction and classification.

3.2.4. Feature extraction

After the candidate extraction step, the feature extraction of each candidate is carried out. The extraction was made for each candidate produced in the 3D CCL algorithm.

First, for each candidate, the 3D coordinates of the centre of mass (CM) were defined and selected as features. Then, the HU scale value and the output from the U-net network of these central coordinates were selected as features, both raw values. Finally, the total volume in mm^3 of each candidate was calculated, along with the larger distance along the three axes (x, y, z), which was obtained by setting the value of the median x, y and z and finding the maximum distances in the other axis. In the end, nine features (CM coordinates (x, y, z), CM HU value, CM output U-net value, volume, larger distance in axes x, y, z) were used in the classification step, which were normalised using a standard scaler, leaving the values between 0 and 1. The selected features were chosen based on other works that used machine learning for PE classifications, such as Cano-Espinosa et al. (2020), Özkan et al. (2014), Bouma et al. (2009) and Liang and Bi (2007)

3.2.5. Classification

To make the classification between PE and non-PE candidates, a combination of classifiers was trained and tested. In all tests regarding the FP classification step, the same hold-out division of the data sets was used (train, validation and test).

The classifiers KNN, SVM (with linear and Gaussian), random forest, Naive Bayes and a multi-layer perceptron, with some early testing using the training and validation set, some algorithms that performed poorly were discarded, keeping only the KNN and SVM with Gaussian kernel and random forest. The next step was to find the optimal parameters and hyper-parameters for each classifier. For the KNN classifier, a range of neighbours were tested using the first 100 odd numbers beginning from 3. For the SVM classifier, a grid search was performed, varying the C (0.001, 0.01, 0.1, 1, 10) and gamma values (0.001, 0.01, 0.1, 1), for the random forest, another grid search was done, using the range of parameters for the maximum depth (200–2000, increasing 200 at each step) and the number of estimators (100–500 increasing 40 at each step). For both grid searches, the optimal parameters were found by means with fivefold cross validation. All training of the classification step was done with a combination of all three data sets after being segmented by the U-net network.

4. Results and discussion

The results are presented with a two-step method. A segmentation technique uses a U-net model to detect candidates and a machine learning algorithm for PE classification. Moreover, the tests included two approaches varying the data distribution and the hyper-parameters.

4.1. Segmentation results

4.1.1. Model-centric approach

The development of the U-net segmentation network and all the training process used the PyTorch library (Paszke et al. (2019)). We tested the three data sets separately using a hold-out division of 70% for training, 10% for validation and 20% for testing. The number of epochs used was 100, and the highest IoU was chosen for the test step.

For training, the optimiser used was Adam with an initial learning rate of 1×10^{-4}, which was reduced by 1×10^{-2} every time the loss reached a plateau. The loss function used was the dice and the IoU with a threshold of 0.5 of the output values. The size of batches per iteration of training chosen was 16 images, and each image having five channels of neighbouring slices.

The experiments had many hyper-parameters tested in each data set: the HU window scale, the optimiser, the initial learning rate and a method (or absence) for reducing the learning rate during training, the number of batches for training and the preprocessing steps (normalisation and scaling). However, only the parameters cited showed a satisfactory result. In Masoudi et al. (2018) data set, the network got a 0.68 mean IoU, 0.72 dice score and recall of 0.82 in the test set. Figure 2 shows the changes in the loss. It is observable that the validation loss quickly decreases as values below the training loss. There are two possibilities for this to have happened. The first one is due to the amount of data used since the plot shows the mean loss during an epoch; another one is that there are more patients in the training set than in the validation set. While in the CAD-PE test set the proposed algorithm achieved a mean IoU of 0.79, 0.81 mean dice score, and 0.89 recall, the training and validation losses can be observed in Figure 2, showing that with more data, the curves behave more conventionally. In the HC data set, a mean IoU of 0.75, dice of 0.78 and recall of 0.78 were achieved (Table 3).

4.1.2. *Data-centric approach*

Another experimental approach was to use different combinations of the selected data sets for training and testing the U-net. The network used was the same as the model-centric tests. Varying just in some hyper-parameters for training with a higher amount of data.

We tested two combinations of the data sets: one with the CAD and the FUMPE data sets for training/validation and the HC data set for the test, and another training with the CAD and HC data sets and testing in the FUMPE data set.

In training, the SGD optimiser had better results with a constant learning rate of 1×10^{-4}, 16 images in each batch and different HU window scales.

