

# Brain Anomaly Segmentation, A Medical Image Application

Daniel Xavier<sup>[pg50310]</sup> and Diogo Rebelo<sup>[pg50327]</sup>

<sup>1</sup> University of Minho, Informatics Department, 4710-057 Braga, Portugal

<sup>2</sup> email: {pg50310, pg50327}@alunos.uminho.pt

**Abstract.** O presente artigo tem como objetivo explorar os modelos generativos profundos na área de Imagem Médica. Estes modelos são uma classe de algoritmos de *Machine Learning* que utilizam redes neurais para aprender a criar novas amostras a partir de dados existentes. Eles têm sido amplamente utilizados em aplicações de visão computacional, incluindo a geração de imagens realistas, e têm mostrado resultados promissores em aplicações médicas, como a geração de imagens de ressonância magnética e tomografia computadorizada. Então, são analisados vários modelos generativos profundos, sendo discutidos os principais desafios e oportunidades associados ao uso desses modelos em aplicações de Imagem Médica. Além disso, são ainda apresentados alguns exemplos de aplicações bem sucedidas no mundo real, que se socorrem da utilização destes mesmos modelos. Em suma, este trabalho visa apresentar uma visão geral das principais técnicas de modelos generativos profundos em aplicações médicas e as suas potenciais implicações no futuro da medicina.

**Keywords:** Modelos Generativos Profundos · Imagem Médica · *Machine Learning* · Redes Neurais · Ressonância Magnética · Tomografia Computadorizada · Redes Adversárias Generativas · *Variational Autoencoders* · Medicina

## 1 Metodologia

A metodologia utilizada na pesquisa de referências bibliográficas em vários bancos de dados envolveu uma busca sistemática por artigos científicos, livros e estudos relevantes relacionados com o tema de investigação em questão. Então, Utilizaram-se vários bancos de dados como *Google Scholar*, *IEEE Xplore*, *ScienceDirect*, *Scopus*, MDPI, bem como outras fontes relevantes.

Para a pesquisa, foram selecionadas palavras-chave relacionadas ao tema em questão e combinadas com operadores booleanos (AND e OR), para refinar os resultados da busca. Foram considerados apenas artigos publicados há menos de 5 anos. Além disso, foram analisadas as referências bibliográficas dos artigos selecionados para identificar outras fontes relevantes que poderiam ter sido perdidas nas buscas iniciais. Todos os artigos foram avaliados de acordo com critérios pré-estabelecidos, como relevância, atualidade e integridade das fontes.

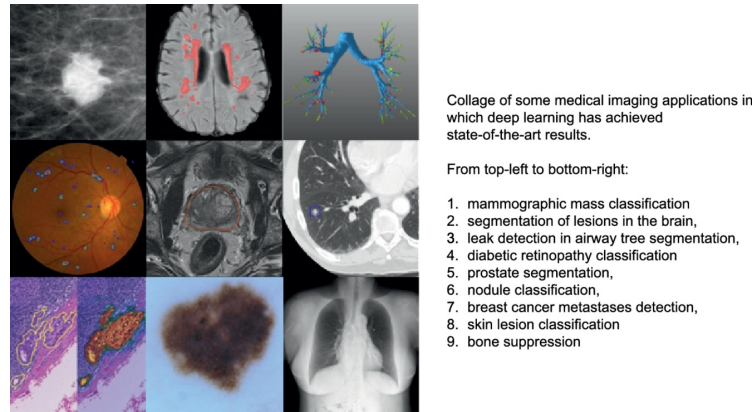
O caráter descritivo e exploratório da metodologia sugere que o presente artigo procurará descrever e explorar conceitos, ideias e teorias relacionadas com tema de estudo, com uma breve análise das contribuições práticas dos artigos analisados.

## 2 Introdução

Nos últimos anos, houve um grande impacto dos modelos generativos em várias áreas da Inteligência Artificial e *Machine Learning*. Esses modelos visam adquirir a capacidade de aprender e inferir conhecimento a partir de um conjunto de dados de treino, procurando encontrar uma distribuição de dados que seja o mais semelhante possível à distribuição real dos dados. Para alcançar esse objetivo, são usadas redes neurais, que são modelos computacionais capazes de reconhecer padrões e aprender continuamente. Este treino é possível devido à existência de recursos computacionais mais rápidos na atualidade. Os modelos são assim usados para aproximar distribuições complexas através de grandes volumes de dados e produzir conteúdo altamente realista, como imagens médicas.

Neste artigo, estuda-se a utilização de modelos generativos profundos, como *Generative Adversarial Networks* e *Variational Autoencoders*. Para além disso, apresenta-se um resumo do estado atual deste tipo de modelos em imagens médicas, incluindo características importantes, desafios atuais e futuras direções de pesquisa, investigando-se sobre de que forma estes modelos podem contribuir para a imagem médica.

Como mote, na figura seguinte, é possível verificar a utilização primordial destes algoritmos na área da imagem médica, desde a supressão de ossos, até à deteção do cancro da mama.



**Fig. 1.** Aplicações do *Deep Learning* na Imagem Médica [21]

### 3 *Deep Generative Models* (DGM)

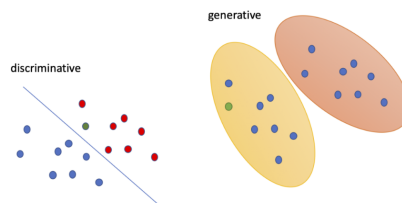
A definição e caracterização deste tipo de modelos torna-se necessária para a posterior discussão sobre a sua aplicabilidade.

#### 3.1 Generativos vs Discriminativos

Os modelos Generativos procuram encontrar padrões que manifestem os diferentes atributos ou categorias dentro de um determinado *input*. Através destes modelos, é possível ainda criar novos dados, todavia, os mesmos revelam-se bastante frágeis perante a presença de *outliers*, os quais podem afetar significativamente o modelo [25] [31].

Por sua vez, os modelos Discriminativos procuram, consoante um conjunto de dados, distinguir a que categoria os mesmos pertencem, conseguindo expor os limites entre as diferentes classes de um *dataset* [36].

