

BrainScanNet: Diagnósticos Inteligentes Para Tumores Cerebrais Usando IA

1st Gabriel Carvalho
CCSA - UFPE
Recife, Brasil
gdc2@cin.ufpe.br

2nd Leandro Freitas
CIn - UFPE
Recife, Brasil
lllf@cin.ufpe.br

3rd Lucas Sales
CIn - UFPE
Recife, Brasil
lfasm@cin.ufpe.br

Abstract—Este relatório apresenta a implementação e avaliação de um modelo de redes neurais convolucionais (CNN) para a classificação de tumores cerebrais. O modelo foi treinado para identificar três tipos de tumores: glioma, meningioma e pituitário, além de classificar imagens de cérebros sem tumores. Utilizou-se o dataset Brain Tumor MRI Dataset disponível no Kaggle, e o desenvolvimento foi realizado utilizando a linguagem de programação Python. O estudo busca comparar diferentes abordagens e técnicas de regularização para otimizar o desempenho do modelo, avaliando as métricas de perda, acurácia e F1-Score.

Index Terms—CNN, Brain Tumor, Kaggle, Python, Acurácia, F1-Score.

I. INTRODUCTION

Segundo o Instituto Nacional do Câncer (INCA), "O cérebro e a medula espinhal formam o Sistema Nervoso Central (SNC), e os tumores do SNC são resultantes do crescimento de células anormais nos tecidos dessas localizações. O câncer do SNC representa de 1,4% a 1,8% de todos os tumores malignos no mundo, sendo que cerca de 88% dos tumores de SNC ocorrem no cérebro." Diante dessa realidade, o presente estudo aborda três tipos de tumores cerebrais: glioma, meningioma e tumores pituitários, além de dos cérebros sem tumores.

Gliomas, meningiomas e tumores pituitários são tipos distintos de tumores cerebrais, com etiologias, sintomas e tratamentos específicos. Os gliomas originam-se do próprio cérebro, os meningiomas das meninges, e os tumores pituitários da glândula pituitária. A Figura 1 ilustra as diferenças morfológicas e anatômicas entre esses tumores.

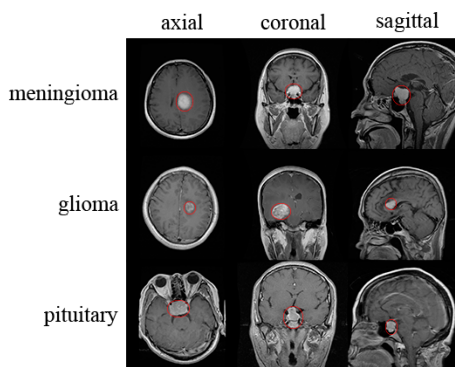


Fig. 1. Exemplos de Tumores e Regiões Marcadas

Como ilustrado na Figura 1, os tumores meningioma e pituitário podem aparecer em regiões anatômicas bastante próximas, especialmente em imagens bidimensionais. O meningioma se localiza nas meninges, que revestem o cérebro, enquanto o tumor pituitário ocorre na glândula pituitária, localizada na base do cérebro. Assim, caso um meningioma se encontre perto da base, pode ser confundido com um tumor pituitário.

Essa proximidade anatômica e as dificuldades associadas à classificação visual de tais tumores evidenciam a complexidade do diagnóstico, especialmente para pacientes ainda sem sintomas aparentes. Historicamente, a detecção precoce desses tumores tem sido um desafio. Portanto, o objetivo deste estudo é desenvolver e avaliar um modelo de Redes Neurais Convolucionais (CNN) para classificar os tumores mencionados, visando melhorar a precisão do diagnóstico e permitir a detecção precoce.

II. BASE DE DADOS

A primeira etapa para a construção do modelo envolveu a seleção e importação do banco de dados. Utilizou-se o **Brain Tumor MRI Dataset**, disponível no Kaggle, que contém imagens de ressonância magnética (MRI) rotuladas de tumores cerebrais. O dataset pode ser acessado através do link: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>.

