

Universidade Federal de São Carlos – UFSCar
Centro de Ciências Exatas e de Tecnologia – CCET
Departamento de Computação – DC
Curso de Bacharelado em Engenharia de Computação – EC

Guilherme Gomes Arcencio

**Adaptação dos algoritmos ROCKET e
MiniRocket para classificação de tumores
cerebrais**

São Carlos
2024

Guilherme Gomes Arcencio

Adaptação dos algoritmos ROCKET e MiniRocket para classificação de tumores cerebrais

Trabalho de conclusão de curso apresentado ao Curso de Bacharelado em Engenharia de Computação do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos para obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Alexandre Magalhães Levada

São Carlos

2024

Este trabalho é dedicado à minha avó Marta, que não pôde me acompanhar até o presente, mas sempre levou em seu coração o meu futuro.

Agradecimentos

Agradeço aos meus pais, Anderson e Fernanda, pelo amor, carinho e apoio em todas as dimensões da vida; a toda a minha família, por me acompanhar na longa jornada universitária; aos meus amigos Gabriel e Nathaelly, pelos incontáveis momentos de alegria e pelas experiências e ensinamentos que me orientaram na graduação; ao professor Alexandre Levada, por nortear e acompanhar o desenvolvimento deste trabalho.

“O presente é de luta, o futuro é nosso.”
(Ernesto “Che” Guevara)

Resumo

Um dos principais métodos de diagnóstico de tumores cerebrais são as imagens de ressonância magnética da cabeça de um paciente. A identificação do tipo de tumor a partir dessas imagens é importante para que este seja tratado de forma adequada, e um dos meios mais relevantes de automatizar esse procedimento é através do uso de Aprendizado de Máquina, em especial o aprendizado profundo e as redes neurais convolucionais. No entanto, esses modelos podem ser custosos em termos de tempo e quantidade de dados para treino e, portanto, este trabalho propôs a adaptação de dois algoritmos da tarefa de classificação de séries temporais para a classificação de imagens: o ROCKET e o MiniRocket, que não empregam aprendizado profundo. Foi mostrado que, apesar dos procedimentos experimentais superficiais apresentados na maior parte dos trabalhos da literatura, os dois algoritmos propostos exibem desempenho tão bom quanto os modelos do estado da arte e que, além disso, o MiniRocket leva apenas uma fração do tempo de treinamento das redes convolucionais para atingir ótimos resultados.

Palavras-chaves: tumor cerebral, ressonância magnética, classificação de imagens, aprendizado de máquina.

Abstract

One of the main methods of brain tumor diagnosis is the magnetic resonance imaging of a patient's head. Identifying a tumor's type using those images is important for its adequate treatment, and one of the most relevant means of automating this process is the use of Machine Learning, in particular deep learning and convolutional neural networks. However, those models may be expensive in terms of training time and data. Therefore, this work proposed the adaptation of two time series classification algorithms to the image classification task: ROCKET and MiniRocket, which do not employ deep learning. We showed that, in spite of simplistic experimental procedures presented throughout most of the literature, both proposed algorithms exhibit state of the art performance and, furthermore, MiniRocket requires a fraction of convolutional networks' training time to achieve great results.

Keywords: brain tumor, magnetic ressonance, image classification, machine learning.

Lista de ilustrações

Figura 1 – Exemplos de imagens de ressonância magnética de pacientes acometidos por (a) meningioma, (b) glioma e (c) tumor pituitário.	17
Figura 2 – Ilustração do mecanismo de classificação do ROCKET.	24
Figura 3 – Ilustração do mecanismo de geração de filtros convolucionais do MiniRocket.	25
Figura 4 – Adaptações dos filtros convolucionais do ROCKET e do MiniRocket para entradas bidimensionais.	28
Figura 5 – Matrizes de confusão dos algoritmos (a) ROCKET, (b) MiniRocket e (c) CNN, amostradas do primeiro teste de validação cruzada de 5 pastas separadas por paciente.	32
Figura 6 – Exemplos de imagens de tumores cerebrais classificadas correta e incorretamente pelo MiniRocket.	33
Figura 7 – Acurácia e tempo de execução do (a) ROCKET e (b) MiniRocket em função do número de filtros e do número de características, respectivamente.	34

Lista de tabelas

Tabela 1 – Resultados dos testes de validação cruzada em 10 pastas com pacientes presentes nos conjuntos de treino e de teste simultaneamente.	31
Tabela 2 – Resultados dos testes de validação cruzada em 5 pastas com pacientes apenas no conjunto de treino ou de teste de forma mutuamente exclusiva.	31

Sumário

1	INTRODUÇÃO	17
1.1	Contexto e motivação	17
1.2	Objetivos	18
1.3	Organização	19
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	Aprendizado de máquina	21
2.2	Trabalhos anteriores	22
2.3	ROCKET e MiniRocket	23
3	METODOLOGIA	27
3.1	Obtenção dos dados	27
3.2	Adaptação dos algoritmos	27
3.3	Procedimentos experimentais	28
4	RESULTADOS, ANÁLISE E DISCUSSÃO	31
5	CONCLUSÃO	35
	REFERÊNCIAS	37

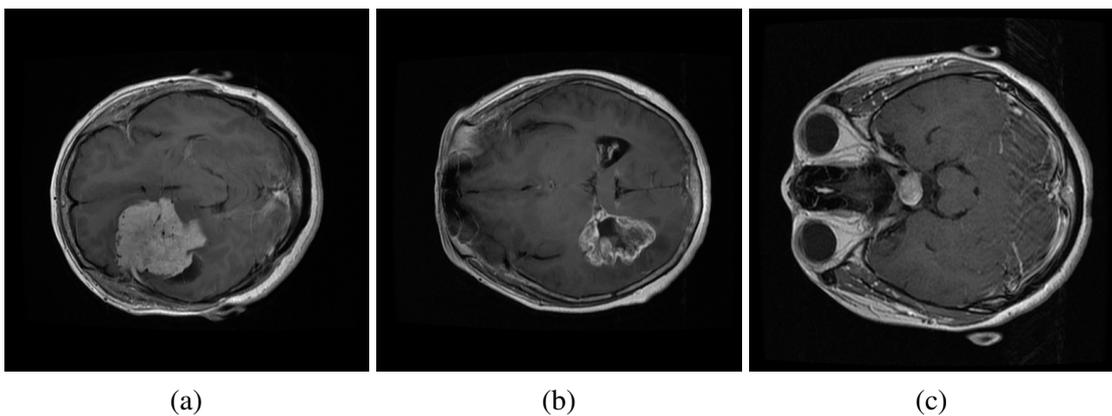
1 Introdução

1.1 Contexto e motivação

Um tumor cerebral é uma patologia associada à reprodução anormal e descontrolada de células no cérebro. Dependendo da região acometida, a doença pode causar fraqueza, disfasia, desorientação, mudanças de personalidade ou convulsões (LAPOINTE; PERRY; BUTOWSKI, 2018). Entre os diversos tipos de tumor cerebral estão os gliomas, originados nas células da glia, os meningiomas, originados nas meninges, e os tumores pituitários, originados na glândula pituitária (LOUIS et al., 2016; ASA et al., 2022).