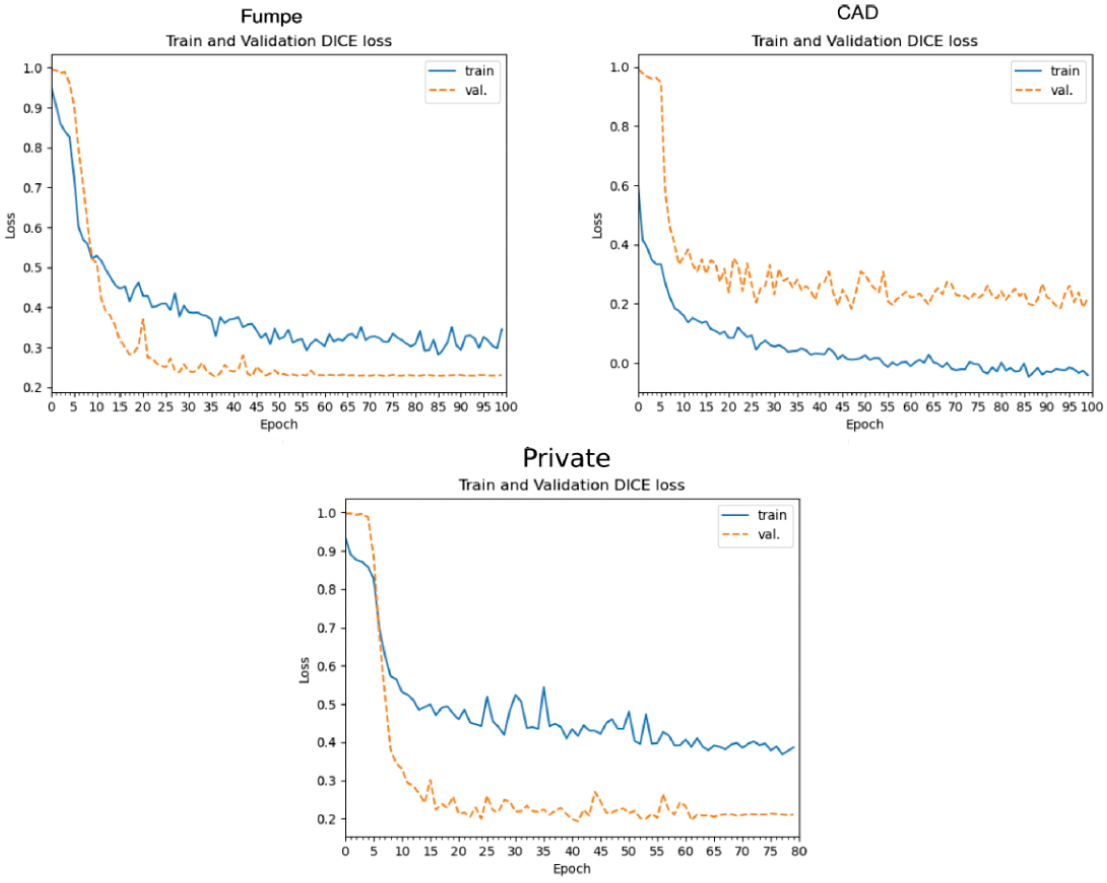


Figure 2. **Top left**: Dice loss over the epochs during the U-net training step for the data set (Masoudi et al. 2018). **Top right**: Dice loss over the epochs during the U-net training step for the data set (González et al. 2020). **Bottom**: Dice loss during training of U-net model in the private data set. In the loss of both the FUMPE data set and the private data set, the validation set had a lower loss than the training set. the possible explanation is that due to the low amount of data in those data sets, the validation set is less representative and easiest to classify than the training set.

Table 3. Segmentation results in the model-centric approach separately.

Data set	Dice	IoU	Recall
FUMPE	0.72	0.68	0.82
CAD-PE	0.81	0.79	0.89
Private	0.78	0.75	0.78

Table 4. Results from training the model with the two whole data sets and testing in the HC data set.

Data set	Dice	IoU	Recall
CAD+FUMPE (test HC)	0.78	0.76	0.86
CAD+HC (test FUMPE)	0.74	0.70	0.86

Table 5. Segmentation of the proposed U-net (Long et al. 2021).

Data set	AP50	AP75	AP
P Mask *R*-CNN	81.55	41.87	41.43
Proposed	0.1875	0.06	0.06

Table 6. Segmentation of the proposed U-net (Cano-Espinosa et al. 2020).

Work	Sensibility
Cano-Espinosa et al.	0.68
Proposed	0.82

Table 4 shows results achieved with each combination. The combination of CAD and FUMPE data sets and testing in the HC data set present slightly better results.

4.1.3. Model comparison

Two papers were selected, which used only public data sets for their validation. The results from the proposed model used the same data as the original works. The first comparison was with the paper by Long et al. (2021), in which the authors had found EPs using a model based on the Mask *R*-CNN. The metric used to evaluate the model was the average precision with different IoU thresholds on the FUMPE data set. Table 5 shows the comparison results.

The second comparison work is from Cano-Espinosa et al. (2020), which used only the CAD data set. The authors also used a U-net-based model for segmentation, reporting the sensibility achieved. Table 6 presents the results.

It is observable that the model had a worse result when comparing APs to the work of Long et al. (2021) and it achieved a better sensibility when compared to the work of Cano-Espinosa et al. (2020). The amount of data used in each experiment could explain the results. The FUMPE data set has a PE distribution less uniform than CAD-PE. This difference could explain the better results achieved using the CAD-PE data set.

4.2. Classification results

Each candidate from the output of the network was intersected with the ground truth mask. If there was an intersection (candidate mask $\times$ ground truth $\neq 0$), the candidate was considered a true positive. The components without any were marked as false positives. These components were fed to the false-positive reduction phase, applying all stages of preprocessing and feature extraction before the classification training. The classification obtained an accuracy of 83% in the CAD-PE dataset, 74% in FUMPE and 79% HC. The same training, validation and testing distribution of each dataset were used for the false-positive reduction step.

An analysis was made by the size of the components, comparing them to the ground truth (Figure 3). The network had difficulties in finding small components, presenting a high false-positive rate. In some cases, the expected number of true positives of a size interval is greater than the number of actual components, and this occurs because the method considers the components to be true positives if there is an intersection of the candidate with a real component, meaning that there are cases where PEs have been found but not marked in their entirety, resulting in a lower volume prediction. The segmentation model did not produce many false-positive PEs in the private data set. This situation can be occured by a better distribution of larger PEs in this data set.

The data set used was shown to be another factor that directly interferes in the number of false positives found, and it also affects in which volume interval these false positives are found. This behaviour is believed to happen due to the number of elements in each size range. Figure 4 shows the effectiveness of the model in reducing false-positive components in all three data sets, especially the components of smaller volume, but it makes some mistakes while classifying larger components and classifying true components as false positives among the higher volumes. The proposed method also appeared to be unable to detect smaller PEs, and even in data sets with greater representation of them (CAD-PE and private), the model was not able to detect much smaller PEs. The private data set seemed to better fit in the segmentation stage, producing an already low amount of false positives, which explains the poor performance of the data set in the classification phase, where there were not that many false positives to remove.

5. Discussion

Even though some segmentation (Figure 5(a)) achieved good results, the model still needs further optimisations and some hyper-parameters can still be adjusted. The network seems to be able to precisely distinguish PE from within the exams, but it can produce some false negatives (Figure 5(b,c,d)), seeming unable to clots too small or with an odd shape. Some false positives have also been observed in the test set (Figure 5(e,f)), but some techniques, such as 3D representation (Tajbakhsh et al. (2015)) of the input or normalisation of the HU window, can help mitigate this problem (Liu et al. (2020)), these false positives show the need for a second FP removal network specialising in this task.

The results between different data sets were also not that different from each other. The best overall CAD data set results show that the model performance scales with the amount of data provided for training, which has also been confirmed in experiments with a combination of data sets. It was also observed that some small adjustments in the network were needed when training with a small amount of data, especially in the Fumpe data set. The optimiser, which gave the best results, was the Adam. While the SGD did not mark any voxels as true positive, as there is a high amount of true

