



UNIVERSIDADE D  
COIMBRA

FACULDADE  
DE  
MEDICINA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

CARLOS ALBERTO MARQUES GOMES

***Inteligência Artificial em imagem médica: amiga ou  
adversária?***

ARTIGO DE REVISÃO

ÁREA CIENTÍFICA DE IMAGIOLOGIA

Trabalho realizado sob a orientação de:  
LUÍS MIGUEL CATARINO CURVO SEMEDO

FEVEREIRO/2020

# INTELIGÊNCIA ARTIFICIAL EM IMAGEM MÉDICA: AMIGA OU ADVERSÁRIA?

Artigo de revisão

Trabalho final do 6º ano do Mestrado Integrado em Medicina, com vista à  
atribuição do grau de Mestre em Medicina

## Aluno

Carlos Alberto Marques Gomes

Aluno do 6º ano do Mestrado Integrado em Medicina

Faculdade de Medicina, Universidade de Coimbra, Coimbra, Portugal

carlos.alberto.gomes@outlook.pt

## Orientador

Prof. Doutor Luís Miguel Catarino Curvo Semedo

Professor Auxiliar, Regente da Unidade Curricular Opcional de Ecografia Clínica,

Faculdade de Medicina, Universidade de Coimbra, Coimbra, Portugal – Azinhaga de  
Santa Comba, 3000-354, Coimbra

lmcsemedo@fmed.uc.pt

## ÍNDICE

---

<b>RESUMO</b> .....	<b>3</b>
<b>ABSTRACT</b> .....	<b>5</b>
<b>LISTA DE ABREVIATURAS</b> .....	<b>7</b>
<b>INTRODUÇÃO</b> .....	<b>8</b>
<b>MATERIAL E MÉTODOS</b> .....	<b>11</b>
<b>RESULTADOS</b> .....	<b>12</b>
1. MACHINE LEARNING E DEEP LEARNING: SÍNTESE DE BASES E CONCEITOS .....	12
2. IA NA RADIOLOGIA: APLICAÇÕES ATUAIS E FUTURAS .....	16
2.1 <i>Sistemas CAD</i> .....	17
2.2 <i>Imagem torácica</i> .....	17
2.3 <i>Senologia</i> .....	20
2.4 <i>Neuroimagem</i> .....	22
2.5 <i>Triagem de emergências</i> .....	24
2.6 <i>Outras aplicações, radiômica e NLP</i> .....	25
3. LIMITAÇÕES, DESAFIOS E PERSPETIVAS PARA A IA NA RADIOLOGIA .....	28
3.1 <i>Limitações técnicas</i> .....	28
3.2 <i>Dados: treino e validação dos algoritmos</i> .....	28
3.3 <i>Segurança e confiança</i> .....	31
3.4 <i>Legislação, ética e validação clínica</i> .....	32
<b>DISCUSSÃO E CONCLUSÃO</b> .....	<b>36</b>
<b>AGRADECIMENTOS</b> .....	<b>41</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>42</b>

## RESUMO

---

A inteligência artificial (IA) é um ramo da ciência computorizada, que surgiu há cerca de 50 anos e que tem vindo a sofrer importantes desenvolvimentos nas últimas duas décadas. A sua aplicação na Medicina tem ganho notável protagonismo em diversas áreas, de que destacamos a da imagem médica. A base do funcionamento da IA assenta no reconhecimento de padrões por parte de sistemas informáticos que, através de uma leitura prévia de um elevado volume de dados que lhes são fornecidos, recorrem a algoritmos matemáticos para cumprirem a sua tarefa. Os termos *machine learning* (ML) e *deep learning* (DL) referem-se a dois subtipos de IA com potencial aplicação na área da radiologia. Deste modo, a IA prevê e permite o diagnóstico de diversas patologias e entidades clínicas, complementando os métodos de imagem 'convencionais'. Atualmente, o tema concentra elevada controvérsia uma vez que há quem considere que o futuro da imagem médica poderá estar indissociado da aplicação de sistemas de IA, os quais poderão substituir os médicos radiologistas.

Este artigo de revisão pretende efetuar uma breve descrição dos fundamentos teóricos associados ao conceito de IA, reconhecer e descrever as suas aplicações na área da imagem médica e analisar os benefícios e inconvenientes dos sistemas de IA na radiologia, tendo em conta os limites associados à sua implementação clínica.

Foi efetuada uma pesquisa bibliográfica na base de dados *PubMed*, foram selecionados artigos que cumprissem os objetivos propostos e procedeu-se à sua revisão e organização de acordo com os subtemas considerados pertinentes.

A era dos *big data* e a investigação crescente da IA aplicada à imagem médica têm permitido demonstrar o seu elevado potencial nos vários subcampos da radiologia (detecção, segmentação, classificação, quantificação), em diversas patologias com destaque para as torácicas e pulmonares, mamárias e neuropsiquiátricas. Para além do seu uso na interpretação, os sistemas de IA têm sido alvo de intenso estudo para a melhoria da aquisição de imagem, para a organização de listas de trabalho, e para a análise radiómica e radiogenómica.

Apesar dos obstáculos relacionados com o acesso a grandes bases de dados e com as implicações técnicas, de regulação e ético-legais, esta tecnologia tem todas as ferramentas para que possa ser muito útil na prática clínica. Contudo, a capacidade para estes sistemas ocuparem o lugar de um radiologista experiente na interpretação de exames imagiológicos e nas restantes funções de um médico é algo que poderá estar

ainda aquém das expectativas dos mais entusiastas, que creem nas máquinas como substitutas dos radiologistas.

**Palavras-chave:** Inteligência artificial, imagem médica, diagnóstico assistido por computador, *machine learning*, *deep learning*.

## ABSTRACT

---

Artificial intelligence (AI) is a branch of computer science, which has emerged about 50 years ago and has undergone important developments over the past two decades. Its application in medicine has gained a notable prominence in several areas, including medical imaging. The functioning of AI is based on pattern recognition by computer systems that use mathematical algorithms to perform their task through a previous reading of a large quantity of data provided to them. Machine learning (ML) and deep learning (DL) are subtypes of AI with potential application in the field of radiology. Thus, AI predicts and allows the diagnosis of several pathologies and clinical entities, complementing the 'conventional' imaging methods. Currently, the subject is highly controversial since some consider the future of medical imaging to be inseparable from the application of AI systems, which may replace radiologists, while others believe AI can never suffice without radiologists.

This review intends to provide a brief description of the theoretical foundations associated with the concept of AI, to recognize and describe its applications in the area of medical imaging and to analyze the benefits and drawbacks of AI systems in radiology, taking into account the limits associated with their clinical implementation.

A bibliographic research was performed on the PubMed database and the articles that got selected met the proposed objectives and were reviewed and organized according to the sub-themes considered relevant.

The big data era and the growing investigation of AI applied to medical imaging have demonstrated great value in the various subfields of radiology (detection, segmentation, classification, quantification) and, particularly, in thoracic and pulmonary, breast and neuropsychiatric diseases. In addition to their use in image interpretation, AI systems have been widely studied to improve image acquisition, organization of work lists and radiomic and radiogenomic analysis.

In spite of the obstacles to access large databases and the technical, regulatory and ethical-legal implications, this technology has potential to be useful in clinical practice. However, the possibility of this systems taking the place of a radiologist, trained in imaging interpretation and in the remaining functions of a physician, still remains unattainable, contradicting the most enthusiastic, who believe in machines as replacers of radiologists.

**Keywords:** Artificial intelligence, medical imaging, computer assisted diagnosis, machine learning, deep learning.

## LISTA DE ABREVIATURAS

---

**ANN** – *Artificial neural network*

**AVC** – Acidente vascular cerebral

**BI-RADS** – *Breast Imaging Reporting and Data System*

**CAD** – Diagnóstico assistido por computador

**CADe** – *Computer-aided detection*

**CADx** – *Computer-aided diagnosis*

**CDR** – *Cancer detection rate*

**CNN** – *Convolutional neural networks*

**COR** – Característica de operação do recetor

**DA** – Doença de Alzheimer

**DL** – *Deep learning*

**DPI** – Doenças pulmonares intersticiais

**EAM** – Enfarte agudo do miocárdio

**EM** – Esclerose múltipla

**EMR** – Registos médicos eletrónicos

**FDA** – *Food and Drug Administration*

**GPU** – Unidade de processamento gráfico

**IA** – Inteligência artificial

**ML** – *Machine learning*

**NLP** – *Natural language processing*

**PACS** – *Picture archiving and communication systems*

**PEA** – Perturbação do espectro do autismo

**PMA** – *Premarket approval*

**RM** – Ressonância magnética

**TC** – Tomografia computadorizada

**VA** – Vias aéreas



## INTRODUÇÃO

---

Ao longo de toda a História que o Homem se debate constantemente com a busca da sua evolução e, sobretudo, da do mundo em que vive. No que concerne à ciência, atualmente, há quem considere a inteligência artificial (AI) o *ex-libris* dessa longa e incessante procura.

Embora o termo, IA, tenha sido descrito há cerca de 60 anos atrás, se recuarmos ao século XVIII, altura em que Jacques de Vaucanson inventou o *The Flute Player* com um reportório de 12 canções, verificamos que a ideia de conceber algo mecânico e realista que, de certo modo, pudesse mimetizar as ações humanas, vem-se desenvolvendo há mais de duas centenas de anos. Na primeira metade do século XX, quando a palavra *robot* foi introduzida na literatura, deu-se início a grandes pesquisas e avanços tecnológicos, levando a que, associado ao desenvolvimento dos robôs, surgisse o conceito de IA. Em 1948, William Gray Water criou o primeiro robô autónomo e eletrónico, *Machina Speculatrix*, permitindo que se percebesse que a conexão entre um pequeno número de neurónios podia originar diferentes comportamentos. Assim, em 1955, John McCarthy descreveu a inteligência artificial como “a ciência e a engenharia de fazer máquinas inteligentes” e, em 1956, foi reconhecida como uma área de pesquisa interdisciplinar, numa conferência realizada no Dartmouth College, nos Estados Unidos (1).

Inicia-se, assim, uma nova fase tecnológica, com o desenvolvimento da inteligência artificial nas mais diversas áreas (finanças, reconhecimento de objetos e de linguagem, reconhecimento facial, jogos informáticos, robótica...) (2, 3). De igual modo, a medicina tornou-se um campo aberto onde a IA tem sido alvo de uma intensa investigação.

Todos os dias, na prática clínica, os médicos vêm-se obrigados a tomar decisões, e este facto foi a chave para que, a partir da década de 70, emergisse uma comunidade médico-científica dedicada à pesquisa em IA, onde vários investigadores tentavam desenvolver sistemas de resolução de problemas e de apoio à decisão clínica, com base em *hierarchical and assoiational networks* e modelos heurísticos, para confirmar hipóteses de diagnóstico de acordo com o estado fisiopatológico do doente. Vários *workshops* e encontros internacionais sobre a IA na medicina se seguiram, tendo o primeiro dos quais, *Knowledge-based Systems in Biomedicine*, sido realizado em Nova Jérсия, na Universidade de Rutgers, em Junho de 1975. Estes encontros viriam a servir para que os investigadores não só apresentassem os seus projetos, mas também os

discutissem, comparando ideias e métodos, e encontrassem formas e abordagens práticas de testar os sistemas (4).

A introdução da informática na imagem médica levou a um grande avanço na aquisição de imagem que abrangeu a medicina nuclear, a radiologia de intervenção, com imagens de subtração digital, permitindo ainda o aparecimento da tomografia computadorizada (TC), nos anos 70, e da ressonância magnética (RM), nos anos 80. Mais tarde, surgiram sistemas que possibilitaram o arquivamento de imagens - *picture archiving and communication systems* (PACS). Isto contribuiu não só para formar volumes de imagens cada vez maiores, mas também para facilitar o seu acesso em locais remotos, usando redes informáticas externas. O crescente aumento dos PACS e o advento dos registos médicos eletrónicos - *electronic health records* (EMR) - impulsionaram o acesso a grandes quantidades de dados, clínicos e imagiológicos, numa era que se prolonga até aos dias de hoje, a era dos "big data" (5, 6).

O conceito de "fazer pensar as máquinas" de modo a que estas sejam capazes de reconhecer padrões, através da leitura prévia de um elevado volume de dados que lhe são fornecidos, é a base do funcionamento da IA. Na radiologia, essa elevada quantidade de dados são as imagens. No início do século XXI, a chegada do diagnóstico assistido por computador (CAD), usado na mamografia para a deteção de cancro da mama, revelou-se a primeira aplicação da IA na área da imagem médica (3). Mas é em 2013, com a criação do supercomputador *IBM Watson*, capaz de analisar dados e imagens com grande velocidade e precisão, fornecendo a sua interpretação e apoio à decisão clínica através de um processamento cognitivo, que se dá uma revolução na comunidade médico-científica, que ao longo da última década, devido ao fornecimento de sistemas computacionais poderosos por parte de diversas instituições, tem feito grandes avanços no uso da IA na radiologia (7). Baseada no pressuposto do reconhecimento de padrões de imagem, a IA prevê e permite o diagnóstico de diversas patologias e entidades clínicas, complementando os métodos de imagem 'convencionais'. Este reconhecimento é feito à custa de algoritmos matemáticos, nomeadamente *machine learning* (ML) e *deep learning* (DL), noções abordadas mais adiante.

Apesar do seu grande potencial, atualmente, o tema provoca acesa discussão, havendo quem considere que o futuro da imagem médica está indissociado da aplicação de sistemas de IA, os quais poderão atuar sinergicamente (ou competitivamente?) com os médicos radiologistas.

Com este trabalho, pretende efetuar-se uma breve descrição dos fundamentos teóricos associados aos termos "inteligência artificial", "*machine learning*" e "*deep learning*";

reconhecer e descrever as principais aplicações da IA na área da imagem médica com especial destaque para a aplicabilidade clínica; produzir uma análise crítica dos benefícios e inconvenientes dos sistemas de IA aplicados à radiologia; oferecer uma perspectiva futura acerca do potencial impacto da IA na área da imagem médica, abordando ainda os problemas éticos e legais que advém da utilização destes sistemas.

## MATERIAL E MÉTODOS

---

Para a elaboração desta revisão bibliográfica, entre Junho de 2019 e Novembro de 2019, foi realizada uma pesquisa eletrónica e seleção de artigos na base de dados *Pubmed*.