A classificação correta de um tumor cerebral durante o diagnóstico é essencial para o seu tratamento adequado. No entanto, a biópsia desses tumores não é realizada antes de uma cirurgia definitiva (BYRNE; DWIVEDI; MINKS, 2022). Desse modo, soluções de diagnóstico auxiliado por computador para classificação de tumores cerebrais são uma ferramenta importante para a oncologia, destacando-se dentre elas o uso de imagens de ressonância magnética (MRI, do inglês *magnetic resonance imaging*) obtidas das cabeças dos pacientes, como as mostradas na Figura 1.

Figura 1 – Exemplos de imagens de ressonância magnética de pacientes acometidos por (a) meningioma, (b) glioma e (c) tumor pituitário.



Fonte: Cheng et al. (2015)

A interpretação dessas imagens, contudo, depende da experiência do profissional responsável, o que pode levar a erros caso as imagens não apresentem elementos discriminatórios o suficiente para uma decisão acertada (AFSHAR; PLATANIOTIS; MOHAMMADI, 2019). Assim, técnicas para o processamento automatizado dessas imagens fazem-se necessárias para a classificação rápida e precisa dos tumores cerebrais.

O Aprendizado de Máquina (AM), em particular o aprendizado profundo, é uma das

principais formas de automatizar essa classificação. Métodos de AM utilizam um conjunto de imagens de MRI de tumores já classificados para identificar padrões e classificar novas imagens. Nesse sentido, as redes neurais convolucionais (CNN, do inglês *convolutional neural networks*) são os modelos mais utilizados para essa tarefa, visto que apresentam grande poder de generalização e não requerem intervenção humana durante o aprendizado, como a delimitação da região do tumor ou a escolha manual de hiper-parâmetros (BADŽA; BARJAKTAROVIĆ, 2020).

Entretanto, o uso de aprendizado profundo e CNNs também traz consigo desafios. As redes neurais profundas exigem grandes conjuntos de dados para que convirjam (KHAN et al., 2022) e *hardware* poderoso para que possam ser treinadas em tempo hábil. Ainda que abordagens como *data augmentation* e aprendizado por transferência amenizem esses problemas, elas podem levar a outras dificuldades, como a possibilidade de aumento do sobre-ajuste aos dados de treino e a necessidade de um modelo pré-treinado, respectivamente.

A partir de uma conjuntura semelhante em outro domínio – a classificação de séries temporais – foi desenvolvido o algoritmo ROCKET (do inglês *Random Convolutional Kernel Transform*), que alcançou excelentes resultados utilizando apenas uma fração do tempo de execução de algoritmos de aprendizado profundo (DEMPSTER; PETITJEAN; WEBB, 2020). Outras versões do algoritmo foram propostas em seguida, como o MiniRocket (DEMPSTER; SCHMIDT; WEBB, 2021) e o MultiRocket (TAN et al., 2022).

Tanto o ROCKET quanto os seus sucessores utilizam, como principal mecanismo, as convoluções. Assim, eles podem ser facilmente convertidos para o domínio da classificação de imagens, bastando apenas utilizar filtros bidimensionais em vez de unidimensionais. Nessa perspectiva, uma abordagem de AM clássico para a classificação de tumores cerebrais que atinja resultados tão bons quanto o estado da arte do aprendizado profundo mostra-se como uma alternativa interessante para diversos casos de uso.

1.2 Objetivos

O objetivo geral deste trabalho é desenvolver e testar adaptações dos algoritmos ROCKET e MiniRocket para o domínio da classificação de imagens de ressonância magnética de tumores cerebrais, comparando-as, em acurácia e eficiência, a abordagens de aprendizado profundo do estado da arte. O desenvolvimento dessa proposta dar-se-á através dos seguintes objetivos específicos.

- Revisão dos modelos propostos na literatura;
- Adaptação dos mecanismos do ROCKET e MiniRocket para entradas bidimensionais (imagens);
- Implementação dos algoritmos através da biblioteca PyTorch (ANSEL et al., 2024);

- Realização de testes comparativos entre os algoritmos adaptados e uma CNN do estado da arte;
- Disponibilização de todo o código desenvolvido em licença aberta.

1.3 Organização

O restante deste trabalho está organizado da seguinte forma. O Capítulo 2 aborda as bases teóricas do Aprendizado de Máquina e dos algoritmos ROCKET e MiniRocket, além das propostas da literatura para a classificação de tumores cerebrais. O Capítulo 3 descreve os métodos elaborados para a adaptação e a validação dos dois algoritmos propostos. No Capítulo 4, são mostrados e discutidos os resultados obtidos a partir dos experimentos realizados e o desempenho dos modelos estudados. Por fim, o Capítulo 5 sumariza o trabalho desenvolvido, traça conclusões a partir dos resultados alcançados e propõe caminhos para futuras abordagens na área.

2 Fundamentação teórica

2.1 Aprendizado de máquina

O Aprendizado de Máquina é um campo da Inteligência Artificial que estuda algoritmos que utilizam dados coletados para aprender a realizar determinadas tarefas sem programação explícita (MITCHELL, 1997). No contexto de AM, a classificação de tumores cerebrais encaixa-se no paradigma de aprendizado supervisionado, no qual um modelo que relaciona entradas e saídas é construído a partir de dados de entrada cujas saídas já são conhecidas.

Uma das mais poderosas abordagens de AM é o aprendizado profundo e as redes neurais artificiais, modelos que emulam o sistema nervoso humano e consistem em unidades de processamento – denominadas neurônios – organizadas em várias camadas. A saída de um neurônio é usada como entrada por outros através da aplicação de pesos e vieses, aprendidos a partir dos dados no algoritmo de *backpropagation*.

Quando os dados de entrada de uma tarefa de AM são imagens, as redes neurais convolucionais são os modelos mais comumente usados. Para capturar padrões e características relevantes de uma imagem, as CNNs empregam várias camadas de convolução, operação matemática que, para um sinal de entrada x e outro sinal h , denominado filtro, produz um terceiro sinal y , dado por

$$y[i, j] = (x * h)[i, j] = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} h[i, j] \cdot x[i - k_1][j - k_2]. \quad (2.1)$$

O resultado de cada convolução passa por uma camada de amostragem responsável por reduzir o tamanho do sinal recebido. As camadas de convolução e amostragem alternam-se em sequência, seguidas, por fim, por uma ou mais camadas de neurônios totalmente conectados, as quais são responsáveis por rotular a imagem original.

