

Influência de ruído e realce na classificação de tumores cerebrais em redes profundas

Karina de Carvalho Fróes

Trabalho de Conclusão de Curso
MBA em Inteligência Artificial e Big Data

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Influência de ruído e realce na
classificação de tumores cerebrais em
redes profundas

Karina de Carvalho Fróes

Karina de Carvalho Fróes

Influência de ruído e realce na classificação de tumores cerebrais em redes profundas

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dr. Fernando Pereira dos Santos

USP - São Carlos

2022

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

F925i Fróes, Karina
Influência de ruído e realce na classificação de
tumores cerebrais em redes profundas / Karina Fróes;
orientador Fernando Santos. -- São Carlos, 2022.
43 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2022.

1. Deep learning. 2. redes neurais
convolucionais. 3. detecção de tumores cerebrais. I.
Santos, Fernando, orient. II. Título.

RESUMO

FRÓES, K. C. **Influência de ruído e realce na classificação de tumores cerebrais em redes profundas.** 2022. 43 f. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2022.

Tumores cerebrais representam aproximadamente 4% das mortes por câncer no Brasil. Atualmente não existe um teste único específico para seu diagnóstico, que quanto mais cedo for feito, leva a maiores chances de sobrevivência do paciente. O emprego de redes neurais convolucionais, partes do ramo de *deep learning*, tem mostrado resultados importantes para a área de detecção de doenças. Assim, no presente trabalho, emprega-se as redes Resnet50, Inception V2, CheXNet, *auto-enconders* e algoritmo SVM para classificar 3 tipos de tumores cerebrais. As redes são desenvolvidas com e sem técnicas de *fine tuning*, extração de características e classificação com SVM. Para a apuração do desempenho dos algoritmos foram introduzidas modificações como ruídos e realce aos dados originais, verificando-se o impacto na classificação final dos tumores. Como resultado, observou-se que a rede Inception V2 com *fine tuning* obteve os melhores resultados na predição, com 95,6% de acurácia no conjunto de dados originais e 94,2% nos dados realçados, demonstrando a robustez da rede.

Palavras-chave: detecção de tumores cerebrais; *deep learning*; redes neurais convolucionais; Resnet50; Inception V2; CheXNet; *auto-enconders*.

ABSTRACT

FRÓES, K. C. **Influence of noise and enhancement in the classification of brain tumors in deep neural networks.** 2022. 43 f. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2022.

Brain tumors represent approximately 4% of cancer deaths in Brazil. Currently, there is no single specific test for its diagnosis, which the earlier is performed, the greater the chances of patient survival. The use of convolutional neural networks, part of deep learning, has shown important results in the area of disease detection. Thus, in the present work, the networks Resnet50, Inception V2, CheXNet, auto-encoders and SVM algorithm are used to classify 3 types of brain tumors. The networks are developed with and without fine tuning techniques, feature extraction and classification with SVM. To determine the performance of the algorithms modifications were introduced, such as noise and enhancement on the original data, verifying the impact on the final classification of tumors. As a result, it was observed that the Inception V2 network with fine tuning obtained the best results in the prediction, with 95,6% accuracy in the original dataset and 94,2% in the modified data, demonstrating the robustness of the network.

Keywords: brain tumor detection; *deep learning*; convolutional neural networks; Resnet50; Inception V2; CheXNet; auto-enconders.

LISTA DE FIGURAS

Figura 1 – Representação das áreas de Inteligência Artificial.....	16
Figura 2 – Representação da arquitetura da Resnet50.....	21
Figura 3 – Representação da arquitetura da Inception V2.....	22
Figura 4 – Representação de <i>data augmentation</i>	23
Figura 5 – Representação de um <i>auto-encoder</i>	24
Figura 6 – Tipos de tumor cerebral e cérebro saudável.....	29
Figura 7 – Exemplo de adição de ruído na imagem.....	30
Figura 8 – Fluxograma de processamento.....	31
Figura 9 – Gráfico da <i>Loss</i> por épocas.....	38

LISTA DE TABELAS

Tabela 1 – Resumo dos métodos discutidos.....	28
Tabela 2 – Estratégias de processamento do <i>dataset</i>	30
Tabela 3 – Resultados dados originais.....	33
Tabela 4 – Resultados dados com realce.....	33
Tabela 5 – Resultados dados com ruído.....	34
Tabela 6 – Resultados dados com ruído e realce.....	34
Tabela 7 – Resultados <i>Auto-encoder</i>	35
Tabela 8 – Resultados Resnet dados originais.....	35
Tabela 9 – Resultados Resnet dados com realce.....	35
Tabela 10 – Resultados Resnet dados com ruído.....	36
Tabela 11 – Resultados Inception V2 dados originais.....	36
Tabela 12 – Resultados Inception V2 dados com realce.....	36
Tabela 13 – Resultados Inception V2 dados com ruído.....	36
Tabela 14 – Resultados CheXNet dados originais.....	37
Tabela 15 – Resultados CheXNet dados com realce.....	37
Tabela 16 – Resultados CheXNet dados com ruído.....	37

SUMÁRIO

1 INTRODUÇÃO.....	16
1.1 Contextualização.....	16
1.2 Justificativa e Motivação.....	17
1.3 Questões de Pesquisa e Objetivos.....	18
2 REFERENCIAL TEÓRICO.....	19
2.1 Redes Neurais Convolucionais.....	19
2.1.1 Rede Resnet50.....	20
2.1.2 Rede Inception V2.....	21
2.1.3 Rede CheXNet.....	22
2.2 Técnicas para melhorar convergência.....	22
2.2.1 <i>Data Augmentation</i>	22
2.2.2 <i>Fine Tuning</i>	23
2.2.3 Realce e Ruídos.....	23
2.3 Técnicas para reduzir dimensionalidade.....	24
2.3.1 <i>Auto-encoders</i>	24
2.3.2 <i>Principal Component Analysis</i>	25
2.4 Classificação com <i>Support Vector Machine</i> (SVM).....	25
3 TRABALHOS RELACIONADOS.....	26
4 DADOS, MÉTODOS E MÉTRICAS DE VALIDAÇÃO.....	29
4.1 Descrição do <i>dataset</i>	29
4.2 Pré-processamento dos dados.....	30
4.3 Treinamento das redes neurais.....	31
4.4 Métricas de validação dos dados.....	32
5 RESULTADOS.....	33
5.1 Resultado dos dados com realce e ruído.....	33
5.2 Resultado com <i>auto-encoders</i>	34
5.3 Resultado com <i>fine tuning</i>	35
5.4 Interpretação geral dos modelos.....	38
6 CONCLUSÃO.....	40
REFERÊNCIAS.....	41

1 INTRODUÇÃO

1.1 Contextualização

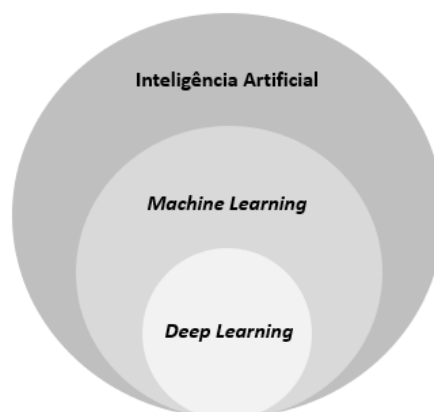
Tumores cerebrais atingem em média 11 mil brasileiros por ano e representam aproximadamente 4% das mortes por câncer no país segundo o Instituto Nacional do Câncer (INCA). Hoje não há maneira específica reconhecida para prevenção desses tumores e não há um único teste específico para o seu diagnóstico, embora no caso de suspeita de sua existência, o procedimento usual seja solicitar a realização de exames de imagem para apoiar a decisão médica.

Atualmente considera-se como tumor cerebral o crescimento anormal de células em alguma estrutura dentro do crânio. Sua classificação é feita com base no tipo celular e no grau do tumor, que varia de I a IV, sendo que os graus I e II representam tumores que crescem mais lentamente e que são menos propensos a se infiltrar nos tecidos cerebrais adjacentes e os graus III e IV justamente o oposto.

Um dos fatores de grande valia para o sucesso do tratamento do tumor cerebral é a sua identificação precoce. Dessa maneira, mostra-se necessário o desenvolvimento de métodos para auxílio na detecção rápida e eficaz de sua existência. Uma abordagem eficiente é o estudo a partir de modelos de *deep learning* para extração de padrões e classificação de dados reais coletados de indivíduos.