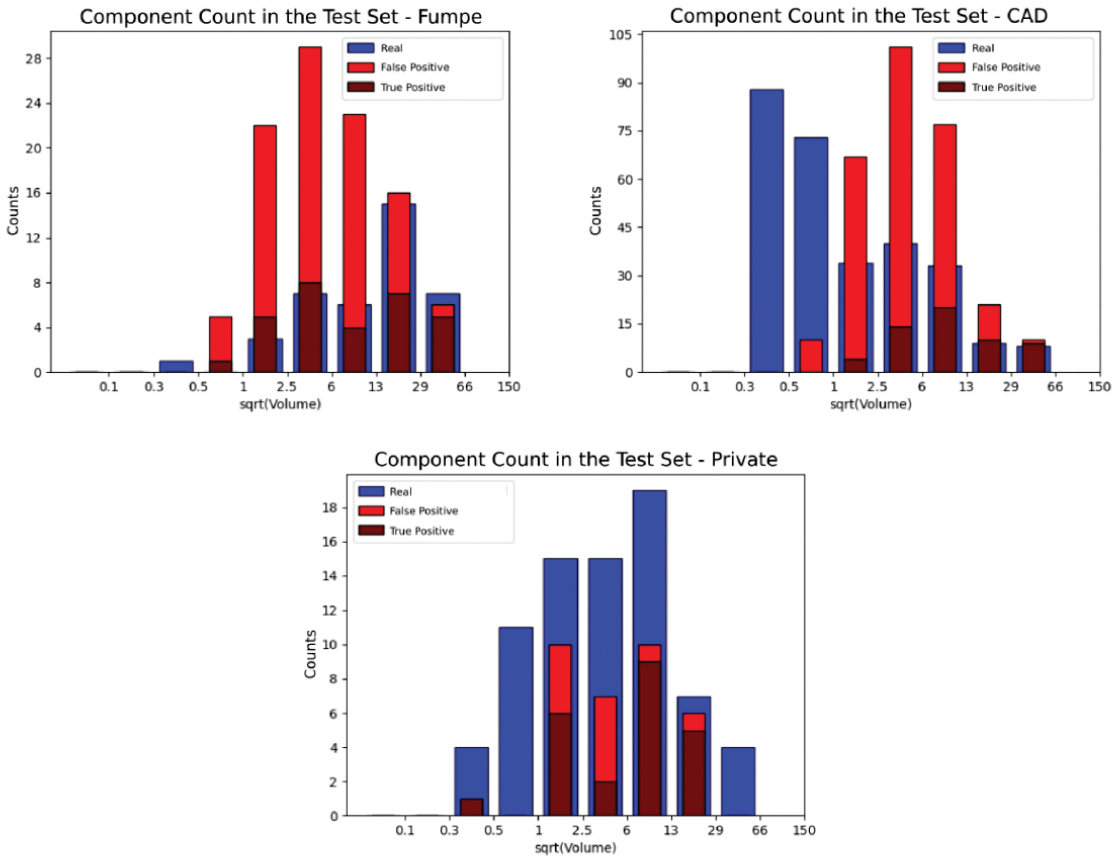


Figure 3. Graphs showing the square root count (*y*-axis) of the volumes (*x*-axis) of the components (PEs) found in the test set of the three data sets before the false-positive classification method step. In blue, the true PEs (i.e. marked by the radiologist); in dark red, the true-positive (TP) PEs segmented by the U-net network; and in bright red, the false-positive PEs produced by the network.

negatives in each mask, training with a 'slower' optimiser such as the SGD can never achieve a good result and be stuck in a local minimum. Although training the network with a larger amount of data has a similar problem, but using the Adam optimiser, the loss drops drastically in the first few epochs, still it does not reach a good result. The model was also easily trained with different data sets and combinations, therefore proving that it can give good results in real-world applications.

6. Conclusion

This paper proposes a new method based on another state-of-the-art PE detection. Although able to find PE candidates and correctly classify them as false positives and true positives, it still has some limitations. The amount of false negatives, or PEs that the U-net network segmentation could not find, especially smaller PEs, shows that the network is not optimised to differentiate between small noise and other thrombus artefacts. Moreover, the number of false positives is still significant, even before the false-positive reduction step.

The size and number of thrombi in the database directly influences the capacity of the model. It is possible to improve the results of the first step of the method with the U-net network using more tests of different hyper-parameters. The network has proven to be able to distinguish PEs within the exams, even though it is produced some false positives. The model also proved to have difficulties in detecting small and more peripheral thrombi.

6.1. Future works

In order to improve the accuracy of the model in finding small components, other hyper-parameters of the network related to convolution and channel reduction can be explored, for example, squeeze-and-excitation blocks as in the work of Huang et al. (2020), performing model training with different combinations of databases can also increase the rate of found components. Models with 3D convolution operations also show an increase in the network segmentation rate as in Cano-Espinosa et al. (2020) and Yang et al. (2019). Other targeting networks can also serve as an alternative, such as Mask *R*-CNN. In the

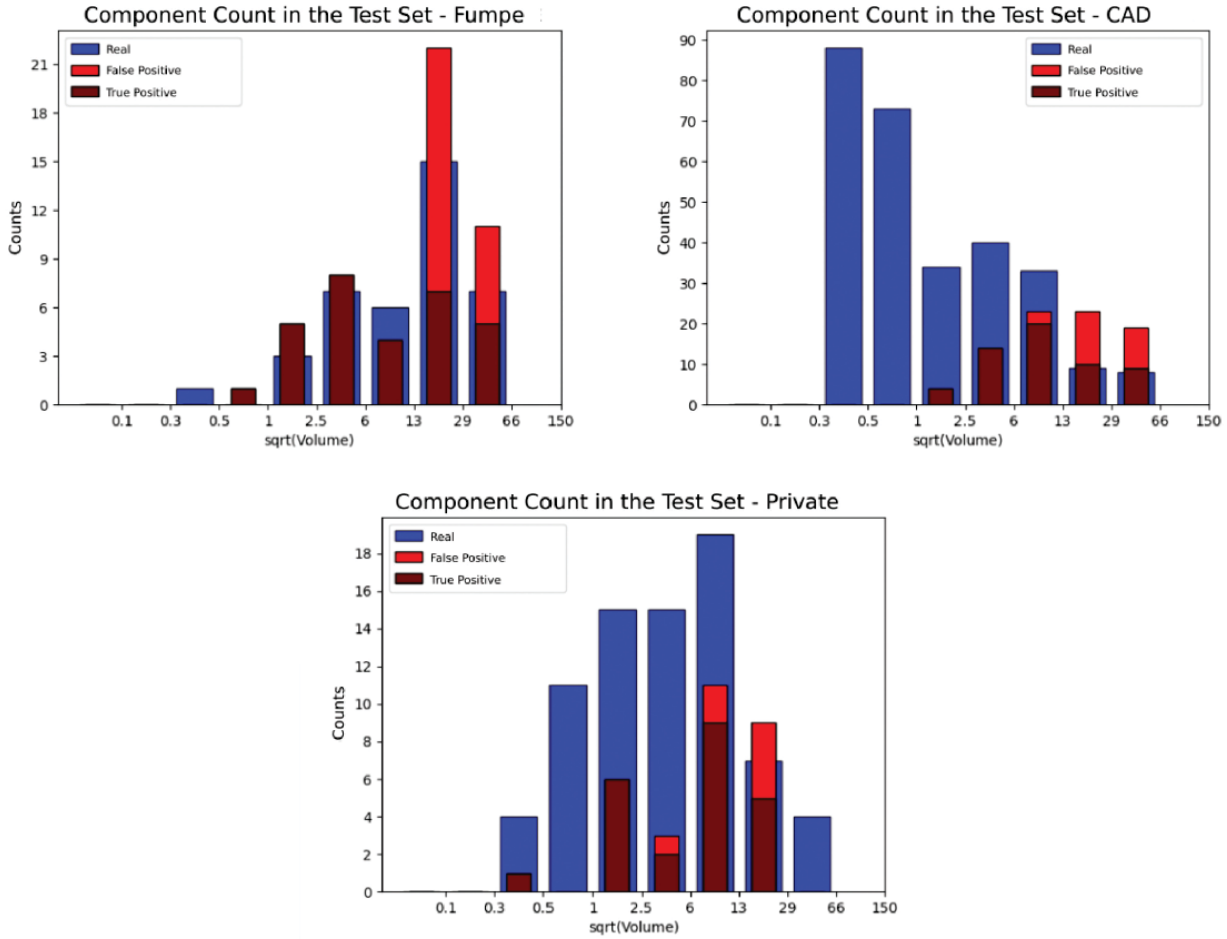


Figure 4. Graphs showing the square root count (*y*-axis) of the volumes (*x*-axis) of the components (PEs) found in the test set of the three data sets after applying the false-positive classification method. In blue, the true PEs (i.e. marked by the radiologist); in dark red, the true-positive (TP) PEs segmented by the U-net network; and in bright red, the false-positive (FP) PEs produced by the network.