Apesar de seu grande sucesso em diversas aplicações, o uso de CNNs pode trazer desvantagens. O treinamento de um modelo pode levar dias e exigir enormes quantidades de dados rotulados. Além disso, quanto maior é a quantidade de neurônios na rede neural – que, no geral, deve aumentar para que se modele problemas mais complexos –, maior é o risco de sobre-ajuste ao conjunto de dados de treino (SHRESTHA; MAHMOOD, 2019). Portanto, é importante verificar se, para uma dada tarefa, algoritmos mais simples, que não lancem mão de aprendizado profundo, são suficientes para alcançar o resultado desejado.

2.2 Trabalhos anteriores

Ao longo dos últimos anos, diversas abordagens e arquiteturas de CNN foram propostas para a classificação de tumores cerebrais e testadas em conjuntos de MRI contendo diferentes tipos de tumores, como o conjunto alcunhado pela literatura como “*Figshare*” (CHENG et al., 2015), composto por imagens de meningiomas, gliomas e tumores pituitários. Alguns trabalhos também aplicaram modelos de aprendizado profundo ao “*Whole Brain Atlas*” (JOHNSON; BECKER, 1999), um conjunto de dados de classificação binária, isto é, cada imagem no conjunto é classificada pela presença ou ausência de um tumor cerebral, tarefa também chamada de detecção de tumores cerebrais. Bases de dados semelhantes estão disponíveis publicamente no Kaggle^{1,2}.

Abiwinanda et al. (2019) propuseram uma arquitetura simples de CNN para atacar o problema: três camadas convolucionais, com filtros de tamanho 5 por 5 e amostragens de tamanho 2 por 2, seguidas de uma camada totalmente conectada. O modelo inclui o uso de *dropout* para reduzir o impacto de sobre-ajuste. Os autores também utilizaram três etapas de pré-processamento: redimensionamento das imagens, aplicação de filtro gaussiano e equalização por histograma, aumentando o contraste em cada imagem. A abordagem proposta alcançou resultados bons, ainda que não significativamente superiores ao estado da arte até então, como acurácia acima de 90% e área abaixo da curva (AUC, do inglês *area under curve*) muito próxima a 1. Uma arquitetura ainda mais compacta foi desenvolvida por Badža e Barjaktarović (2020), com métricas de desempenho tão boas quanto as anteriores.

Deepak e Ameer (2019), identificando limitações nos métodos de classificação de tumores cerebrais apresentados até então, como desempenho insatisfatório, exigência de intervenção manual e escassez de dados para treinamento, propuseram o uso de aprendizado por transferência. Nesse sentido, a GoogLeNet, uma rede neural pré-treinada em um grande conjunto de imagens, foi adaptada e reajustada para o conjunto *Figshare*. A saída das camadas convolucionais da rede foi usada como atributos de entrada para classificadores clássicos, como o *k-Nearest Neighbors* (kNN) e a *Support Vector Machine* (SVM). O aprendizado por transferência também foi aplicado por Khan et al. (2022), com uso da rede VGG16, e por Asiri et al. (2023), adaptando os modelos ResNet50 e U-Net para detecção e segmentação semântica dos tumores, respectivamente. Todas essas abordagens exibiram acurácia acima de 90%.

Outra solução para o problema da baixa quantidade de dados disponíveis é o uso de *data augmentation*. Essa técnica aplica transformações morfológicas nas imagens, como rotações, translações e aproximações, para gerar novos dados a partir do mesmo fenômeno representado. A aplicação de *data augmentation* na classificação de tumores cerebrais é encontrada em diversos trabalhos da literatura (ASIF et al., 2022).

¹ <<https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection>>

² <<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>>

Por fim, Ravinder et al. (2023) apresentaram um novo método para classificação de tumores cerebrais que utiliza redes neurais em grafos (GNN, do inglês *graph neural networks*). Uma imagem pode ser considerada um grafo em que os nós representam os *pixels* e as arestas representam as adjacências entre os *pixels*. A representação em grafos permite que se induza um modelo de AM a partir das relações entre diferentes sub-regiões de uma imagem. Com a transformação das imagens em matrizes de adjacência de grafos, os autores utilizaram-nas como entrada para uma GNN cuja saída era outra matriz de adjacência. Esta última é usada, então, como entrada de uma CNN convencional, responsável pela classificação final das imagens originais. Dentre as diferentes arquiteturas propostas, a melhor avaliada alcançou mais de 95% de acurácia.

Mesmo com métricas excelentes obtidas pelos modelos apresentados, a maior parte dos trabalhos citados não realizou validação cruzada para avaliação dos algoritmos. São ainda menos frequentes as publicações que, ao separar conjuntos de treino e teste, evitaram que um mesmo paciente tivesse suas imagens em ambos, o que pode fazer com que o modelo sob teste aprenda os padrões morfológicos do paciente em vez do tumor, ocorrendo, assim, sobre-ajuste ao conjunto de dados utilizado para avaliação. Dessa maneira, os valores das métricas alcançadas pelas propostas da literatura podem estar inflados e irreais, demandando reprodução de testes para comparações justas.

2.3 ROCKET e MiniRocket

Tendo em vista a alta complexidade computacional dos algoritmos do estado da arte de classificação de séries temporais, Dempster, Petitjean e Webb (2020) desenvolveram o ROCKET, um algoritmo de classificação que utiliza filtros convolucionais aleatórios e funções de amostragem para transformar as séries de entrada em tabelas atributo-valor. O modelo alcançou acurácias tão altas quanto o então estado da arte em tempos de execução muito menores.

Cada filtro convolucional do ROCKET é gerado a partir da escolha aleatória de diversos parâmetros, estruturados da seguinte forma:

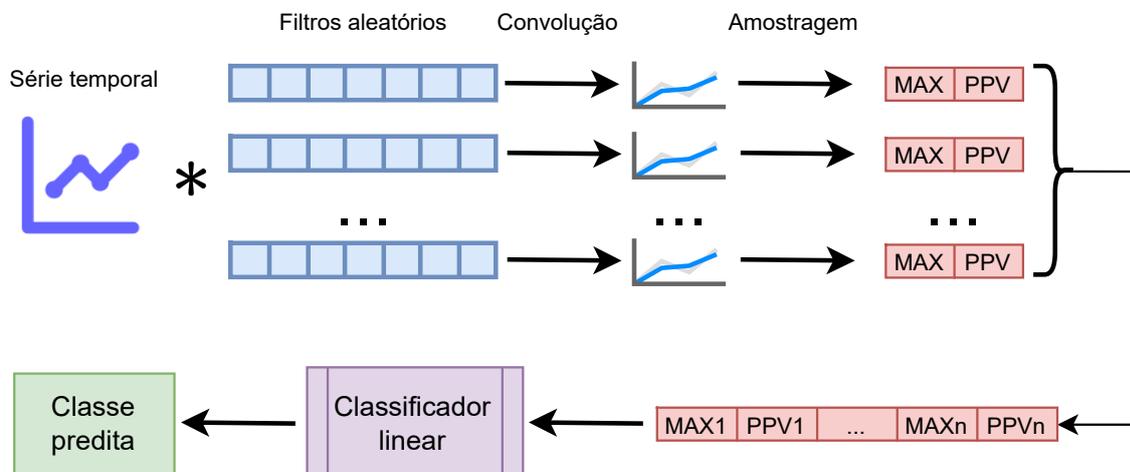
- O tamanho ou comprimento l corresponde ao número de pesos que compõem o filtro. É escolhido dentre o conjunto $\{7, 9, 11\}$, com probabilidades iguais;
- Os pesos $w \in W$ são os coeficientes utilizados na operação de convolução, como descrito na Equação 2.1. São amostrados da distribuição normal padrão, $w' \sim \mathcal{N}(0, 1)$, e, então, centralizados em torno da média das amostras: $w = w' - \bar{W}$;
- O viés b é subtraído de todos os valores do sinal resultante da convolução. É amostrado a partir da distribuição uniforme $b \sim \mathcal{U}(-1, 1)$;