Deep learning é uma parte do campo de *machine learning*, que por sua vez é parte do campo de inteligência artificial, que tem por objetivo funcionar como o próprio cérebro humano e assim poder simular suas ações. Seu grande destaque de uso são as tarefas de reconhecimento de fala, identificação de imagem, análise de textos, tradução automática e previsões.

Figura 1 – Representação das áreas de Inteligência Artificial



Neste trabalho, foi implementada uma solução baseada em redes neurais convolucionais para classificação de tumores cerebrais a partir de um *dataset* disponibilizado no Kaggle, com 7022 imagens de ressonâncias magnéticas de cérebros humanos acometidos com tumores e de pacientes saudáveis.

Esses dados foram classificados com as redes Resnet50, Inception V2 e CheXNet, que foram ajustadas para comparar a melhor acurácia com técnicas de *fine tuning* e métodos de classificação que se utilizam de dados de camadas congeladas da rede neural. Além disso, foram feitos ajustes para mitigar desbalanceamentos e análise dos impactos de ruídos, de maneira a contribuir para o maior entendimento e precisão na classificação desses tumores.

1.2 Justificativa e Motivação

A identificação humana de doenças costuma ser um processo demorado, custoso e nem sempre tão acurado, focado em análises de pequenas porções de informações sobre os pacientes. Dessa maneira, a descoberta de métodos mais baratos e acurados para detecção de doenças é um dos grandes benefícios que podem ser obtidos com técnicas sofisticadas de *machine learning*, que conseguem trabalhar com complexos e grandes volumes de dados. Além disso, tais métodos mostram-se capazes de fomentar a detecção prematura de doenças bem como a descoberta e desenvolvimento de novas terapias.

Embora na literatura estejam cada vez mais sendo publicados estudos *de machine learning* aplicados à dados biológicos, a área ainda possui muitas limitações e desafios que podem ser explorados, como a dificuldade em encontrar dados rotulados que possibilitem o treinamento dos modelos de aprendizagem. Outro fator importante é que a divulgação de dados reais de pacientes ainda é um assunto sensível e não há um padrão universal para divulgação deles, o que dificulta o avanço nesses estudos. Além disso, por se tratar de problemas do mundo real, espera-se uma confiança elevada nos resultados obtidos, sendo necessário grande número de testes para apresentação de um baixo risco de classificação incorreta dos dados.

Dado que um dos principais desafios de sistemas supervisionados é definir um conjunto de treinamento que seja representativo e conciso, estudos recentes apontam a necessidade de soluções que atinjam bom desempenho de classificação utilizando poucos dados rotulados. Neste contexto, é proposto o uso de técnicas de extração de características e de ajustes

normalizativos para o desenvolvimento de uma solução capaz de classificar corretamente dados de tumores cerebrais.

1.3 Questões de Pesquisa e Objetivos

Neste trabalho espera-se que com a rotulação de um conjunto relativamente significativo de dados sobre tumores cerebrais, o modelo de aprendizado proposto consiga classificar corretamente os dados como pertencentes ou não à classe de interesse. Para nortear este cenário, tem-se como principal interesse responder as seguintes perguntas:

Q1 “ *Qual modelo apresenta maior desempenho na classificação correta dos dados?* “

Q2 “ *Qual o efeito de se adicionar ruído aos dados no que diz respeito a classificação?* “

Diante destas questões de pesquisa são definidos os seguintes objetivos para o desenvolvimento deste trabalho:

- Mapear algoritmos de aprendizado em redes neurais disponíveis na literatura, analisando suas lacunas e desempenhos na classificação de diferentes cenários (dados originais sem pré-processamento e pré-processados), de modo a encontrar justificativas para o desempenho deles quando aplicados a base de estudo deste projeto.
- Verificar possíveis métodos de pré-processamento dos dados para obtenção de melhores resultados para os algoritmos estudados.
- Analisar o impacto de ruído e realce nos dados em relação a classificação.

A partir dos itens propostos espera-se obter um bom modelo para avaliação dos diferentes tumores que possa servir para predição de dados futuros.

2 REFERENCIAL TEÓRICO

Neste trabalho utilizou-se diferentes modelos de *deep learning* para a tarefa de classificação de tumores cerebrais. A motivação para o uso desta sub-área da inteligência artificial é baseada no grande destaque que esses algoritmos têm revelado em alguns setores como o reconhecimento de fala, reconhecimento de padrões, análise de textos e sentimentos, traduções automáticas, entre outros.

O *deep learning* é o aprendizado baseado em redes neurais profundas, que recebe esse adjetivo devido ao fato de os algoritmos geralmente apresentarem um elevado número de camadas. Essas camadas são responsáveis por aplicar transformações não lineares nos dados de entrada, criar e passar adiante representações abstratas dos dados para as próximas camadas e gerar um modelo estatístico como saída. Dessa maneira, o computador é capaz de aprender através do seu próprio processamento de dados.

Como desvantagem tem-se que esse processo de aprendizado requer acesso a grandes quantidades de dados para treinamento e alto poder de processamento computacional, geralmente utilizando GPUs e TPUs. Além disso, pode-se enfrentar problemas de *overfitting* e de dissipação do gradiente, fenômeno em que o gradiente apresenta a tendência de diminuir à medida que se move para trás através das camadas da rede, no movimento chamado de retropropagação.

Mais adiante serão abordados os conceitos necessários para o entendimento do trabalho, como a definição de redes neurais e breves introduções sobre as redes utilizadas, técnicas para melhora de convergência e técnicas de redução de dimensionalidade dos dados.

2.1 Redes Neurais Convolucionais

Redes neurais convolucionais (CNN - do inglês *Convolutional Neural Network*) são algoritmos de aprendizado profundo criados inicialmente em 1998 por Fukushima[1], cuja arquitetura é baseada no córtex de animais. Dessa maneira, as redes são dotadas de vários neurônios, sendo que cada neurônio tem seu respectivo peso e viés, além de algumas camadas em sua organização, como camadas de input e de output e camadas ocultas de convolução, *pooling*, camadas totalmente conectadas e de perda/classificação[2].

Essas redes são amplamente utilizadas em várias vertentes, podendo ser destacadas as tarefas de classificação de imagens[3], processamento de voz e de linguagem natural,

reconhecimento de atividade humana[4], reconhecimento de expressões faciais humanas[5], modelagem de *features*[6], detecção de objetos[7], entre outros.

Historicamente, seus maiores desafios têm sido sua grande demanda por dados para input na rede, classificação distorcida ao se adicionar técnicas de introdução de ruído, problemas de *overfitting* e explosão de gradiente[8].

Para o alto nível de precisão alcançado por essas redes, que hoje são consideradas o estado da arte, vale ressaltar o funcionamento de suas camadas ocultas:

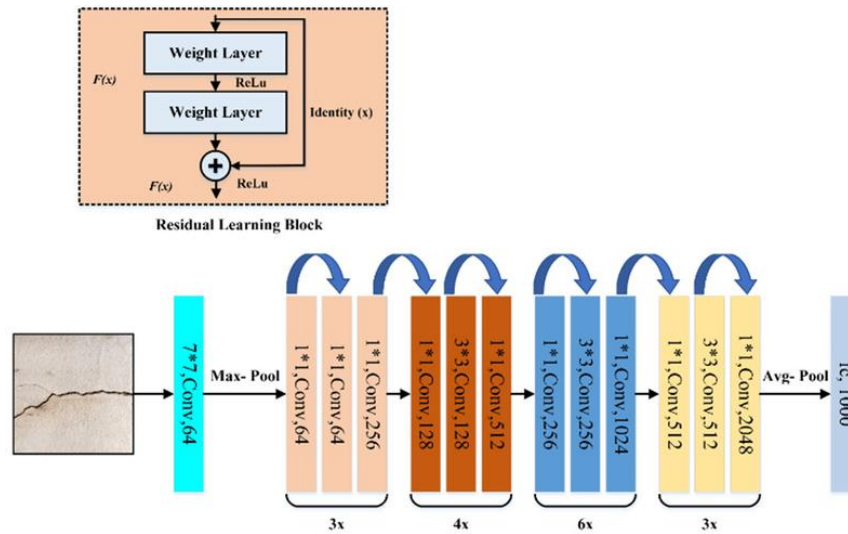
- **Camada de convolução:** sua função é a de transformar informações através de filtros espaciais lineares denominados de kernels. Ao longo da convolução, o kernel se desloca ao longo da imagem e vai multiplicando e somando os valores sobrepostos, de acordo com um número de passos determinado previamente, de maneira que ao final temos uma projeção dos dados de entrada.
- **Camada de *Pooling*:** sua função é a de reduzir a dimensionalidade dos dados através de algumas funções, como por exemplo a *max pooling*[9].
- **Camada totalmente conectada:** sua função é a de receber o *output* da camada anterior, aplicar funções de transformação nos dados de acordo com os pesos gerados na rede e passar esses dados para a próxima camada. Em outras palavras, ela conecta cada neurônio de input a cada neurônio de *output*.
- **Camada de perda/classificação:** sua função é a de determinar o quanto os resultados do treinamento diferem do resultado esperado, de maneira a guiar o treinamento da rede.