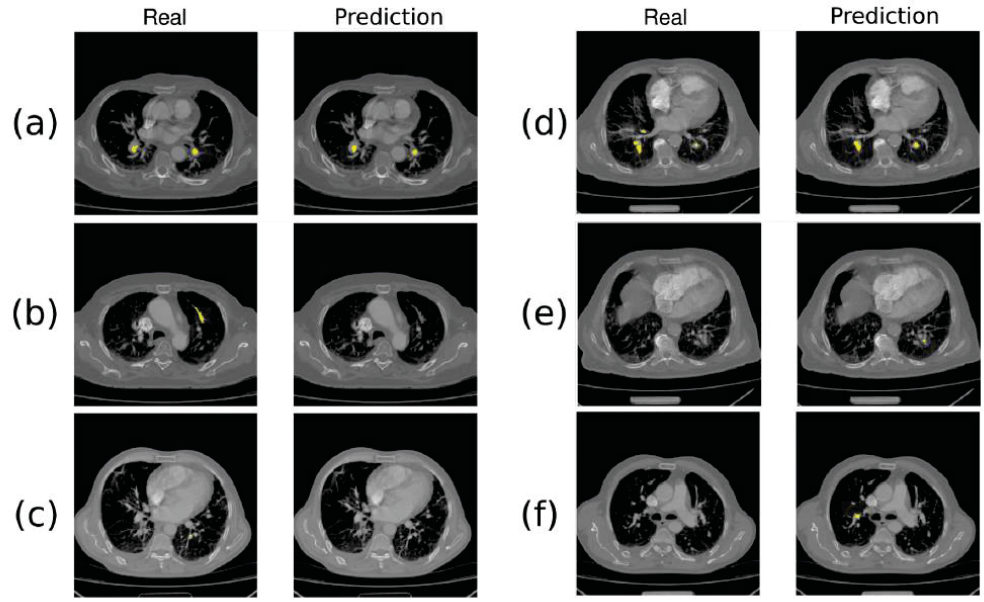


Figure 5. Ground truth and prediction results samples from the test set of Masoudi et al. (2018), in (a) an example of a correctly labelled image, in (b) and (c) two examples where the network did not find PE when there are in the ground truth, in (d) only parts of the image were correctly segmented and in (e) and (f) two examples of a false-positive label produced by the network.

work of Long et al. (2021), the author uses a variation of the *R*-CNN Mask, called P Mask *R*-CNN, which is optimised to finding small objects.

Regarding the number of false positives, approaches such as Tajbakhsh et al. (2015) can increase the success rate of this step. In it, the author produced a single projection of the PE, called 2.5D, using eigenvalues and eigenvectors. An image was generated within two channels: one longitudinal and the other transversal, of each thrombus, also explained in Tajbakhsh et al. (2019). With this representation, it is possible to classify thrombi with other image classification networks, facilitating training with the possibility of using pre-trained weights, e.g. networks such as InceptionNet (Szegedy et al. 2017), ResNext (Xie et al. 2016), EfficientNet (Tan and Le 2019) and AmoebaNet (Real et al. 2019), all with a hit rate above 90% in the Imagenet database (Deng et al. 2009).

Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES). The author LFO would like to thank NVIDIA Corporation for the donation of the Titan XP graphics cards used in experiments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the CAPES.

ORCID

Lucas F. Oliveira http://orcid.org/0000-0002-8198-0877

References

Bouma H, Sonnemans JJ, Vilanova A, Gerritsen FA. 2009. Automatic detection of pulmonary embolism in CTA images. IEEE Trans Med Imaging. 28 (8):1223–1230. doi:10.1109/TMI.2009.2013618.

Cano-Espinosa C, Cazorla M, González G. 2020. Computer aided detection of pulmonary embolism using multi-slice multi-axial segmentation. Appl Sci. 10(8):2945. doi:10.3390/app10082945.

Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L 2009. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops); Los Alamitos, CA, USA; IEEE Computer Society. p. 248–255. doi:10.1109/CVPR.2009.5206848.

Girshick R, Donahue J, Darrell T, Malik J 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Columbus, Ohio; p. 580–587. doi:10.1109/CVPR.2014.81.

Goldhaber SZ, Bounameaux H. 2012. Pulmonary embolism and deep vein thrombosis. Lancet. 379(9828):1835–1846. doi:10.1016/S0140-6736(11)61904-1.

González G, Jimenez-Carretero D, Rodrguez-López S, Cano-Espinosa C, Cazorla M, Agarwal T, Agarwal V, Tajbakhsh N, Gotway MB, Liang J, et al. 2020. Computer aided detection for pulmonary embolism challenge (CAD-PE). arXiv preprint arXiv:200313440.

Huang SC, Kothari T, Banerjee I, Chute C, Ball RL, Borus N, Huang A, Patel BN, Rajpurkar P, Irvin J, et al. 2020. Penet—a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. Npj Digit Med. 3(1):1–9. doi:10.1038/s41746-019-0211-0

Huisman MV, Barco S, Cannegieter SC, Le Gal G, Konstantinides SV, Reitsma PH, Rodger M, Noordegraaf AV, Klok FA. 2018. Pulmonary embolism. Nat Rev Dis Primers. 4(18028). doi:10.1038/nrdp.2018.28.

Jha AK, Larizgoitia I, Audera-Lopez C, Prasopa-Plaizier N, Waters H, Bates DW. 2013. The global burden of unsafe medical care: analytic modelling of observational studies. BMJ Qual Saf. 22(10):809–815. doi:10.1136/bmjqs-2012-001748.

Kligerman SJ, Mitchell JW, Sechrist JW, Meeks AK, Galvin JR, White CS. 2018. Radiologist performance in the detection of pulmonary embolism. J Thorac Imaging. 33(6):350–357. doi:10.1097/RTI.0000000000000361.

Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, Huisman MV, Humbert M, Jennings CS, Jiménez D, et al. 2019. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC). Eur Heart J. 41(4):543–603. Available from. doi:10.1093/eurheartj/ehz405.