O objetivo foi recolher bibliografia relativa aos conceitos básicos de IA direcionada à imagem médica, às suas presentes e possíveis futuras aplicações na radiologia e às vantagens e obstáculos que advêm da sua utilização.

Numa primeira fase, foi efetuada uma pesquisa com o recurso à equação “((((“Diagnostic Imaging”[Majr]) OR “Diagnosis, Computer-Assisted”[Majr])) AND “Artificial Intelligence”[Majr]”, tendo sido incluídos artigos de revisão, revisões sistemáticas e meta-análises redigidos em língua portuguesa e inglesa, publicados entre o ano de 2009 e Novembro de 2019. Obteve-se um total de 208 artigos. Posteriormente, os artigos foram selecionados de acordo com a pertinência do título e *abstract*, pelo que foram excluídos 142 artigos. De seguida, foi efetuada a leitura dos artigos e, de acordo com a sua relevância para o tema, permaneceram 34 artigos. Numa segunda fase, utilizando a mesma equação “((((“Diagnostic Imaging”[Majr]) OR “Diagnosis, Computer-Assisted”[Majr])) AND “Artificial Intelligence”[Majr]”, foi efetuada outra pesquisa onde foram incluídos artigos de revisão narrativa, artigos de revisão sistemática, meta-análises e artigos científicos publicados entre Junho de 2018 e Outubro de 2019, redigidos em língua portuguesa e inglesa. Obteve-se um total de 999 artigos. Tendo em conta o método anterior, foram excluídos 910 artigos, de acordo com a pertinência do título e *abstract*. Após a leitura integral dos artigos, permaneceram 53 artigos considerados relevantes para a presente revisão.

Pontualmente, alguns dos artigos resultantes desta pesquisa serviram como fonte de novas referências pelas citações que apresentaram, dada a sua pertinência. Foram assim adicionados 13 artigos, perfazendo um total de 100 artigos que constituíram a base de redação desta revisão.

Para complementar a informação recolhida, foram também consultados o *website* da FDA e o *website* [medicalfuturist.com](http://medicalfuturist.com).

A gestão das referências bibliográficas do trabalho foi realizada através do gestor online *EndNote Web*, com o sistema de citação de Vancouver.

## RESULTADOS

---

### 1. Machine learning e deep learning: síntese de bases e conceitos

A IA, nomeadamente a aplicada à área da radiologia, baseia-se num conceito fundamental: *machine learning* (ML). Este termo refere-se a um campo no qual o computador (ou o sistema) consegue “pensar” sem que esteja programado. Isto difere da programação convencional, em que os sistemas desempenhavam uma determinada tarefa seguindo códigos delineados, ou seja, eram programados. O reconhecimento dos padrões assenta no facto dos algoritmos usarem modelos matemáticos que permitem como que um processamento de auto-decisão para a resolução de problemas. Para que sejam bem-sucedidos na sua função, os sistemas são treinados com dados (imagens médicas, no caso da radiologia). Cada algoritmo de ML contém um modelo estatístico, ou conjunto de hipóteses sobre os dados observados (dados de treino), que é igualmente aplicado aos novos dados, sobre os quais são feitas previsões (8).

Existem dois tipos de ML: *supervised ML* e *unsupervised ML*. No primeiro, o investigador treina a máquina: são inseridos no sistema dados *inputs* e dados *outputs* devidamente rotulados pelo investigador; o algoritmo é capaz de comparar as características (*features*) e aprender as relações entre cada um dos componentes; depois de ensinado e com base no conhecimento adquirido *à priori*, o sistema gera uma resposta, quando apenas os *inputs* (agora, não rotulados) são inseridos. O processo de *unsupervised ML* é mais complexo, já que os dados não são rotulados nem conhecidos: o investigador não fornece ao sistema qualquer tipo de informação; assim, o algoritmo usa apenas os modelos estatísticos e matemáticos, isto é, o seu próprio método, sendo capaz de determinar quantos grupos de *outputs* irão ser gerados, comparando as diferenças existentes nas características de cada um dos dados *inputs* inseridos e, aprendendo a distingui-los; este modelo permite a aquisição de novo conhecimento, já que os *outputs* gerados poderão ser os mesmos, diferentes, ou até em maior quantidade do que aqueles que seriam gerados por um método humano (7). A análise de um esquema com base em figuras geométricas poderá ser elucidativa para a compreensão destes conceitos (Fig. 1 e Fig. 2).

Existem vários tipos de algoritmos de ML: *neural networks*, *k-nearest neighbors*, *support vector machines*, *decision trees*, *random forest*, *the naive Bayes algorithm* e algoritmos de *deep learning* (DL) (9). Os mais comumente utilizados (e com maior potencial) são os que se baseiam em redes neuronais artificiais (*artificial neural network*, ANN): as ANN são compostas por centenas de unidades funcionais interconectadas (neurónios

artificiais ou nodos, em que cada neurónio corresponde a um elemento de processamento). Cada unidade funcional (nodo) recebe como *input* um conjunto de valores que traduzem um determinado conjunto de características (*features*) dos dados, sendo que cada valor é multiplicado por um peso numérico (*weight*). O neurónio artificial faz a soma ponderada desses pesos, a que se chama soma ponderada de evidências, e passa-a através do algoritmo (função de ativação não-linear) que determina a saída para o neurónio seguinte. O treino do algoritmo serve para que seja encontrado o melhor peso numérico, um valor ótimo, de modo a ser possível extrair o melhor sinal de decisão. Tudo isto mimetiza o que ocorre no cérebro humano (3, 10).

DL é um algoritmo específico de ML que engloba uma rede neuronal artificial de aprendizagem profunda com diferentes níveis de abstração. A vantagem destes algoritmos é que as suas ANN estão organizadas em múltiplas camadas. Tipicamente, as ANN tradicionais têm menos de 5 camadas de redes neuronais, em profundidade; os algoritmos de DL possuem um número de camadas muito superior (>20), tornando a rede mais profunda e permitindo um nível cada vez maior de abstração à medida que o número de camadas aumenta: *deep neural network learning*. A primeira camada recebe os *inputs* que são transferidos às camadas interiores (*hidden layers*), até à última camada que produz um *output* (Fig. 3). Quanto maior for o número de camadas, maior é a acuidade e melhor o desempenho do sistema (2, 3, 9).

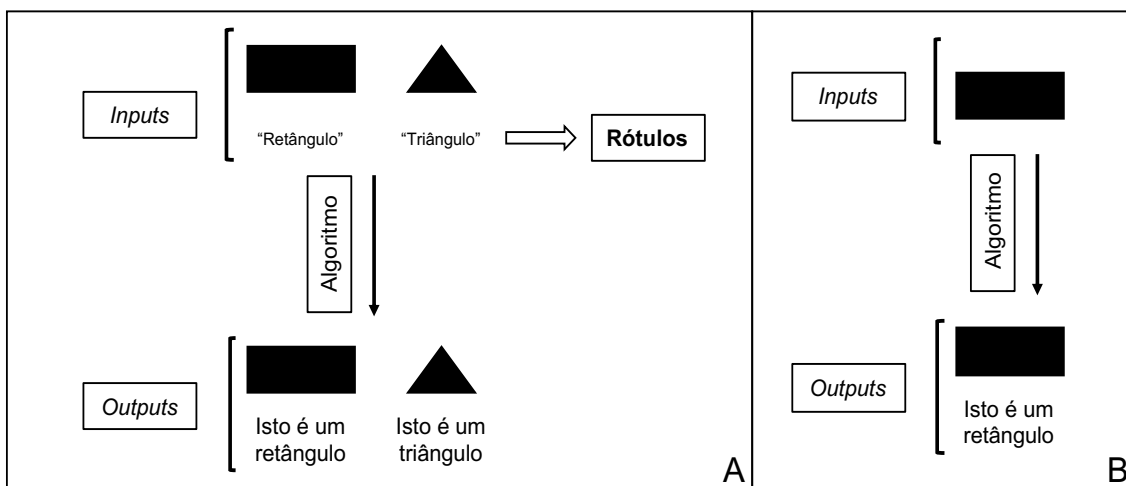


Fig. 1: **A.** *Supervised ML* – treino do algoritmo. O operador fornece ao sistema dados rotulados e o algoritmo aprende as relações entre cada um dos componentes; **B.** *Supervised ML* – execução. Depois de treinado, o algoritmo é capaz de classificar os dados inputs, agora não rotulados. Adaptado de: Nawrocki T, Maldjian PD, Slasky SE, Contractor SG. “Artificial Intelligence and Radiology: Have Rumors of the Radiologist’s Demise Been Greatly Exaggerated?”

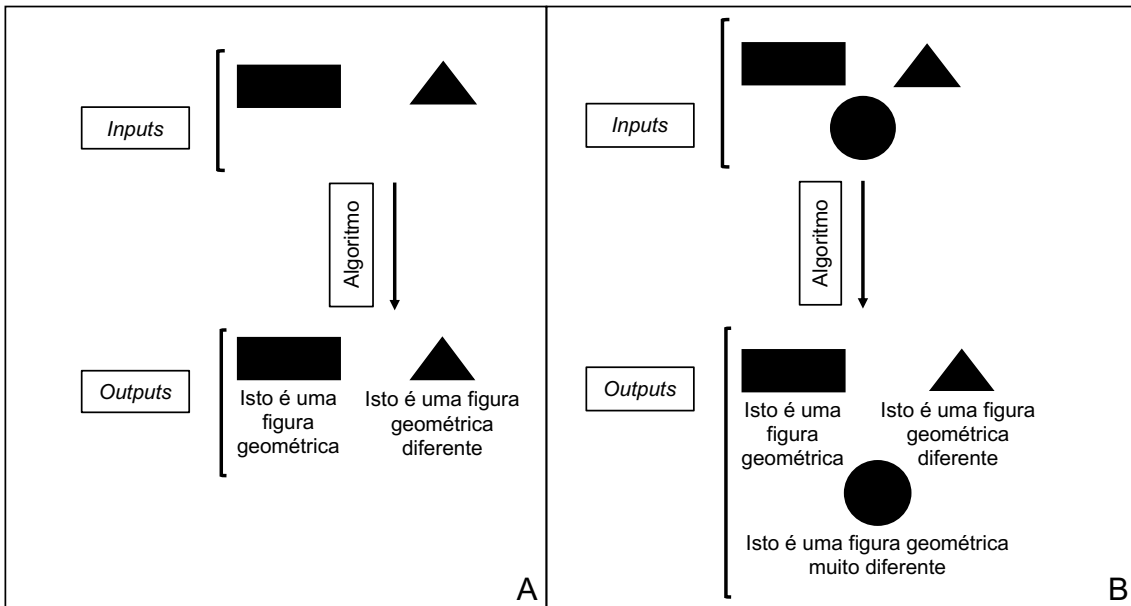


Fig. 2: **A.** *Unsupervised ML* – treino do algoritmo. O operador fornece ao sistema dados não rotulados e o algoritmo reconhece as diferenças entre os dados, determinando o número de grupos de outputs a ser gerado. **B.** *Unsupervised ML* – execução. Depois de treinado, o algoritmo é capaz de classificar os inputs com base no seu próprio método, podendo classificar dados não presentes na base de dados de treino. *Adaptado de: Nawrocki T, Maldjian PD, Slasky SE, Contractor SG. “Artificial Intelligence and Radiology: Have Rumors of the Radiologist’s Demise Been Greatly Exaggerated?”.*

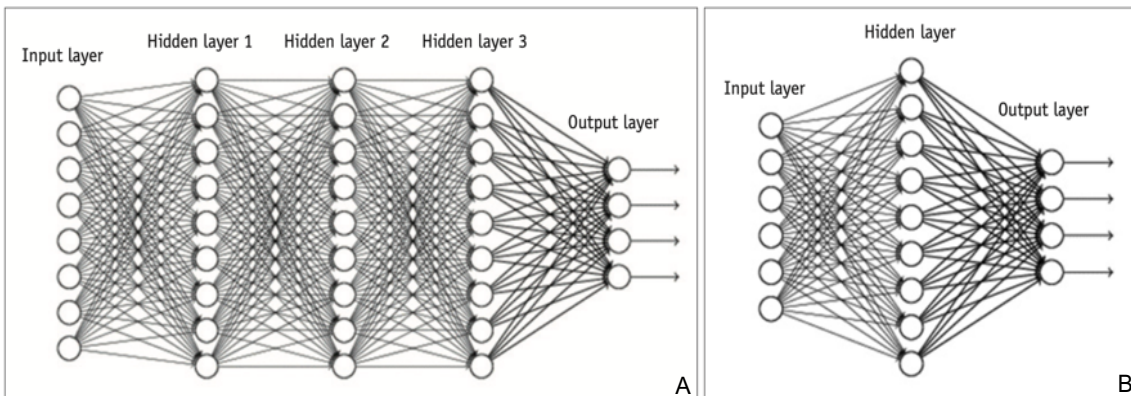


Fig. 3: Ilustração da diferença entre uma rede de DL e uma rede tradicional de ANN’s. **A.** Rede de DL com 3 camadas profundas (podendo ter muitas mais). **B.** Rede neuronal típica com apenas 1 camada. Fonte: Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. “Deep Learning in Medical Imaging: General Overview.”

Para a recolha dos dados, é necessário um bom conhecimento das características (*features*) subjacentes para extrair aquelas que são mais relevantes. Isto pode ser extremamente difícil quando existe uma grande coleção de dados e quando esses dados são bastante heterogêneos, como é o caso das imagens médicas. Uma das grandes vantagens em utilizar métodos de DL é a sua capacidade inerente para aprender características complexas diretamente a partir dos dados apresentados em bruto. Isto permite a definição de um sistema que não depende da extração manual de características (*handcrafted features*), que são na sua maioria necessárias noutras técnicas de ML. Por outro lado, tendo a capacidade de aprender as características relevantes dos dados, o processamento de DL permite que não sejam apenas estudadas as especificidades que o investigador acharia importantes. No entanto, isto faz com que um elevado número de parâmetros tenha de ser analisado e compreendido pelo algoritmo. Deste modo, os algoritmos de DL necessitam de uma maior quantidade de dados para serem treinados, em comparação com a quantidade necessária para os algoritmos tradicionais de ML. Para além disto, necessitam de um elevado poder informático (*hardware*), para que seja possível acomodar as muitas camadas de ANN's, e de mais tempo de treino. O aparecimento das unidades de processamento gráfico (GPU), que integram circuitos eletrónicos especificamente projetados para manipular dados de imagem, atualmente disponíveis a um baixo preço, tem permitido um processamento mais rápido e poderoso (2, 7, 11).