- A dilatação d controla o espaçamento entre as amostras obtidas do sinal de entrada. É amostrada em escala exponencial, fazendo $d = \lfloor 2^x \rfloor$, $x \sim \mathcal{U}(0, A)$, sendo $A = \log_2 \frac{l_{\text{entrada}}-1}{l_{\text{filtro}}-1}$;
- O preenchimento estende o sinal de entrada com zeros para que o filtro possa ser aplicado em suas bordas. A sua presença ou ausência é decidida aleatoriamente, com iguais probabilidades.

Ao contrário das camadas convolucionais encontradas em CNNs, os pesos dos filtros do ROCKET não precisam ser aprendidos, o que reduz a complexidade computacional do algoritmo. Além disso, a variedade estrutural dos filtros, que possuem não apenas pesos diferentes, mas também tamanhos e dilatações diversos, permite que várias características e padrões dos dados de entrada sejam captados pelas convoluções.

Do resultado da aplicação de cada filtro convolucional a uma série de entrada são extraídos dois atributos: o valor máximo e a proporção de valores positivos (PPV, do inglês *proportion of positive values*). Assim, com k filtros – por padrão, 10000 –, o ROCKET produz $2k$ características para cada instância, as quais são utilizadas como dados de entrada de um classificador linear. Esse procedimento é melhor ilustrado na Figura 2.

Figura 2 – Ilustração do mecanismo de classificação do ROCKET.



Fonte: Autor

De modo a tornar o algoritmo mais determinístico e mais rápido em grandes conjuntos de dados, Dempster, Schmidt e Webb (2021) publicaram uma versão reformulada do ROCKET, denominada MiniRocket. Esse novo modelo obteve acurácias muito próxima às alcançadas pelo ROCKET, mas a um menor custo computacional e sob menor efeito da aleatoriedade.

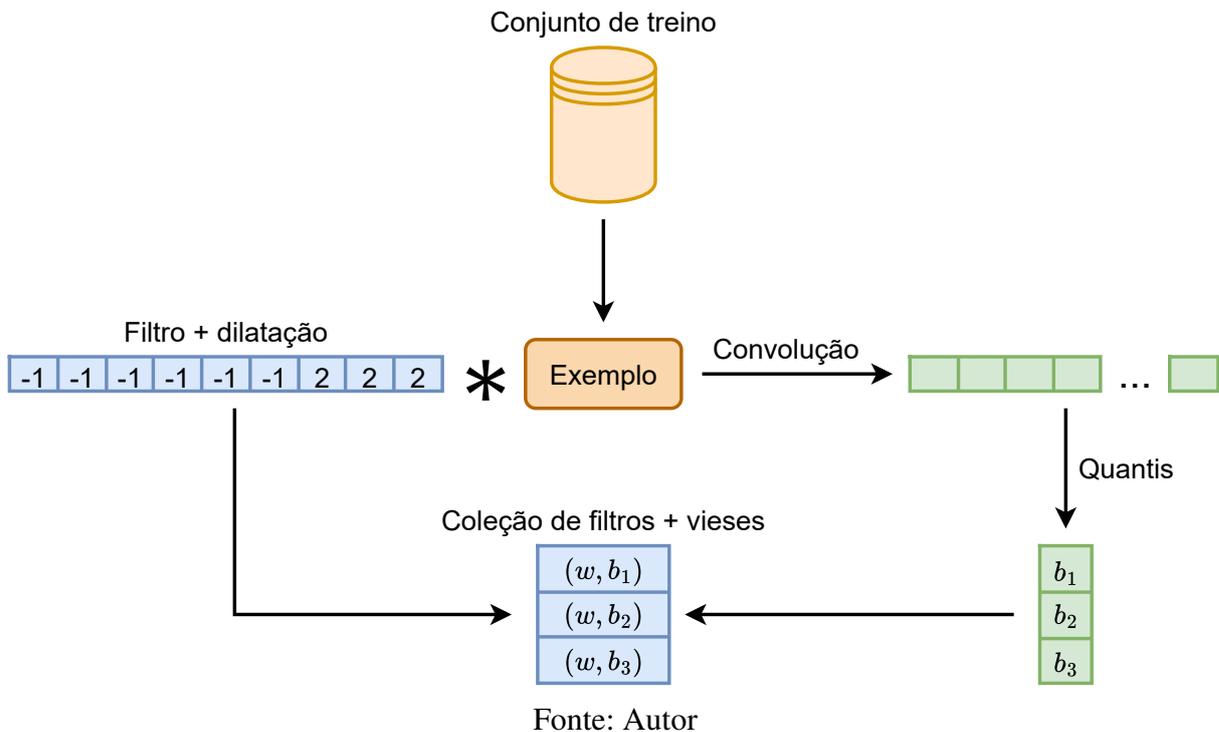
O MiniRocket, diferente de seu antecessor, utiliza apenas filtros convolucionais de tamanho 9, cujos pesos são restritos a apenas dois valores, canonicamente -1 e 2, cujo número de ocorrências também é fixo, canonicamente 6 e 3, de forma que a soma dos pesos seja sempre zero. Assim, o filtro será sensível apenas à magnitude relativa dos valores de entrada, tornando

a convolução invariante à adição ou subtração de valores constantes à entrada. Além disso, as restrições impostas mantêm o número total de filtros gerados pequeno: apenas 84.

As dilatações são escolhidas a partir do conjunto $\{[2^0], \dots, [2^a]\}$, sendo que $a = \log_2 \frac{l_{\text{entrada}}-1}{l_{\text{filtro}}-1}$. Todas as combinações de filtro e dilatação são utilizadas, mas, como o número de dilatações aumenta com o tamanho da série de entrada, este é limitado a no máximo 32. A presença ou ausência de preenchimento com zeros é alternada para cada combinação de filtro e dilatação, de forma que, em geral, metade delas utilizem preenchimento.

Para a escolha dos vieses de cada filtro, uma instância do conjunto de treino é escolhida aleatoriamente e o filtro é aplicado a ela. Da convolução resultante, calculam-se os valores de um conjunto de quantis, que serão utilizados como vieses para o cálculo das características. Esse processo é ilustrado na Figura 3.