Liang J, Bi J 2007. Computer aided detection of pulmonary embolism with tobogganing and mutiple instance classification in CT pulmonary angiography. In: Biennial International Conference on Information Processing in Medical Imaging; Kerkrade, The Netherlands; Springer. p. 630–641.

Liu W, Liu M, Guo X, Zhang P, Zhang L, Zhang R, Kang H, Zhai Z, Tao X, Wan J, et al. 2020. Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. Eur Radiol:1–9. doi:10.1007/s00330-020-06699-8.

Long K, Tang L, Pu X, Ren Y, Zheng M, Gao L, Song C, Han S, Zhou M, Deng F. 2021. Probability-Based mask R-CNN for pulmonary embolism detection. Neurocomputing. 422:345–353. doi:10.1016/j.neucom.2020.10.022.

Masoudi M, Pourreza HR, Saadatmand-Tarzjan M, Eftekhari N, Zargar FS, Rad MP. 2018. A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. Sci Data. 5(1):180180. doi:10.1038/sdata.2018.180.

Özkan H, Osman O, Şahin S, Boz AF. 2014. A novel method for pulmonary embolism detection in CTA images. Comput Methods Programs Biomed. 113(3):757–766. doi:10.1016/j.cmpb.2013.12.014.

Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, et al. 2019. Pytorch: an imperative style, high-performance deep learning library. In: Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox E, and Garnett R, editors. Advances in neural information processing systems 32. Red Hook, NY, USA: Curran Associates, Inc.; p. 8024–8035.

Rajan D, Beymer D, Abedin S, Dehghan E 2020. Pi-Pe: a pipeline for pulmonary embolism detection using sparsely annotated 3d CT images. In: Machine Learning for Health Workshop; Virtual. p. 220–232.

Real E, Aggarwal A, Huang Y, Le QV 2019. Regularized evolution for image classifier architecture search. In: Proceedings of the AAAI conference on artificial intelligence; Honolulu, Hawaii, USA; vol. 33. p. 4780–4789.

Ronneberger O, Fischer P, Brox T 2015. U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention; Munich, Germany; Springer. p. 234–241.

Sadigh G, Kelly AM, Cronin P. 2011. Challenges, controversies, and hot topics in pulmonary embolism imaging. Am J Roentgenol. 196 (3):497–515. doi:10.2214/AJR.10.5830.

Szegedy C, Ioffe S, Vanhoucke V, Alemi AA 2017. Inception-V4, inception-ResNet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence; San Francisco, California, USA; ACM. p. 4278-–4284.

Tajbakhsh N, Gotway MB, Liang J 2015. Computer-Aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention; Munich, Germany; Springer. p. 62–69.

Tajbakhsh N, Shin JY, Gotway MB, Liang J. 2019. Computer-Aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation. Med Image Anal. 58:101541. doi:10.1016/j.media.2019.101541.

Tan M, Le QV. 2019. Efficientnet: rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:190511946.

Xie S, Girshick R, Dollár P, Tu Z, He K. 2016. Aggregated residual transformations for deep neural networks. arXiv preprint arXiv:161105431.

Yang X, Lin Y, Su J, Wang X, Li X, Lin J, Cheng KT. 2019. A two-stage convolutional neural network for pulmonary embolism detection from CTPA images. IEEE Access. 7:84849–84857.

6 DISCUSSÃO

Na revisão da literatura realizada, encontramos apenas duas bases de dados publicas com exames de angioTC de tórax de pacientes com TEP. Uma delas derivada de uma competição de programas para o diagnóstico de TEP, denominada CAD-PE (GONZALES, 2020), contêm exames coletados em seis hospitais de Madrid e região. A outra, desenvolvida na Ferdowsi Universidade de Mashad, no Irã, denominada FUMPE, contêm exames de dois centros médicos (MASOUDI, M. et al., 2018). Apesar de haver pontos em comum, notam-se algumas diferenças entre estas e a desenvolvida.

Na elaboração da base de dados buscamos trazer uma amostra realista dos exames feitos em um hospital terciário e em uma clínica de imagem. Para isto, fomos rigorosos com os critérios de inclusão e exclusão pré-estabelecidos e, após, definir de forma arbitrária a data inicial do período de coleta, que foi retrospectiva, selecionamos todos os exames sucessivos que se enquadraram nos critérios. A definição da metodologia buscou evitar ao máximo vieses de seleção. As informações publicadas das demais bases de dados não detalham o processo de seleção dos exames. As informações públicas apenas revelam que na base CAD-PE a coleta de exames também foi retrógrada e, na base FUMPE, foram priorizados exames com trombos periféricos. A justificativa dos autores para esta priorização é a maior dificuldade diagnóstica para os CADs oferecida por trombos nesta topografia.

Quanto ao tamanho das bases de dados, a com maior número de exames é a CAD-PE, contendo 91 exames com TEP, totalizando 24.624 imagens, com as dimensões dos voxels variando entre 0.52 e 0.92 mm. A base FUMPE, contém 35 exames, sendo 33 com TEP, totalizando 8.792 imagens, com as dimensões dos voxels entre 0.52 e 0.78 mm. A base de dados desenvolvida contém 40 exames com TEP, totalizando 14.463 imagens, com as dimensões dos voxels entre 0.57 e 1.0 mm. Apesar do número de exames disponíveis nos arquivos digitais HC-UFPR e da clínica DAPI serem muito superior aos 40 utilizados, os limitantes do tamanho da base foram o tempo necessário para marcar cada exame e o cronograma do projeto. Somando as três bases, após a publicação da nova base, os desenvolvedores terão disponíveis 164 exames com marcação semântica, ao nível do pixel, para utilizar no desenvolvimento de novas soluções.

Outro ponto complementar das bases é a variação dos aparelhos utilizados para obtenção as imagens. Recentemente a revista Radiology, da Sociedade Norte América de Radiologia (RSNA), publicou um guia para produção de trabalhos com inteligência artificial (BLUEMKE et al., 2020), no qual é sugerido que os programas sejam desenvolvidos utilizando imagens de diferentes fabricantes. A CAD-PE contém exames realizados com tomógrafos SIEMENS Somaton Sensation, 40 canais. A FUMPE contém exames realizados com tomógrafos da PHILIPS e Neusoft Medical System Co., 16 canais. Buscando aumentar a variabilidade de exames disponíveis, a nova base de dados contém exames realizados em aparelhos diferentes das bases já publicadas, sendo 20 exames obtidos com o aparelho Toshiba Aquilion de 64 canais (HC-UFPR) e 20 no GE Revolution de 512 canais (clínica DAPI).