III. ANÁLISE EXPLORATÓRIA

Para avaliar a qualidade dos dados, foi realizada uma análise exploratória detalhada, com o objetivo de identificar a proporção entre a base de treinamento e a base de teste, assim como a distribuição das classes nos dados de treinamento. Além disso, foram verificadas as dimensões das imagens, a quantidade de canais, a distribuição do tamanho das imagens em KB, e a análise estatística dos pixels, incluindo o desvio padrão e a média.

A. Proporção Dataset

Na Figura 2, apresentamos a divisão dos dados entre treinamento e teste, além da distribuição das classes no conjunto de treinamento. Vale destacar que as proporções do conjunto de teste não serão discutidas nesta etapa, pois ele sofrerá modificações devido ao uso de técnicas de data augmentation,

e sua análise ocorrerá apenas na fase final do processo de avaliação.

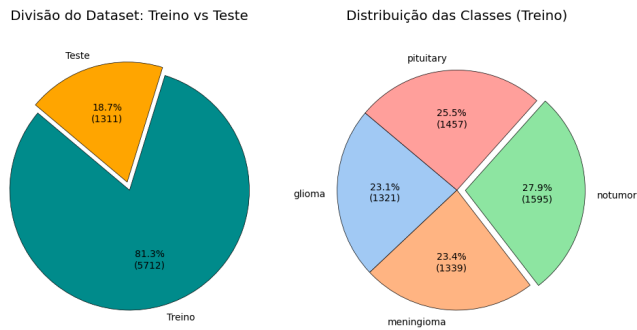


Fig. 2. Gráfico de Pizza da Distribuição do Dataset

Como mostrado no gráfico, o banco de dados está dividido da seguinte forma: 18,7% (1.311 imagens) para teste e 81,3% (5.712 imagens) para treino, totalizando 6.523 imagens. Essa divisão é bem equilibrada. Os dados de teste serão apenas analisados em termos de quantidade, a fim de evitar qualquer viés nesta etapa da análise.

Além disso, a distribuição das classes no conjunto de treinamento é bastante equilibrada entre as quatro categorias: 27,5% (1.595 imagens) de ressonâncias sem tumor, 25,5% (1.457 imagens) com tumor pituitário, 23,4% (1.339 imagens) com meningioma e 23,1% (1.321 imagens) com glioma. Com isso, não foi necessário aplicar técnicas de balanceamento de classes, pois as distribuições são suficientemente homogêneas.

B. Amostragem de Ressonâncias

Abaixo, apresentamos uma amostra com 3 imagens por classe do banco de dados de treinamento.

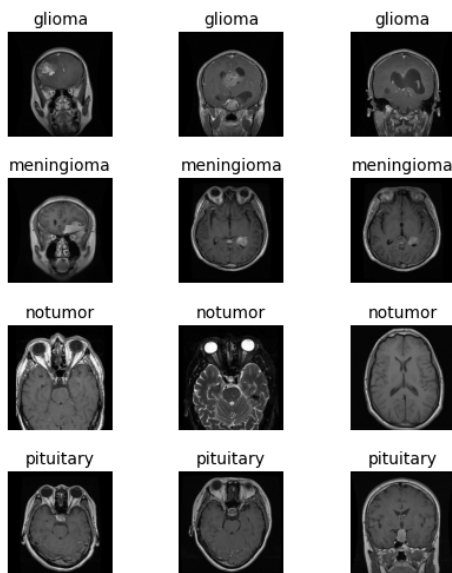


Fig. 3. Amostragem de Ressonâncias

Na figura 3, observamos três exemplos representativos de cada classe. É importante notar que as ressonâncias não foram realizadas de maneira padronizada, resultando em variações na posição e no formato das imagens. Portanto, é possível que a classe com maior variabilidade nas imagens apresente uma acurácia inferior. Caso isso aconteça, serão propostas melhorias na etapa de experimentação do modelo para lidar com essa questão.

A seguir, vamos analisar mais detalhadamente os metadados das imagens, com o objetivo de identificar técnicas de pré-processamento que possam contribuir para a melhoria do desempenho do modelo de redes neurais convolucionais (CNN).