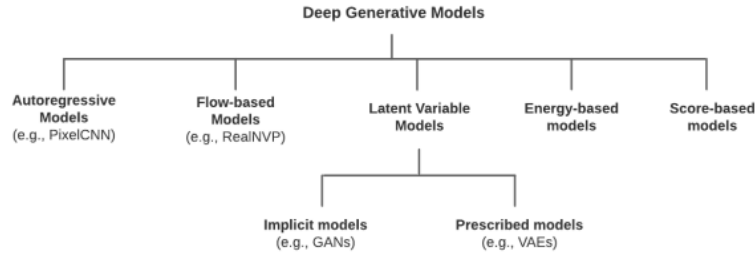
Estes modelos, apesar de serem mais robustos à presença de *outliers* manifestam-se mais suscetíveis ao problema da classificação incorreta dos dados: por exemplo, em situações de fronteira entre duas classes, alguns pontos singulares de dados podem ficar incorretamente classificados. São exemplos de modelos Discriminativos: Árvores de Decisão, Regressão Logística, *Random Forest* e *Support Vector Machine* [7] [37] [2].



**Fig. 2.** Comparação dos modelos [16]

#### 3.2 Tipos de Modelos Generativos

Deste modo, os DGM podem ser classificados como pertencendo a um dos cinco grupos principais, sendo eles os seguintes:



**Fig. 3.** Tipos de Modelos Generativos Profundos [35]

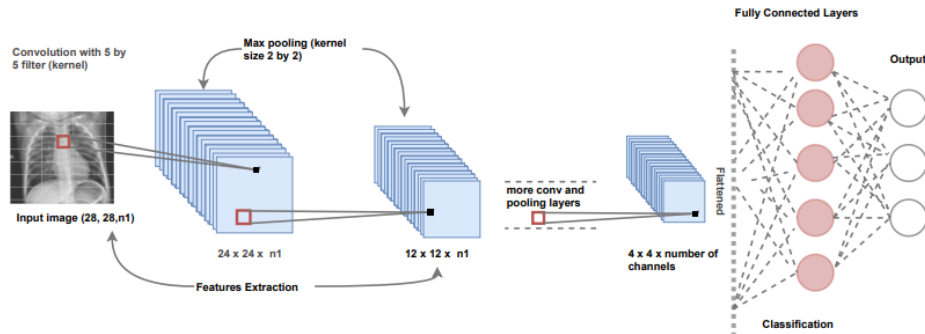
A classe que normalmente é mais usada em imagem médica é a dos *Latent Variable Models*, sendo essa mesma analisada em detalhe posteriormente. Nesta classe, inclui-se *Generative Adversarial Networks (GANs)* e *Variational Autoencoders (VAEs)*. De forma sucinta, estes *Latent Variable Models* utilizam camadas de (des)codificação para aprender representações latentes dos dados de entrada [9].

A grande vantagem destes modelos é que são capazes de capturar a variabilidade intrínseca das imagens médicas, que podem ter diferenças significativas em termos de posição, orientação, escala e outras características, tornando desafiadora a tarefa de geração de imagens mais realistas. Além disso, estes modelos podem ser facilmente combinados com outras técnicas de *deep learning*, como *convolutional neural networks (CNNs)*, que são também muito utilizadas em imagens médicas para extrair recursos discriminativos das imagens [34].

Contudo, também os outros grupos de *Deep Generative Models*, como modelos autoregressivos ou baseados em fluxo, podem ser aplicáveis em imagens médicas para tarefas específicas, como geração de sequência de imagens ou segmentação de imagens.

## 4 Latent Variable Models

Numa primeira fase, é importante caracterizar as CNNs usadas nestes modelos generativos profundos. Este tipo de redes têm a capacidade de capturar a representação subjacente das imagens usando imagens parcialmente ligadas, camadas e partilha de pesos. Muitas arquiteturas CNN consistem num pequeno número de camadas de convolução, seguidas por funções de ativação, e camadas de *pooling* para a amostragem das imagens, como pode ser visto na Figura 4. A aplicação repetida de filtros (*kernels*) à imagem de entrada resulta num mapa de ativações (*feature maps*), que indicam locais de interesse na imagem de entrada [18] [4] [30].



**Fig. 4.** Diagrama esquemático do modelo CNN com arquitetura arbitrária. [18]

Estas CNNs não são necessariamente modelos generativos profundos, mas as variações como GANs e VAEs, que usam CNNs como componentes podem ser consideradas como tal [6].

#### 4.1 Generative Adversarial Networks (GANs)

**Caraterização do Modelo** Estas redes foram introduzidas por *Ian Goodfellow* com outros investigadores como um novo quadro para estimar modelos generativos através de um processo adversário. Os GANs consistem em duas redes neurais diferentes que são treinadas separadamente. Um dos dois modelos é um **Gerador**, uma rede neuronal que captura a distribuição de dados e gera uma imagem falsa. O outro é um **Discriminador**, que funciona como um classificador que diz se a sua entrada é uma imagem real ou falsa. Estas redes atuam assim uma contra a outra, derivando daí o nome *adversarial*.

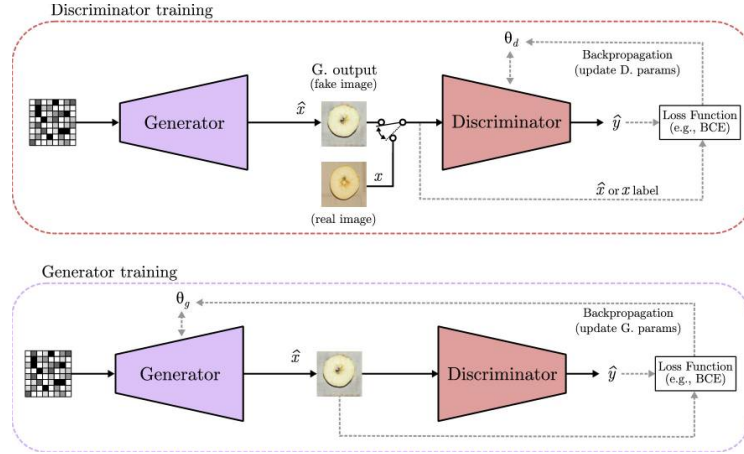


Fig. 5. Modelo de uma Rede Adversária Generativa (GAN) [10]

De um modo mais técnico, numa análise à figura anterior, observa-se uma estrutura básica de uma GAN, onde  $\theta_g$  e  $\theta_d$  representam os parâmetros de formação do gerador e do discriminador, respetivamente. Uma GAN é treinada utilizando um conjunto de imagens reais, onde cada uma delas é representada como  $x$  características. Além disso, um exemplo falso gerado por  $G$  é representado como  $\hat{x}$ .