Quanto ao método de marcação de exames utilizado, as bases diferem. As bases CAD-PE e FUMPE utilizaram métodos semiautomatizados para a marcação de exames. Na base CAD-PE, após a marcação semiautomatizada, a revisão dos exames foi manual. A base desenvolvida foi marcada com o método manual, por escolha dos responsáveis pela marcação. Esta escolha pode ser vista como uma possível limitação, devido a eventual variação entre os marcadores. Buscando reduzir esta variabilidade, todos os exames foram marcados por médicos com igual grau de treinamento na radiologia e revisados por um único radiologista.

Ainda em relação à marcação dos exames, outro ponto de grande importância é o grau de conhecimento da equipe envolvida na marcação dos exames para o reconhecimento correto desta patologia, uma vez que, é esta habilidade que se busca reproduzir por um programa computacional. A base CAD-PE, foi marcada por diferentes radiologistas, sendo os primeiros 40 exames marcados por radiologistas com pelo menos 15 anos de experiência e os demais exames marcados por um único radiologista habilitado, cuja experiência não foi divulgada. Na base FUMPE a marcação foi realizada por um radiologista com 5 anos de experiência em radiologia torácica e revisada por outro com 18 anos de experiência. Na nova base, todos os exames foram inicialmente avaliados por radiologistas certificados, com grau de experiência variável, e classificados como contendo TEP, que foi confirmado por um radiologista torácico com 32 anos de experiência. Após esta etapa, as marcações foram realizadas por médicos residentes do terceiro ano de sua formação e revisadas e realizadas eventuais

correções por um único médico radiologista certificado, com um ano de experiência, visando padronizar as marcações. A utilização de mais de um marcador, como realizado na CAD-PE e na nova base, pode gerar uma variação nos critérios de marcação, pelo potencial subjetividade de cada marcador. Na nova base, a revisão de todas as marcações por um único radiologista buscou reduzir esta variação.

Assim como a base FUMPE, a nova base apresenta análises adicionais de alterações por imagem relacionadas a repercussões cardiovasculares do TEP que podem ter impacto no manejo clínico dos pacientes.

Quanto aos arquivos digitais referentes às marcações ao nível do pixel e aos exames, a nova base segue o mesmo formato das duas bases anteriores, com um arquivo correspondendo ao gabarito, em que os pixels representando o TEP foram marcados. Esta escolha objetivou facilitar o uso conjunto da nova base e das duas outras bases públicas no programa desenvolvido pelo DINF- UFPR (OLESCKI et al, 2022) e por demais grupos de pesquisa que venham a utilizá-la após sua publicação.

7 CONCLUSÃO

1- Com este projeto foi desenvolvida uma base de dados com uma amostra realista de exames de angiotomografia de tórax com as marcações necessárias para o desenvolvimento de programas de aprendizado de máquinas.

2- A base foi utilizada primeiramente no desenvolvimento de um programa computadorizado para o diagnóstico de TEP pelo DINF- UFPR e, em breve, será tornada pública para que outros grupos possam utilizá-la no desenvolvimento de seus projetos.

REFERÊNCIAS

WHITE, R. H. The epidemiology of venous thromboembolism. **Circulation**, v. 23, p.107, 2003. https://doi.org/10.1161/01.CIR.0000078468.11849.66

TERRA-FILHO, M. Diretrizes da SBPT: Recomendações para o manejo da tromboembolia pulmonar, **J. bras. pneumol.** v. 36, n. 55, p. 1–3, 2010. https://doi.org/10.1590/S1806-37132010001300001

DALEN, J. E. Pulmonary Embolism: What Have We Learned Since Virchow?, **CHEST**,v.122, p.1440-1456, 2016. https://doi.org/10.1378/chest.122.4.1440

ROY, P.; MEYER, G. Annals of Internal Medicine Appropriateness of Diagnostic Management and Outcomes of suspected pulmonary embolism, **Ann. Intern. Med.**, p.144-157, 2006. https://doi.org/10.7326/0003-4819-144-3-200602070-00003

KONSTANTINIDES, S.V.; TORBICKI, AGNELLI, G. Task Force for the Diagnosis and Management of Acute Pulmonary Embolism of the European Society of Cardiology (ESC). ESC Guidelines on the diagnosis and management of acute pulmonary embolism. **Eur. Heart J.**, v.35, p.3033–3069, 2014. https://doi.org/10.1093/eurheartj/ehn310

REMY-JARDIN, M., PISTOLESI, M., GOODMAN, L. R. Management of suspected acute pulmonary embolism in the era of CT angiography: a statement from the Fleischner Society. **Radiology***,* v.245, p.315–329, 2007. https://doi.org/10.1148/radiol.2452070397

CONRAD, W. et al, CT Angiography of Pulmonary Embolism: Diagnostic Criteria and Causes of Misdiagnosis. **RadioGraphics**, v.24, p.219–1238, 2004. https://doi.org/10.1148/rg.245045008

MAIZLIN, Z.V. et al. Computer-aided detection of pulmonary embolism on CT angiography: initial experience. **J. Thorac. Imaging**, v.22, n.4, p. 324–329, 2007. https://doi.org/10.1097/RTI.0b013e31815b89ca

CHOY, G. et al. Current applications and future impact of machine learning in radiology. **Radiology**, v. 288, p.318–328, 2018. https://doi.org/10.1148/radiol.2018171820

REMY-JARDIN, M. et al. Machine Learning and Deep Neural Network Applications in the Thorax, **J Thorac Imaging**. v.35, p 40-48, 2020. https://doi.org/10.1097/RTI.0000000000000492

COLAK, E., et al. The RSNA Pulmonary Embolism CT Dataset. **Radiology: Artificial Intelligence**, v. 3, n. 2., 2021. https://doi.org/10.1148/ryai.2021200254

GONZÁLEZ, G. et al. Computer Aided Detection for Pulmonary Embolism Challenge (CAD-PE). Disponível em: < https://arxiv.org/abs/2003.13440>. Acesso em: 30 mar., 2020.

MASOUDI, M.et al. A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. **Nature Sci. Data**, v. 5, 2018. https://doi.org/10.1038/sdata.2018.180

COLAK, E. et al. The RSNA Pulmonary Embolism CT Dataset. Colak et al. Radiology: Artificial Intelligence. v.3, 2021. https://doi.org/10.1148/ryai.2021200254

BLUEMKE D. A. et al. Assessing Radiology Research on Artificial Intelligence: A Brief Guide for Authors, Reviewers, and Readers. **Radiology**, v. 294, p. 487-489, 2019. https://doi.org/10.1148/radiol.2019192515

OLESCKI, G. et al. A two step workflow for pulmonary embolism detection using deep learning and feature extraction, **Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization**, 2022. https://doi.org/10.1080/21681163.2022.2060866

HUISMAN, M. et al. Pulmonary embolism. **Nat Rev Dis Primers**, v.4, 2018. https://doi.org/10.1038/nrdp.2018.28

BECKMAN, M. G. et al. Venous thromboembolism: a public health concern. **Am J Prev Med.**, v. 4, p. 495-501, 2010. doi: 10.1016/j.amepre.2009.12.017

RASKOB, G.E. et al. Thrombosis: a major contributor to global disease burden. **Arterioscler Thromb Vasc Biol.**, v.34, p. 2363–2371, 2014. https://doi.org/10.1161/ATVBAHA.114.304488

KONSTANTINIDES, S. V. et al. ESC Scientific Document Group, 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS): The Task Force for the diagnosis and management of acute pulmonary embolism of the European Society of Cardiology (ESC), **European Heart Journal**, v. 41, p. 543-603, 2020. https://doi.org/10.1093/eurheartj/ehz405

HOUNSFIELD, G. N. Computerized transverse axial scanning (tomography). 1. Description of system. **The British Journal of Radiology**, v.46, p.1016-1022. https://doi.org/10.1259/0007-1285-46-552-1016

SEERAM, E. Computed Tomography: A Technical Review. **Radiol Technol.**, jan., 2018. PMID: 29298954.