Associado ao elevado potencial destes sistemas para a interpretação de dados nas mais diversas áreas, os algoritmos têm ainda a capacidade de diminuir a sua suscetibilidade ao erro à medida que vão lendo e analisando mais dados. Uma vez que a resposta da rede com base na sua configuração atual é registada (ou seja, o peso numérico para a soma ponderada de evidência é registado), aquando da passagem de um novo conjunto de valores *input* pelo algoritmo, a previsão real gerada é comparada quantitativamente com a previsão estimada, tendo em conta a leitura prévia, através de uma função de erro, denominada função de perda. Repetidamente e, utilizando um processo de retorno, o sistema ajusta cada parâmetro do algoritmo (ou peso numérico) para minimizar a função de perda. Assim, a máquina aprende com o seu próprio erro, ficando cada vez mais apta para desempenhar a sua tarefa ao longo do tempo. Isto faz com que o processo de ML seja contínuo e não termine, o que reforça a ideia de que são necessárias grandes quantidades de dados para o seu desenvolvimento (12).

Em suma, o desempenho do sistema é proporcional ao número de camadas da rede neuronal, ao número de dados bem rotulados e anotados e ao grau de otimização da função do algoritmo (6).



## 2. IA na radiologia: aplicações atuais e futuras

Fruto do aperfeiçoamento dos algoritmos de DL, do desenvolvimento do poder informático, bem como do acesso a grandes bases de dados, tem sido possível a evolução exponencial da inteligência artificial e a sua utilização na radiologia.

Na análise da imagem médica, os algoritmos de DL utilizam um tipo de ANN específico: *convolutional neural networks* (CNN). Com um design complexo, estas redes neuronais são capazes de analisar e interpretar dados mais sofisticados e complicados, sendo, assim, aplicáveis em áreas de reconhecimento e classificação de imagens (7). A diferença das CNN's para as restantes redes neuronais é o facto das primeiras terem a capacidade de relacionar as características geométricas dos diferentes dados *inputs*, como os bordos, as linhas e as colunas das imagens. Chegado até aqui, o leitor é convidado a fazer uma paragem e a assimilar dois aspetos. Resumidamente, uma rede de DL aplicada a imagens médicas envolve, numa primeira fase, a extração das *features* das imagens em bruto e, numa segunda fase, a aplicação de uma função algorítmica que aprende a correlação entre as *features* e os rótulos das imagens. Ora, devido à existência de um tipo especial de camadas convolucionais, as CNN's diferem das restantes redes de ML por atuarem na primeira fase. Os sistemas aplicam um conjunto de filtros convolucionais a uma imagem em várias posições possíveis, com vários tamanhos de janelas, para abranger estruturas de objetos de diferentes formas e tamanhos. Assim, em vez de escolherem um conjunto fixo de detetores de características, estas redes usam um paradigma de *generate-and-test*, servindo-se de uma variedade de filtros em todas as escalas e resoluções possíveis, para que o aprendiz (a função objetiva) possa julgar o quão útil é ou não qualquer uma das *features* para a discriminação e aprendizagem dos rótulos. Posteriormente, à semelhança do que acontece, então, com os restantes algoritmos de ML, a função aprende a correlação entre as características e os rótulos das imagens (6, 9, 11).

Não obstante existirem algoritmos de DL de deteção e diagnóstico noutras áreas da medicina, como na oftalmologia, dermatologia e anatomia patológica (11, 13), a interpretação dos exames imagiológicos pode envolver milhares de imagens radiológicas, o que se revela uma tarefa difícil, ainda mais quando essas imagens são bastante heterogéneas (ruído, artefactos e variabilidade entre leitor, doente e patologias). O foco das CNN's na análise e interpretação da imagem médica pode, então, ser dividido nas seguintes categorias: deteção de lesões e deteção de doença, classificação de imagens e diagnóstico, segmentação e quantificação (2). Os campos de pesquisa e investigação incidem, sobretudo, na imagem torácica, em senologia e neuroimagem.

## 2.1 Sistemas CAD

Tradicionalmente, um sistema de diagnóstico assistido por computador (*Computer Aided Diagnosis, CAD*) é um sistema usado pelos radiologistas e outros clínicos que auxilia a interpretação de imagens médicas, tendo por base algoritmos de ML. Na prática, um sistema CAD típico serve como um “segundo leitor”, fornecendo informação mais detalhada sobre a região anormal. Existem dois tipos de sistemas CAD: sistemas CAD de detecção (CADE) e sistemas CAD de diagnóstico (CADx). Enquanto que os primeiros se focam na detecção de anomalias, os segundos têm como objetivo a classificação. A utilização de sistemas CAD na mamografia é o paradigma dos sistemas CADE, capazes de detetar anomalias, como assimetrias ou calcificações. Estes sistemas mais tradicionais usavam técnicas de reconhecimento baseadas em algoritmos convencionais de ML, em que as características das imagens eram extraídas manualmente, como já foi referido anteriormente. Porém, estudos revelaram que geravam mais falsos positivos do que os leitores humanos, o que levou a um maior tempo de avaliação e a procedimentos de diagnóstico desnecessários. Assim, o benefício real obtido com o uso de CAD não foi claro. Espera-se que a atual tecnologia de DL possa ajudar a ultrapassar as limitações dos sistemas CAD anteriores, e a alcançar uma maior precisão de detecção. O desenvolvimento de algoritmos de DL tem permitido a criação de sistemas que incorporam CADE e CADx, sendo capazes, não só de detetar, mas também de classificar as anomalias. Um sistema de CADx consiste num pré-processamento de imagem, seguido da extração e seleção adequadas das características que culminam na classificação (3, 11, 14).

## 2.2 Imagem torácica

A segmentação de costelas e da caixa torácica em imagens radiográficas é uma das aplicações de algoritmos de IA na imagem torácica. Vários estudos com algoritmos de ML revelaram a possibilidade da extração das costelas e de outras estruturas ósseas, como as clavículas, das imagens de radiografia (15). Um trabalho recentemente publicado pela American Association of Physicists in Medicine utilizou CNN's para subtrair as estruturas ósseas de imagens radiográficas, e obter imagens de tecidos moles, de forma a melhorar a visualização de lesões, como nódulos pulmonares, muitas vezes indetetáveis pelos radiologistas graças à interposição óssea (16).

Algumas doenças pulmonares podem propagar-se pelas cisuras, que correspondem ao limite dos lobos do pulmão. A sua extração, a partir de imagens de TC, é possível utilizando uma abordagem de ML, o que constitui uma ferramenta útil ao radiologista,

por exemplo, para a melhor caracterização de nódulos pulmonares contidos nas cisuras que, pela forma que por vezes apresentam, podem confundir a natureza maligna ou benigna do nódulo (15).

As vias aéreas (VA) estão implicadas em várias doenças pulmonares: estão afetadas em patologias como bronquiectasias; a medição do seu lúmen e da espessura da parede são ferramentas essenciais para a avaliação do prognóstico da doença pulmonar obstrutiva crónica; a sua segmentação pode ser usada para o planeamento de procedimentos de broncoscopia; e o conhecimento sobre a sua exata localização pode ser usado para melhorar a deteção de anomalias, como nódulos endobrônquicos. Embora os programas baseados em regras tenham, em teoria, um certo potencial em segmentar as VA a partir de imagens de TC (baseado no facto da diferença de densidade entre o ar presente nas VA e o parênquima), o objetivo não é conseguido uma vez que existe uma elevada percentagem de ar também no parênquima, e determinadas doenças, como o enfisema, acrescentam um grande ruído aquando da extração, não permitindo um correto isolamento das VA. A utilização de algoritmos de ML poderá colmatar esta lacuna (15). Um grupo de investigadores (17) construiu um algoritmo otimizado de combinação entre uma abordagem *rule-based* e *unsupervised* ML para a segmentação de pequenos e médios brônquios em imagens de TC torácica, tendo alcançado uma precisão média de 96% na deteção de pontos centrais dos brônquios nos estudos de imagem usados para o treino.

A deteção, a caracterização e a classificação (benigno vs. maligno) de nódulos pulmonares são também um desafio para os radiologistas. Várias são as características a ter em conta: o tamanho, os contornos, a textura, o grau de calcificação, mas também a distinção entre nódulo ou vaso (em imagens de TC, ambos têm a densidade de tecidos moles e, enquanto que os vasos são cilíndricos, os nódulos são esféricos). O principal objetivo é a redução dos falsos positivos, e desde de 2000 que vários grupos têm realizado trabalhos com algoritmos de ML tradicionais no sentido de aumentar a sensibilidade dos métodos de diagnóstico (15). No entanto, o uso de CNN's por vários grupos de investigação tem sido cada vez maior. Por exemplo, este ano, um grupo de investigadores sul-coreanos (18) propôs um algoritmo DL com uma nova arquitetura de CNN para a redução de falsos positivos na classificação de nódulos em imagens de TC, com base na extração das características mais relevantes. Tal como este estudo, vários outros têm sido desenvolvidos. No entanto, apesar de com sucesso demonstrarem a eficácia das CNN's na extração de características morfológicas para a caracterização dos nódulos pulmonares, a redução efetiva de falsos positivos, no que concerne à sua classificação, ainda está longe de ser conseguida, devido não só à grande variabilidade

entre as imagens usadas para o treino dos sistemas, mas também, às diferenças entre os protocolos e entre a arquitetura dos algoritmos usados nos diversos grupos, que não permitem resultados reprodutíveis, quando se efetuam estudos de comparação (19). Por outro lado, as bases de dados para o treino dos algoritmos carecem de um elevado número de imagens de radiografia com qualidade, o que tem levado alguns investigadores ao uso de técnicas de *transfer learning*. Contudo, estudos revelaram que apesar dos modelos serem capazes de determinar a área geral nodular, não conseguem determinar a localização exata dos nódulos (20). Apesar da crescente utilização de redes convolucionais de DL, em 2019, Khan *et al.* propuseram um método de deteção de nódulos pulmonares baseado num algoritmo tradicional de ML com a extração manual de *features* (*support vector machine*) que conseguiu uma sensibilidade de 97,45% e, assim, a redução de falsos positivos (21).

As doenças pulmonares intersticiais (DPI) também despertam interesse, no sentido em que o seu diagnóstico precoce é importante para o aumento da sobrevivência dos doentes. A TC de alta resolução é o exame de eleição: um dos algoritmos de DL proposto, treinado com imagens de TC de alta resolução, conseguiu uma performance ao nível do radiologista na classificação da DPI (22). Em comparação com algoritmos tradicionais de ML, as CNN's revelaram uma acuidade superior na classificação das DPI (23).

Algoritmos para a deteção de outras doenças, como a tuberculose e a pneumonia têm sido alvos de investigação. Hwang *et al.* (2019) propuseram um algoritmo de DL para classificação de radiografias em normais vs. anormais, tendo utilizado para o treino do sistema imagens radiográficas com doenças como, neoplasia maligna do pulmão, pneumonia, tuberculose e pneumotórax. O algoritmo demonstrou uma performance superior à de 3 médicos radiologistas, e quando se comparou a eficácia do diagnóstico por parte dos médicos com o auxílio do algoritmo vs. sem o auxílio do mesmo, esta melhorou substancialmente quando se utilizou a ajuda do sistema (24). Em 2018, Yates *et al.* (25) demonstraram igualmente o potencial da IA na classificação de imagens radiográficas em normais ou anormais, tendo usado um algoritmo de DL que revelou 94,6% de acuidade. Da mesma forma, um grupo de investigadores, na Turquia, utilizou uma CNN para classificar as mesmas doenças em imagens de radiografia, e demonstrou a superior acuidade desta *network* em comparação com outros algoritmos de ML (10). Rajpurkar *et al.* desenvolveram um algoritmo baseado em CNN's que revelou ser eficaz na deteção de 14 diferentes patologias com afeção torácica em imagens radiográficas (onde se incluem, enfisema, atelectasia, e derrame pleural), verificando-se que o sistema foi capaz de alcançar o nível de performance de radiologistas na deteção de 11 das 14 doenças (26).

Existem várias ferramentas de IA que usam *networks* de DL úteis para a deteção de tuberculose (27). Ultimamente, têm sido propostos algoritmos para o diagnóstico precoce de tuberculose multi-resistente. Em 2018, um artigo publicado no *International Journal of Computer Assisted Radiology and Surgery* (28) conseguiu demonstrar, usando algoritmos de ML e DL, que poderá ser possível a extração de características a partir de imagens de radiografias do tórax e, assim, estimar a probabilidade de tuberculose multi-resistente. Gao e Kian, num artigo publicado em Janeiro de 2018 (29), usando um algoritmo de DL em combinação com um algoritmo tradicional de ML, conseguiram demonstrar uma acuidade de 91,11% para a classificação de tuberculose multi-resistente.

Foram, também, publicados estudos que comprovaram a eficácia dos algoritmos de ML na deteção de cardiomegalia, com base na quantificação do índice cardio-torácico (20).

### **2.3 Senologia**

Em 1998, foi aprovado o primeiro sistema CAD acoplado à mamografia digital para a deteção do cancro da mama (14). Ao longo dos anos, a diminuição dos falsos negativos é um dos objetivos fulcrais nos trabalhos de investigação que se vão desenvolvendo, já que a sensibilidade da mamografia é subótima na deteção de cancro da mama no contexto de rastreio. Katzen e Dodelzon (2018) (30) elaboraram um artigo de revisão onde descrevem os resultados de vários estudos que usaram sistemas CAD na mamografia para a deteção de cancro da mama. A adição destes sistemas permitiu o aumento da taxa de deteção de cancro (CDR, *cancer detection rate*) entre 0 e 19,5%. No entanto, apesar da sensibilidade do sistema CAD acoplado à mamografia ter uma variação de sensibilidade entre 60 e 100%, alguns estudos revelaram uma diminuição do valor preditivo positivo na deteção de cancro e um aumento da taxa de biópsias. Já no que diz respeito à deteção de calcificações, a análise dos diferentes estudos revelou uma maior sensibilidade destes sistemas, variando entre 80 e 100%. Quanto à performance, foram mais os estudos que revelaram um aumento da área abaixo da curva COR (curva de característica de operação do recetor) quando foi adicionado um sistema CAD à mamografia, do que os que revelaram uma diminuição. Para além disto, a acuidade dos vários sistemas CAD nos diferentes estudos foram superiores na deteção de cancro em estadios iniciais, relativamente aos estadios mais avançados. A análise das investigações demonstrou que uma das limitações do uso de sistemas CAD na mamografia é, de facto, o aumento de falsos positivos. A pesquisa relacionada com a utilização destes sistemas na tomossíntese mamária tem revelado sensibilidades superiores, próximas de 93%, em comparação com o acoplamento à mamografia digital.