As imagens falsas e reais são alimentadas separadamente no discriminador, cuja função é dizer se são falsas ( $\hat{x}$ ) ou reais ( $x$ ). A saída de  $D$  é representada como  $\hat{y}$ .

A figura seguinte mostra também como um GAN é treinado. Em ambos os casos, a saída do discriminador é comparada com a etiqueta real da imagem, iniciando o processo de *backpropagation* dos parâmetros treináveis  $\theta_g$  ou  $\theta_d$ , dependendo do modelo a ser treinado. O processo de formação é repetido até serem alcançados os resultados desejados. Finalmente, o modelo discriminador pode ser descartado e o gerador é utilizado para criar as imagens falsas como saída final. [10]

Este processo prolongado de treino pode levar a múltiplos problemas. Um dos principais problemas é a incapacidade de convergência, o que significa que o modelo não aprende a gerar imagens que parecem reais. A fim de resolver este problema, a função de perda evoluiu para quantificar a semelhança entre as imagens geradas e as distribuições de dados reais, com base no pressuposto de que quanto mais semelhantes forem as distribuições, melhor será a precisão de geração dos modelos, ou seja, imagens falsas mais diversas e realistas.

A fim de evitar que o gerador produza sempre a mesma saída, é utilizado como entrada um vetor de ruído aleatório. Primeiramente, este vetor era apenas utilizado como semente para gerar diferentes saídas, mas à medida que os modelos se tornaram mais complexos, o ruído transformara-se numa forma de controlar o conteúdo da imagem falsa. Assim, este ruído pode ser combinado com diferentes dados, para selecionar as características da imagem ou qual a classe representada no seu conteúdo. Por exemplo, é possível modificar o ruído para gerar imagens de rostos com diferentes expressões faciais ou cores de cabelo, o que permite o controlo do conteúdo das imagens geradas e torna o modelo mais versátil e útil.

Seguindo o trabalho dos autores originais, os GANs podem ser formalmente descritos como dois modelos baseados num jogo de minimax, como representado na equação abaixo.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

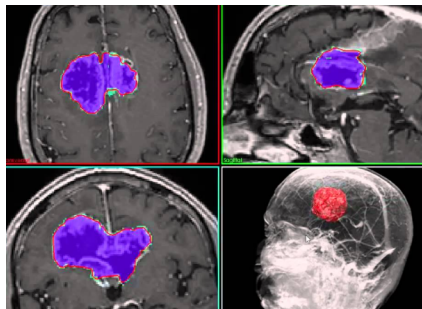
O objetivo é treinar um modelo chamado *Discriminator* D, usando dados (imagens) reais (x) como referência, para que ele consiga maximizar a probabilidade de atribuir a etiqueta correta às imagens geradas pelo *Generator* G, dada a variável de ruído z, enquanto minimiza a hipótese de gerar imagens em G, que não pareçam reais para D. Isto significa que o fim ótimo dos GANs seria a distribuição de dados reais igual à distribuição das imagens geradas.

Compreende-se então que o processo de treino de um GAN corresponde a um jogo de dois modelos, onde cada um dos modelos é treinado separadamente, mas precisa de aprender ao mesmo ritmo. Por um lado, se um gerador completamente treinado for utilizado enquanto um discriminador ingênuo for empregado como classificador, o discriminador não será capaz de distinguir se é uma imagem real ou falsa, porque o gerador está a fazer um trabalho muito bom gerando imagens que parecem reais, não sendo assim capaz de aprender.

Por outro lado, se um Discriminador completamente treinado for utilizado com um Gerador não treinado, o classificador detetará todas as imagens falsas e o gerador não será capaz de aprender a “enganar” o discriminador, tratando-se assim de uma “faca de dois gumes”, que necessita de equilíbrio. É por isso que o processo de formação é feito separadamente, por lotes, onde G aprende a produzir imagens falsas muito boas, e depois D aprende a detetá-las.

**Aplicação do modelo no mundo real** Embora as GAN sejam utilizadas em múltiplas áreas, têm uma grande influência sobre o campo médico, onde a análise e a geração de imagens desempenham um papel muito importante.

Estas redes tiram partido do recente aumento do poder computacional e da enorme disponibilidade de dados médicos, sendo grande parte dos dados gerados aplicada à neuroimagem e neurorradiologia, segmentação cerebral, imagens de acidente vascular cerebral, doenças neuropsiquiátricas, cancro da mama, imagens de tórax, imagens em oncologia, e ultra-som médico.



**Fig. 6.** Detecção de anomalias no cérebro

Apesar de serem geralmente usados para gerar novas imagens (*augmentation*), também podem ser suporte a tarefas de síntese, segmentação, classificação, melhoria ou registo das imagens médicas. Por exemplo na classificação de imagens médicas, pode ser treinado um modelo para gerar imagens sintéticas que representam diferentes categorias de imagens. Isto é, pode ser treinado para gerar imagens de tumores e imagens normais. Essas mesmas imagens podem ser usadas para treinar um classificador convencional, como uma rede neural convolucional, para realizar a classificação.

Na segmentação de imagens médicas, pode ser treinado um modelo para gerar imagens sintéticas que contêm informações sobre as áreas que devem ser segmentadas. Essas imagens podem ser usadas como máscaras de segmentação para segmentar a imagem médica original. Então, treina-se o modelo GAN com imagens de entrada que são mascaradas em regiões de interesse, permitindo que este gere imagens sintéticas com informações sobre essas regiões.