2.1.1 Rede Resnet50

A Resnet50 (*Residual Network 50*) é uma rede neural convolucional que possui 50 camadas de profundidade, admitindo como input de imagem o tamanho de 224x224. Foi criada em 2015 pelos pesquisadores Kaiming He, Xiangyu Zhang, Shaoqing Ren e Jian Sun para a tarefa de reconhecimento de imagens, tendo potencial de aprender funções residuais, ou seja, funções que calculam as diferenças entre os valores observados e preditos pelo modelo.

A rede ganhou grande destaque devido ao fato de conseguir lidar com a dissipação do gradiente. Na prática, isso quer dizer que neurônios nas camadas anteriores aprendem muito mais lentamente que os neurônios nas camadas posteriores[10].

Figura 2 – Representação da arquitetura da Resnet50



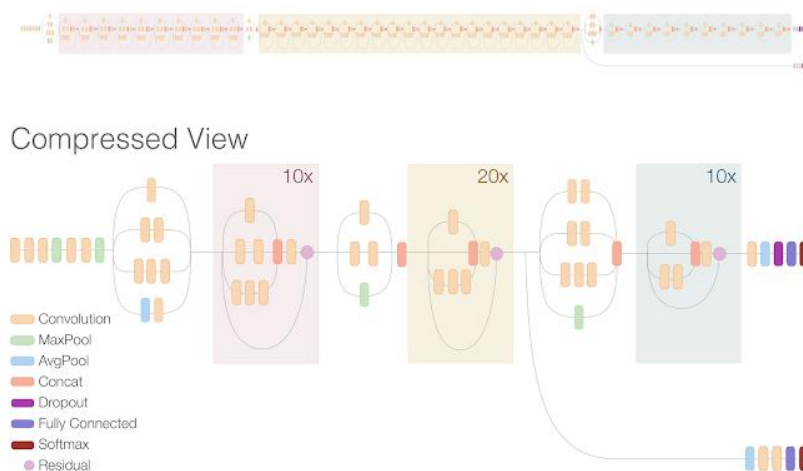
Fonte: Ali et al[11]

2.1.2 Rede Inception V2

A Inception V2 é uma rede neural convolucional de 164 camadas de profundidade, derivada da rede Inception V1, sendo criada com a premissa de aprimorar pontos da última. Nesse sentido, um de seus grandes diferenciais em relação a Inception V1 foi a troca da convolução 5x5 para 2 convoluções 3x3, aumentando a performance de processamento da rede. Além disso, a Inception V2 utiliza-se da troca da fatoração NxN por fatoração 1xN e Nx1, de maneira que uma convolução 3x3 é transformada em uma convolução 1x3 seguida de uma convolução 3x1, que é menos complexa e portanto mais barata computacionalmente.

A rede admite o tamanho de input de imagem 299x299 e trouxe como ganho de sua criação a diminuição do efeito gargalo (*bottleneck*), que ocorre quando diminuimos muito a dimensão do dado, de maneira que se perde informações.

Figura 3 – Representação da arquitetura da Inception V2



Fonte: Improving Inception and Image Classification in TensorFlow[12]

2.1.3 Rede CheXNet

A CheXNet é uma rede que possui 121 camadas de profundidade e que foi baseada originalmente na arquitetura da rede DenseNet-121. Foi criada em 2017 pelos pesquisadores Rajpurkar et al[13] e treinada para detecção de pneumonia através do *dataset* ChestX-ray14, que contém mais de 100 mil imagens frontais de raio-X anotadas com 14 tipos possíveis de patologias, sendo uma rede treinada com imagens em tons de cinza. Seu tamanho de input de imagem é de 224x224.

Na época em que foi criada, atingiu resultados que foram considerados o estado da arte para detecção de pneumonia, de modo que seu uso pode ser benéfico ao se tratar de *datasets* de doenças médicas.

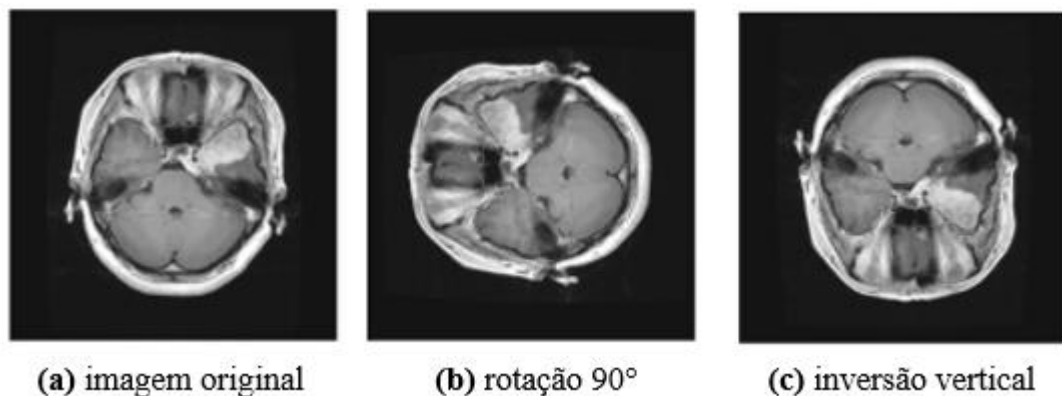
2.2 Técnicas para melhorar convergência

2.2.1 Data Augmentation

Data augmentation é uma estratégia utilizada no problema do desbalanceamento entre as classes, bem como também para evitar *overfitting*. Essa técnica consiste na geração de novas imagens, que pode ser feita de várias formas, como transformação do espaço de cores, mistura de imagens, *random erasing*, aumento do espaço da *feature*, aumentos baseados em redes GAN, transformações geométricas entre outros [14].

Na figura 4, pode-se ver um exemplo de transformação geométrica:

Figura 4 - Representação de *data augmentation*



2.2.2 Fine Tuning

A técnica de *fine tuning* consiste na utilização dos dados de interesse em uma rede neural já conhecida e treinada previamente, de maneira a utilizar todos os seus parâmetros e design, exceto a camada de *output*, que será criada e treinada baseando-se nos dados atuais, com seu respectivo número de classes.

Essa metodologia é utilizada uma vez que as camadas iniciais provem características genéricas a qualquer problema, já as camadas finais apresentam maior especificidade. Dessa maneira, com essa técnica podemos escolher quais camadas finais tem-se o interesse de ajustar para nosso modelo.

Além disso, a técnica propicia o ganho computacional de não recalcular 100% dos pesos para cada problema a ser analisado, sendo também útil para melhorar a generalização de modelos quando se tem poucos dados disponíveis[15].

2.2.3 Realce e Ruídos

A modificação de dados dos *datasets* é uma técnica utilizada para se verificar o comportamento e resiliência dos algoritmos desenvolvidos frente a variações, tendo a possibilidade de aumentar ou diminuir o poder de predição destes. Essa estratégia é considerada inclusive base para a área de *adversarial attacks*, em que pequenas mudanças nos pixels das imagens podem causar mudanças consideradas drásticas no potencial de predição dos algoritmos.

No presente trabalho, um dos intuitos é se verificar o efeito na classificação dos algoritmos ao se introduzir ruídos nos dados originais. Dessa maneira, foram utilizadas técnicas de realce de brilho e contraste das imagens e geração de distorções em imagens escolhidas aleatoriamente no processo.

2.3 Técnicas para reduzir dimensionalidade

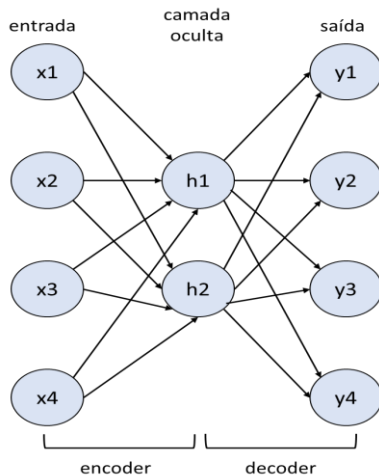
2.3.1 Auto-enconders

Auto-enconder é uma rede neural amplamente difundida para redução de dimensionalidade dos dados, de tal maneira que se seleciona apenas as *features* mais relevantes para o problema em questão, tornando assim a análise de grandes volumes de dados um processo com menor custo computacional.