Ainda assim, a revisão permitiu concluir que os resultados, mesmo que promissores, não são inequívocos quanto ao real benefício e verdadeiro retorno em termos de acuidade de detecção e diagnóstico, usando os sistemas CAD. Em 2017, a FDA aprovou o primeiro sistema de tomossíntese com CADx para o diagnóstico de cancro da mama, contudo, o risco de falsos positivos e de falsos negativos persiste (31).

O desenvolvimento de algoritmos de DL tem também lugar no campo da senologia. Em Março de 2018, um grupo de investigadores deu a conhecer o estado-da-arte quanto à utilização destes algoritmos (32), verificando que a aplicação com o intuito de diagnóstico (vs. detecção e segmentação de lesões) continua a ser superior na modalidade de mamografia, em comparação com a ecografia e RM mamárias, dado o seu potencial de rastreio. De uma forma geral, os algoritmos de DL têm sido usados para melhorar a acuidade dos sistemas CAD. Neste sentido, Cai e a sua equipa demonstraram o elevado potencial de um algoritmo de DL na discriminação de microcalcificações em imagens de mamografia (33). Os investigadores usaram uma CNN que obteve uma acuidade de 87,7%, marginalmente superior ao método de extração manual de características (com base na textura e morfologia das lesões), que obteve uma acuidade de 86,7%. Usando imagens de RM contrastadas, Herent *et al.* desenvolveram um modelo de DL capaz de detetar e caracterizar lesões benignas e malignas, que conseguiu um valor da área abaixo da curva COR para a detecção de lesões malignas de 0,869 e para a detecção de outras lesões benignas de 0,659 (34). A nível ecográfico, um artigo publicado este ano, no Japanese Journal of Radiology, apresentou um algoritmo de DL com uma CNN capaz de distinguir massas benignas de massas malignas em imagens de ecografia, com igual ou melhor performance do que a de 3 radiologistas (35). O modelo alcançou um valor de 0,913 correspondente à área abaixo da curva COR, superior ao dos médicos radiologistas (entre 0,728 e 0,845). Outra investigação levada a cabo por um grupo chinês demonstrou o potencial de um sistema DL em classificar as lesões detetadas em imagens ecográficas em 5 categorias do sistema *Breast Imaging Reporting and Data System* (BI-RADS), já que o algoritmo atingiu uma acuidade de classificação nas categorias 3, 4A, 4B, 4C e 5, entre 0,734 e 0,998 (36).

Embora, relativamente à imagem de mama, também haja uma utilização crescente dos algoritmos de DL, muitos grupos de investigadores ainda utilizam algoritmos mais tradicionais de ML, como *support vector machines*. Uma revisão sistemática (37), ao analisar as publicações entre 2012 e 2017 que usaram diferentes algoritmos de IA nas diversas modalidades de imagem (mamografia, ecografia, RM), permitiu constatar que, devido a diversos fatores (onde se incluem o uso de diferentes bases de dados, o

diferente número de imagens selecionadas entre os estudos, e os diferentes tipos de metodologia para o treino e validação dos sistemas) não é possível fazer uma comparação sustentável entre os diferentes estudos e concluir qual, efetivamente, o melhor tipo de algoritmo para utilizar nos sistemas CAD. Da mesma forma, dado que a RM se revela o método mais sensível para a imagem de mama, Codari *et al.* realizaram uma revisão sistemática com o objetivo de rever a literatura relacionada com a aplicação da IA na imagem de RM mamária, concluindo que, apesar da intensa investigação na última década, a efetividade destes sistemas ainda não é satisfatória (38).

Foi também estudada a possibilidade de diminuir a quantidade de leituras necessárias de exames de rastreio de cancro da mama através da deteção de mamografias sem alterações, excluindo-as, assim, da lista de trabalho. No entanto, não foi colocado de parte o facto de haver resultados positivos que possam escapar ao sistema (39).

## 2.4 Neuroimagem

O campo da neuroimagem é uma área na qual também várias técnicas de IA têm sido aplicadas. Vários estudos têm demonstrado o sucesso da segmentação das estruturas cerebrais com métodos de DL. Em 2018, Roy e a sua equipa propuseram uma CNN capaz de segmentar o cérebro em 20 segundos, a partir de imagens de RM, com uma acuidade elevada quando o método foi experimentado em 5 bases de dados (40). Outro estudo, publicado também em 2018 (41), demonstrou um algoritmo de DL para a segmentação de substância branca em imagens de RM de difusão, no entanto, os resultados não foram efetivos para o objetivo principal, dado que houve falsos positivos e várias reconstruções ficaram incompletas.

Usando como *inputs* várias informações provenientes da RM estrutural cerebral (por exemplo, a densidade ou a espessura cortical da substância cinzenta, ou alterações na substância branca), é possível o estudo de várias doenças, como a doença de Alzheimer (DA). Neste caso, torna-se um desafio saber quais os doentes com incapacidade cognitiva ligeira que irão evoluir para demência, e assim, para o processo neurodegenerativo característico da doença. Vários algoritmos de ML têm sido construídos para estimar a probabilidade da evolução (42). Li *et al.* propuseram um algoritmo com CNN's para a classificação de doentes com DA ou com défice cognitivo ligeiro (43). Têm surgido outros algoritmos com CNN's para a segmentação de estruturas cerebrais subcorticais em imagens de RM, de forma a poupar tempo na sua análise, essencial para a compreensão de várias doenças neuropsiquiátricas, para a monitorização da sua evolução e avaliação dos *outcomes* de tratamento, como a

esquizofrenia, a doença de Parkinson (DP), a perturbação do espectro do autismo (PEA) e a esclerose múltipla (EM) (44). No que diz respeito à PEA, existem algoritmos de ML, com base em alterações nas imagens de RM estrutural cerebral, que obtêm maiores níveis de acuidade diagnóstica, em comparação com a pontuação adquirida no *Autism Diagnostic Observation Schedule*. Na EM, o uso de sistemas de IA permite a segmentação de lesões da substância branca, que são a base do diagnóstico da doença. Estes algoritmos têm também sido propostos para o diagnóstico de DP e de outras doenças neuropsiquiátricas, como a depressão ou a perturbação de hiperatividade e défice de atenção. Dado que este tipo de doenças são complexas e muito heterogêneas, uma vez que, de acordo com os sinais e sintomas, podem conter estados clínicos intermédios na evolução da sua história natural, mais do que obter um diagnóstico de doente vs. não doente, torna-se importante a interpretação dos achados e a probabilidade da evolução da doença. No entanto, verifica-se que os algoritmos são mais eficazes em classificar diferentes doentes com diferentes doenças (DA, PEA, DP, EM) vs. não doentes, do que em estadiar os diferentes estados patológicos de cada doença (42, 45-48).

Os algoritmos de DL têm sido também investigados para a segmentação de tumores, como o glioblastoma. Enquanto que os algoritmos que detetam a lesão identificam locais com potenciais anormalidades, no caso da segmentação, o sistema localiza e segmenta o local anormal, sendo possível calcular o volume da lesão, e analisar a sua morfologia, textura e contornos. Sendo indicadores para avaliar a resposta ao tratamento, são úteis ao radiologista. Ora, se o médico tiver de realizar este procedimento manualmente, há um grande consumo de tempo: os algoritmos de DL com CNN's permitem que tudo isto se faça de uma forma automática e rápida (49, 50). Em 2019, um grupo de investigadores desenvolveu um sistema de IA com uma CNN para a segmentação e avaliação da heterogeneidade vascular do glioblastoma. Os resultados mostraram uma acuidade na segmentação superior a outros métodos de imagem (51). No último ano, outros estudos têm sido desenvolvidos, e a utilização de algoritmos de ML (entre os quais *random forests* e DL) na segmentação de tumores tem demonstrado resultados promissores (52-55). Quanto a outro tipo de tumores, como as metástases cerebrais, Jun *et al.* sugeriram um algoritmo de DL para a aquisição de imagens *black-blood*, a partir de imagens de RM contrastadas. A investigação revelou que a acuidade dos radiologistas na deteção das metástases pela análise de imagens de *black blood* sintéticas foi superior à deteção pela análise das imagens originais de RM (97% para 94%), sugerindo a efetividade dos algoritmos de DL para a correta análise de lesões correspondentes a possíveis metástases cerebrais (56).



## 2.5 Triagem de emergências

A rápida triagem de casos que chegam em risco de vida ao serviço de urgência é também um propósito que seria muito útil ao médico, mas uma grande mais valia para o doente. É o caso de eventos neurológicos, pneumotórax, enfarte agudo do miocárdio (EAM), abdómen agudo e politrauma. Apesar do diagnóstico se basear, na maioria das vezes, em sinais e sintomas, vários sistemas CAD têm sido investigados para uma rápida identificação destas situações. Em 2018, foi desenvolvido um sistema com um algoritmo de DL que permitiu detetar achados neuroradiológicos críticos a partir de imagens de TC, onde se incluem a hemorragia e o acidente vascular cerebral (AVC), por forma a fazer a triagem dos estudos de imagem e ordena-los por ordem de gravidade (57). Da mesma forma, um sistema baseado numa CNN demonstrou uma acuidade superior a 90% na deteção e quantificação de hemorragias intraparenquimatosa, subdural/epidural e subaracnoideia, em imagens de TC não contrastadas (58), e um artigo, publicado em 2018, descreve um algoritmo capaz de segmentar o local de AVC isquémico em imagens de RM de difusão (59). Outros trabalhos demonstraram a eficácia de sistemas baseados em algoritmos de DL para a deteção de hemorragias cerebrais em imagens de TC (60, 61).

Em 2018, foi testada a deteção de EAM a partir de imagens de TC não contrastadas de baixa dose de radiação, usando vários algoritmos de ML com base na análise e extração de características da textura do ventrículo esquerdo. Apesar de ser possível verificar a diferença na textura entre imagens de doentes com EAM e de doentes sem enfarte, o valor da área abaixo da curva ROC atingido (0,78) não foi suficiente para permitir o diagnóstico com certeza (62). Em 2019, um grupo de investigadores suíços desenvolveu um algoritmo que permitiu detetar achados que causam abdómen agudo, como gás livre abdominal ou líquido, em imagens de TC abdominal, com uma sensibilidade de 93% e uma especificidade de 97% (63). Taylor *et al.* propuseram um algoritmo de DL para a deteção automática de médio e grande pneumotórax em imagens radiográficas, de forma a acelerar o seu diagnóstico e tratamento. Embora o algoritmo tenha atingido uma sensibilidade próxima de 80% para a base de dados de treino, quando se testou o sistema com bases de dados externas, a mesma caiu para valores entre cerca de 30 a 40%, o que não permitiu a validação externa (64). Um estudo recente revelou ainda a possível capacidade de um sistema de IA segmentar o rim em ECO-FAST no contexto de trauma, conseguindo diferenciar a estrutura renal da não-renal (65). Os resultados destas investigações, demonstraram que o uso de sistemas de IA poderá ser efetivo na

priorização das listas de trabalho dos radiologistas, para que os casos mais graves sejam rapidamente detetados e tratados.

## **2.6 Outras aplicações, radiômica e NLP**

No campo da imagem músculo-esquelética, é crescente a investigação para métodos de segmentação do músculo esquelético, útil para a cirurgia ortopédica. Assim, publicada em 2018, uma revisão da literatura deu a conhecer vários métodos de DL com CNN's que utilizam imagens em 3D de TC para o efeito (66). Para além da segmentação do músculo, a segmentação do tecido adiposo por algoritmos de ML, tem sido alvo de investigação, especialmente para o estudo da sarcopenia (66).

A aplicação de sistemas de IA também tem sido investigada para o estudo de outros órgãos na região cervical, como a tiróide, na região abdominal, como o fígado e na região pélvica, como a próstata.

Lee e a sua equipa desenvolveram um sistema CAD baseado num algoritmo DL que atingiu uma acuidade de 83% na localização e diferenciação entre nódulos linfáticos metastáticos e carcinoma da tiróide, na ecografia deste órgão (67).

A segmentação do fígado em imagens radiológicas revela-se importante no sentido de acompanhar a resposta ao tratamento de várias patologias que têm sede neste órgão, entre as quais se destaca o carcinoma hepatocelular. Vários algoritmos de IA têm usado CNN's para a segmentação do fígado na imagem médica (68). Em 2018, Chlebus *et al.* propuseram um método de DL para a segmentação de tumores do fígado em imagens de TC (69). Foi ainda demonstrada a possibilidade de estadiar a fibrose hepática através de um algoritmo de DL, com base em imagens de TC, ainda que com uma performance moderada (70).

O rim é também alvo de atenção, sendo que a segmentação do córtex renal com um algoritmo de DL em imagens de TC foi proposta por Couteaux *et al.* num desafio organizado pela Sociedade Francesa de Radiologia, *Journées Francophones de Radiologie*, onde o grupo de investigadores alcançou o primeiro lugar no *ranking* dos participantes (71). A nível da região pélvica, foram demonstrados algoritmos com CNN's para a segmentação da próstata em imagens de RM (72, 73), para a segmentação de tumores do reto em imagens de RM (74, 75) e para a segmentação da parede da bexiga em imagens de urografia por TC, útil para o diagnóstico de neoplasias do urotélio (76). A deteção de cálculos no ureter em imagens de TC, com base num algoritmo de DL com CNN foi igualmente demonstrada, apesar de não ter revelado uma acuidade de

diagnóstico superior à do radiologista (77). Foram também propostos algoritmos para a segmentação da próstata a partir da ecografia (78).