Segue-se alguns exemplos práticos da aplicação de GANs:

- **Ghassemi** e outros investigadores propõem um DCGAN para produzir imagens de RM (ressonância magnética) do cérebro. Os autores treinam o GAN para que o discriminador possa detetar imagens falsas e extrair as suas principais características. Depois, as camadas totalmente ligadas do Discriminador são substituídas por uma camada *softmax* que é novamente treinada. Desta forma, utilizam-no como classificador capaz de detetar meningiomas, gliomas, e tumores pituitários com alta precisão. [13]
- **Zhou** e outros investigadores propõem também um GAN para melhorar as imagens de RM do cérebro para a classificação da doença de Alzheimer. Utilizam scans 2.5D e 3D juntamente com o 3D-GAN, que utiliza convoluções 3D para gerar uma melhor qualidade

- das imagens de MR. Depois, um classificador analisa a imagem sintética, para dizer se o paciente pode ou não sofrer da doença de Alzheimer.[39]
- **Pang** e outros investigadores utilizam um GAN para gerar novos exemplos de imagens de ultra-sons mamários aplicados a doentes com cancro da mama. As imagens sintéticas geradas pelo GAN, juntamente com as imagens reais, são introduzidas numa CNN que classifica as massas encontradas na imagem como benignas e malignas. Este processo permite que a CNN aprenda com um conjunto de dados maior, obtendo assim melhores resultados. [28]
- O recente aparecimento da doença COVID-19 também motivou muitos autores a aplicar GANs para a sua deteção. A principal área de estudo são os efeitos da COVID nos pulmões por meio de pneumonia e inflamação pulmonar. Um dos procedimentos de diagnóstico mais comuns é feito através de raios-X, MR e imagens de TAC, a maioria das quais, como se analisou anteriormente, pode ser usada como entrada de GAN [20].

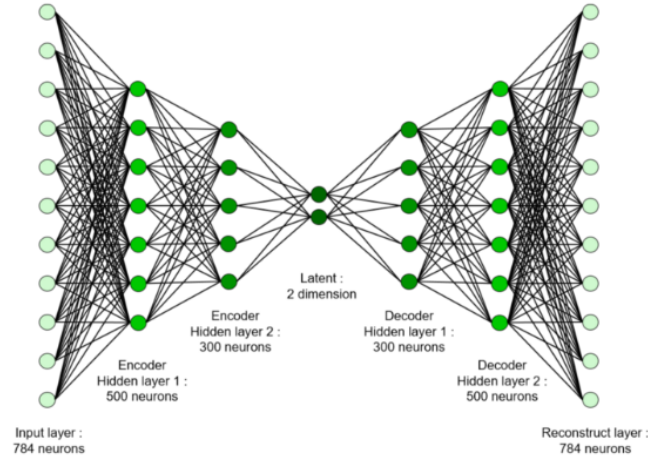
## 4.2 Variational Autoencoders (VAEs)

**Caraterização do modelo** Introduzido por *Diederik P. Kingma e Max Welling* trata-se de uma junção de duas áreas de conhecimento, uma que envolve redes neuronais profundas e outra de modelos probabilísticos.

Os VAE combinam assim dois modelos diferentes, um codificador e um decodificador. Ambos os modelos são redes neurais convolucionais separadas que trabalham em conjunto para aprender diferentes distribuições que representam o input e para transformar as distribuições encontradas em imagens falsas, de modo que, quando a formação termina, o decodificador pode ser usado independentemente para gerar novos exemplos.

Para uma melhor compreensão, antes de abordarmos o conceito de *Variational Autoencoders*, é necessário entender o funcionamento da versão mais simplificada deste modelo.

Neste caso, trata-se do modelo de *Autoencoders (AEs)*, cujo objetivo é criar uma reconstrução dos dados de input, através do uso de uma rede neuronal com várias camadas (Fig.7), tentando sempre minimizar a diferença entre o *input* e o *output*.



**Fig. 7.** Representação gráfica de um *Autoencoder* [1]

O *Encoder* é responsável por comprimir os dados de input de modo a obter uma representação mais pequena destes dados, a qual se dá o nome de *bottleneck* ou *latent space representation*. De seguida, o *Decoder* irá reconstruir os dados comprimidos do *bottleneck*, até obter a dimensão original dos mesmos. A cada iteração o *autoencoder* é cada vez mais otimizado, uma vez que o *bottleneck* seleciona a estrutura principal dos dados e garante a propagação da informação mais relevante dos mesmos, de modo a que possam ser reconstruídos com um grau de maior fidelidade [12].

Outro aspeto crucial dos *autoencoders* trata-se da dimensão do espaço latente/*bottleneck*: quando a dimensão deste *bottleneck* é pequena, mais relevantes serão as características que o *autoencoder* irá aprender sobre os dados. Caso contrário, o *autoencoder* irá apenas aprender a copiar os dados de *input* sem reter nenhuma informação essencial sobre os mesmos.

Assim sendo, o objetivo é descobrir qual o melhor par de *encoder/decoder* (e qual a melhor dimensão de *bottleneck*) que consegue manter o máximo de informação possível no



processo de compressão (*encoding*), mas que tenha o menor erro possível de reconstrução (*decoding*) [22].

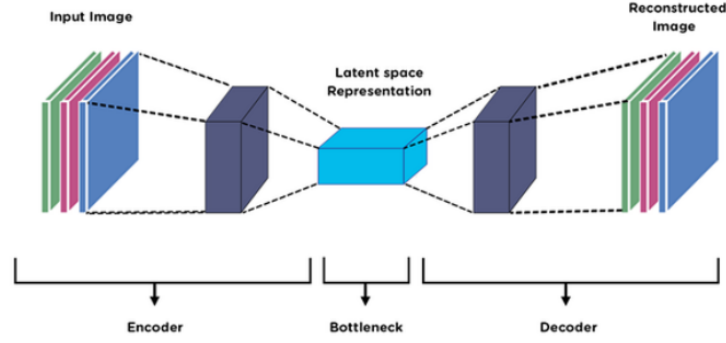


Fig. 8. Arquitetura de um *Autoencoder* [10]

Os VAE partilham a mesma arquitetura codificador-descodificador que os *Autoencoders*, mas o processo de formação é regularizado a fim de alcançar boas propriedades de geração e evitar a sobreposição. Tal é conseguido através da codificação da entrada como uma distribuição pelo espaço latente. Depois, um ponto é amostrado a partir dessa distribuição para ser descodificado e o erro de reconstrução é computado.

Assim a principal diferença é que os VAE utilizam um modelo probabilístico para modelar a distribuição latente dos dados de entrada: enquanto o AE simplesmente comprime os dados numa representação latente e tenta reconstruir os dados de entrada originais a partir dessa representação, o VAE tenta aprender a distribuição latente dos dados e usa essa distribuição para gerar novos exemplos de dados. O VAE, portanto, pode ser usado para tarefas como geração de dados sintéticos, enquanto o AE é mais adequado para tarefas de reconstrução e compressão de dados.