MORITZ, H. A. et al. State-of-the-Art Pulmonary CT Angiography for Acute Pulmonary Embolism. **American Journal of Roentgenology**, v.3, p.495-504, 2017. https://doi.org/10.2214/AJR.16.17202

MASUTANI, Y., MACMAHON, H. & DOI, K. Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis. **IEEE Trans Med Imaging**, v. 21, p. 1517–1523, 2002. https://doi.org/10.1109/TMI.2002.806586

PICHON, E., et al. A novel method for pulmonary emboli visualization from high-resolution CT images. **Proceedings of the SPIE**, v. 5367, p. 161-170, 2004. https://doi.org/10.1117/12.532892

ZHOU, C. et al. Preliminary investigation of computer-aided detection of pulmonary embolism in three-dimensional computed tomography pulmonary angiography images. **Acad. Radio**l., v.12, p.782–792., 2005. https://doi.org/10.1016/j.acra.2005.01.014

BUHMANN, S. et al. Clinical evaluation of a computer-aided diagnosis (CAD) prototype for the detection of pulmonary embolism. Acad. **Radiol.**, v.14, p. 651–658, 2007. https://doi.org/10.1016/j.acra.2007.02.007

SCHOEPF, U. J., et al. Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography. **J Thorac Imagin**, v.22, p.319–323, 2007. https://doi.org/10.1097/RTI.0b013e31815842a9

LIANG, J. & BI, J. Computer Aided Detection of Pulmonary Embolism with Tobogganing and Multiple Instance Classification in CT Pulmonary Angiography, Information **Processing in Medical Imaging**, v. 20, p. 630-641, 2007. https://doi.org/10.1007/978-3-540-73273-0_52

SEBBE, R. Computer-aided Diagnosis of Pulmonary Embolism in Opacified CT Images, dissertation, **Faculte Polytechnique de Mons**, 2006. Disponível em: <http://pul.uclouvain.be/resources/titles/29303100568200/extras/77016similatheseintv2_1005261.pdf >. Acesso em: 18 jul., 2022.

ENGELKE, C. et al. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. **Eur. Radiol**, v.18, p. 298–307, 2008. https://doi.org/10.1007/s00330-007-0770-3

BOUMA, H. Vessel-Diameter Quantification and Embolus Detection in CTA Images, Phd Thesis, **Technische Universiteit Eindhoven**, 2008. Disponível em: <https://research.tue.nl/nl/publications/vessel-diameter-quantification-and-embolus-detection-in-cta-image >. Acesso em: 18 jul., 2022.

DAS, M., et al. Computer-aided detection of pulmonary embolism: influence on radiologists detection performance with respect to vessel segments. **Eur. Radiol.**, v. 18, p. 1350–1355, 2008. https://doi.org/10.1007/s00330-008-0889-x

WALSHAM, A. C., et al. The use of computer-aided detection for the assessment of pulmonary arterial filling defects at computed tomographic angiography. **J. Comput. Assist. Tomogr.**, v. 32, p. 913–918., 2008. https://doi.org/10.1097/RCT.0b013e31815b3ed0

BOUMA, H. et al. Automatic Detection of Pulmonary Embolism in CTA Images, **IEEE Transactions on Medical Imaging**, v.28, p. 1223-1230, 2009. https://doi.org/10.1109/TMI.2009.2013618

DINESH, M. S., et al. Adaptive contrast-based computer aided detection for pulmonary embolism, **Proc. SPIE 7260, Medical Imaging 2009: Computer-Aided Diagnosis**. https://doi.org/10.1117/12.812223

ZHOU, C., et al. Computer-aided detection of pul0monary embolism in computed tomographic pulmonary angiography (CTPA): performance evaluation with independent data sets. **Med Phys**, v.36, p. 3385–3396, 2009. https://doi.org/10.1118/1.3157102

WITTENBERG, R., et al. Computer-assisted detection of pulmonary embolism: evaluation of pulmonary CT angiograms performed in an on-call setting**. Eur Radiol.**, v. 20, p. 801–806, 2010. https://doi.org/10.1007/s00330-009-1628-7

DEWAILLY, M. et al. Computer-aided detection of acute pulmonary embolism with 64-slice multidetector row computed tomography: impact of the scanning conditions and overall image quality in the detection of peripheral clots. **J Comput Assist Tomogr.**, v. 34, p. 23–30, 2010. https://doi.org/10.1097/RCT.0b013e3181b2e383

LEE, C. W., et al. Evaluation of computer aided detection and dual energy software in detection of peripheral pulmonary embolism on dual-energy pulmonary CT angiography. **Eur Radiol.**, v. 21, p.54–62, 2011. https://doi.org/10.1007/s00330-010-1903-7

PARK, S.C., CHAPMAN, B.E. & BIN, Z.B. A multistage approach to improve performance of computer-aided detection of pulmonary embolisms depicted on CT images: preliminary investigation. **IEEE Trans Biomed Eng.**, v.58, p. 1519-1527, 2011. https://doi.org/10.1109/TBME.2010.2063702

BLACKMON, K. N., et al. Computer-aided detection of pulmonary embolism at CT pulmonary angiography: can it improve performance of inexperienced readers? **Eur Radiol.**, v. 21, p.1214–1223, 2011. https://doi.org/10.1007/s00330-010-2050-x

WITTENBERG, R., et al. Impact of image quality on the performance of computer-aided detection of pulmonary embolism. **Am. J. Roentgenol.**, v.196, p. 95-101, 2011. https://doi.org/10.2214/AJR.09.4165

WITTENBERG, R. et al. Acute pulmonary embolism: effect of a computer-assisted detection prototype on diagnosis—an observer study. **Radiology**, v.262, p.305–313, 2012. https:// doi.org/10.1148/radiol.11110372

KLIGERMAN, S. J., et al. Missed pulmonary emboli on CT angiography: assessment with pulmonary embolism–computer-aided detection. **Am. J. Roentgenol.**, v. 202, p. 65–73, 2014. https://doi.org/10.2214/AJR.13.11049