É muito vasta a investigação que atualmente ocorre para que se desenvolvam métodos de IA capaz de cobrir diversas áreas da radiologia. Além da detecção, classificação e segmentação, a melhoria dos métodos de imagem atuais, nomeadamente de TC e RM, através de sistemas de IA também tem sido descrita. Isto inclui a redução do ruído e de artefactos de imagem e a reconstrução e aquisição de imagem (79). Recentemente, existem publicações que demonstram algoritmos que permitem a reconstrução de imagens de TC de baixa dose, de forma a diminuir os artefactos provocados pelo ruído consequentes da baixa dose de radiação, permitindo o aumento da qualidade das imagens (80, 81). Da mesma forma, com um algoritmo de DL, foi demonstrado ser possível a manutenção da qualidade de imagens de RM contrastadas com uma baixa dose de gadolínio, semelhante à qualidade das imagens com a dose convencional (82). Por outro lado, é também possível, através de sistemas de ML, reduzir os artefactos provocados pela implantação de metais nas imagens de TC. Os sistemas são treinados com imagens com e sem implantes metálicos e, posteriormente, fazem a reconstrução de novas imagens com maior qualidade (83, 84). É também possível a avaliação da qualidade de imagens de RM, usando algoritmos de ML (85, 86). Os resultados de um estudo publicado em 2018, sugeriram a possibilidade de se obter imagens de TC a partir de imagens de RM ponderadas em T1, através de um algoritmo de ML (87).

A era dos *big data* permitiu, adicionalmente, o progresso da medicina de precisão. O conceito original envolve o “uso de estratégias de prevenção e tratamento que têm em conta a variabilidade individual” através do acesso a um grande conjunto de dados clínicos, genómicos, imagiológicos e moleculares (88). O funcionamento da medicina de precisão assenta na existência de biomarcadores, sendo que os algoritmos de DL podem ser usados para extrair biomarcadores quantitativos a partir de um elevado número de diferentes tipos de dados médicos. Uma vez que as imagens médicas são conseguidas de forma não invasiva e rotineira na prática clínica, revelam-se uma ferramenta útil para a extração de biomarcadores quantitativos de imagem, processo que se designa de radiómica. Estes biomarcadores são indicadores das características da doença, refletindo propriedades cuja visualização ou medição espontâneas não estão acessíveis ao radiologista, formando uma espécie de fenótipo quantitativo baseado na imagem médica. Por exemplo, no caso de um tumor, depois de segmentado, o algoritmo extrai as suas características, como a forma, o tamanho, a morfologia ou a textura e, posteriormente, as *radiomic features* (ou seja, os biomarcadores quantitativos) são calculadas. Ora, isto permite uma integração de forma

a que as características radiômicas possam ser associadas aos restantes biomarcadores provenientes das diversas fontes de dados para posterior aplicação no desenvolvimento de modelos preditivos para uso na avaliação do risco, triagem, detecção, diagnóstico, prognóstico, resposta terapêutica, risco de recidiva, etc. Vários estudos demonstram associações estatisticamente significativas entre as características radiômicas quantitativas e várias características clínicas, moleculares e genômicas de tumores sólidos, como por exemplo, a mama ou o glioblastoma (3, 89-92). Esta integração de dados tem o nome de radiogenômica. Por exemplo, em relação ao cancro da mama, o reconhecimento de padrões de dados de imagem com algoritmos de DL e a sua correlação com os dados do registo do tumor (através de uma análise radiômica e radiogenômica) permitem que os novos sistemas CADx sejam capazes de incorporar variáveis pertencentes ao doente e ao tumor e, avaliar o risco de cancro (prevendo a probabilidade) quando os mesmos reconhecem um padrão semelhante. Este novo design tem o potencial de maximizar o benefício do rastreio do cancro da mama e de melhorar as questões relacionadas com o sobrediagnóstico e sobretratamento (14).

Para além da interpretação de imagem, os algoritmos de ML revelam potencial no processamento de linguagem natural (*natural language processing*, NLP). De uma forma muito resumida, o NLP consiste num processo de conversão de informação contida na linguagem humana em dados computacionais, de forma a que a mesma seja compreendida por dispositivos tecnológicos. Empregar algoritmos de ML neste processo torna possível, por exemplo, extrair dados clinicamente relevantes de EMR para a construção de relatórios radiológicos padronizados, cuja interpretação varia entre radiologistas e instituições. Eventualmente, a melhor compreensão semântica de texto livre e a criação automática de relatórios padronizados podem servir para a produção de relatórios radiológicos adaptados a diferentes públicos, como o doente, o médico de família ou o cirurgião. Para além disto, os sistemas inteligentes de compreensão de texto podem reconhecer achados importantes, como nódulos ou medidas de uma lesão em determinados órgãos, oferecer terminologia padronizada e fornecer automaticamente recomendações acerca de um determinado estudo imagiológico apropriado a uma dada situação clínica (93). Outros algoritmos de ML são utilizados para o reconhecimento do discurso, o que é vantajoso para a realização e transcrição de áudios de relatórios de imagem médica, diminuindo o tempo necessário à sua elaboração (2, 3).

### **3. Limitações, desafios e perspectivas para a IA na radiologia**

Apesar dos algoritmos de ML e, em especial os de DL, serem muito promissores para a interpretação de dados muito complexos (as imagens médicas), têm ainda várias limitações e passos a dar para que possam ser implementados na prática clínica.

#### **3.1 Limitações técnicas**

Um dos mais importantes desafios à adoção da IA é o poder informático. Nos últimos anos, o desenvolvimento das GPU's permitiu o aumento e o avanço da utilização dos métodos de DL. Estas unidades de processamento permitiram não só o aumento do número de camadas das redes neuronais, mas também uma diminuição do consumo de tempo despendido no treino dos algoritmos. No entanto, para que ainda assim se melhorem os sistemas de ML e de DL, o desenvolvimento computacional continua a ser necessário. Outro aspeto prende-se com a existência de várias arquiteturas de DL que visam a melhoria da análise dos parâmetros de imagem e uma otimização do tempo de treino dos algoritmos. Não obstante as CNN's se revelem a melhor opção, a questão sobre qual será a arquitetura ideal permanece. O tipo e a quantidade de imagens têm de ser considerados e, frequentemente, em modalidades de imagem avançadas e volumétricas, como a TC e a RM, podem existir centenas de imagens por caso e, essas imagens contém informação tridimensional. Uma vez que é necessário um processamento de imagem muito complexo, vários métodos de DL têm sido desenvolvidos para a conversão da informação das imagens 3D em 2D ou, por outro lado, para a extração direta da informação de imagens 3D (CNN's 3D). No entanto, isto torna-se uma limitação, dado o grande tamanho da rede que é requerido e o elevado número de parâmetros envolvidos na extração das *features*. Voltamos ao aspeto inicial e verificamos que o aumento do poder informático volta a ser necessário, ainda que o uso de redes híbridas 2D/3D possa auxiliar a superar esta limitação. Por fim, como a aquisição de imagens não é padronizada, continua a ser um desafio para o computador reconhecer variáveis como o posicionamento, artefactos de movimento e variantes anatómicas (7, 11).

#### **3.2 Dados: treino e validação dos algoritmos**

Como já foi referido anteriormente, para que a tarefa dos sistemas de IA seja bem desempenhada na radiologia, é necessário que seja providenciada uma grande quantidade de dados dos doentes e de imagens médicas para o treino dos algoritmos.

O acesso, as questões relacionadas com a privacidade, proteção, garantia de uniformização e qualidade dos dados são fatores a considerar.

A criação de bases de dados úteis e consistentes requer um elevado investimento de tempo e dinheiro. Neste sentido, a colaboração entre as várias instituições que detém os conjuntos de dados para a construção das suas bases e para o acesso e partilha dos mesmos é essencial. Por outro lado, o acesso direto aos dados, por parte dos investigadores e das instituições, requer que a garantia da sua proteção seja assegurada, tópico que será analisado adiante. Atualmente, existem bases de dados com milhares de imagens que são usadas para o treino dos algoritmos. Assim, à medida que estas vão sendo criadas, a disponibilidade necessária deixa de ser o maior obstáculo. Porém, torna-se caro e difícil conseguir um elevado número de imagens médicas bem rotuladas e com boas anotações para treinar corretamente os algoritmos, de forma a que eles consigam reproduzir resultados sustentáveis. Este é o maior desafio. Transformar os relatórios de texto (que descrevem os achados e a interpretação das imagens pelos radiologistas) em anotações precisas ou rótulos estruturados de forma automatizada requer métodos sofisticados da chamada mineração de texto (processo de extração de um determinado conjunto de dados das informações consideradas relevantes para um dado propósito; no caso das bases de dados para os algoritmos de interpretação de imagem médica, consiste em filtrar as informações mais importantes dos relatórios radiológicos para a criação de rótulos e anotações para as imagens de treino). Neste campo, também os métodos de DL são atualmente usados. Por outro lado, a classificação das imagens médicas é, normalmente, feita de forma binária, ou seja, normal vs. anormal ou, objeto vs. fundo. No entanto, ambas as classes podem ser manifestamente heterogéneas: por exemplo, a categoria normal engloba tecido completamente normal, mas também contém várias categorias de achados benignos, que podem ser raros e podem ocasionalmente incluir uma grande variedade de artefactos de imagem. Isto pode levar a que os sistemas excluam bem as subclasses normais mais comuns, mas não excluam as subclasses mais raras. Uma solução seria transformar o algoritmo num sistema multiclasse, fornecendo-lhe anotações detalhadas de todas as subclasses possíveis, mas mais uma vez, a questão da disponibilidade limitada do tempo dos especialistas para rotular os dados torna o processo inviável. Aliado a este ponto, há o facto de que existem muito menos imagens anormais do que normais, o que se revela também uma limitação para o treino dos algoritmos, sendo que saber lidar com este desequilíbrio de classes (normal vs. anormal) é também motivo de investigação (64, 94, 95).

Outro fator a ter em conta é o desempenho da generalização do sistema. Num estudo publicado em 2018, Zech *et al.* não conseguiram demonstrar a generalização consistente de uma arquitetura CNN de um algoritmo de DL criado para a detecção de pneumonia quando testaram a rede com imagens de hospitais diferentes daqueles que detinham os dados usados para o treino (96). Diferenças na forma (qualidade, existência de ruídos e artefactos) e no equipamento utilizado para a aquisição de imagem, diferenças no processamento e na rotulagem/anotação dos dados e, distintas prevalências de doença entre os dados de treino e de teste podem contribuir para a falha da generalização. Ou seja, se as amostras usadas para o treino forem de uma distribuição de dados muito diferente daquela que é encontrada no mundo real (isto é, se não forem representativas da população alvo), então o desempenho da generalização da rede neuronal será menor do que o esperado.

Adicionalmente, os algoritmos de IA são suscetíveis a um fenómeno que tem o nome de *overfitting* (sobreajuste) cujo risco aumenta se o número de características extraídas for maior do que o conjunto de dados usado. Acontece que, se esse conjunto for pequeno, o sistema aprende a fazer as suas previsões com base em características idiossincráticas dos dados de treino, o que impede o ajuste de forma confiável a novos dados introduzidos à posteriori. Naturalmente, este risco diminui se os dados de treino forem suficientemente elevados em número de forma a que as suas características básicas estejam representadas, o que enfatiza a necessidade de grandes quantidades de dados para o treino dos algoritmos (12).

Neste sentido, existem alternativas que têm o potencial para superar as limitações relacionadas com a disponibilidade dos dados. Um dos métodos é o de *transfer learning*, considerado eficiente, especialmente, quando se dispõe de bases de dados relativamente limitadas. Comparando com a forma de aprendizagem “a partir do zero” que é a seguida na maioria dos algoritmos, com esta técnica, os parâmetros da rede aprendidos pelas funções do algoritmo treinado com um elevado número de dados são transferidos para um novo modelo. Considerando que alguns dados ou tarefas estão relacionados, é possível partilhar os parâmetros entre uma rede previamente treinada e outra ainda não treinada, de forma a melhorar o desempenho da segunda. Como tal, dispendo de menos dados, o treino do novo algoritmo será mais fácil. Um exemplo é a transferência dos parâmetros da rede de um algoritmo de classificação treinado com imagens não médicas para outro que será treinado com imagens médicas (20, 97). Outra técnica utilizada é aquilo a que se chama de *augmenting*, e consiste em maximizar a utilidade dos dados usados para o treino dos algoritmos. Visto que esses dados otimizados se tornam difíceis de obter num contexto clínico em que as imagens geradas

têm muitas vezes ruídos e artefactos, o uso de técnicas de “*morphing*” de imagem, nas quais a aparência de um objeto de interesse é ligeiramente modificada, pode aumentar o conjunto de dados de treino em vários múltiplos: há como que um aumento dos dados. Alguns exemplos incluem inclinar ligeiramente a imagem, pequenas modificações no contraste ou na resolução, inverter ou rodar a imagem, ajustar o zoom e alterar a localização de um achado dentro de uma imagem. Estas estratégias criam uma aparência ligeiramente diferente do mesmo objeto, mas podem tornar um algoritmo ML muito mais robusto com um conjunto de dados elaborados relativamente pequeno. Isto é importante na radiologia, onde é difícil obter bons dados. Apesar de existirem várias investigações que usam estes dois métodos, as redes neuronais de ML e DL standard são as que continuam a ter melhor desempenho, e o problema permanece (8).

Ainda no sentido de colmatar estas falhas relacionadas com a disponibilidade de dados bem rotulados, ultimamente, os investigadores têm manifestado interesse crescente no método de *unsupervised* ML. Esta técnica permite o treino dos algoritmos com dados (imagens) não rotulados, muito mais fáceis de obter do que os dados rotulados que requerem uma maior intervenção humana. Por outro lado, a aprendizagem a partir de dados cujos rótulos são desconhecidos parece ser mais eficiente, já que se aproxima mais do verdadeiro conceito de inteligência artificial e da própria inteligência humana (95).

### **3.3 Segurança e confiança**

Sabemos, então, que os métodos de ML utilizam representações hierárquicas interconectadas dos dados para gerarem previsões/soluções. Ora, a forma como os algoritmos atuam, ainda que possam existir imensas publicações que o tentem fazer, torna-se difícil de explicar, e as próprias previsões são difíceis de interpretar, bem como as *features* que são extraídas são, por vezes, invisíveis a olho nu. Há como que um mistério quanto ao funcionamento claro da inteligência artificial. Isto contribui para aquilo a que os especialistas chamam de “*black box*”, surgindo a seguinte questão: como podemos ter a certeza de que estes sistemas são seguros e como podemos confiar neles? Nesta perspetiva, para o doente, a responsabilidade será do radiologista. No entanto, se a máquina falha num determinado diagnóstico, há um problema (11, 94, 98). E isto leva-nos ao próximo ponto.