A arquitetura completa pode ser vista na figura seguinte. O codificador é uma CNN cujo objetivo é codificar dados de alta dimensão numa representação de baixa dimensão, sendo capaz de encontrar uma representação da imagem e coloca-a num espaço latente.

Uma amostra do espaço latente consiste num conjunto de meios e desvios padrão, portanto em vez de termos um só vetor a representar o *bottleneck*, temos dois vetores separados: um que representa a média da distribuição e outro que representa o desvio padrão da mesma. Estes são tomados como entrada para o decodificador, onde os dados são reconstruídos a partir da representação de baixa dimensão. É uma rede neural convolucional transposta (aumenta dimensionalidade), que recolhe o vetor da amostra e produz uma imagem falsa [29].

Esta imagem falsa é então comparada com a imagem real introduzida como entrada para o codificador através de uma função de perda (*reconstruction loss*), a partir da qual o gradiente é calculado. Finalmente, a etapa de *backpropagation* é realizada desde a última camada do decodificador até à primeira camada do codificador [10].

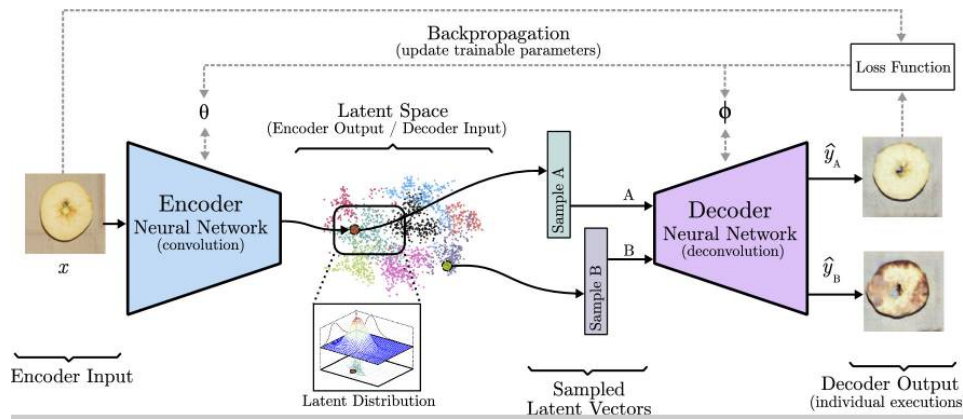


Fig. 9. Arquitetura de um *Variational Autoencoder* [10]

**Aplicação do modelo no mundo real** Deste modo, os VAEs têm sido utilizados com sucesso na área de imagem médica devido à sua capacidade de aprender uma representação latente de alta qualidade dos dados de imagem. Isso é especialmente importante em imagens médicas, onde a qualidade e a interpretabilidade dos dados são fundamentais para o diagnóstico e tratamento.

São especialmente úteis em casos em que há uma escassez de dados de imagem, pois permite que os médicos e pesquisadores gerem novas imagens sintéticas para análise e diagnóstico.

Além disso, os VAEs são capazes de realizar tarefas de segmentação de imagem, que é a identificação e separação de diferentes regiões de interesse.

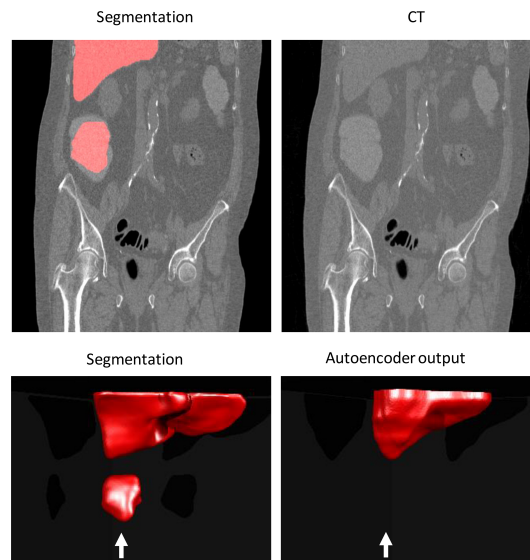
Os VAEs também são capazes de aprender uma representação latente contínua dos dados de imagem, o que significa que podem gerar imagens interpoladas entre duas imagens reais, algo útil em imagens médicas, onde é necessário analisar a evolução temporal de uma doença ou tratamento.

Por sua vez, também nestes modelos a segmentação de imagem é um dos objetivos mais pretendidos. Neste caso, gera-se uma nova imagem referenciando a original na entrada e destacando áreas de interesse. A imagem “falsa” tem nova informação, identificando objetivos específicos como características anormais ou diferenciação de partes.

Seguem-se alguns exemplos práticos da aplicação dos VAEs:

- **Uzunova** e outros investigadores utilizam *conditional variational autoencoders* (CVAEs), juntamente com conjuntos de dados médicos 2D e 3D, como uma abordagem não supervisionada para a deteção de patologias. Concluem que os CVAEs são capazes de aprender uma representação adequada dos dados aplicados à deteção de patologias em imagens 2D, enquanto a representação complexa dos dados 3D prejudica o processo de reconstrução. Ou seja, o modelo proposto é capaz de enriquecer a representação 3D, todavia, é incapaz de reconstruir adequadamente o *input* [38].
- **Akrami** e outros investigadores analisam como os modelos VAE são afetados por um processo de pré-treino aplicado à deteção de lesões cerebrais através de imagens 2D de Ressonância Magnética do cérebro. Os autores chegam à conclusão de que o seu modelo sugerido (VAE robusto), que utiliza uma nova função de perda, atinge um desempenho superior utilizando a aprendizagem por transferência através do pré-treino, do que o VAE básico [5].

A figura seguinte ilustra um exemplo de uma segmentação errada solucionada por um *autoencoder*. Neste caso, o cólon teve uma atenuação homogênea incomum semelhante à do fígado e foi erroneamente segmentado como fígado. Depois de passar pelo VAE, a área problemática não foi reconstruída pelo *decoder* e desapareceu, corrigindo-se o erro e assinalada esta segmentação como sendo problemática, pelo que se verificam assim técnicas complementares [32].