O algoritmo consiste no aprendizado não supervisionado de uma simplificação da representação dos dados originais, de tal maneira que estes são mapeados para um espaço de dimensão reduzida. Para se chegar a essa representação, o método faz a minimização da função de perda entre os dados originais e os dados decodificados[16].

Sua estrutura consiste basicamente em 3 seções de neurônios: a camada de entrada, uma camada oculta contendo menos neurônios do que a camada anterior e uma camada de saída, composta pelo mesmo número de neurônios da camada de entrada. O *encoder* é a parte que fará a transformação dos dados de maneira a comprimi-los e o *decoder* é a parte responsável por recuperar os dados dessa representação[17].

Figura 5 - Representação de um *auto-enconder*



2.3.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) é um algoritmo não supervisionado utilizado na redução da dimensionalidade dos dados. Sua estratégia é projetar o espaço das *features* para espaços de dimensões reduzidas de acordo com o valor especificado no algoritmo. É um método de simples implementação e oferece como benefício um baixo custo computacional para utilização, além da possibilidade de se fazer a seleção de *features*.

2.4 Classificação com Support Vector Machine (SVM)

Support Vector Machine(SVM) é um algoritmo supervisionado utilizado para análise e classificação de dados linearmente e não-linearmente separáveis, tendo uma taxa de aprendizagem rápida mesmo em densos conjuntos de dados[18].

Seu funcionamento nos casos linearmente separáveis consiste na construção de um hiperplano ótimo que separa 2 diferentes grupos de vetores de *features* em uma margem máxima. Já nos casos não linearmente separáveis, o SVM mapeia a variável de input para um alto espaço dimensional de *features* através de uma função kernel, que pode ser por exemplo a função de kernel gaussiana[19].

É um dos algoritmos supervisionados mais utilizados e como as redes neurais convolucionais, pode ser utilizado para reconhecimento de expressões faciais humanas, processamento de voz e reconhecimento de imagens, além de funções como detecção de genes, classificação de textos, análises químicas[20], entre outros.

No entanto, é necessário levar em consideração alguns pontos antes de adotar esse algoritmo, uma vez que encontrar uma função kernel adequada não é um trabalho fácil e o modelo final construído pelo SVM não é facilmente interpretável, ou seja, não é trivial demonstrar o que foi aprendido pela rede para se chegar ao resultado final demonstrado.

3 TRABALHOS RELACIONADOS

Ao longo dos anos, com o avanço do conhecimento e compartilhamento de dados, vários estudos foram publicados no que diz respeito a detecção de tumores através de imagens de ressonância magnética. Esses estudos variam em relação aos algoritmos, parâmetros e abordagens, sendo que no presente trabalho foram considerados apenas os trabalhos relacionados a tumores localizados no cérebro, destacando os trabalhos que são considerados o estado da arte.

Nesse sentido, podemos citar como marco inicial das análises a partir algoritmos para detecção de tumores cerebrais o trabalho conduzido em 2015 por Cheng et al[21], que também estudou tumores do tipo glioma, meningioma e pituitário. Os autores se utilizaram de técnicas de dilatação de imagens para aumentar a região do tumor e adicionalmente testaram uma metodologia com posterior divisão dessas regiões em sub-regiões em formato de anéis. Para esses dois processos, é utilizado três métodos de extração de *features*: histograma de intensidade, matriz de co-ocorrência de níveis de cinza e modelo de *bag-of-words*(BoW). Os resultados de acurácia obtidos para esses métodos, na ordem em que aparecem, foram da ordem de 82,31%, 84,75% e 88,19%. Ao se adicionar a estratégia de sub-regiões em formato de anéis, os mesmos métodos atingiram resultados de 87,54%, 89,72% e 91,28% respectivamente. Os mesmos autores publicaram novo estudo em 2016[22] adicionado um passo extra no processo que consistia na aplicação de um kernel Fisher, que levou a obtenção de uma precisão de 94.68%.

Outros autores como Hemanth et al[23] fizeram modificações, especialmente nos pesos, nas redes neurais *Counter Propagation Neural Network* (CPN) e Kohonen neural network (KNN) para classificar 4 tipos de tumores cerebrais. Além do pré-processamento das imagens para níveis de cinza entre 0 e 1, um dos passos propostos no processo consiste na extração de oito features baseadas no histograma de primeira ordem e na matriz de co-ocorrência de níveis de cinza. Ao final do processo, obteve-se o resultado de 92% de acurácia para a rede modificada de *Counter Propagation* e 98% de acurácia para a rede modificada de Kohonen.

Entre 2019 e 2020 outros artigos importantes para o estado da arte nesse problema foram publicados atingindo acurácias superiores a 94%. Um desses trabalhos foi conduzido por Guamei et al[24], que propõe o uso de um método de extração híbrido de *features* com regularização de *extreme learning machines*(RELM) para classificação. As imagens

inicialmente são pré-processadas por uma regra de normalização min-max para contrastar as partes relevantes das imagens e posteriormente é utilizado método híbrido para extração de *features*, que acaba na RELM para classificação atingindo acurácia de 94,23%.

Anaraki et al[25] propôs um método baseado em redes neurais convolucionais e algoritmo genético para classificação dos tumores a partir de imagens de ressonância magnética. No estudo, as redes neurais são evoluídas através do algoritmo genético, que escolhe os melhores parâmetros para a rede como número de camadas de convolução, número de filtros, etc. Inicialmente, há o pré-processamento das imagens para se adequarem ao intervalo [0,1], bem como o uso de *data augmentation* para geração de novos dados e para melhora da performance da rede. Adicionalmente utiliza-se o método de regularização *dropout*, que remove nós da camada totalmente conectada na fase de treinamento, o que também ajuda a diminuir o *overfitting*. Por fim, outro método utilizado é o de *Bootstrap Aggregating*, que faz a junção das imagens de treino e validação e seleciona aleatoriamente partes desse conjunto, reduzindo o erro de generalização. Toda essa estratégia levou ao resultado de 94,2% de acurácia.

Swati et al[26] e Deepak and Ameer[27] utilizaram redes neurais já treinadas para a classificação dos tumores: VGG19 e GoogleNet respectivamente, adotando a estratégia de *transfer learning*, que consiste na propagação do conhecimento já adquirido por um modelo previamente treinado para ser usado em novos dados não vistos pelo modelo. Os dois trabalhos adotam estratégias de pré-processamento baseadas na normalização min-max para transformar os dados para o intervalo [0,1] e utilizam o processo de validação cruzada com 5 folds. Os resultados obtidos para acurácia são 94.82% para o primeiro e 98% para o segundo.

Uma estratégia um pouco distinta das apresentadas anteriormente foi introduzida no estudo de Ghassemi et al[28]. Nesse trabalho, há o uso de redes neurais profundas, normalização dos dados no intervalo [-1,1], utilização de *data augmentation*, mas o grande diferencial foi o uso de redes GAN para aumentar o tamanho do *dataset*, atingindo ao final do estudo acurácia de 95.6%.

Por fim, o melhor resultado obtido foi o do modelo apresentado por Kokkalla et al[29] que utiliza a rede neural Inception ResNet V2. Um dos pontos vantajosos dessa rede neural é que ela possui uma rede convolucional de resíduo, que é capaz de controlar o problema de dissipação do gradiente. A partir dessa estratégia obteve-se 99,69% de acurácia.

Para ajudar na visualização e entendimento, pode-se observar na tabela 1 um comparativo entre os métodos discutidos:

Tabela 1 – Resumo dos métodos discutidos

Referência	Abordagem	Acurácia
Cheng et al [21]	Sub-regiões + BoW	91,28%
Anaraki et al [25]	CNN + Algoritmo genético	94,20%
Gumaei et al [24]	Extração de features + RELM	94,23%
Cheng et al [22]	Sub-regiões + BoW + kernel Fisher	94,68%
Swati et al [26]	VGG19 + Fine tuning	94,82%
Ghassemi et al [28]	Rede neural profunda + Redes GAN	95,60%
Deepak e Ameer [27]	GoogleNet + Transfer learning	98,00%
Hemanth et al [23]	Kohonen neural network modificada	98,00%
Kokkalla et al [29]	Inception ResNet V2	99,69%

Um ponto a ser destacado é que estes trabalhos utilizam o mesmo contexto de tumores cerebrais, mas não utilizam necessariamente os mesmos *datasets*, de maneira que a comparação direta entre eles não é adequada. Ainda assim, a partir deles pode-se observar como o uso de redes neurais têm provido resultantes relevantes para a área.