Figura 3 – Ilustração do mecanismo de geração de filtros convolucionais do MiniRocket.



Por fim, o MiniRocket descarta o uso de valor máximo como característica extraída, lançando mão apenas da PPV. Para um dado número de características desejadas – por padrão, 10000 –, um número adequado de vieses é gerado para cada combinação de filtro e dilatação. Como no ROCKET, as características geradas são usadas como entrada para um classificador linear.

3 Metodologia

3.1 Obtenção dos dados

Para testar e avaliar o desempenho dos algoritmos desenvolvidos, foram utilizadas, como dados de *benchmarking*, as imagens de ressonância magnética do conjunto *Figshare* (CHENG et al., 2015), no qual encontram-se imagens de meningiomas, gliomas e tumores pituitários. Esse conjunto, utilizado amplamente na literatura, permitiu a reprodutibilidade completa dos experimentos e a comparação dos resultados obtidos com trabalhos anteriores.

O conjunto também oferece índices de validação cruzada de forma que imagens obtidas de um mesmo paciente não estejam no subconjunto de treino e no de teste simultaneamente. Dessa forma, evita-se o sobre-ajuste dos modelos aos pacientes e as métricas de desempenho obtidas tornam-se mais realistas. Para facilitar o carregamento dos dados, as imagens do conjunto, originalmente em arquivos Matlab, foram armazenadas como vetores de inteiros de 16 bits em arquivos NumPy, e estes comprimidos com a ferramenta Gzip.

3.2 Adaptação dos algoritmos

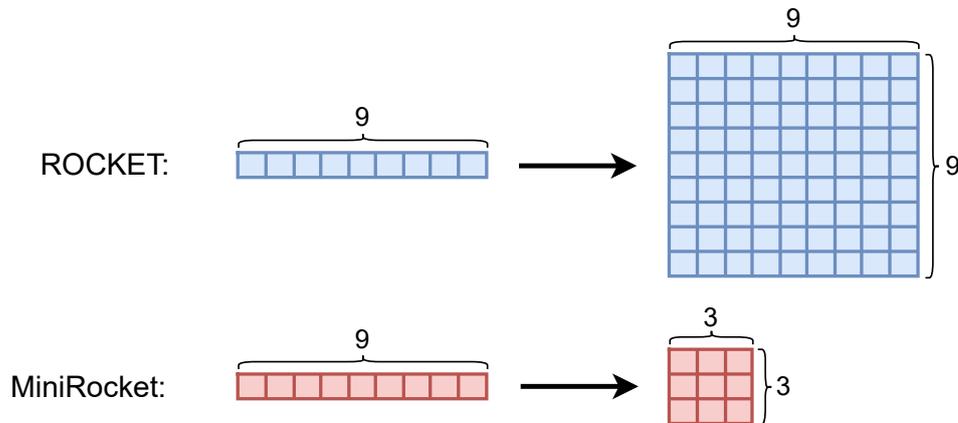
A partir dos mecanismos descritos na seção 2.3, os algoritmos ROCKET e MiniRocket foram alterados para classificar entradas bidimensionais. Para tanto, os filtros convolucionais passaram a ter duas dimensões, e seus parâmetros aleatórios replicados. No caso do ROCKET adaptado, os filtros gerados são quadrados de lado 7, 9 ou 11. Os pesos, o viés, a dilatação e o preenchimento de cada filtro são amostrados de forma idêntica ao algoritmo original, sendo os dois últimos replicados entre as duas dimensões.

Já a adaptação bidimensional do MiniRocket apresenta apenas filtros quadrados de lado 3, de forma a manter o número total de configurações de pesos em 84, visto que a conversão dos filtros originais em quadrados de lado 9 aumentaria exponencialmente o número de configurações possíveis, tornando inviável o uso determinístico do modelo. O restante dos parâmetros de cada filtro é escolhida de forma idêntica ao algoritmo original, incluindo a geração de vieses a partir dos quantis do resultado da convolução do filtro por um exemplo do conjunto de treino.

A sutil diferença entre as adaptações bidimensionais dos filtros do ROCKET e as do MiniRocket é ilustrada na Figura 4. O cálculo da convolução das imagens pelos filtros é o mesmo que o definido pela Equação 2.1.

Os algoritmos adaptados foram implementados através da biblioteca PyTorch. A ferramenta foi escolhida devido à possibilidade e facilidade de uso de unidades de processamento gráfico (GPUs, do inglês *graphical processing unit*) para o cálculo das convoluções, o que reduz

Figura 4 – Adaptações dos filtros convolucionais do ROCKET e do MiniRocket para entradas bidimensionais.



Fonte: Autor

drasticamente o tempo de execução dos algoritmos. Além disso, todas as imagens de entrada foram redimensionadas para 256 por 256 *pixels*, padronizando o tamanho das instâncias do conjunto de dados.

3.3 Procedimentos experimentais

Visando a comparação justa entre modelos e a reprodução de resultados anteriores, a rede neural convolucional proposta por Badža e Barjaktarović (2020) foi reimplementada, também em PyTorch, e testada nas mesmas condições dos novos algoritmos. Esse modelo foi escolhido como representante experimental do aprendizado profundo devido a sua alta acurácia, compacidade e reprodutibilidade.

Para comparar o desempenho e a complexidade dos algoritmos desenvolvidos àqueles de modelos de aprendizado profundo, foram realizados dois experimentos de validação cruzada no conjunto *Figshare*. A primeira configuração experimental foi a classificação das imagens através de validação cruzada em 10 pastas, isto é, o conjunto foi dividido em dez partições, e cada partição foi classificada por um modelo treinado a partir das nove restantes. Nesse primeiro procedimento, um mesmo paciente poderia ter suas imagens presentes nos conjuntos de treino e teste de forma simultânea.

Em seguida, os algoritmos foram testados novamente a partir dos índices de validação cruzada em 5 pastas disponibilizados pelo conjunto *Figshare*. Nesse caso, nenhum paciente esteve presente simultaneamente nos conjuntos de treino e teste, de modo a tornar as métricas de desempenho mais realistas, como discutido na seção 2.2.

Os algoritmos ROCKET e MiniRocket foram testados com diversos números de filtros convolucionais e características geradas, incluindo o valor proposto pelos autores, 10000. Cada

modelo, incluindo a CNN reproduzida, foi testado cinco vezes em cada um dos procedimentos e a média aritmética de suas métricas de desempenho foram consideradas, de forma a levar em conta a sua aleatoriedade. Além disso, a implementação dos algoritmos e testes, juntamente com todos os resultados, foi disponibilizada no GitHub¹.