### 3.4 Legislação, ética e validação clínica

Apesar do rápido desenvolvimento da IA, um dos desafios prende-se com o facto do conhecimento clínico ser limitado por parte dos especialistas em IA e do conhecimento em IA ser limitado por parte dos radiologistas. É necessário que haja uma colaboração dos radiologistas para que possam adquirir conhecimento base sobre estes algoritmos, de forma a que consigam integrá-los no fluxo de trabalho, e isto pode implicar alterações no programa do internato em radiologia (99). Da mesma forma, os criadores dos sistemas têm de alojar formas de saber como ocorre a natural prática clínica para conceberem algoritmos que satisfaçam as necessidades dessa mesma prática. A envolvimento do radiologista no processo de criação pode também contribuir para que haja uma atenuação do problema referido anteriormente: a confiança dos doentes nestes sistemas (94).

Contudo, a existência de aprovação regulamentada que permita às máquinas fazer o trabalho de radiologistas experientes é um outro grande obstáculo. Nesta perspetiva, vale a pena ter em conta a aprovação por parte da *Food and Drug Administration* (FDA). Quando se pretende a comercialização e a utilização médica de um produto, é necessário que o mesmo demonstre equivalência ou superioridade em relação ao similar que já se encontra no mercado. Isto é possível pela obtenção de equivalência substancial pelo método 510(k) que confere autorização, por parte da FDA, para a comercialização. Caso não exista um comparador do produto no mercado, o comerciante necessita de obter uma aprovação de pré-comercialização (*premarket approval*, PMA). Os padrões de PMA são mais rigorosos do que a aprovação 510(k) e, geralmente, requerem uma revisão científica e regulatória, muitas vezes com necessidade de ensaios clínicos, para avaliar a segurança e eficácia dos dispositivos médicos. Todo o processo é mais caro e demorado, daí o método 510(k) ser preferido. Para além disto, as modificações a um dispositivo pré-comercializado podem exigir um suplemento de PMA ou uma nova submissão de PMA. Ora, deparamo-nos, então, com dois aspetos: a aprovação inicial de um algoritmo de ML para ser implementado no mercado e uma potencial nova aprovação de um algoritmo que possa ter que ser alterado, que exigiria mais tempo e dinheiro (100). Isto tem implicações significativas para os sistemas de IA, uma vez que, por um lado, sabe-se que os algoritmos melhoram à medida que leem mais dados, surgindo a dúvida se devem ou não permanecer no mercado estáticos e com os códigos congelados (8); por outro lado, até há pouco tempo atrás, não existia nenhum dispositivo autorizado e usado como um substituto dos radiologistas, o que fazia com que os fabricantes tivessem de despender muito tempo e dinheiro para tentarem obter uma PMA. Contudo, existem já sistemas de IA utilizados

na prática clínica em procedimentos médicos aprovados pela FDA. Em Junho de 2019, um artigo publicado no site *The Medical Futurist* deu a conhecer o número e a área de aplicação de algoritmos aprovados até ao momento (101). Perante os milhares de artigos publicados ao longo da última década, e mesmo com uma enorme quantidade de algoritmos disponíveis, são relativamente poucos os aprovados pela instituição americana, o que reflete a dificuldade em protocolar e em validar os sistemas. No entanto, é possível verificar que, após a aprovação, em 2017, do primeiro algoritmo de IA para a deteção de fibrilação auricular, o número de aprovações tem aumentado a cada ano, e a tendência é que sejam cada vez mais. Embora na Europa, exista a Agência Europeia do Medicamento, a FDA é o único órgão regulador com instrumentos eficientes capazes de conferir credibilidade e precisão aos algoritmos para fins médicos. Da lista disponibilizada salientam-se as seguintes aplicações dos sistemas de IA na imagem médica: utilização de tele-ecografia que permite que um médico num local remoto consiga visualizar a realização do exame e, ao mesmo tempo, trocar pareceres com o radiologista operador; análise de RM cardíaca; análise de nódulos da tiróide em ecografia; deteção e diagnóstico de cancro da mama com sistema CADx para imagens de RM; deteção de AVC em imagens de TC; deteção de lesões neoplásicas no fígado e nos pulmões em imagens de TC e RM; interpretação de RM cerebral; deteção de fraturas do punho em imagens de radiografia; auxílio na análise do fluxo sanguíneo cerebral através de um sistema de eco-doppler transcraniano; quantificação da concentração de ferro no fígado; triagem de doentes com hemorragia cerebral aguda e deteção de embolismo pulmonar em imagens de TC; classificação da densidade mamária e deteção de cancro da mama em imagens de mamografia; auxílio na deteção de derrame pleural em radiografias do tórax; redução de ruído em imagens de TC; melhoria da qualidade na aquisição imagens de RM. Existem ainda algoritmos aprovados nas áreas de cardiologia, psiquiatria, endocrinologia, geriatria, oftalmologia, neurologia, ortopedia, anatomia patológica e oncologia. A radiologia e a cardiologia destacam-se como sendo as especialidades onde já existem mais algoritmos de ML aprovados. De facto, tanto a visão computacional como a imagem médica têm todos os instrumentos para que a IA possa prosperar nestas áreas. Em 2018, a FDA aprovou o primeiro sistema de IA para a deteção de retinopatia diabética que pode ser usado para diagnóstico sem a intervenção de um clínico humano. Podemos considerar esse momento como um ponto de viragem. Se, até aí os fabricantes dos algoritmos não tinham um produto no mercado ao qual pudessem comparar os seus sistemas, a partir desse momento, não só os mesmos passaram a ser comercializados, como também o regulamento de aprovação da FDA passou a prever requisitos necessários para novas submissões de PMA de sistemas equivalentes. Para além disto, em junho do mesmo

ano, a FDA reclassificou os dispositivos de análise de imagem médica com sistemas CAD em classe II, em vez de classe III (ou seja, passou a considera-los dispositivos de médio risco, e não de alto risco) (14). Espera-se que daqui em diante haja cada vez menos dificuldade na obtenção da aprovação.

Para além da regulação, permanecem desafios éticos e legais. A *black box* e a falta de transparência na forma como os resultados são obtidos têm implicações médico-legais. Estas prendem-se, essencialmente, com o facto de quem assumirá a responsabilidade no momento em que as técnicas de ML e DL falharem a interpretação dos exames imagiológicos. Os erros da IA podem surgir quando fatores de confusão presentes nos dados são relacionados com entidades patológicas, levando a uma classificação errónea. Por exemplo, se um algoritmo de DL falha na identificação de um nódulo pulmonar, a responsabilidade será do fabricante do algoritmo ou do radiologista? Perante as circunstâncias atuais, a responsabilidade seria do médico. Contudo, se o sistema é aprovado e classificado como tendo autonomia suficiente para o diagnóstico e interpretação da imagem médica, sendo substituto do radiologista nessa tarefa, isto pode levar a processos judiciais, e terá de haver uma legislação que defina claramente o culpado nestas eventuais situações adversas. Mesmo com a definição dessa legislação, uma decisão judicial desfavorável poderia alterar todo o panorama das aplicações da IA na radiologia e prejudicar gravemente o progresso desta tecnologia (2, 7). As questões éticas também são levantadas e, este ano, num artigo publicado no *Canadian Association of Radiologists Journal*, vários membros pertencentes a associações de radiologia, onde se incluem o *American College of Radiology*, a *European Society of Radiology*, a *Radiologic Society of North America*, a *Society for Imaging Informatics in Medicine*, a *European Society of Medical Imaging Informatics*, a *Canadian Association of Radiologists*, e a *American Association of Physicists in Medicine* revelaram a preocupação quanto a uma série de aspetos éticos a ter em conta para a adoção da IA na radiologia (102). Os investigadores consideraram existirem 3 pontos-chave onde deve incidir a análise: ética na manipulação dos dados para o treino das imagens, ética no treino dos algoritmos e, por fim, ética na prática clínica que envolve os sistemas. Tal como já mencionado anteriormente, impõem-se os fatores de propriedade, privacidade e proteção dos dados para o treino dos algoritmos e, obviamente, o consentimento informado dos indivíduos aos quais pertencem as imagens. A objetividade e a transparência devem ser garantidas, e deve ser minimizada a lacuna entre as instituições que têm e aquelas que não têm os recursos para gerir grandes conjuntos de dados e a capacidade para fornecer direitos morais e de acesso aos mesmos. Por outro lado, os radiologistas e os fabricantes têm, em si, inerente o

dever de usarem os dados apenas para extrair informação sobre os doentes e as suas doenças, e melhorar a prática clínica, e não para os utilizarem de forma não ética, como a sua comercialização. Além disto, há que considerar o viés que pode ser obtido no treino dos algoritmos, que advém não só do facto da máquina poder ser enganada por diferenças subtis nos dados ou devido às diferentes técnicas de obtenção dos mesmos, mas também o que resulta da anotação/rotulagem dos dados e da extração das suas características, processos esses que podem ser influenciados pelo preconceito e pela discriminação relacionadas com o sexo, a raça, a orientação sexual e fatores ambientais do foro social e económico que, no limite, podem levar a uma promoção de apenas alguns subconjuntos de dados. Assim, o viés humano tem implicação também na forma como os algoritmos são treinados. Para além disto, é importante ter a maior transparência possível quanto à forma como as decisões dos algoritmos são tomadas, e isto leva-nos à questão da *black box*: se um algoritmo falhar ou contribuir para um evento clínico adverso, é necessária capacidade para entender o porquê do erro e a forma como o sistema lá chegou. Porém, é discutível até que ponto um modelo deve ser transparente, uma vez que demasiada transparência pode torná-lo mais suscetível a ataques maliciosos ou comprometer a privacidade ao revelar dados pessoais ocultos em conjuntos de dados subjacentes. Quanto à prática clínica, Geis *et al.* consideram que se deve ter em conta o facto de haver o risco dos clínicos humanos privilegiarem as decisões tomadas pelas máquinas, ignorando dados ou decisões humanas contrárias, o que pode levar à não perceção ou desconsideração de uma verdadeira falha no algoritmo e, no limite, à implementação de uma decisão mal tomada pela máquina, mesmo havendo evidências contrárias. Por último, há que ter em conta que hospitais mais pequenos e com poucos departamentos de investigação podem não ter a tecnologia, capacidade e recursos suficientes para implementar e gerir sistemas complexos de IA. Certamente, parte dos sistemas de IA na radiologia serão (e são) propriedade e desenvolvidos por grandes entidades académicas ou privadas de saúde, companhias de seguros ou grandes empresas com conhecimentos técnicos e de ciência de dados, mas pouco conhecimento histórico no domínio da radiologia, o que pode exacerbar as desigualdades no acesso à IA.

## DISCUSSÃO E CONCLUSÃO

---

Através de uma síntese da literatura científica mais atual, este artigo pretende elucidar o leitor acerca dos conceitos fundamentais de IA, nomeadamente ML e DL, e dar a conhecer o estado da arte quanto às aplicações atuais dos sistemas de IA na radiologia, bem como a sua perspetiva futura e os limites impostos à implementação dos algoritmos na prática clínica.

É notório que a investigação em IA na radiologia tem aumentado de uma forma exponencial ao longo dos últimos anos. Dado que extraem automaticamente as *features* mais relevantes dos dados, atualmente, os algoritmos de DL (especialmente os que contemplam CNN's) dominam o plano de pesquisa e têm encontrado grande aceitação nos vários sub-campos da análise de imagem médica que incluem a segmentação, a deteção, a quantificação e a classificação. Diversos estudos têm incorporado estes algoritmos nos sistemas CAD de modo a que a performance na deteção, classificação e diagnóstico de várias doenças sejam melhoradas, com especial foco para a diminuição de falsos positivos. Sendo utilizadas as diferentes modalidades de imagem da radiologia, os campos da imagem torácica, de mama e, neuroimagem são aqueles nos quais os investigadores têm trabalhado mais, salientando-se o cancro (pulmão, mama e sistema nervoso) como a principal doença alvo nos artigos publicados. A correta segmentação de órgãos e lesões e, a sua classificação para o diagnóstico em diferentes tipos de imagem, usando algoritmos de ML e DL, é um dos objetivos primordiais. Outros estudos revelam ainda potenciais aplicações da IA noutras doenças pulmonares, mamárias, neuropsiquiátricas, e em áreas como a imagem músculo-esquelética ou, mais recentemente, a emergência médica. Quanto a esta última, a triagem de doentes com achados que os colocam em risco de vida (como a hemorragia cerebral, o EAM, ou o pneumotórax), através da priorização das listas de trabalho, é um desafio para que seja possível a sua rápida monitorização, diagnóstico e tratamento. A redução do ruído e dos artefactos de imagem e, a aquisição e reconstrução de imagem através destes algoritmos (por exemplo, a obtenção de imagens de TC a partir de imagens de RM), também têm sido investigadas e revelado resultados promissores. O alcance dos sistemas de IA pode ir para além das 'simples' tarefas de segmentação, deteção e classificação. A utilização de algoritmos de DL na análise radiómica e radiogenómica, para a extração de biomarcadores quantitativos, poderá revolucionar a abordagem quanto ao prognóstico e tratamento dos doentes e de muitas doenças, e prever os *outcomes* de doença com elevada acuidade. Por outro lado, a compreensão dos fenótipos baseados em imagens e genética, e a sua integração nos regimes de rastreio

e tratamento têm o potencial para, por exemplo, evitar biópsias desnecessárias, e fornecer biópsias virtuais quando as biópsias reais não são passíveis de ser realizadas.

O rápido desenvolvimento da IA a que se assistiu na última década mostra que esta tecnologia veio para ficar, e isto deve-se aos avanços nos algoritmos, ao aumento do poder computacional e ao acesso a grandes bases de dados. Contudo, existem ainda muitas limitações e desafios a considerar para a implementação dos seus sistemas na prática clínica diária. As dificuldades prendem-se, essencialmente, com implicações técnicas, treino dos algoritmos e validação dos dados e, constrangimentos relacionados com a segurança, confiança, regulamentação, legislação e ética no uso dos algoritmos de IA.

O advento das GPU's permitiu a utilização crescente de algoritmos de DL, dada a possibilidade de aumentar o número de camadas nas redes neuronais e de menos tempo ser despendido no treino dos algoritmos. No entanto, o aumento do poder informático continua a ser necessário e, as limitações técnicas relacionadas com as diferenças de arquitetura das CNN's entre as diferentes investigações, e com a complexidade das imagens radiológicas têm de ser superadas.