Outras abordagens com menores acurácias na classificação de tumores cerebrais podem ser encontradas em trabalhos de 2018 de Ismael e Qader[30], que extraíram *features* de imagens de ressonância magnética através da transformada discreta de *wavelet* e filtros Gabor para posterior uso em uma rede perceptron multicamadas, atingindo acurácia de 91,9%, Abir et al[31], que utilizou a abordagem de redes neurais probabilísticas e atingiu acurácia de 83,33% e Afshar et al[32], que utilizou redes neurais de cápsulas(CapsNet) e atingiu acurácia de 86,56%.

4 DADOS, MÉTODOS E MÉTRICAS DE VALIDAÇÃO

4.1 Descrição do *dataset*

Os dados utilizados para o desenvolvimento desse trabalho são provenientes do *dataset* chamado “*brain tumor dataset*”, originalmente disponibilizado pelo repositório Figshare, mas que também pode ser encontrado no repositório público do Kaggle sob o nome de “*Brain Tumor MRI Dataset*” ou através do link: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>.

O *dataset* é constituído de 7023 imagens de ressonâncias magnéticas de cérebros humanos, sendo rotulados em 3 classes distintas de tipos de tumor: glioma, meningioma, pituitário, além de uma classe de imagens de cérebros saudáveis.

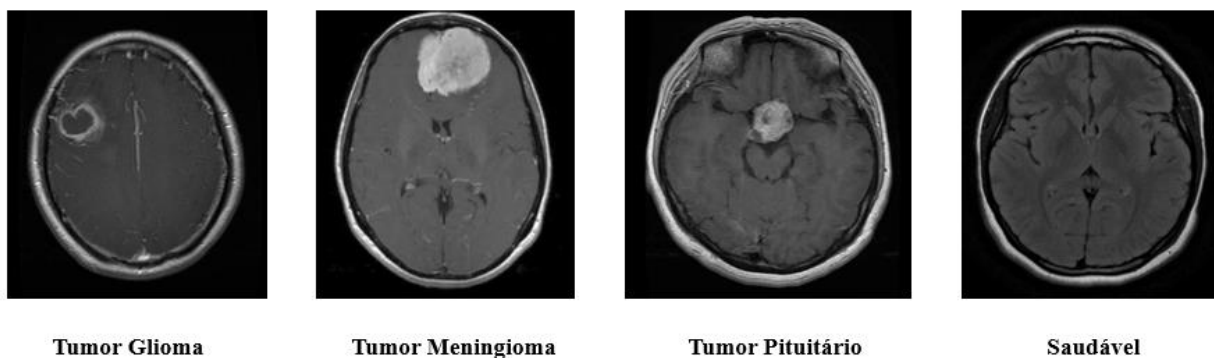
Para ser considerado do tipo glioma o tumor precisa ocorrer nas células gliais e para ser do tipo meningioma deve ocorrer nas meninges. Já os tumores pituitários são assim classificados pois tem sua origem na glândula pituitária.

No *dataset* as classes estão distribuídas da seguinte maneira:

- Glioma: 1621;
- Meningioma: 1645;
- Pituitário: 2000;
- Saudável: 1757.

Na figura 6 pode-se observar exemplos de ressonâncias magnéticas para cada tipo de tumor descrito e um exemplo de um cérebro saudável:

Figura 6 – Tipos de tumor cerebral e cérebro saudável



4.2 Pré-processamento dos dados

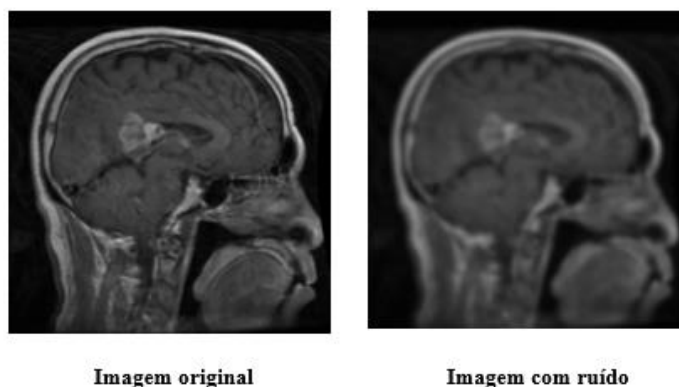
Para se testar a robustez dos algoritmos implementados, este trabalho utilizou-se de diferentes estratégias para avaliação dos resultados, de forma que a etapa de pré-processamento dos dados é distinta de acordo com o método escolhido. A tabela 2 visa demonstrar a diferença entre os métodos aqui abordados:

Tabela 2 – Estratégias de processamento do *dataset*

Método	Realce da imagem	Adição de Ruído na imagem
1	-	-
2	X	-
3	-	X
4	X	X

Para os métodos com realce utilizou-se a estratégia de realçar o brilho e o contraste de todas as imagens em valor escolhido aleatoriamente na curva normal de intervalo $[0.8, 1.2]$, que a partir de testes foi o intervalo que propiciou impacto mais significativo para os experimentos. Já para os métodos de ruído, foram realizadas distorções em imagens escolhidas ao acaso no processo, a partir da adição de um valor aleatório entre 1 e 12, com passo de tamanho 2 na função *GaussianBlur*, disponibilizada na biblioteca cv2 do Python.

Figura 7 – Exemplo de adição de ruído na imagem



Após esta etapa inicial de pré-processamento, também foi feito *data augmentation* de maneira a tentar reduzir *overfitting*, sendo geradas 10% de imagens novas para cada classe, através da rotação em 90 ou 180 graus das imagens selecionadas aleatoriamente. Por se tratar de um conjunto de dados que não é tão desbalanceado, a introdução de novos dados visa também buscar maior representatividade para o conjunto de treinamento.

Após esses pré-processamentos, as imagens foram divididas entre treinamento, validação e teste, numa proporção de 70% para treino, 15% para validação e 15% para teste. Como a intenção desse trabalho é avaliar as métricas para redes neurais distintas, para cada algoritmo as imagens foram re-dimensionadas para atender as necessidades específicas da rede em questão, por exemplo a Resnet50 e a CheXNet que utilizam como input de entrada imagens 224x224 e a Inception V2 que utiliza input 299x299.

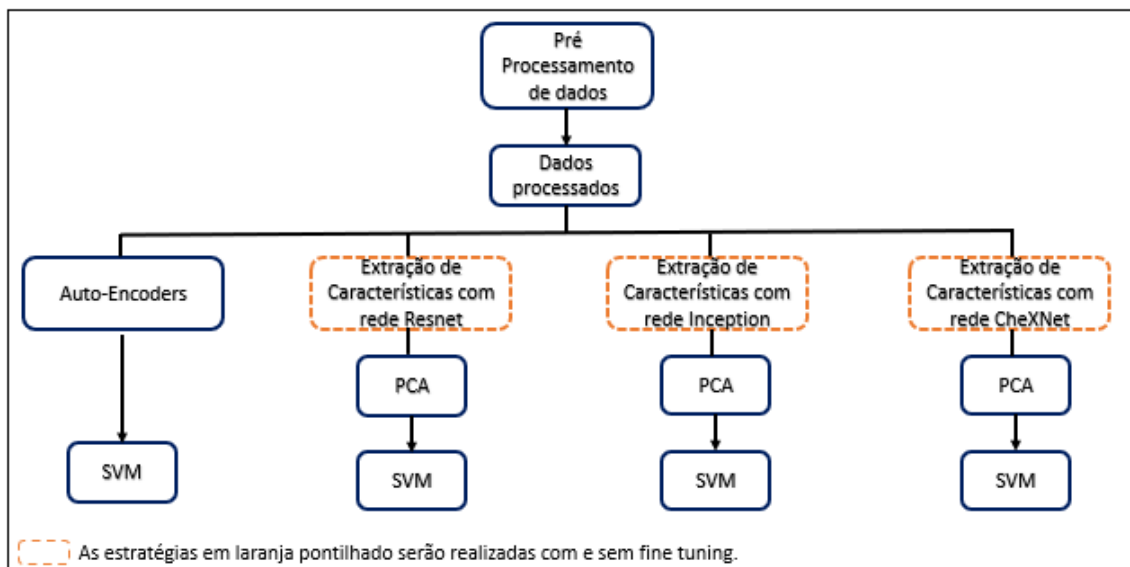
4.3 Treinamento das redes neurais

Após o pré-processamento dos dados, a performance foi obtida com SVM para as redes neurais com *auto-encoders* e nas redes Resnet50, InceptionV2 e CheXNet.