ÖZKAN H., et al. A novel method for pulmonary embolism detection in CTA images. **Comput Methods Programs Biomed.**, v.113, p. 757-766, 2014. https://doi.org/10.1016/j.cmpb.2013.12.014

LAHIJI, K. et al. Improved accuracy of pulmonary embolism computer-aided detection using iterative reconstruction compared with filtered back projection. **Am J Roentgenol**, v.203, p.763–771, 2014. https://doi.org/10.2214/AJR.13.11838

TAJBAKHSH N., GOTWAY M.B. & LIANG J. Computer-Aided Pulmonary Embolism Detection Using a Novel Vessel-Aligned Multi-planar Image Representation and Convolutional Neural Networks. Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015. **MICCAI 2015. Lecture Notes in Computer Science**, v. 9350, p.62-69, 2015. https://doi.org/10.1007/978-3-319-24571-3_8

OZKAN, H, et al. Automatic Detection of Pulmonary Embolism in CTA Images Using Machine Learning. **Elektronika Ir Elektrotechnika**, v.23, p. 63-67,2017. https://doi.org/10.5755/j01.eie.23.1.17585

TAJBAKHSH, N., et al. Computer-aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation. **Med Image Anal.**, v. 58, p. 101541, 2019. https://doi.org/10.1016/j.media.2019.101541

RAJAN, D., et al. Pi-PE: A pipeline for pulmonary embolism detection using sparsely annotated 3D CT images. **Electrical Engineering and Systems Science**, 2019. Disponível em: < https://arxiv.org/pdf/1910.02175v3.pdf >. Acesso em: 18 jul., 2022.

YANG, X. et al., A Two-Stage Convolutional Neural Network for Pulmonary Embolism Detection from CTPA Images, **IEEE Access**, v.7, p. 84849-84857, 2019. https://doi.org/10.1109/ACCESS.2019.2925210

CANO-ESPINOSA, C., CAZORLA, M. & GONZÁLEZ, G. Computer Aided Detection of Pulmonary Embolism Using Multi-Slice Multi-Axial Segmentation. **Appl Sci.**, v.10, p. 2945, 2020. https://doi.org/10.3390/app10082945

HUANG, S. C., et al. Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection. **Sci Rep**, v.10, p.22147, 2020. https//doi.org/10.1038/s41598-020-78888-w

YU, C. -Y., CHENG, Y. -C. & KUO, C. Early Pulmonary Embolism Detection from Computed Tomography Pulmonary Angiography Using Convolutional Neural Networks, 2020 Joint 9th International Conference on Informatics, **Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)**, 2020. https://doi.org/10.1109/ICIEVicIVPR48672.2020.9306659

HUANG, S., et al. PENet—a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. **npj Digit. Med.**, v.3, p.61, 2020. https://doi.org/10.1038/s41746-020-0266-y

LIU, W., et al. Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. **Eur Radiol.**, v.30, p. 3567-3575, 2020. https://doi.org/10.1007/s00330-020-06699-8

WEIKERT, T., et al. Automated detection of pulmonary embolism in CT pulmonary angiograms using an AI-powered algorithm. **Eur. Radiol.**, v.30, p.6545–6553, 2020. https://doi.org/10.1007/s00330-020-06998-0

VAINIO, T., et al. Performance of a 3D convolutional neural network in the detection of hypoperfusion at CT pulmonary angiography in patients with chronic pulmonary embolism: a feasibility study. **Eur Radiol Exp**, v.5, 2021. https://doi.org/10.1186/s41747-021-00235-z

LI, X., et al. Preliminary study on artificial intelligence diagnosis of pulmonary embolism based on computer in-depth study. **Annals of translational medicine**, v.9, p.838, 2021. https://doi.org/10.21037/atm-21-975

LONG, K. et al. Probability-based Mask R-CNN for pulmonary embolism detection, **Neurocomputing**, v. 422, p. 345-353, 2021 https://doi.org/10.1016/j.neucom.2020.10.022

HUHTANEN, H., et al. Automated detection of pulmonary embolism from CT-angiograms using deep learning. **BMC Med Imaging**, v.22, p.43, 2022. https://doi.org/10.1186/s12880-022-00763-z

YADLAPALLI, P., et al. Segmentation of Pulmonary Embolism Using Deep Learning, **2022 International Conference for Advancement in Technology (ICONAT)**, 2022. https://doi.org/10.1109/ICONAT53423.2022.9726048

CHARTRAND, G. et al. Deep learning: A primer for radiologists. **RadioGraphics**, v.37, p. 2113-2131, 2017. https://doi.org/10.1148/rg.2017170077

YAO, A., D.et al. Deep Learning in Neuroradiology: A Systematic Review of Current Algorithms and Approaches for the New Wave of Imaging Technology. **Radiology: Artificial Intelligence**, v.2, 2020.https://doi.org/10.1148/ryai.2020190026

MURPHY, K. P. Machine learning: a probabilistic perspective. Cambridge: The MIT Press, 2012.

LEE, J.G. et al. Deep Learning in Medical Imaging: General Overview. **Korean J Radiol**, v.4, p.570-584, 2017. https://doi.org/10.3348/kjr.2017.18.4.570

BOESCHIMAGE, G. ANNOTATION: NEW TECHNOLOGIES AND TOOLS IN 2022, **VISO.AI, 2022**. Disponível em: < https://viso.ai/computer-vision/image-annotation/>. Acesso em: 26 jul., 2022.

ZHANG, Y. Semantic-Based Visual Information Retrieval, London, **IRM Press**, 2006.

POKHREL, S. Exploring Data Labeling and the 6 Different Types of Image Annotation, **Xailient**, 2020. Disponível em: <https://xailient.com/blog/exploring-data-labeling-and-the-6-different-types-of-image-annotation/>. Acesso em: 26 jul., 2022.

PAIVA, O. A., PREVEDELLO, L. M. The potential impact of artificial intelligence in radiology. **Radiol Bras**, v.50, p.5-6, 2017. https://doi.org/10.1590/0100-3984.2017.50.5e1

GROOT, P., BIRLUTIU, A., HESKES, T. Learning from multiple annotators with gaussian processes. **Proceedings ICANN**, 2011. https://doi.org/10.1007/978-3-642-21738-8_21

LEITÃO, C. A. Análise dos efeitos da redução da dose sobre a detecção de nódulos pulmonares em tomografias de tórax de baixa e de ultrabaixa dose para rastreamento de câncer de pulmão (um estudo em fantoma). **Dissertação (Mestrado em Medicina Interna– Programa de pós graduação em Medicina Interna e Ciências da Saúde, Universidade Federal do Paraná, Curitiba (PR), 2020**.). Disponível em: < https://acervodigital.ufpr.br/handle/1884/70714 >. Acesso em: 21 jun., 2022.

YUSHKEVICH, P. A. et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability**. Neuroimage**, v.31, p.1116-28, 2006. https://doi.org/10.1016/j.neuroimage.2006.01.01