**Fig. 10.** Exemplo do uso de um *Autoencoder* na Segmentação de Imagens [32]



### 4.3 Generative Adversarial Networks vs. Variational Autoencoders

Confirma-se que os GANs estão mais difundidos nos trabalhos de medicina do que os VAE. Como salientado por autores, os GANs e VAEs são complementares, sendo as vantagens dos GANs desvantagens dos VAEs e vice-versa. Esta é a razão pela qual muitos autores optam por não escolher um modelo em particular, utilizando um *modelo híbrido* que possui simultaneamente o melhor dos dois modelos [19].

Assim, são inúmeros os exemplos práticos que podem ser aplicados diretamente a imagens médicas, como por exemplo o de **Larsen** que acrescenta um discriminador a uma estrutura VAE, alterando-a para um VAE-GAN. O VAE é utilizado como um gerador cuja saída passa pelo discriminador para detetar se se trata de uma imagem real ou não. O modelo proposto atinge uma geração de imagem de alta qualidade com a possibilidade de trabalhar sobre as características do espaço latente [8] [26].

O trabalho de **Nakao** também combina um VAE e um GAN para construir um sistema de deteção de anomalias de radiografia torácica não supervisionada. A sua proposta é capaz de detetar anomalias com alta precisão, mas não é capaz de as diagnosticar. Citando os autores, esta abordagem não substitui o médico humano, sendo antes uma ferramenta para ajudar a detetar lesões e prevenir o erro da visão [24].

## 5 Desafios e progresso futuro

Tendo em conta a informação exposta anteriormente, e considerando que a aprendizagem profunda está continuamente em desenvolvimento, é seguro dizer que novos modelos e arquiteturas melhoradas virão à luz em anos futuros, oferecendo resultados ainda melhores até por uma melhor manipulação dos dados ou uma melhor otimização dos hiperparâmetros.

Até agora, os GANs são os modelos mais utilizados, mas os autores relatam que podem ser alcançados bons resultados utilizando modelos híbridos que fundem GANs e VAEs. A prática mais comum é usar as capacidades de codificação VAE para localizar cada exemplo no espaço latente, e depois usar a capacidade do GAN para produzir uma imagem de alta qualidade a partir da codificação VAE. Afinal, ainda há muito a fazer em termos de modelos generativos, e a combinação de diferentes modelos poderia oferecer um enorme impulso à investigação médica [11].

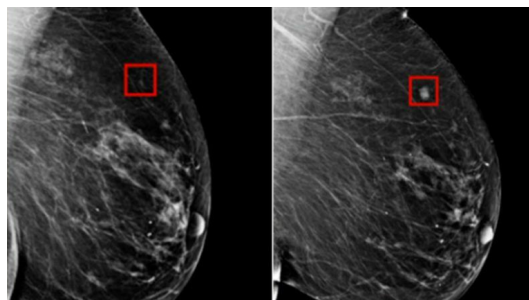
Depois de analisar o trabalho interno de diferentes modelos e múltiplos trabalhos de diferentes autores, fica claro que os algoritmos de aprendizagem profunda proporcionam elevados benefícios à análise da imagem médica. No entanto, novos desafios e obstáculos estão a surgir. Muitos autores salientam que a falta de grandes conjuntos de dados específicos de imagem de doença dificulta singularmente o trabalho de formação de um modelo. A isto acresce a complexa e morosa tarefa de etiquetar cada uma das imagens. Este processo só pode ser feito por pessoal médico especializado, limitando grandemente a quantidade de dados que podem ser gerados num pequeno espaço de tempo.

Num futuro próximo, para além de melhorar as arquiteturas existentes, implementar-se-ão novas técnicas para continuar a expandir a aplicação de modelos de aprendizagem profunda e permitir melhores resultados. Um dos últimos avanços na visão informática em medicina é a inclusão de **Transformers**. Estes baseiam-se no mecanismo de auto-atendimento e oferecem uma alternativa às simples redes neurais convolucionais e recorrentes através da simbolização de grupos de *pixels* que representam um conceito semântico da imagem, e mais tarde ligando-os para classificar características maiores. Em publicações recentes, os *Transformers* provam obter resultados promissores ao trabalhar com imagens médicas na deteção e classificação de lesões cutâneas, aplicando segmentação semântica, e melhorando a qualidade das imagens de Ressonância Magnética [15] [27] [33].

Além disso, estão a ser desenvolvidas novas alternativas aos GANs e VAEs como **modelos de Difusão**. Os modelos de difusão são uma técnica baseada na pontuação, onde o treino é realizado através da adição gradual de ruído seguindo uma distribuição gaussiana, e da aprendizagem de como reverter a mesma. Em trabalhos futuros, os modelos híbridos poderiam tirar benefícios dos modelos de difusão, fundindo-os com outras alternativas generativas anteriormente apresentadas [23] [17].

O recente aparecimento de coleções de dados com um ponto de vista evolutivo, como o desenvolvido por *Gomez*, prova que os modelos generativos poderiam não só focar numa única imagem, mas também cobrir diferentes estados através de um período de tempo. Este processo é semelhante à geração de vídeos e sequências de imagens. Uma investigação adicional poderia desenvolver este campo e oferecer grandes benefícios no campo médico [14].

Por fim, recentemente foram divulgadas as seguintes imagens por investigadores do MIT que mostram como a IA detetou o cancro da mama de uma mulher 4 anos antes do seu desenvolvimento [3].



**Fig. 11.** Cancro da mama detetado 4 anos antes de se desenvolver [3]

O poder e o impacto deste tipo de tecnologias representam as milhões de vidas que são e poderão a vir a ser salvas num futuro próximo à medida que há soluções cada vez melhores e eficazes para a deteção e para o tratamento.

Mas será que os técnicos e médicos que analisam as ressonância magnéticas estão treinados para procurar sinais como o pré-cancro? Ou será que procuram apenas um conjunto específico de anomalias que se concentram mais no diagnóstico do que na medicina preventiva? Num futuro próximo, mudarão a estratégia e deverão assim procurar indicadores e sinais preventivos, e não apenas sinais que confirmem um diagnóstico.