Para o método com *auto-encoders*, que visa manter as características relevantes dos dados de entrada, é aplicado o classificador SVM para cálculo da acurácia do modelo proposto. Para os demais testes, utilizou-se a penúltima camada das redes antes das camadas completamente conectadas para se fazer a extração das características mais significativas para os dados. Essa camada é escolhida uma vez que as próximas camadas que serão completamente conectadas são muito especializadas nos dados em que foram treinadas previamente, de maneira que não serão muito úteis para dados de tarefas novas.

Essa estratégia de extração de características será feita sob as redes sem e com *fine tuning* para comparação. Após esse processo, a classificação dos dados é dada através do modelo SVM aplicado sob os dados já modificados pela técnica de PCA, que visa diminuir a dimensão das características extraídas pelas redes.

Figura 8 – Fluxograma de processamento



Para se realizar o *fine tuning* é necessário definir parâmetros como *batchsize*, *optimizer*, número de épocas e *learning rate* para o algoritmo, dado que cada tarefa pode atingir melhores resultados com valores diferentes. O valor adequado para esses parâmetros pode ser definido realizando-se a análise incremental dos valores e observando os resultados obtidos de acordo com o critério de interesse. Para os experimentos, o *setup* de treinamento foi definido como tendo número máximo de épocas = 40, *learning-rate* = $1e^{-4}$ e momentum = 0.9.

4.4 Métricas de validação dos dados

As análises dos resultados obtidos pelas diferentes estratégias são baseadas na comparação das principais métricas de validação conhecidas na literatura: acurácia, precisão, revocação e F1-Score.

A métrica de acurácia visa indicar a performance geral do modelo adotado, avaliando a proximidade entre o resultado experimental e o valor real de referência. Em outras palavras, ela determina o grau de exatidão: quanto maior seu valor, mais próximo se está do valor real.

A métrica de precisão visa indicar o grau de variação ocasionado por medições diferentes, dessa maneira entende-se como mais preciso um processo que apresente menor variabilidade entre os valores. Essa métrica é muito usada em situações em que faz sentido penalizar com maior força casos de falsos positivos - casos em que o teste prediz determinada característica de interesse presente, porém na realidade ela é ausente – por serem considerados mais prejudiciais do que falsos negativos - casos em que o teste prediz determinada característica de interesse ausente, porém na realidade ela é presente.

A revocação por sua vez é utilizada no cenário contrário ao da precisão: nesses casos, é mais interessante penalizar casos de falsos negativos, que são mais prejudiciais do que falsos positivos. Seu intuito é avaliar a capacidade do modelo em detectar com sucesso resultados classificados como positivos.

Por fim, a medida F1-Score é a média harmônica entre revocação e precisão, podendo ser entendida como uma medida para apontar a confiabilidade da acurácia.

5 RESULTADOS

Conforme mencionado nos capítulos anteriores foram realizados diferentes experimentos com 4 redes neurais: *auto-encoders*, Resnet, Inception e CheXNet. Cada rede foi testada com versões dos dados originais e versões com dados alterados, bem como também foi utilizada a estratégia de realizar o *fine tuning* nas redes para se verificar o impacto nos resultados.

Apesar do *dataset* não apresentar desbalanceamento expressivo realizou-se processos para transformação dessas informações conforme mencionado na seção de metodologia no intuito de promover maior variabilidade e se testar a robustez dos métodos. Dessa maneira, os experimentos foram feitos de maneira independente e podem ser comparados de acordo com as métricas de acurácia, precisão, revocação e F1-Score nas seções posteriores.

5.1 Resultado dos dados com realce e ruído

Nesse experimento o intuito era avaliar o impacto da modificação dos dados originais, seja com técnicas de realce ou de ruído, na performance final das redes. Para tanto, as 4 redes neurais foram testadas com os dados originais e posteriormente com os dados modificados para realizar a comparação.

O processo realizado consistiu na extração de características, que foram então submetidas a técnica de PCA, para por último ser realizada a classificação com o algoritmo SVM. Dessa forma, tem-se os resultados mostrados nas tabelas 3, 4, 5 e 6:

Tabela 3 – Resultados dados originais

Dados Originais				
Rede Neural	Acurácia (%)	Precisão (%)	Revocação (%)	F1-score (%)
Resnet	83,7%	84,1%	83,7%	83,8%
Inception V2	91,9%	92,2%	92,0%	92,0%
CheXNet	88,3%	88,6%	88,3%	88,4%

Tabela 4 – Resultados dados com realce

Dados com Realce				
Rede Neural	Acurácia (%)	Precisão (%)	Revocação (%)	F1-score (%)
Resnet	83,5%	83,7%	83,4%	83,5%
Inception V2	91,4%	91,7%	91,4%	91,5%
CheXNet	88,1%	88,3%	88,1%	88,2%

Tabela 5 – Resultados dados com ruído

Dados com Ruído				
Rede Neural	Acurácia (%)	Precisão (%)	Revocação (%)	F1-score (%)
Resnet	78,3%	78,6%	78,3%	78,4%
Inception V2	90,5%	90,9%	90,5%	90,6%
CheXNet	87,2%	87,5%	87,3%	87,3%

Tabela 6 – Resultados dados com ruído e realce

Dados com Ruído e Realce				
Rede Neural	Acurácia (%)	Precisão (%)	Revocação (%)	F1-score (%)
Resnet	79,3%	79,6%	79,6%	79,5%
Inception V2	89,3%	89,7%	89,3%	89,4%
CheXNet	86,1%	86,2%	86,1%	86,1%

De acordo com as tabelas apresentadas temos que em desempenho geral as redes Inception V2 e CheXNet atingem melhores resultados que as demais, sendo o melhor resultado alcançado se mantendo os dados originais com a rede Inception, alcançando acurácia máxima de 91,9%.

Porém, o grande ponto de interesse do trabalho era a avaliação do impacto de alterações nos dados originais. Assim, de acordo com os resultados demonstrados, temos como melhor resultado a acurácia de 91,4% no *dataset* com realce e 90,5% no *dataset* com ruído, que nos mostra diferenças percentuais de 0,5% e 1,4% respectivamente em relação ao desempenho alcançando com os dados originais na Inception V2. Temos então que as redes se comportaram bem frente a essas alterações, pois os resultados não variaram de maneira muito expressiva entre experimentos, mostrando a robustez dos algoritmos desenvolvidos.

5.2 Resultado com *Auto-enconders*

Para esse experimento, as imagens foram dimensionadas com *input* 256x256 e passadas para o classificador. A primeira camada de convolução usa filtro de 32 para gerar os *features maps*, que são passados para uma camada da *maxpooling* que diminui as dimensões da imagem por um fator de 2. Essa imagem é então passada para a segunda camada de convolução que usa filtro de 64 e então chega a segunda camada de *maxpooling*, que também diminui as dimensões da imagem por um fator de 2. Em seguida, esse output passa pela camada de *flatten* e vai para

a camada totalmente conectada. Esse *setup* é feito para os dados originais, com realce, com ruído e com realce e ruído juntos para comparação

Na tabela 7 pode-se verificar os resultados alcançados:

Tabela 7 – Resultados Auto-encoder

Auto-encoder				
Rede Neural	Acurácia (%)	Precisão (%)	Revocação (%)	F1-score (%)
Dados originais	81,4%	81,6%	81,4%	81,4%
Dados com realce	81,0%	81,2%	81,0%	81,1%
Dados com ruído	77,5%	77,5%	77,5%	77,6%
Dados com ruído e realce	76,6%	76,6%	76,6%	76,6%

Observa-se através dos dados que o melhor resultado para todas as métricas foi obtido nos dados originais, atingindo acurácia de 81,4%. Já nos dados modificados com realce e ruído a acurácia ficou em 76,6%, o que representa uma diferença de 4,8% em relação ao melhor resultado. Outro ponto a se destacar é a baixa diferença percentual em relação aos resultados obtidos com a rede Resnet sem *fine tuning*.