¹ <<https://github.com/GuiArcenio/classificacao-tumor-cerebral>>

4 Resultados, análise e discussão

Após a realização dos testes de validação cruzada, foram calculados os valores de acurácia, precisão, revocação (ou *recall*) e medida F1 obtidos pela CNN, pelo ROCKET de 10000 filtros e pelo MiniRocket de 10000 características. Esses resultados, acompanhados pelo de tempo de execução do teste de cada modelo, são mostrados na Tabela 1, na validação cruzada com pastas separadas a nível de instância, e na Tabela 2, na validação cruzada com pastas separadas a nível de paciente. Devido à característica multiclasse do problema, a precisão, a revocação e a F1 apresentadas são suas macro-médias entre classes.

Tabela 1 – Resultados dos testes de validação cruzada em 10 pastas com pacientes presentes nos conjuntos de treino e de teste simultaneamente.

Modelo	Acurácia (%)	Macro precisão (%)	Macro revocação (%)	Macro F1 (%)	Tempo de execução (s)
CNN	86,39 ± 1,89	86,57 ± 1,14	83,55 ± 3,23	84,39 ± 2,61	408,22 ± 43,49
ROCKET	96,44 ± 0,13	96,22 ± 0,12	95,93 ± 0,21	96,07 ± 0,16	406,31 ± 38,05
MiniRocket	93,77 ± 0,08	92,88 ± 0,09	94,38 ± 0,04	93,50 ± 0,07	140,45 ± 0,02

Tabela 2 – Resultados dos testes de validação cruzada em 5 pastas com pacientes apenas no conjunto de treino ou de teste de forma mutuamente exclusiva.

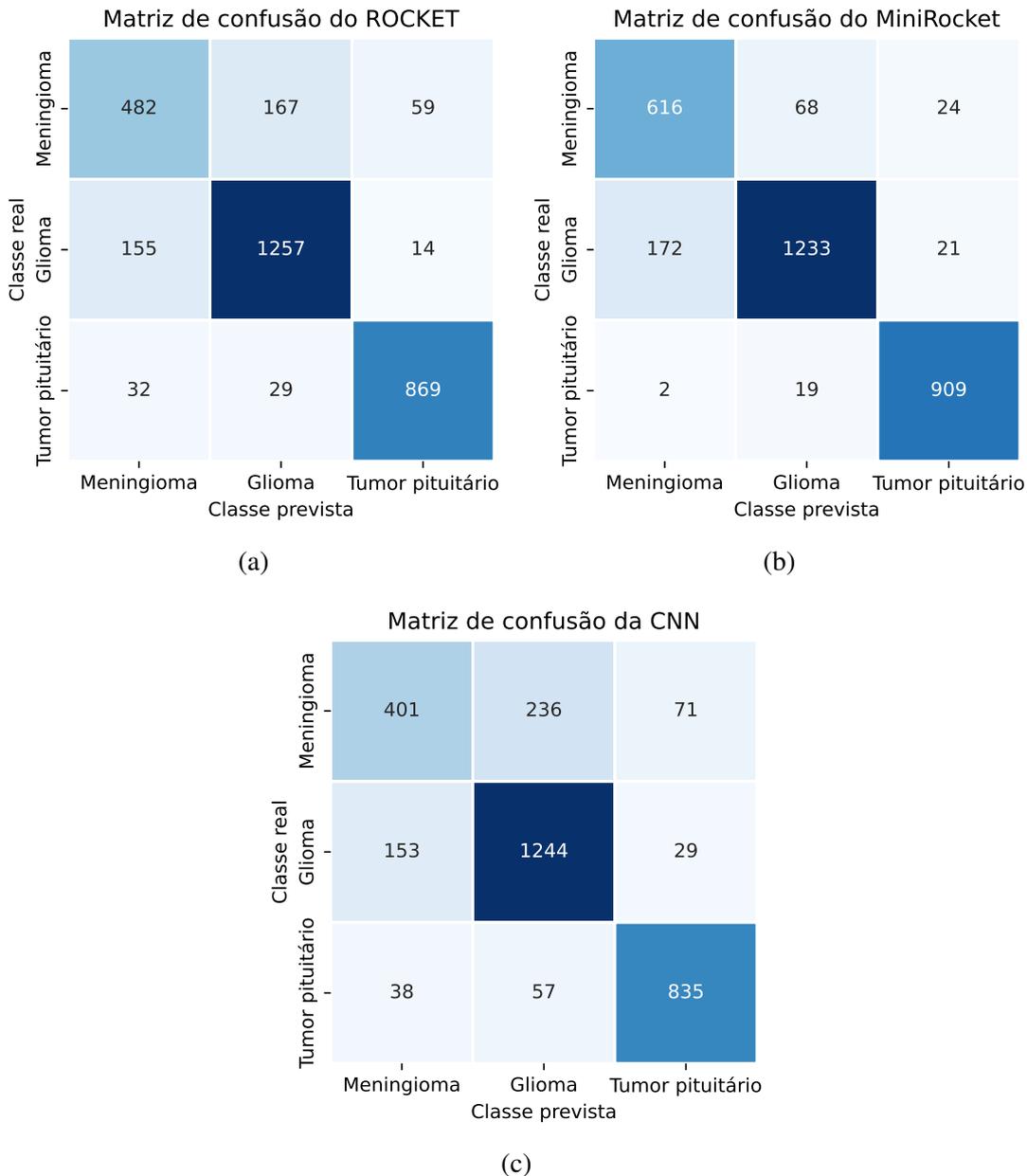
Modelo	Acurácia (%)	Macro precisão (%)	Macro revocação (%)	Macro F1 (%)	Tempo de execução (s)
CNN	81,31 ± 0,95	79,86 ± 0,89	78,77 ± 1,78	79,07 ± 1,34	387,08 ± 33,25
ROCKET	85,03 ± 0,36	83,44 ± 0,42	82,96 ± 0,54	83,14 ± 0,51	355,16 ± 0,69
MiniRocket	89,93 ± 0,24	88,80 ± 0,25	90,29 ± 0,23	89,42 ± 0,24	144,31 ± 0,03

Em primeiro lugar, é possível observar a redução geral das métricas quando os conjuntos de treino e de teste não compartilham pacientes, indicando a presença de sobre-ajuste dos modelos quando este não é o caso e a possibilidade de métricas infladas nos trabalhos onde esse problema não é tratado. Não obstante, em ambos os experimentos, o ROCKET e o MiniRocket apresentaram métricas maiores que aquelas da rede convolucional, sendo que o MiniRocket levou muito menos tempo para alcançar resultados muito semelhantes, e até superiores, aos dos outros modelos.

Dado que os valores de acurácia ficaram próximos aos de precisão, revocação e F1, não há indicativo de desempenho desigual entre as três classes. Para melhor avaliar essa questão, as matrizes de confusão dos três modelos são mostradas na Figura 5.

De acordo com as matrizes, houve pouca diferença de desempenho dos modelos entre as três classes. No entanto, há um número significativo de casos onde houve confusão entre meningiomas e gliomas, o que pode ser explicado pelas similaridades morfológicas e de localização entre os dois tumores, as quais não são compartilhadas pelo tumor pituitário, que é restrito a uma glândula específica. Para melhor ilustrar esse fenômeno, exemplos de imagens classificadas pelo MiniRocket são mostrados na Figura 6.