A verdade é que o papel do homem é fundamental: *Uma máquina consegue fazer o trabalho de 50 homens ordinários. Nenhuma máquina consegue fazer o trabalho de um homem extraordinário.* - Steve Jobs

## 6 Conclusão

Em suma, a utilização de *Deep Generative Models* na análise de imagem médica tem mostrado um grande potencial (ainda emergente) para melhorar a precisão e eficácia do diagnóstico, bem como para ajudar os profissionais de saúde a compreender melhor a natureza das doenças, sendo os VAEs e os GANs os modelos que causam o maior impacto no terreno, pois são modelos capazes de aprender a partir de grandes conjuntos de dados, gerando novas imagens médicas com alta precisão e fidelidade de doenças que não são frequentemente encontradas.

Além disso, estes modelos têm sido utilizados para gerar imagens médicas sintéticas, que podem ser usadas para treinar outros modelos de análise de imagem ou para fins educacionais. No entanto, ainda existem alguns desafios a serem enfrentados, como o tamanho e a complexidade dos dados, a necessidade de grandes conjuntos de treino e o tempo e os recursos necessários para treinar esses modelos.

Ainda assim, a crescente adoção de *Deep Generative Models* na imagem médica tem o potencial de revolucionar a forma como as doenças são diagnosticadas e tratadas, e deve continuar a ser uma área de pesquisa ativa e emocionante.

## Siglas e Acrónimos

- **DGM** - *Deep Generative Model*
- **GAN** - *Generative Adversarial Networks*
- **D** - *Discriminator*
- **G** - *Generator*
- **VAE** - *Variational Autoencoders*
- **RM** - Ressonância Magnética
- **TAC** - Tomografia Computadorizada
- **CNN** - *Convolutional Neural Networks*
- **IA** - Inteligência Artificial

## References

1. Extreme rare event classification using autoencoders in keras - processminer, <https://processminer.com/autoencoders-in-keras/>
2. Machine learning: Discriminative and generative - tony jebara - google books,
3. Watch: See how ai detected woman's breast cancer 4 years before it developed - cnn video, <https://edition.cnn.com/videos/health/2023/03/07/artificial-intelligence-breast-cancer-detection-mammogram-cnntm-vpx.cnn>
4. Editorial latent variable modelling (2008). <https://doi.org/10.1177/0962280207081235>
5. Akrami, H., Joshi, A.A., Li, J., Aydore, S., Leahy, R.M.: Brain lesion detection using a robust variational autoencoder and transfer learning. *Proceedings. IEEE International Symposium on Biomedical Imaging* **2020**, 786 (4 2020). <https://doi.org/10.1109/ISBI45749.2020.9098405>, <https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC7831448/> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7831448/?report=abstract>
6. Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Essen, B.C.V., Awwal, A.A., Asari, V.K.: A state-of-the-art survey on deep learning theory and architectures. *Electronics* 2019, Vol. 8, Page 292 **8**, 292 (3 2019). <https://doi.org/10.3390/ELECTRONICS8030292>, <https://www.mdpi.com/2079-9292/8/3/292/htm> <https://www.mdpi.com/2079-9292/8/3/292>
7. Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.F.M., West, M., Bishop, C.M., Lasserre, J.: Generative or discriminative? getting the best of both worlds. *BAYESIAN STATISTICS* **8**, 3–24
8. Boesen, A., Larsen, L., Sønderby, S.K., Larochelle, H., Winther, O., Dk, O.: Autoencoding beyond pixels using a learned similarity metric (6 2016), <https://proceedings.mlr.press/v48/larsen16.html>
9. Carrivick, L., Prabhu, S., Goddard, P., Rossiter, J.: Unsupervised learning in radiology using novel latent variable models (2005). <https://doi.org/10.1109/CVPR.2005.357>
10. Celard, P., Iglesias, E.L., Sorribes-Fdez, J.M., Romero, R., Vieira, A.S., Borrajo, L.: A survey on deep learning applied to medical images: from simple artificial neural networks to generative models. *Neural Computing Applications* **35**, 2291 (1 2023). <https://doi.org/10.1007/S00521-022-07953-4>, <https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC9638354/> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9638354/?report=abstract>
11. Doi, K.: Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics* **31**, 198–211 (2007). <https://doi.org/10.1016/J.COMPAMEDIMAG.2007.02.002>
12. El-Shafai, W., El-Rabaie, E.S.M., Magdy, A., El-Nabi, S.A., El-Rabaie, E.S.M., Ali, A.M., Soliman, N.F., Algarni, A.D., El-Samie, F.E.A., Author, C.: Efficient deep-learning-based autoencoder denoising approach for medical image diagnosis image enhancement view project speech processing view project efficient deep-learning-based autoencoder denoising approach for medical image diagnosis . <https://doi.org/10.32604/cmc.2022.020698>, <https://www.researchgate.net/publication/355176798>
13. Ghassemi, N., Shoeibi, A., Rouhani, M.: Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images. *Biomedical Signal Processing and Control* **57**, 101678 (3 2020). <https://doi.org/10.1016/J.BSPC.2019.101678>
14. Gomez, T., Feyeux, M., Boulant, J., Normand, N., David, L., Paul-Gilloteaux, P., Fréour, T., Mouchère, H.: A time-lapse embryo dataset for morphokinetic parameter prediction. *Data in Brief* **42** (6 2022). <https://doi.org/10.1016/J.DIB.2022.108258>
15. He, K., Gan, C., Li, Z., Rekik, I., Yin, Z., Ji, W., Gao, Y., Wang, Q., Zhang, J., Shen, D.: Transformers in medical image analysis. *Intelligent Medicine* **3**, 59–78 (2 2023). <https://doi.org/10.1016/J.IMED.2022.07.002>
16. Hill, D.L.G., Batchelor, P.G., Holden, M., Hawkes, D.J.: Medical image registration. *MEDICINE AND BIOLOGY Phys. Med. Biol* **46**, 1–45 (2001), [www.iop.org/Journals/pbPII:S0031-9155](http://www.iop.org/Journals/pbPII:S0031-9155)
17. Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Hacihaliloglu, I., Merhof, D., Softmax, D.: Diffusion models for medical image analysis: A comprehensive survey (2022), <https://github.com/amirhossein-kz/Awesome-Diffusion-Models-in-Medical-Imaging>
18. Lee, Y., Jun, D., Kim, B.G., Lee, H.: Enhanced single image super resolution method using lightweight multi-scale channel dense network. *Sensors* 2021, Vol. 21, Page 3351 **21**, 3351 (5 2021). <https://doi.org/10.3390/S21103351>, <https://www.mdpi.com/1424-8220/21/10/3351/htm> <https://www.mdpi.com/1424-8220/21/10/3351>
19. Li, F., Huang, W., Luo, M., Zhang, P., Zha, Y.: A new vae-gan model to synthesize arterial spin labeling images from structural mri. *Displays* **70** (12 2021). <https://doi.org/10.1016/J.DISPLA.2021.102079>
20. Li, Z., Zhang, J., Li, B., Gu, X., Luo, X.: Covid-19 diagnosis on ct scan images using a generative adversarial network and concatenated feature pyramid network with an attention mechanism. *Medical Physics* **48**, 4334–4349 (8 2021). <https://doi.org/10.1002/MP.15044>, <https://onlinelibrary.wiley.com/doi/full/10.1002/mp.15044>
21. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis.