5.3 Resultado com *fine tuning*

Já para os experimentos com *fine tuning* nas redes manteve-se fixos os parâmetros do número máximo de épocas = 40, *learning-rate* = $1e^{-4}$ e momentum = 0.9 com o intuito de fazer a comparação entre os diferentes modelos. Além disso, para todas as redes testou-se os *solvers* Adam, SGD e RMS Prop.

Os resultados observados são mostrados nas tabelas 8, 9 e 10:

Tabela 8 – Resultados Resnet dados originais

Resnet - Dados Originais			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	15	16	14
Acurácia	86,4%	89,3%	84,7%
Precisão	86,5%	89,5%	84,8%
Revocação	86,4%	89,3%	84,7%
Tempo (min.)	18	14	15

Tabela 9 – Resultados Resnet dados com realce

Resnet - Dados com Realce			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	15	17	13
Acurácia	84,1%	87,3%	83,1%
Precisão	84,2%	87,6%	83,2%
Revocação	84,1%	87,4%	83,1%
Tempo (min.)	16	18	14

Tabela 10 – Resultados Resnet dados com ruído

Resnet - Dados com Ruído			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	16	18	14
Acurácia	81,3%	85,1%	79,7%
Precisão	81,4%	85,3%	79,7%
Revocação	81,3%	85,2%	79,6%
Tempo (min.)	18	21	16

Para o *fine tuning* na rede Resnet o melhor resultado alcançado na acurácia foi de 89,32%, com os dados originais e *solver* SGD e o pior resultado de acurácia foi de 79,73%, com os dados com ruído e *solver* RMS.

Tabela 11 – Resultados Inception V2 dados originais

Inception V2 - Dados Originais			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	9	11	8
Acurácia	94,3%	95,6%	92,2%
Precisão	94,5%	95,7%	92,3%
Revocação	94,4%	95,6%	92,2%
Tempo (min.)	40	51	39

Tabela 12 – Resultados Inception V2 dados com realce

Inception V2 - Dados com Realce			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	11	12	9
Acurácia	93,9%	94,2%	91,0%
Precisão	94,1%	94,3%	91,3%
Revocação	93,9%	94,2%	91,1%
Tempo (min.)	47	53	41

Tabela 13– Resultados Inception V2 dados com ruído

Inception V2 - Dados com Ruído			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	13	14	11
Acurácia	92,3%	93,4%	88,4%
Precisão	92,4%	93,7%	88,5%
Revocação	92,3%	93,4%	88,4%
Tempo (min.)	55	56	49

Para o *fine tuning* na rede Inception o melhor resultado alcançado na acurácia foi de 95,64%, com os dados originais e *solver* SGD e o pior resultado de acurácia foi de 88,43%, com os dados com ruído e *solver* RMS.

Tabela 14– Resultados CheXNet dados originais

CheXNet - Dados Originais			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	9	12	7
Acurácia	90,7%	91,1%	88,7%
Precisão	90,8%	91,3%	88,8%
Revocação	90,7%	91,1%	88,7%
Tempo (min.)	35	39	31

Tabela 15– Resultados CheXNet dados com realce

CheXNet - Dados com Realce			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	9	11	8
Acurácia	89,4%	90,5%	86,7%
Precisão	89,4%	90,7%	86,8%
Revocação	89,3%	90,5%	86,6%
Tempo (min.)	33	36	30

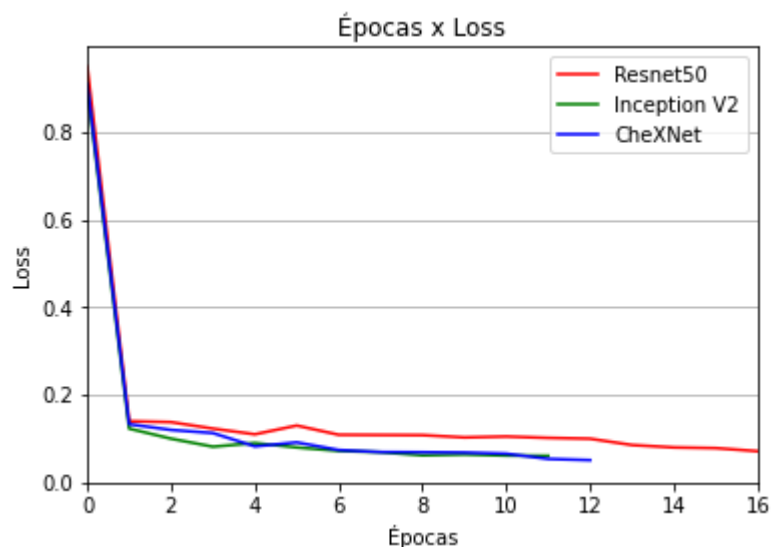
Tabela 16– Resultados CheXNet dados com ruído

CheXNet - Dados com Ruído			
Rede Neural	Adam	SGD	RMS Prop
Batch size	10	10	10
Max. Épocas	10	11	7
Acurácia	87,1%	89,1%	84,4%
Precisão	87,3%	89,3%	84,6%
Revocação	87,2%	89,1%	84,4%
Tempo (min.)	35	37	29

Para o *fine tuning* na rede CheXnet o melhor resultado alcançado na acurácia foi de 91,12%, com os dados originais e *solver* SGD e o pior resultado de acurácia foi de 84,41%, com os dados com ruído e *solver* RMS.

O gráfico da *loss* por épocas para as 3 redes pode ser observado na figura 9:

Figura 9 – Gráfico da *Loss* por épocas



5.4 Intepretação geral dos modelos

Por fim, comparando-se todos os experimentos, concluí-se que o melhor modelo foi o da rede Inception V2 com *fine tuning* e *solver* SGD, que atingiu acurácia de 95,6%. O ajuste da rede com o *fine tuning* propiciou um ganho percentual de 3,7% na acurácia ao se comparar esse resultado com o resultado da rede sem a técnica. Estes valores corroboram os dados apresentados no capítulo 2 de trabalhos relacionados, em que o melhor resultado foi obtido ao se empregar a mesma rede, atingindo acurácia de 99,69%.

Já o pior resultado foi obtido ao se utilizar os dados com realce e ruído juntos na rede Resnet sem *fine tuning*, atingindo acurácia de 79,3%. Ao se comparar esse mesmo cenário de teste, porém com *fine tuning* e *solver* SGD obtemos uma diferença percentual de 5,4% na acurácia, mostrando o ganho que se obtém ao adaptar a rede com a técnica. Como o processo de *fine tuning* recalcula os pesos para a camada de output, a rede tem a possibilidade de aprender características relevantes para a tarefa que está sendo desempenhada, podendo impactar positivamente nos resultados.

Ademais, podemos ver que os melhores desempenhos para todos os cenários propostos são obtidos com as redes Inception V2 e CheXNet e os piores resultados com as redes Resnet e *auto-enconders*, sendo que a diferença percentual entre os valores para todas as métricas das últimas não chegou a passar de 3%.

Outro ponto a ser destacado são os melhores desempenhos em relação as métricas de precisão e revocação, que são métricas importantes a serem consideradas dependendo do objetivo almejado. Por exemplo, se é desejado um modelo que não apresente alto valor de falsos positivos, tem-se que o modelo da rede Inception V2 com *fine tuning* é o mais adequado, pois atingiu precisão de 92,2% com os dados originais e 89,7% com os dados realçados e com ruído. Ou seja, este modelo é o qual atingiu melhor desempenho no sentido de não classificar erroneamente uma pessoa saudável como sendo uma pessoa portadora de tumor cerebral.

Da mesma forma, caso o interesse seja em um modelo que não apresente alto valor de falsos negativos, o modelo da rede Inception V2 com *fine tuning* também seria o mais adequado, uma vez que atingiu 92% de revocação com os dados originais e 89,3% com os dados realçados e com ruído. Ou seja, este modelo é o qual atingiu melhor desempenho no sentido de não classificar erroneamente uma pessoa portadora de tumor cerebral como sendo uma pessoa saudável.

Por fim, é importante ressaltar que a escolha de valores para cada parâmetro de treinamento das redes como *batchsize*, *solver*, entre outros, é uma tarefa considerada difícil e que pode influenciar diretamente em fatores importantes do resultado, como por exemplo o poder de generalização das redes.

6 CONCLUSÃO

O objetivo principal deste trabalho era a criação e comparação de modelos de redes neurais que atingissem boa performance para identificação de tumores cerebrais, uma vez que o aprendizado propiciado por essas redes é um importante e eficaz aliado nessa causa.