Figura 5 – Matrizes de confusão dos algoritmos (a) ROCKET, (b) MiniRocket e (c) CNN, amostradas do primeiro teste de validação cruzada de 5 pastas separadas por paciente.

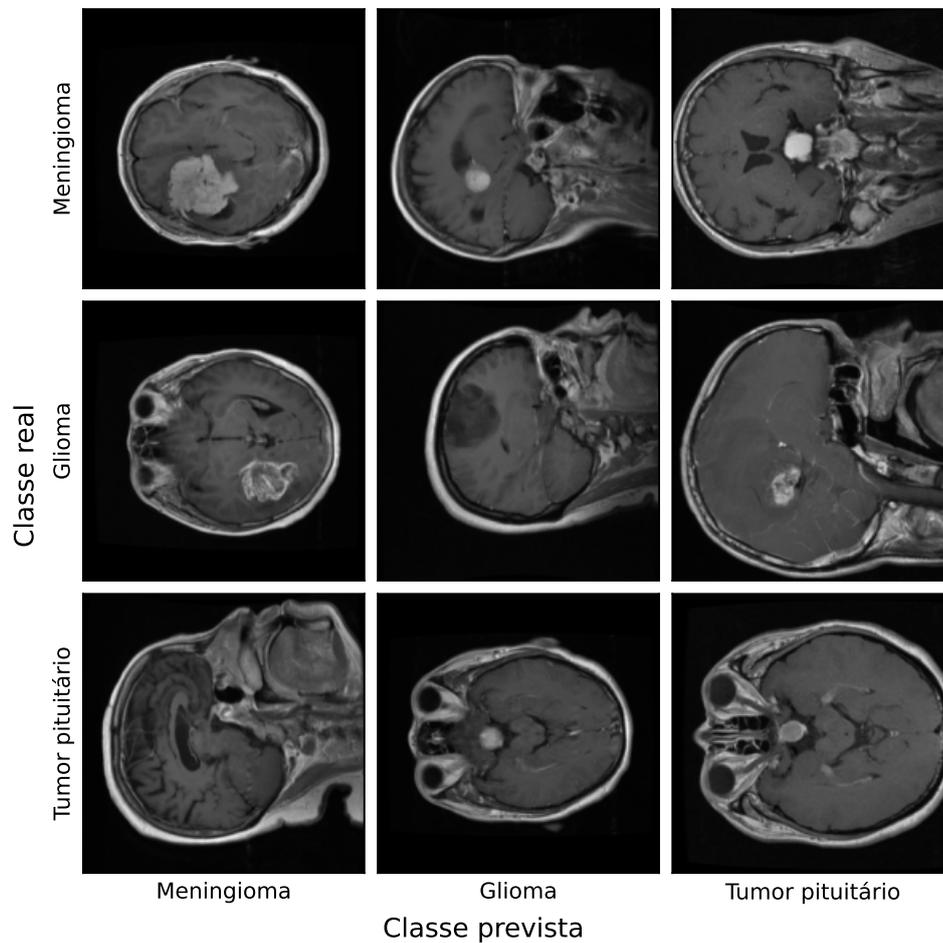


Fonte: Autor

Os resultados dos testes do ROCKET e MiniRocket com diferentes números de filtros e características são mostrados na Figura 7. Ainda que o ROCKET com quantidade reduzida de filtros apresente acurácia consideravelmente pior que a de sua versão completa, seu tempo de execução também é muito menor. No caso do MiniRocket, a perda de alguns pontos percentuais de acurácia é compensada por uma redução em mais de 7 vezes do tempo de execução.

Em geral, ambos os algoritmos propostos mostraram excelente desempenho na classificação de tumores cerebrais, com métricas iguais ou superiores às obtidas por modelos de aprendizado profundo, mesmo considerando os valores possivelmente inflados na literatura,

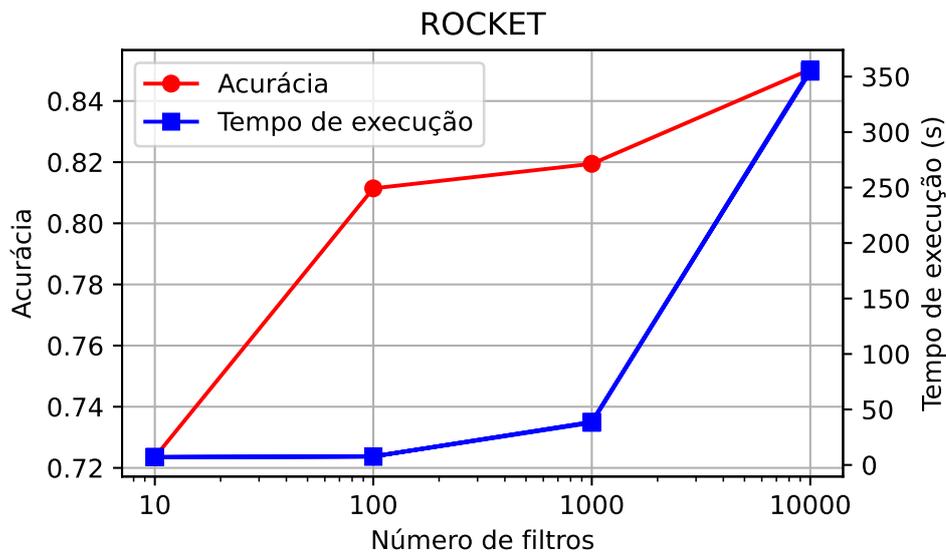
Figura 6 – Exemplos de imagens de tumores cerebrais classificadas correta e incorretamente pelo MiniRocket.



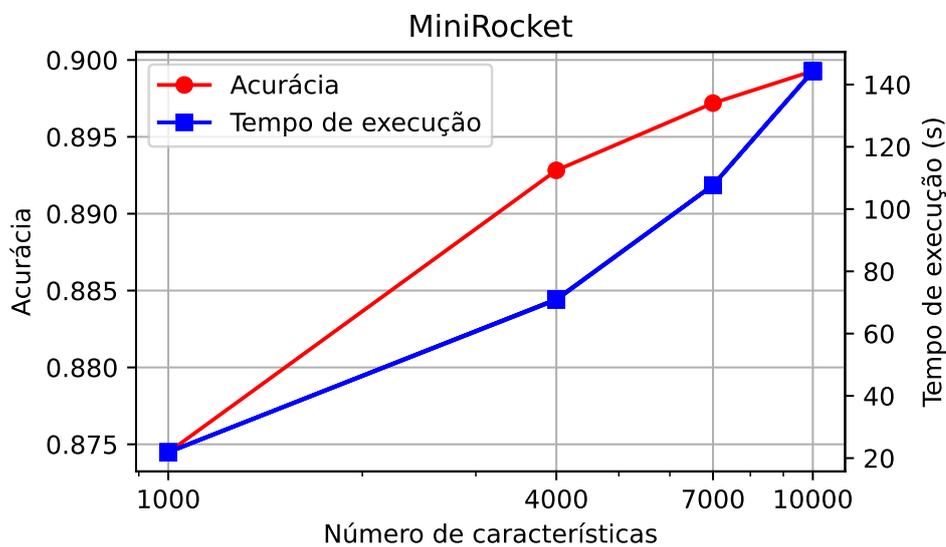
Fonte: Cheng et al. (2015)

oriundos de testes pouco cuidadosos. Além disso, o MiniRocket exige muito menos tempo computacional que as redes neurais para alcançar esses resultados, permitindo ainda grande redução de sua complexidade computacional ao custo de uma pequena perda de seu poder de classificação.