C. Resolução das Imagens

Um modelo de redes neurais exige que suas entradas possuam um tamanho fixo. No caso de um modelo de rede neural convolucional, a altura e largura das imagens também devem ser constantes. Para identificar possíveis padrões nas imagens de entrada, apresentamos abaixo o gráfico de dispersão de todas as imagens do conjunto de treinamento.

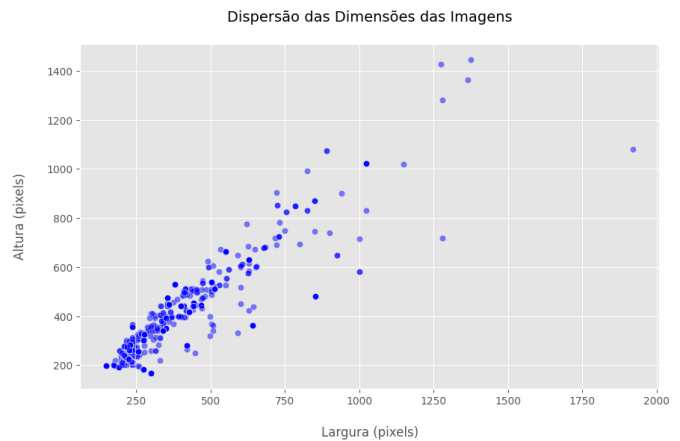


Fig. 4. Gráfico de Dispersão da Resolução das Imagens

No gráfico acima, observamos a dispersão das resoluções das imagens. É possível perceber que a maioria das imagens apresenta dimensões na faixa de altura entre [200, 600] e largura entre [200, 500], mas há alguns outliers com resoluções significativamente maiores.

Dado que este estudo propõe o uso de uma arquitetura de rede neural convolucional (CNN), que exige considerável poder computacional, será necessário reduzir as dimensões das imagens. Essa redução não só ajudará a diminuir a quantidade de operações de convolução nas primeiras camadas da rede, mas também contribuirá para a padronização das entradas do modelo. Na próxima seção, será definida uma abordagem de redimensionamento que equilibre a redução dos parâmetros da rede sem comprometer o desempenho do modelo.

D. Canal de Cores

Além da transformação da resolução das imagens, será aplicada a conversão do canal de cores de RGB para escala de cinza, resultando em apenas um canal. Essa conversão

contribuirá para a otimização das operações de convolução na primeira camada da rede, reduzindo a complexidade computacional sem comprometer a qualidade das imagens.

No caso das ressonâncias magnéticas, apesar de estarem no formato RGB, sua visualização é essencialmente em tons de preto e branco. Portanto, a conversão para escala de cinza não acarretará perda de informação relevante para o modelo.

E. Tamanho das Imagens

Para identificar possíveis outliers ou inconsistências no conjunto de dados, apresentamos abaixo o boxplot que mostra o tamanho das imagens em KB.

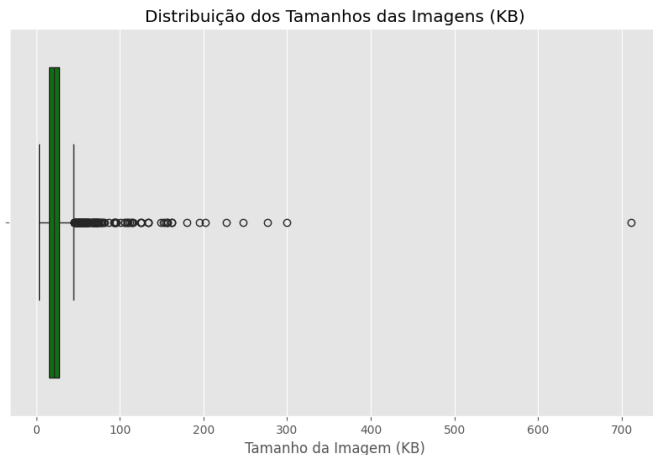


Fig. 5. Distribuição do Tamanho das Imagens em KB

No boxplot acima, podemos observar que o terceiro quartil está aproximadamente em 40 KB, o que indica que 75% (4.284 imagens) do conjunto de treinamento possuem um tamanho igual ou inferior a esse valor.