- Medical Image Analysis **42**, 60–88 (12 2017). <https://doi.org/10.1016/J.MEDIA.2017.07.005>, <http://arxiv.org/abs/1702.05747> <http://dx.doi.org/10.1016/j.media.2017.07.005>
22. Liu, X., Zhang, L., Guo, Z., Han, T., Ju, M., Xu, B., Liu, H.: Medical image compression based on variational autoencoder. *Mathematical Problems in Engineering* **2022** (2022). <https://doi.org/10.1155/2022/7088137>
23. Muzaffer“ozbey, M.M., Dalmaz, O., Dar, S.U., Bedel, H.A., Aban“ozturk, S.A., Ukur, T.C.: Unsupervised medical image translation with adversarial diffusion models <https://github.com/icon-lab/SynDiff>.
24. Nakao, T., Hanaoka, S., Nomura, Y., Murata, M., Takenaga, T., Miki, S., Watadani, T., Yoshikawa, T., Hayashi, N., Abe, O.: Unsupervised deep anomaly detection in chest radiographs. *Journal of Digital Imaging* **34**, 418–427 (4 2021). <https://doi.org/10.1007/S10278-020-00413-2>, <https://link.springer.com/article/10.1007/s10278-020-00413-2>
25. Ng, A.Y., Jordan, M.I.: On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in Neural Information Processing Systems* **14** (2001)
26. Niu, Z., Yu, K., Wu, X.: Lstm-based vae-gan for time-series anomaly detection. *Sensors* 2020, Vol. 20, Page 3738 **20**, 3738 (7 2020). <https://doi.org/10.3390/S20133738>, <https://www.mdpi.com/1424-8220/20/13/3738/htm> <https://www.mdpi.com/1424-8220/20/13/3738>
27. Nvidia, A.H., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Daguang, N., Nvidia, X.: Unetr: Transformers for 3d medical image segmentation (2022), <https://monai.io/research/unetr>
28. Pang, S., Du, A., Orgun, M.A., Wang, Y., Yu, Z.: Tumor attention networks: Better feature selection, better tumor segmentation. *Neural Networks* **140**, 203–222 (8 2021). <https://doi.org/10.1016/J.NEUNET.2021.03.006>
29. Pu, Y., Gan, Z., Henao, R., Yuan, X., Li, C., Stevens, A., Carin, L.: Variational autoencoder for deep learning of images, labels and captions. *Advances in Neural Information Processing Systems* **29** (2016)
30. Rabe-Hesketh, S.: Classical latent variable models for medical research. *Statistical Methods in Medical Research* **17**, 5–32 (2008). <https://doi.org/10.1177/0962280207081236>
31. Raina, R., Shen, Y., Ng, A.Y., McCallum, A.: Classification with hybrid generative/discriminative models. *Advances in Neural Information Processing Systems* **16** (2003)
32. Sandfort, V., Yan, K., Graffy, P.M., Pickhardt, P.J., Summers, R.M.: Use of variational autoencoders with unsupervised learning to detect incorrect organ segmentations at ct. *Radiology: Artificial Intelligence* **3** (7 2021), <https://pubs.rsna.org/doi/10.1148/ryai.2021200218>
33. Shamshad, F., Khan, S., Zamir, S.W., Khan, M.H., Hayat, M., Khan, F.S., Fu, H.: Transformers in medical imaging: A survey <https://github.com/fahadshamshad/awesome-transformers-in-medical-imaging>.
34. Su, C., Su, C., Liu, T., Chen, C., Wang, H., Hu, J., Zhou, Y., Xiong, B., Cao, X.: Analysis of pre-processing methods for lossless compression of multi-component medical images based on latent variable models. *Frontiers in Optics + Laser Science* 2022 (FIO, LS) (2022), paper JW5B.53 p. JW5B.53 (10 2022). <https://doi.org/10.1364/FIO.2022.JW5B.53>, <https://opg.optica.org/abstract.cfm?uri=LS-2022-JW5B.53>
35. Tomczak, J.M.: Deep generative modeling. *Deep Generative Modeling* (2022). <https://doi.org/10.1007/978-3-030-93158-2>
36. Tu, Z.: Learning generative models via discriminative approaches
37. Tu, Z., Narr, K.L., Dollár, P., Dinov, I., Thompson, P.M., Toga, A.W.: Brain anatomical structure segmentation by hybrid discriminative/generative models. *IEEE TRANSACTIONS ON MEDICAL IMAGING* **27**, 495 (2008). <https://doi.org/10.1109/TMI.2007.908121>, <http://ieeexplore.ieee.org>.
38. Uzunova, H., Schultz, S., Handels, H., Ehrhardt, J.: Unsupervised pathology detection in medical images using conditional variational autoencoders. *International Journal of Computer Assisted Radiology and Surgery* **14**, 451–461 (2019). <https://doi.org/10.1007/s11548-018-1898-0>, <https://doi.org/10.1007/s11548-018-1898-0>
39. Zhou, X., Qiu, S., Joshi, P.S., Xue, C., Killiany, R.J., Mian, A.Z., Chin, S.P., Au, R., Kolachalama, V.B.: Enhancing magnetic resonance imaging-driven alzheimer’s disease classification performance using generative adversarial learning. *Alzheimer’s Research and Therapy* **13**, 1–11 (12 2021). <https://doi.org/10.1186/S13195-021-00797-5>, <https://link.springer.com/articles/10.1186/s13195-021-00797-5>, <https://link.springer.com/article/10.1186/s13195-021-00797-5>