Figura 7 – Acurácia e tempo de execução do (a) ROCKET e (b) MiniRocket em função do número de filtros e do número de características, respectivamente.



(a)



(b)

Fonte: Autor

5 Conclusão

Neste trabalho, foram propostos dois algoritmos de Aprendizado de Máquina clássico, ROCKET e MiniRocket, adaptados do domínio de séries temporais, para a classificação de tumores cerebrais em imagens de ressonância magnética. Os modelos publicados de melhor desempenho nessa tarefa – em geral, redes neurais convolucionais – são custosos e seus resultados apresentados podem ser irrealistas em razão de testes demasiado rudimentares.

A partir da avaliação com validação cruzada e criação de conjuntos de treino e teste com pacientes diferentes, foi mostrado que os modelos propostos demonstram desempenho muito próximo ao de algoritmos mais complexos que envolvem aprendizado profundo. Além disso, o MiniRocket é também mais eficiente que outros modelos do estado da arte, levando muito menos tempo que uma CNN simples para atingir valores de acurácia equivalentes ou superiores. A implementação de todos os algoritmos estudados e desenvolvidos foi disponibilizada em licença aberta.

Entre as sugestões para trabalhos futuros estão a reprodução de testes dos modelos propostos na literatura, levando em conta a sensibilidade do conjunto de dados a sobre-ajuste aos pacientes, o estudo dos efeitos de técnicas de pré-processamento das imagens, como realce de contraste, no desempenho do ROCKET e do MiniRocket e a otimização e aperfeiçoamento da implementação em GPU dos dois algoritmos. Ademais, o bom desempenho dos modelos propostos em um domínio específico pode sugerir sua adequação a tarefas mais gerais de classificação de imagens.

Referências

- ABIWINANDA, N. et al. Brain tumor classification using convolutional neural network. In: *World Congress on Medical Physics and Biomedical Engineering 2018*. Singapore: Springer Nature Singapore, 2019. p. 183–189. Citado na página 22.
- AFSHAR, P.; PLATANIOTIS, K. N.; MOHAMMADI, A. Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In: *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. U.K.: IEEE, 2019. p. 1368–1372. Citado na página 17.
- ANSEL, J. et al. PyTorch 2: Faster machine learning through dynamic python bytecode transformation and graph compilation. In: *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*. New York, NY, USA: Association for Computing Machinery, 2024. p. 929–947. Citado na página 18.
- ASA, S. L. et al. Overview of the 2022 WHO classification of pituitary tumors. *Endocrine Pathology*, Springer, v. 33, p. 6–26, 2022. Citado na página 17.
- ASIF, S. et al. Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from MR images. *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., v. 10, p. 34716–34730, 2022. Citado na página 22.
- ASIRI, A. A. et al. Brain Tumor Detection and Classification Using Fine-Tuned CNN with ResNet50 and U-Net Model: A Study on TCGA-LGG and TCIA Dataset for MRI Applications. *Life*, Multidisciplinary Digital Publishing Institute (MDPI), v. 13, p. 1449, 2023. Citado na página 22.
- BADŽA, M. M.; BARJAKTAROVIĆ, M. Č. Classification of brain tumors from MRI images using a convolutional neural network. *Applied Sciences*, MDPI AG, v. 10, p. 1999, 2020. Citado 3 vezes nas páginas 18, 22 e 28.
- BYRNE, J.; DWIVEDI, R.; MINKS, D. Tumours of the brain. In: *Recommendations for cross-sectional imaging in cancer management*. The Royal College of Radiologists, 2022. Disponível em: <<https://www.rcr.ac.uk/our-services/all-our-publications/clinical-radiology-publications/recommendations-for-cross-sectional-imaging-in-cancer-management-third-edition>>. Acesso em: 24 abr. 2024. Citado na página 17.
- CHENG, J. et al. Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLOS ONE*, Public Library of Science, v. 10, 2015. Citado 4 vezes nas páginas 17, 22, 27 e 33.
- DEEPAK, S.; AMEER, P. Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine*, Elsevier Ltd, v. 111, p. 103–345, 2019. Citado na página 22.
- DEMPSTER, A.; PETITJEAN, F.; WEBB, G. I. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery*, Springer, v. 34, p. 1454–1495, 2020. Citado 2 vezes nas páginas 18 e 23.

DEMPSTER, A.; SCHMIDT, D. F.; WEBB, G. I. MiniRocket: A very fast (almost) deterministic transform for time series classification. In: . [S.l.]: Association for Computing Machinery, 2021. p. 248–257. Citado 2 vezes nas páginas 18 e 24.

JOHNSON, K. A.; BECKER, J. A. *Whole Brain Atlas*. 1999. Disponível em: <<https://www.med.harvard.edu/AANLIB>>. Acesso em: 29 nov. 2023. Citado na página 22.

KHAN, M. S. I. et al. Accurate brain tumor detection using deep convolutional neural network. *Computational and Structural Biotechnology Journal*, Elsevier B.V., v. 20, p. 4733–4745, 2022. Citado 2 vezes nas páginas 18 e 22.

LAPOINTE, S.; PERRY, A.; BUTOWSKI, N. A. Primary brain tumours in adults. *The Lancet*, Lancet Publishing Group, v. 392, p. 432–446, 2018. Citado na página 17.

LOUIS, D. N. et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica*, Springer Verlag, v. 131, p. 803–820, 2016. Citado na página 17.

MITCHELL, T. *Machine Learning*. New York, NY: McGraw-Hill Professional, 1997. (McGraw-Hill series in computer science). Citado na página 21.

RAVINDER, M. et al. Enhanced brain tumor classification using graph convolutional neural network architecture. *Scientific Reports*, Nature Research, v. 13, p. 14938, 2023. Citado na página 23.

SHRESTHA, A.; MAHMOOD, A. Review of deep learning algorithms and architectures. *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., v. 7, p. 53040–53065, 2019. Citado na página 21.

TAN, C. W. et al. MultiRocket: multiple pooling operators and transformations for fast and effective time series classification. *Data Mining and Knowledge Discovery*, Springer, v. 36, p. 1623–1646, 2022. Citado na página 18.