No entanto, também é possível identificar a presença de vários outliers, com algumas imagens atingindo até 700 KB. Essas imagens de maior tamanho serão tratadas durante o pré-processamento, em que serão convertidas para escala de cinza (reduzindo para um único canal) e redimensionadas para um tamanho uniforme de largura e altura. Esse processo de padronização ajudará a reduzir a variação no tamanho das imagens, além de otimizar o desempenho computacional do modelo.

F. Análise de Ruído e Brilho

No gráfico a seguir, será possível observar a distribuição do brilho e do ruído das imagens após a conversão para o canal de cinza. Isso nos ajudará a identificar se há imagens inconsistentes no banco de dados.

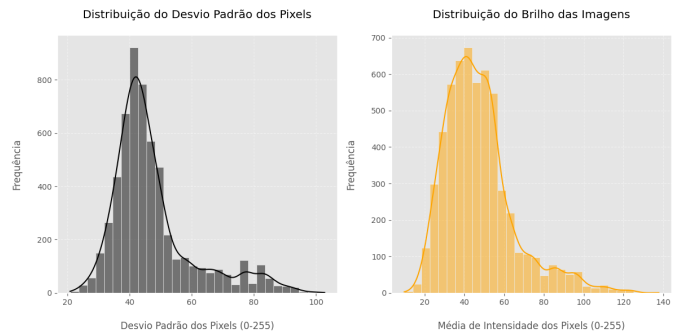


Fig. 6. Histogramas de Ruído e Brilho

Nos histogramas apresentados, podemos observar a distribuição do desvio padrão dos pixels das ressonâncias (em preto, à esquerda) e a distribuição da média dos pixels, que representa o brilho das imagens (em laranja, à direita). Ambas as distribuições apresentam assimetria positiva, o que indica que a maioria das imagens tem valores de desvio padrão e média entre 20 e 60, embora algumas observações apresentem valores significativamente mais elevados.

O desvio padrão dos pixels reflete o nível de contraste das imagens. A maior parte das ressonâncias possui contraste baixo a moderado, situando-se nos intervalos de [20, 40] e [40, 60], respectivamente. No entanto, há um pico de aproximadamente 150 observações com desvio padrão em torno de 80, sugerindo a presença de imagens com alto contraste.

O histograma de brilho segue um comportamento semelhante ao do desvio padrão. Valores baixos de brilho indicam imagens predominantemente escuras, com grande parte dos pixels na cor preta, sugerindo uma menor visibilidade do tecido cerebral. Já imagens com altos níveis de brilho correspondem a ressonâncias em que o tecido cerebral ocupa uma maior área da imagem, apresentando tons mais intensos de branco.

A seguir, serão apresentadas as imagens com os maiores e menores valores de desvio padrão e brilho para uma análise visual mais detalhada.

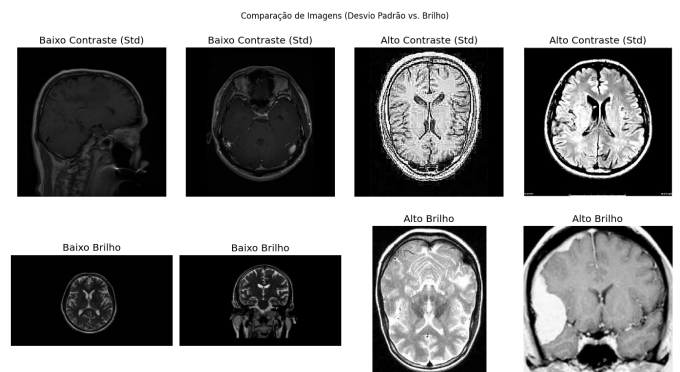


Fig. 7. Imagens com Maior e Menor Ruído e Brilho

Na figura acima, podemos observar oito exemplos de ressonâncias magnéticas: duas com baixo brilho e contraste e

duas com alto brilho e contraste.

As ressonâncias de baixo contraste apresentam poucos detalhes estruturais, dificultando a distinção das regiões cerebrais. Essas imagens tendem a ser mais escuras, com menos variação nos tons de cinza. Em contrapartida, as imagens de alto contraste exibem uma maior riqueza de informações sobre os tecidos cerebrais, mas podem conter ruídos, possivelmente resultantes de artefatos de captura ou processamento.