O acesso a grandes conjuntos de dados para o treino dos sistemas tem deixado de ser um problema major graças à construção crescente de bases de dados. Porém, a anotação correta dos mesmos é um dos maiores obstáculos, e a mineração de dados para gerar rótulos e anotações estruturadas de uma forma automatizada desempenha um papel importante neste aspeto. A variabilidade nos achados entre as diferentes patologias (intra- e inter-individual), a variação anatómica entre as populações subjacentes e, os ruídos e artefactos, levam a que as imagens médicas sejam bastante heterogéneas. Aliado a isto, há que ter em conta o facto de haver mais imagens classificadas como normais do que como anormais, o que se torna uma dificuldade para o treino dos algoritmos. A padronização na aquisição das imagens e a garantia da generalização dos sistemas têm de ser asseguradas e, o risco de sobreajuste tem de ser encarado. Todos estes fatores enfatizam a necessidade de um elevado número de imagens médicas bem etiquetadas. Técnicas como *transfer learning* e *data augmenting*, e a adoção do método de *unsupervised* ML podem ajudar a minimizar estes problemas, ainda assim, há um longo caminho a percorrer. Neste sentido, a colaboração multi-institucional é fulcral para a criação e partilha de dados úteis para o treino dos sistemas.

A IA é complexa e carrega potenciais armadilhas e preconceitos inerentes. As questões relacionadas com a segurança e confiança nestes sistemas e, com a legislação, ética e regulamentação podem ser consideradas, de longe, como o maior desafio com que a comunidade médica e científica se deparam. A caixa negra e a falta de transparência

associada à forma como os algoritmos geram as suas próprias previsões coloca em causa a legitimidade dos sistemas para atuarem na clínica. Deste modo, a integração e aquisição de conhecimento sobre IA por parte dos clínicos e, de conhecimento clínico por parte dos fabricantes dos algoritmos, poderá contribuir para dissolver esta incerteza e aumentar a confiança dos utilizadores. Contudo, é necessária uma regulamentação por entidades competentes que confirmem credibilidade às máquinas que, em teoria, estão habilitadas a realizar a tarefa de um radiologista treinado. Alguns dos sistemas de IA para várias aplicações referidas neste artigo encontram-se já aprovadas pela FDA. Se, até ao momento a validação por parte da instituição americana era algo quase impossível, o número de aprovações crescente nos últimos dois anos e as alterações já efetuadas à regulamentação, que colocam menos entrave no acesso dos sistemas ao mercado, leva-nos a crer que, ao contrário do que se pensava há pouco tempo atrás, a dificuldade para a obtenção de aprovação seja menos complicada, embora continuem a haver requisitos situados numa linha muito estreita de regulação que tenham de ser preenchidos.

Ainda que o potencial da IA seja a diminuição do erro, a introdução inevitável do viés humano (devido ao facto da maioria dos sistemas utilizar *supervised ML*) leva a que o uso generalizado de máquinas inteligentes e autónomas na radiologia possa aumentar o risco de erros sistemáticos e, conseqüentemente, a possibilidade de danos com alto impacto e a ampliação de questões éticas e legais complexas. Quanto à tutela da responsabilidade final perante uma eventual situação adversa, permanece a ausência de legislação aplicável, urgindo a necessidade de se encontrar um consenso neste tema. Da mesma forma, a garantia da proteção dos dados usados para o treino dos algoritmos para que os interesses dos indivíduos a quem pertencem os dados sejam salvaguardados deve ser assegurada; no entanto, estamos perante uma situação não isenta de riscos, já que a demasiada transparência quanto ao modo de atuação dos sistemas de IA pode torna-los suscetíveis a ataques maliciosos e levar à revelação dos dados. Para além disto, dado os elevados custos que acarretam, existe ainda o perigo de haver desigualdades no acesso aos sistemas de IA, cujo potencial de diagnóstico é maior do que os métodos tradicionais, privilegiando determinadas populações, em detrimento das mais vulneráveis. Tudo isto destaca a importância de que o uso ético da IA na radiologia deve promover o bem-estar, minimizar os danos e garantir que os benefícios e prejuízos sejam distribuídos entre as partes interessadas de forma justa, já que, como é óbvio, não existe nem existirá um sistema totalmente perfeito. A implementação da IA deve respeitar os direitos e as liberdades, incluindo a dignidade e privacidade, devendo ser projetada para a transparência e máxima confiabilidade.

Assim, verificamos a necessidade para o desenvolvimento de legislação e códigos de ética e prática para a integração da IA na clínica que promovam o seu uso para o bem comum, e impeçam a sua utilização para fins não éticos e de proveito individual, o que requer entendimento entre diversas entidades médicas (nomeadamente, as associações de radiologia) e científicas e, o legislador.

Chegando até aqui, é preciso admitir que os fabricantes dos sistemas e os radiologistas estão a reconhecer os aspetos éticos e legais sobre a IA ao mesmo tempo que a inventam e tentam implementar; e o desenvolvimento e as alterações provocadas pela IA, e a consequente resposta da sociedade médica, estão a evoluir a uma velocidade difícil de acompanhar, entender e, sobretudo, difícil de administrar. Como tal, havendo pouca experiência no uso da IA em diversos ambientes clínicos, levar a cabo a investigação adequada para entender como melhor implementar estes sistemas no fluxo de trabalho clínico é um dos objetivos que deve ser priorizado. Não é só a performance dos sistemas em cumprirem a tarefa da deteção ou da classificação que deve ser tida em conta, e o desempenho da máquina não deve ser avaliado apenas pela comparação com os melhores especialistas, como acontece na maioria da literatura atual. Ao mesmo tempo, o comportamento ético e segurança, bem como a reprodutibilidade, o verdadeiro ganho de tempo e o retorno em termos de custos são fatores a ser testados e considerados antes da aplicação clínica. A validação destes sistemas a uma grande escala populacional é uma necessidade que ainda os estudos atuais não foram capazes de satisfazer.

Com base em tudo o que foi anteriormente referido e discutido, estamos quase em condições de responder à questão colocada no título. Todavia, olhemos ainda aos seguintes aspetos. As funções de um radiologista vão além da elaboração de relatórios de imagens médicas. Incluem também, a comunicação de descobertas e a sua discussão, a garantia e melhoria da qualidade dos exames, o papel importante na educação médica, e outros procedimentos que não podem ser realizados por um computador. A capacidade de fornecer uma boa interpretação e integração de resultados complexos, aliada ao julgamento médico e à sabedoria que um radiologista experiente possui é difícil de quantificar e, ainda mais, difícil de simular com um *software* informático. Estas metodologias de análise semiológica das características das imagens num contexto anatómico e fisiopatológico específico foram sendo refinadas ao longo dos séculos de pensamento médico e científico, e permitem gerar soluções que ainda estão inacessíveis à IA. Uma vez que a cada dia existem muitas descobertas relacionadas com os aspetos das doenças e um vasto conjunto de diagnósticos que um radiologista treinado é capaz de distinguir, serão necessários milhares de algoritmos



que funcionem bem e, cada um deles irá necessitar de uma base de dados bem definida e validada. Além disso, vários criadores de *software* irão desenvolver vários sistemas e é um desafio integrar todos os diferentes sistemas nas mais diversas aplicações. No entanto, existem resultados que não podem ser desvalorizados e a aposta no desenvolvimento desta tecnologia deve e está a ser no sentido de um percurso em que os radiologistas e os sistemas de IA caminharão lado-a-lado e, não em desigualdade.

Em suma, embora as máquinas cometam erros, é provável que tomem decisões mais eficientes e com mais consistência do que os humanos e, em alguns casos, contradigam os radiologistas e sejam comprovadamente corretas; os sistemas de IA irão ampliar o conhecimento dos radiologistas, extraindo ainda mais informações para fazer previsões melhores ou inteiramente novas sobre os doentes e as suas doenças; as conclusões sobre as imagens serão retiradas por radiologistas humanos em conjunto com máquinas inteligentes e autónomas. Esta entreaajuda permitirá libertar o radiologista para o desempenho de tarefas para as quais, geralmente, tem menos disponibilidade, como a participação num maior número de reuniões multidisciplinares de decisão terapêutica e a maior dedicação a tarefas de *management* e de investigação clínica. Tudo isto contribuirá para uma maior relação do radiologista com o doente e a sua doença.

Permanece a dúvida se haverá algum dia em que a performance dos sistemas não necessitará da vigilância do radiologista. A IA é um mundo quase esotérico que, numa determinada perspectiva, impõe limites à mente e ao pensamento humano. Todavia, é certo que estamos perante o início de uma mudança de paradigma na imagem médica e, mais concretamente, na radiologia. Mesmo que, porventura, haja quem seja mais cético quanto a todos estes avanços, as adaptações já ocorrem e continuarão a ser implementadas. De facto, um dos motores da evolução é, justamente, a constante adaptação.

Finalmente, é possível concluir que a IA tem todas as ferramentas para ser amiga. Amiga do radiologista e amiga do doente. Mas, lembremo-nos de que por mais inteligentes que possam ser as máquinas, estas não trabalham num cenário que habitualmente um clínico atua, isto é, num contexto socioeconómico de doentes e familiares de doentes. Para já, esta não é uma tarefa da inteligência artificial, mas sim da humana, e talvez este seja o maior desafio que os radiologistas e médicos em geral terão de enfrentar ou, pelo contrário, o maior obstáculo que a própria IA terá de superar.

## **AGRADECIMENTOS**

---

Ao Professor Doutor Luís Semedo, pela prontidão com que aceitou orientar o meu trabalho, pela disponibilidade permanente, pelo apoio, espírito crítico e pelos esclarecimentos facultados ao longo deste percurso.

Aos meus pais, pilares na minha vida e na minha formação, pelo apoio incondicional e pelos valores e ensinamentos que me transmitem diariamente.

À minha irmã, Isabel, por estar sempre presente.

À minha família, por todos os conselhos, motivação e dedicação.

Por fim, aos meus amigos, a família que escolhi, por todas as partilhas e vivências, conselhos, palavras de incentivo e pela compreensão e carinho diários.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

1. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017;69s:S36-s40.
2. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, et al. Deep Learning in Radiology. *Acad Radiol*. 2018;25(11):1472-80.
3. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol*. 2017;18(4):570-84.
4. Kulikowski CA. An Opening Chapter of the First Generation of Artificial Intelligence in Medicine: The First Rutgers AIM Workshop, June 1975. *Yearb Med Inform*. 2015;10(1):227-33.
5. Mayo RC, Leung J. Artificial intelligence and deep learning - Radiology's next frontier? *Clin Imaging*. 2018;49:87-8.
6. Syeda-Mahmood T. Role of Big Data and Machine Learning in Diagnostic Decision Support in Radiology. *J Am Coll Radiol*. 2018;15(3 Pt B):569-76.
7. Nawrocki T, Maldjian PD, Slasky SE, Contractor SG. Artificial Intelligence and Radiology: Have Rumors of the Radiologist's Demise Been Greatly Exaggerated? *Academic Radiology*. 2018;25(8):967-72.
8. Kohli M, Prevedello LM, Filice RW, Geis JR. Implementing Machine Learning in Radiology Practice and Research. *AJR Am J Roentgenol*. 2017;208(4):754-60.
9. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine Learning for Medical Imaging. *Radiographics*. 2017;37(2):505-15.
10. Abiyev RH, Ma'aitah MKS. Deep Convolutional Neural Networks for Chest Diseases Detection. *J Healthc Eng*. 2018;2018:4168538.
11. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical Image Analysis using Convolutional Neural Networks: A Review. *J Med Syst*. 2018;42(11):226.
12. Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, et al. Demystification of AI-driven medical image interpretation: past, present and future. *Eur Radiol*. 2019;29(3):1616-24.
13. Karako K, Chen Y, Tang W. On medical application of neural networks trained with various types of data. *Biosci Trends*. 2019;12(6):553-9.

14. Gao Y, Geras KJ, Lewin AA, Moy L. New Frontiers: An Update on Computer-Aided Diagnosis for Breast Imaging in the Age of Artificial Intelligence. *AJR Am J Roentgenol*. 2019;212(2):300-7.
15. van Ginneken B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiol Phys Technol*. 2017;10(1):23-32.
16. Zarshenas A, Liu J, Forti P, Suzuki K. Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. *Med Phys*. 2019;46(5):2232-42.
17. Moses DA, Dawes L, Sammut C, Zrimec T. Peripheral bronchial identification on chest CT using unsupervised machine learning. *Int J Comput Assist Radiol Surg*. 2018;13(9):1379-95.
18. Kim BC, Yoon JS, Choi JS, Suk HI. Multi-scale gradual integration CNN for false positive reduction in pulmonary nodule detection. *Neural Netw*. 2019;115:1-10.
19. Murphy A, Skalski M, Gaillard F. The utilisation of convolutional neural networks in detecting pulmonary nodules: a review. *Br J Radiol*. 2018;91(1090):20180028.
20. Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. *Biomed Eng Online*. 2018;17(1):113.
21. Khan SA, Nazir M, Khan MA, Saba T, Javed K, Rehman A, et al. Lungs nodule detection framework from computed tomography images using support vector machine. *Microsc Res Tech*. 2019;82(8):1256-66.
22. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *Lancet Respir Med*. 2018;6(11):837-45.
23. Kim GB, Jung KH, Lee Y, Kim HJ, Kim N, Jun S, et al. Comparison of Shallow and Deep Learning Methods on Classifying the Regional Pattern of Diffuse Lung Disease. *J Digit Imaging*. 2018;31(4):415-24.
24. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, et al. Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Netw Open*. 2019;2(3):e191095.