Assim, realizou-se o treinamento e teste de redes com diferentes arquiteturas e com diferentes estratégias, obtendo-se o melhor resultado com técnicas de *fine tuning* aplicadas a rede Inception V2. Entretanto, é necessário levar em consideração que o processo de *fine tuning* exige certo recurso computacional, uma vez que a camada de output será treinada com base nos dados da tarefa atual, que normalmente são volumosos. A depender do tempo disponível para o projeto, pode ser interessante se fazer uma análise se há disponível outras estratégias de processamento mais leves que gerem bons resultados.

Outro ponto a se testar para o uso dessas redes é a eficiência que elas demonstram frente a ruídos, que pode variar de acordo com as configurações feitas e a tarefa pretendida. Para a tarefa desenvolvida nesse trabalho os resultados foram pouco impactados percentualmente, que provaram a robustez das redes no cenário estudado.

Como mencionado ao longo do trabalho, o aperfeiçoamento e a empregabilidade das redes neurais é um espaço de crescimento ao longo dos anos graças aos ótimos resultados que vem sendo demonstrados e certamente novas técnicas serão criadas e poderão ser utilizadas para resultados cada vez mais assertivos.

7 REFERÊNCIAS

- [1] FUKUSHIMA, K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. In: **Biological Cybernetics**, v. 36, p. 193–202, 1980.
- [2] DHILLON, A.; VERMA, G.K. Convolutional neural network: a review of models, methodologies and applications to object detection. In: **Progress in Artificial Intelligence**, v. 9, p. 85–112, 2020.
- [3] XIE, S.; YANG, T.; WANG, X.; LIN, Y. Hyper-class augmented and regularized deep learning for fine-grained image classification. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**, p. 2645–2654, 2015.
- [4] PARK, S.U.; PARK, J.H.; AL-MASNI, M.A.; AL-ANTARI, M.A.; UDDIN, Z.; KIM, T. A depth camera-based human activity recognition via deep learning recurrent neural network for health and social care services. In: **Procedia Computer Science**, v. 100, p. 78–84, 2016.
- [5] ZHAO, X.; SHI, X.; ZHANG, S. Facial expression recognition via deep learning. In: **IETE Technical Review**, v. 32, p.347–355, 2015.
- [6] FLOYD, M.W.; TURNER, J.T.; AHA, D.W. Using deep learning to automate feature modeling in learning by observation: a preliminary study. In: **AAAI Spring Symposium Series**, 2017.
- [7] TANG, C.; FENG, Y.; YANG, X.; ZHENG, C.; ZHOU, Y. The object detection based on deep learning. In: **4th International Conference on Information Science and Control Engineering (ICISCE)**, p. 723–728, 2017.
- [8] JOSHI, S.; VERMA, D.K.; SAXENA, G.; PARAYE, A. Issues in Training a Convolutional Neural Network Model for Image Classification. In: **International Conference on Advances in Computing and Data Sciences(ICACDS)**, p. 282-293, 2019.
- [9] ARORA, D.; GARG, M.; GUPTA, M. Diving deep in Deep Convolutional Neural Network. In: **2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)**, p. 749-751, 2020.
- [10] Deep Learning Book - Capítulo 34 – O Problema da Dissipação do Gradiente.
Disponível em:
<https://www.deeplearningbook.com.br/o-problema-da-dissipacao-do-gradiente/>.
Acesso em: 21 jun. 2022.
- [11] ALI, L.; ALNAJJAR, F.; JASSMI, H.A.; GOCHO, M.; KHAN, W.; SERHANI, M.A. Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. In: **Sensors**, no. 5: 1688, 2021.

- [12] Improving Inception and Image Classification in TensorFlow. In: ai.googleblog. Disponível em: <https://ai.googleblog.com/2016/08/improving-inception-and-image.html> Acesso em: 30 jul. 2022.
- [13] RAJPURKAR, P.; IRVIN, J.; ZHU, K.; YANG, B.; MEHTA, H.; DUAN, T.; DING, D.; BAGUL, A.; LANGLLOTZ, C.; SHPANSKAYA, K.; LUNGREN, M.; NG, A. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, In: **arXiv:1711.05225**, p. 3-9, 2017.
- [14] SHORTEN, C.; KHOSHGOFTAAR, T. A survey on Image Data Augmentation for Deep Learning. In: **Journal of Big Data**, v. 6, 2019.
- [15] Computer Vision - Fine-Tuning. In: d2l.ai. Disponível em: https://d2l.ai/chapter_computer-vision/fine-tuning.html/. Acesso em: 10 out. 2021.
- [16] MENG, Q.; CATCHPOOLE, D.; SKILLICOM, D.; KENNEDY, P. Relational autoencoder for feature extraction. In: **International Joint Conference on Neural Networks (IJCNN)**, p.364-371, 2017.
- [17] WU, H.; WU, J.; DING, J. Extract Generalization Ability from Convolutional Neural Networks, In: **International Joint Conference on Neural Networks (IJCNN)**, p. 1-6, 2018.
- [18] PARVEEN; SINGH, A. Detection of brain tumor in MRI images, using combination of fuzzy c-means and SVM, In: **2nd International Conference on Signal Processing and Integrated Networks (SPIN)**, p. 98-102, 2015.
- [19] LIU, H.; XIAO, X.; LI, Y.; MI, Q.; YANG, Z. Effective Data Classification via Combining Neural Networks and SVM, In: **Chinese Control And Decision Conference (CCDC)**, p. 4006-4009, 2019.
- [20] DAI, H. Research on SVM improved algorithm for large data classification, In: **IEEE 3rd International Conference on Big Data Analysis (ICBDA)**, p. 181-185, 2018.
- [21] CHENG, J.; HUANG, W.; CAO, S.; YANG, R.; YANG, W.; YUN, Z.; WANG, Z.; FENG; Q. Enhanced performance of brain tumor classification via tumor region augmentation and partition. In: **PLoS ONE 10(12)**, 2015.
- [22] CHENG, J.; YANG, W.; HUANG, M.; HUANG, W.; JIANG, J.; ZHOU, Y.; YANG, R.; ZHAO, J.; FENG, Y.; FENG; Q. Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation. In: **PLoS ONE 11(6)**, 2016.
- [23] HEMANTH, D.; VIJILA, C.; SELVAKUMAR, A.; ANITHA, J. Performance Improved Iteration-Free Artificial Neural Networks for Abnormal Magnetic Resonance Brain Image Classification, In: **Neurocomputing**, v. 130, p. 98-107, 2014.

- [24] GUMAEI, A.; HASSAN, M. M.; HASSAN, M. R.; ALELAIWI, A.; FORTINO, G. A Hybrid Feature Extraction Method With Regularized Extreme Learning Machine for Brain Tumor Classification, In: **IEEE Access**, v. 7, p. 36266-36273, 2019.
- [25] ANARAKI, A.; AYATI, M.; KAZEMI, F.; Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms, In: **Biocybernetics and Biomedical Engineering**, v. 39, p. 63-74, 2019.
- [26] SWATI, Z.; ZHAO, Q.; KABIR, M.; ALI, F.; ALI, Z.; AHMED, S.; LU, J. Brain tumor classification for MR images using transfer learning and fine-tuning, In: **Computerized Medical Imaging and Graphics**, v. 75, p. 34-46, 2019.
- [27] DEEPAK, S.; AMEER, P. Brain tumor classification using deep CNN features via transfer learning, In: **Computers in Biology and Medicine**, v. 111, 2019.
- [28] GHASSEMI, N.; SHOEIBI, A.; ROUHANI, M. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on MR images, In: **Biomedical Signal Processing and Control**, v. 57, 2020.
- [29] KOKKALLA, S., KAKARLA, J., VENKATESWARLU, I.; SINGH, M. Three-class brain tumor classification using deep dense inception residual network. In: **Soft Computing**, v. 25, p. 8721–8729, 2021.
- [30] ISMAEL, M.; ABDEL-QADER, I. Brain tumor classification via statistical features and back-propagation neural network. In: **IEEE International Conference on Electro/Information Technology (EIT)**, p. 252–257, 2018.
- [31] ABIR, T.; SIRAJI, J.; AHMED, E.; KHULNA, B. Analysis of a novel MRI based brain tumour classification using probabilistic neural network (PNN). In: **International journal of scientific research in science, engineering and technology**, v. 4, p. 69–75, 2018.
- [32] AFSHAR, P.; MOHAMMADI, A.; PLATANIOTIS, K. Brain Tumor Type Classification via Capsule Networks, In: **25th IEEE International Conference on Image Processing (ICIP)**, p. 3129-3133, 2018.