Em relação ao brilho, as imagens com baixo brilho apresentam uma grande área predominantemente escura, onde o fundo preto ocupa a maior parte da imagem. Já nas imagens de alto brilho, a maior parte da composição visual é formada pelo tecido cerebral, tornando os detalhes mais evidentes.

Para mitigar a variação na distribuição do desvio padrão e da média de intensidade, podemos aplicar a normalização Z-score, o que ajudará a reduzir discrepâncias no brilho e contraste dos pixels, tornando as imagens mais consistentes para o modelo.

IV. PRÉ-PROCESSAMENTO DOS DADOS

Como mostrado na Figura 2, os dados estão bem balanceados, mas foram identificados dois desafios principais: a falta de padronização na resolução das imagens e a presença de algumas com brilho e contraste excessivos. Para lidar com esses problemas, as imagens foram redimensionadas para uma resolução uniforme de 224×224 pixels. Além disso, as imagens foram normalizadas utilizando o Z-score, com média de 0.1854 e desvio padrão de 0.1793. Esse processo de normalização ajuda a garantir maior uniformidade nos valores dos pixels e a reduzir o impacto das variações extremas de iluminação. Para aumentar a robustez do modelo e introduzir variabilidade no conjunto de treinamento, foi aplicada uma rotação aleatória de até 5 graus em cada imagem.

A. Treinamento e Validação

Das 5.712 imagens de treinamento, foi proposta uma divisão de 80% (4.570 imagens) para treinamento e 20% (1.142 imagens) para validação. Essa divisão permite monitorar o desempenho do modelo ao longo do treinamento, enquanto a validação oferece uma estimativa precisa da sua capacidade de generalização. Além disso, utilizou-se um tamanho de lote (batch size) de 32 para acelerar o treinamento e introduzir estocasticidade no cálculo do gradiente.

B. Transformações na Base de Teste

Para garantir consistência e evitar vieses, as mesmas transformações aplicadas à base de treinamento foram replicadas para a base de teste. Isso incluiu o redimensionamento das imagens para 224×224 , a normalização Z-score com média de 0.1854 e desvio padrão de 0.1793, e a conversão das imagens para escala de cinza. Essas etapas garantem que as imagens de teste tenham as mesmas características das imagens usadas no treinamento, permitindo uma avaliação justa do modelo.

V. MODELO BRAINSCANET

Conforme descrito na introdução, o objetivo do nosso modelo é classificar imagens de ressonância magnética em quatro classes: sem tumor, meningioma, glioblastoma e pituitário. Para isso, propomos um modelo de rede neural convolucional densa, com a seguinte estrutura:

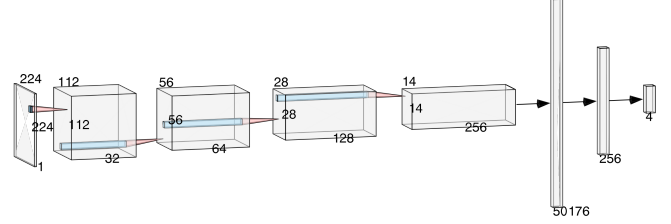


Fig. 8. Estrutura do modelo BrainScanNet

O modelo é composto por quatro blocos convolucionais, onde cada camada convolucional é seguida de Batch Normalization, função de ativação ReLU e max-pooling.

Em cada camada convolucional, utilizamos um kernel de tamanho 3×3 e um padding de 1.

Cada bloco dobra o número de filtros convolucionais, o que permite ao modelo aprender características mais complexas e detalhadas das imagens conforme as camadas se tornam mais profundas.

Após a extração das características nas camadas convolucionais, o modelo segue com camadas totalmente conectadas (Fully Connected). A primeira camada contém 50% de dropout e 50176 neurônios, enquanto a camada intermediária possui 256 neurônios, responsáveis por sintetizar características globais, como bordas, texturas e formas. Por fim, a camada final tem 4 neurônios, correspondendo às classes que desejamos classificar.