25. Yates EJ, Yates LC, Harvey H. Machine learning "red dot": open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification. *Clin Radiol*. 2018;73(9):827-31.
26. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med*. 2018;15(11):e1002686.
27. Dande P, Samant P. Acquaintance to Artificial Neural Networks and use of artificial intelligence as a diagnostic tool for tuberculosis: A review. *Tuberculosis (Edinb)*. 2018;108:1-9.
28. Jaeger S, Juarez-Espinosa OH, Candemir S, Poostchi M, Yang F, Kim L, et al. Detecting drug-resistant tuberculosis in chest radiographs. *Int J Comput Assist Radiol Surg*. 2018;13(12):1915-25.
29. Gao XW, Qian Y. Prediction of Multidrug-Resistant TB from CT Pulmonary Images Based on Deep Learning Techniques. *Mol Pharm*. 2018;15(10):4326-35.
30. Katzen J, Dodelzon K. A review of computer aided detection in mammography. *Clin Imaging*. 2018;52:305-9.
31. US Food and Drug Administration. Premarket Approval (PMA); [Visitado a 18 de outubro de 2019] Disponível em: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=380594>
32. Burt JR, Torosdagli N, Khosravan N, RaviPrakash H, Mortazi A, Tissavirasingham F, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *Br J Radiol*. 2018;91(1089):20170545.
33. Cai H, Huang Q, Rong W, Song Y, Li J, Wang J, et al. Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms. *Comput Math Methods Med*. 2019;2019:2717454.
34. Herent P, Schmauch B, Jehanno P, Dehaene O, Saillard C, Balleyguier C, et al. Detection and characterization of MRI breast lesions using deep learning. *Diagn Interv Imaging*. 2019;100(4):219-25.
35. Fujioka T, Kubota K, Mori M, Kikuchi Y, Katsuta L, Kasahara M, et al. Distinction between benign and malignant breast masses at breast ultrasound using deep learning method with convolutional neural network. *Jpn J Radiol*. 2019;37(6):466-72.

36. Huang Y, Han L, Dou H, Luo H, Yuan Z, Liu Q, et al. Two-stage CNNs for computerized BI-RADS categorization in breast ultrasound images. *Biomed Eng Online*. 2019;18(1):8.
37. Yassin NIR, Omran S, El Houbay EMF, Allam H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput Methods Programs Biomed*. 2018;156:25-45.
38. Codari M, Schiaffino S, Sardanelli F, Trimboli RM. Artificial Intelligence for Breast MRI in 2008-2018: A Systematic Mapping Review. *AJR Am J Roentgenol*. 2019;212(2):280-92.
39. Rodriguez-Ruiz A, Lang K, Gubern-Merida A, Teuwen J, Broeders M, Gennaro G, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol*. 2019;29(9):4825-32.
40. Guha Roy A, Conjeti S, Navab N, Wachinger C. QuickNAT: A fully convolutional network for quick and accurate segmentation of neuroanatomy. *Neuroimage*. 2019;186:713-27.
41. Wasserthal J, Neher P, Maier-Hein KH. TractSeg - Fast and accurate white matter tract segmentation. *Neuroimage*. 2018;183:239-53.
42. Mateos-Perez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. *Neuroimage Clin*. 2018;20:506-22.
43. Li F, Liu M. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Comput Med Imaging Graph*. 2018;70:101-10.
44. Dolz J, Desrosiers C, Ben Ayed I. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *Neuroimage*. 2018;170:456-70.
45. Pagnozzi AM, Conti E, Calderoni S, Fripp J, Rose SE. A systematic review of structural MRI biomarkers in autism spectrum disorder: A machine learning perspective. *Int J Dev Neurosci*. 2018;71:68-82.
46. Sakai K, Yamada K. Machine learning studies on major brain diseases: 5-year trends of 2014-2018. *Jpn J Radiol*. 2019;37(1):34-72.

47. Walter M, Alizadeh S, Jamalabadi H, Lueken U, Dannlowski U, Walter H, et al. Translational machine learning for psychiatric neuroimaging. *Prog Neuropsychopharmacol Biol Psychiatry*. 2019;91:113-21.
48. Zhang X, He L, Chen K, Luo Y, Zhou J, Wang F. Multi-View Graph Convolutional Network and Its Applications on Neuroimage Analysis for Parkinson's Disease. *AMIA Annu Symp Proc*. 2018;2018:1147-56.
49. Villanueva-Meyer JE, Chang P, Lupo JM, Hess CP, Flanders AE, Kohli M. Machine Learning in Neurooncology Imaging: From Study Request to Diagnosis and Treatment. *AJR Am J Roentgenol*. 2019;212(1):52-6.
50. Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep Learning in Neuroradiology. *AJNR Am J Neuroradiol*. 2018;39(10):1776-84.
51. Juan-Albarracin J, Fuster-Garcia E, Garcia-Ferrando GA, Garcia-Gomez JM. ONCOhabitats: A system for glioblastoma heterogeneity assessment through MRI. *Int J Med Inform*. 2019;128:53-61.
52. Ma C, Luo G, Wang K. Concatenated and Connected Random Forests With Multiscale Patch Driven Active Contour Model for Automated Brain Tumor Segmentation of MR Images. *IEEE Trans Med Imaging*. 2018;37(8):1943-54.
53. Iqbal S, Ghani Khan MU, Saba T, Mehmood Z, Javaid N, Rehman A, et al. Deep learning model integrating features and novel classifiers fusion for brain tumor segmentation. *Microsc Res Tech*. 2019;82(8):1302-15.
54. Laukamp KR, Thiele F, Shakirin G, Zopfs D, Faymonville A, Timmer M, et al. Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI. *Eur Radiol*. 2019;29(1):124-32.
55. Hoseini F, Shahbahrani A, Bayat P. An Efficient Implementation of Deep Convolutional Neural Networks for MRI Segmentation. *J Digit Imaging*. 2018;31(5):738-47.
56. Jun Y, Eo T, Kim T, Shin H, Hwang D, Bae SH, et al. Deep-learned 3D black-blood imaging using automatic labelling technique and 3D convolutional neural networks for detecting metastatic brain tumors. *Sci Rep*. 2018;8(1):9450.
57. Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med*. 2018;24(9):1337-41.

58. Chang PD, Kuoy E, Grinband J, Weinberg BD, Thompson M, Homo R, et al. Hybrid 3D/2D Convolutional Neural Network for Hemorrhage Evaluation on Head CT. *AJNR Am J Neuroradiol*. 2018;39(9):1609-16.
59. Zhang R, Zhao L, Lou W, Abrigo JM, Mok VCT, Chu WCW, et al. Automatic Segmentation of Acute Ischemic Stroke From DWI Using 3-D Fully Convolutional DenseNets. *IEEE Trans Med Imaging*. 2018;37(9):2149-60.
60. Majumdar A, Brattain L, Telfer B, Farris C, Scalera J. Detecting Intracranial Hemorrhage with Deep Learning. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:583-7.
61. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet*. 2018;392(10162):2388-96.
62. Mannil M, von Spiczak J, Manka R, Alkadhi H. Texture Analysis and Machine Learning for Detecting Myocardial Infarction in Noncontrast Low-Dose Computed Tomography: Unveiling the Invisible. *Invest Radiol*. 2018;53(6):338-43.
63. Winkel DJ, Heye T, Weikert TJ, Boll DT, Stieltjes B. Evaluation of an AI-Based Detection Software for Acute Findings in Abdominal Computed Tomography Scans: Toward an Automated Work List Prioritization of Routine CT Examinations. *Invest Radiol*. 2019;54(1):55-9.
64. Taylor AG, Mielke C, Mongan J. Automated detection of moderate and large pneumothorax on frontal chest X-rays using deep convolutional neural networks: A retrospective study. *PLoS Med*. 2018;15(11):e1002697.
65. Marsousi M, Plataniotis KN, Stergiopoulos S. Kidney Detection in 3-D Ultrasound Imagery via Shape-to-Volume Registration Based on Spatially Aligned Neural Network. *IEEE J Biomed Health Inform*. 2019;23(1):227-42.
66. Kamiya N. Muscle Segmentation for Orthopedic Interventions. *Adv Exp Med Biol*. 2018;1093:81-91.
67. Lee JH, Baek JH, Kim JH, Shim WH, Chung SR, Choi YJ. Deep Learning-Based Computer-Aided Diagnosis System for Localization and Diagnosis of Metastatic Lymph Nodes on Ultrasound: A Pilot Study. *Thyroid*. 2018;28(10):1332-8.



68. Zhou LQ, Wang JY, Yu SY, Wu GG, Wei Q, Deng YB, et al. Artificial intelligence in medical imaging of the liver. *World J Gastroenterol*. 2019;25(6):672-82.
69. Chlebus G, Schenk A, Moltz JH, van Ginneken B, Hahn HK, Meine H. Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Sci Rep*. 2018;8(1):15497.
70. Yasaka K, Akai H, Kunimatsu A, Abe O, Kiryu S. Deep learning for staging liver fibrosis on CT: a pilot study. *Eur Radiol*. 2018;28(11):4578-85.
71. Couteaux V, Si-Mohamed S, Renard-Penna R, Nempont O, Lefevre T, Popoff A, et al. Kidney cortex segmentation in 2D CT with U-Nets ensemble aggregation. *Diagn Interv Imaging*. 2019;100(4):211-7.
72. Zavala Bojorquez JA, Jodoin PM, Bricq S, Walker PM, Brunotte F, Lalande A. Automatic classification of tissues on pelvic MRI based on relaxation times and support vector machine. *PLoS One*. 2019;14(2):e0211944.
73. Wang B, Lei Y, Tian S, Wang T, Liu Y, Patel P, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Med Phys*. 2019;46(4):1707-18.
74. Wang J, Lu J, Qin G, Shen L, Sun Y, Ying H, et al. Technical Note: A deep learning-based autosegmentation of rectal tumors in MR images. *Med Phys*. 2018;45(6):2560-4.
75. Wang M, Xie P, Ran Z, Jian J, Zhang R, Xia W, et al. Full convolutional network based multiple side-output fusion architecture for the segmentation of rectal tumors in magnetic resonance images: A multi-vendor study. *Med Phys*. 2019;46(6):2659-68.
76. Gordon MN, Hadjiiski LM, Cha KH, Samala RK, Chan HP, Cohan RH, et al. Deep-learning convolutional neural network: Inner and outer bladder wall segmentation in CT urography. *Med Phys*. 2019;46(2):634-48.
77. Langkvist M, Jendeberg J, Thunberg P, Loutfi A, Liden M. Computer aided detection of ureteral stones in thin slice computed tomography volumes using Convolutional Neural Networks. *Comput Biol Med*. 2018;97:153-60.
78. Brattain LJ, Telfer BA, Dhyani M, Grajo JR, Samir AE. Machine learning for medical ultrasound: status, methods, and future opportunities. *Abdom Radiol (NY)*. 2018;43(4):786-99.

79. Higaki T, Nakamura Y, Tatsugami F, Nakaura T, Awai K. Improvement of image quality at CT and MRI using deep learning. *Jpn J Radiol.* 2019;37(1):73-80.
80. Wang Y, Liao Y, Zhang Y, He J, Li S, Bian Z, et al. Iterative quality enhancement via residual-artifact learning networks for low-dose CT. *Phys Med Biol.* 2018;63(21):215004.
81. Shan H, Zhang Y, Yang Q, Kruger U, Kalra MK, Sun L, et al. 3-D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning From a 2-D Trained Network. *IEEE Trans Med Imaging.* 2018;37(6):1522-34.
82. Gong E, Pauly JM, Wintermark M, Zaharchuk G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. *J Magn Reson Imaging.* 2018;48(2):330-40.
83. Zhang Y, Yu H. Convolutional Neural Network Based Metal Artifact Reduction in X-Ray Computed Tomography. *IEEE Trans Med Imaging.* 2018;37(6):1370-81.
84. Huang X, Wang J, Tang F, Zhong T, Zhang Y. Metal artifact reduction on cervical CT images by deep residual learning. *Biomed Eng Online.* 2018;17(1):175.
85. Kustner T, Gatidis S, Liebgott A, Schwartz M, Mauch L, Martirosian P, et al. A machine-learning framework for automatic reference-free quality assessment in MRI. *Magn Reson Imaging.* 2018;53:134-47.
86. Graham MS, Drobnjak I, Zhang H. A supervised learning approach for diffusion MRI quality control with minimal training data. *Neuroimage.* 2018;178:668-76.
87. Xiang L, Wang Q, Nie D, Zhang L, Jin X, Qiao Y, et al. Deep embedding convolutional neural network for synthesizing CT image from T1-Weighted MR image. *Med Image Anal.* 2018;47:31-44.
88. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med.* 2015;372(9):793-5.
89. Giger ML. Machine Learning in Medical Imaging. *J Am Coll Radiol.* 2018;15(3 Pt B):512-20.
90. Shaikh F, Franc B, Allen E, Sala E, Awan O, Hendrata K, et al. Translational Radiomics: Defining the Strategy Pipeline and Considerations for

Application-Part 1: From Methodology to Clinical Implementation. *J Am Coll Radiol.* 2018;15(3 Pt B):538-42.

91. Shaikh F, Franc B, Allen E, Sala E, Awan O, Hendrata K, et al. Translational Radiomics: Defining the Strategy Pipeline and Considerations for Application-Part 2: From Clinical Implementation to Enterprise. *J Am Coll Radiol.* 2018;15(3 Pt B):543-9.

92. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom KW, et al. Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches. *AJNR Am J Neuroradiol.* 2018;39(2):208-16.

93. Lakhani P, Prater AB, Hutson RK, Andriole KP, Dreyer KJ, Morey J, et al. Machine Learning in Radiology: Applications Beyond Image Interpretation. *J Am Coll Radiol.* 2018;15(2):350-9.

94. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys.* 2019;29(2):102-27.

95. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60-88.

96. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Med.* 2018;15(11):e1002683.

97. Lakhani P, Gray DL, Pett CR, Nagy P, Shih G. Hello World Deep Learning in Medical Imaging. *J Digit Imaging.* 2018;31(3):283-9.

98. Fazal MI, Patel ME, Tye J, Gupta Y. The past, present and future role of artificial intelligence in imaging. *Eur J Radiol.* 2018;105:246-50.

99. Tajmir SH, Alkasab TK. Toward Augmented Radiologists: Changes in Radiology Education in the Era of Machine Learning and Artificial Intelligence. *Acad Radiol.* 2018;25(6):747-50.

100. Recht M, Bryan RN. Artificial Intelligence: Threat or Boon to Radiologists? *J Am Coll Radiol.* 2017;14(11):1476-80.

101. The Medical Futurist. FDA Approvals For Smart Algorithms In Medicine In One Giant Infographic; 2019 [Visitado a 29 de outubro de 2019] Disponível em: <https://medicalfuturist.com/fda-approvals-for-algorithms-in-medicine/>.

102. Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. *Radiology*. 2019;293(2):436-40.