A. Hiperparâmetros do Modelo

Para o treinamento do modelo, optamos por utilizar a função de otimização Adam, com uma taxa de aprendizado inicial de 0.001. Além disso, empregamos o mecanismo ReduceLROnPlateau, que reduz a taxa de aprendizado sempre que o modelo apresenta estagnação no aprendizado. Como o treinamento foi realizado em uma placa de vídeo T4, e devido a limitações de tempo de uso, foi possível treinar o modelo por apenas 25 épocas. No entanto, mesmo com esse número reduzido de épocas, o modelo obteve bons resultados, como pode ser observado a seguir.

VI. DESEMPENHO DO MODELO

Após 25 épocas de treinamento, o modelo BrainScanNet apresentou as seguintes estatísticas no final do processo:

Métrica	Valor
Train Loss	1.78%
Val Loss	18.99%
Train Acc	99.32%
Val Acc	95.88%

TABLE I

DESEMPENHO DO MODELO AO FINAL DO TREINAMENTO

Podemos observar que, ao final do treinamento, o modelo apresentou um train loss baixo de 1.78%, indicando que ele foi capaz de aprender muito bem os dados de treinamento. No entanto, o validation loss foi relativamente alto, atingindo 18.99%, o que sugere uma possível ocorrência de overfitting. Essa hipótese, no entanto, não é confirmada completamente, pois o validation accuracy foi bastante elevado, atingindo 95.88%, o que indica que o modelo já era muito eficaz na classificação durante a validação.

Portanto, embora o modelo tenha apresentado um ótimo desempenho geral, a grande discrepância entre o train loss e o validation loss sugere que ele pode estar se especializando excessivamente nos dados de treinamento, comprometendo sua capacidade de generalização para novos dados.

A seguir, apresentamos o gráfico de desempenho ao longo do treinamento:

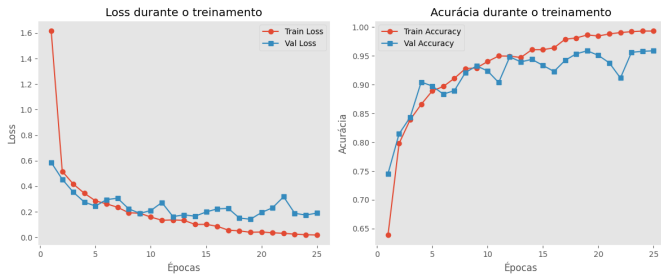


Fig. 9. Desempenho do modelo durante o treinamento

Analisando o gráfico de loss (à esquerda), podemos observar que, ao longo do treinamento, a perda para os dados de treinamento diminui constantemente. No entanto, ao observar os dados de validação, percebemos uma tendência de queda com maior ruído, estabilizando-se por volta da época 11. Isso sugere que o modelo não estava apresentando uma melhoria contínua para novas observações. Essa tendência também é visível no gráfico à direita, onde a acurácia de validação se mantém lateralizada a partir da época 11 e apresenta algumas quedas. Para verificar se o modelo está sofrendo de overfitting, dado que a acurácia de treinamento foi de 99.32%, vamos avaliar os dados de teste na seção a seguir.

VII. ETAPA DE TESTE

Para investigar se o nosso modelo sofreu overfitting, realizamos uma análise utilizando os dados de teste.

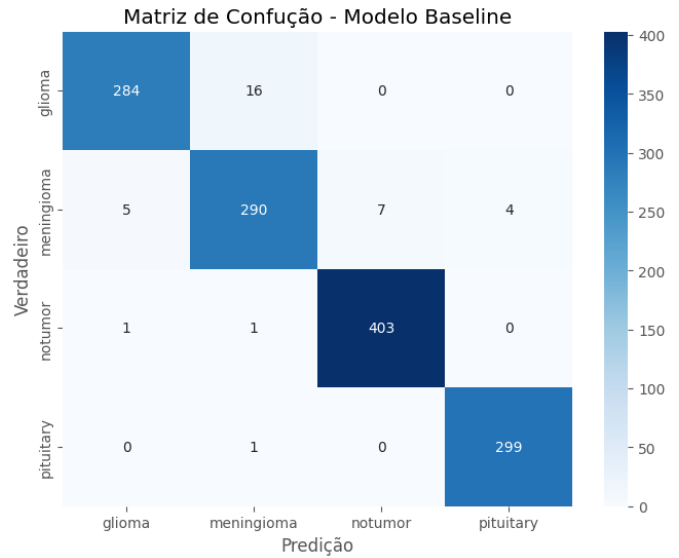


Fig. 10. Matriz de confusão do modelo

Na matriz de confusão do modelo baseline, observamos que as classes com maior número de erros foram glioma e meningioma, com 16 imagens classificadas incorretamente nos dados de teste.

Por outro lado, para indivíduos sem tumor, apenas dois casos foram erroneamente classificados como glioma ou meningioma, indicando que o modelo tem bom desempenho na diferenciação entre tumores e a ausência deles.

Em relação ao pituitário, o modelo errou apenas uma classificação, confundindo-o com meningioma. Esse erro pode ser explicado, pois como mencionado anteriormente, em alguns casos, um meningioma localizado na região inferior do cérebro pode ser confundido com um tumor pituitário.

Além da matriz de confusão, utilizamos a métrica F1-Score para avaliar o equilíbrio da classificação do modelo. O F1-Score é uma métrica que pondera tanto a acurácia quanto o recall:

- **Acurácia:** Indica a proporção de classificações corretas do modelo em relação ao total de classificações realizadas (considerando tanto os verdadeiros positivos quanto os falsos positivos).
- **Recall:** Mede a capacidade do modelo de identificar corretamente as imagens de uma determinada classe (considerando verdadeiros positivos e falsos negativos).

Foi observado que o modelo BrainScanNet obteve um F1-Score de aproximadamente 97,32%, o que indica que o modelo possui um ótimo equilíbrio entre a classificação correta das imagens e a minimização de classificações errôneas.

Portanto, podemos concluir que o nosso modelo apresenta uma boa performance para novas observações.

A. Visão do Modelo

A imagem a seguir ilustra como o modelo está visualizando e atribuindo peso às regiões das imagens nunca vistas. Isso

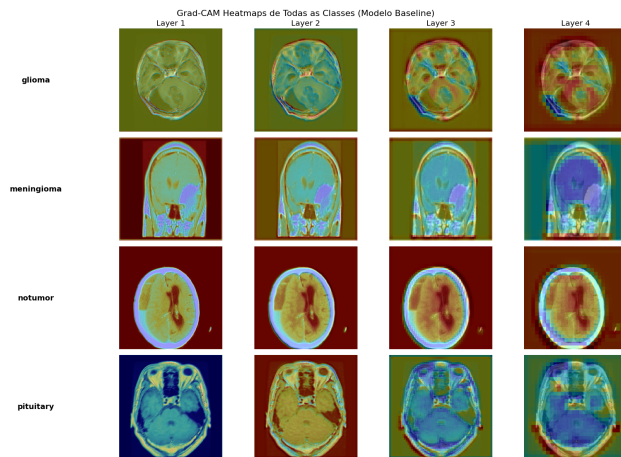


Fig. 11. Visualização das regiões de destaque do modelo

nos proporciona uma visão do que está acontecendo em cada camada convolucional.

Como pode ser visto, as duas primeiras camadas convolucionais tentam delimitar e entender a estrutura geral da tomografia, considerando que as imagens podem ter diferentes ângulos da cabeça. Já a terceira e quarta camadas atribuem pesos diferentes às regiões internas do cérebro, buscando identificar pixels relevantes para a classificação.

VIII. CONCLUSÃO

Ao longo deste trabalho, observamos que, conforme afirmado pela INC em 2024, não há atualmente um mecanismo eficaz para identificar tumores cerebrais em pacientes assintomáticos. No entanto, os resultados obtidos com o modelo BrainScanNet demonstram que redes neurais convolucionais podem ser uma grande aliada na saúde pública. O desempenho do BrainScanNet foi bastante promissor em novas observações, e o modelo pode ser facilmente implementado para auxiliar médicos na tomada de decisões, além de permitir a identificação precoce de tumores cerebrais em pacientes. Dessa forma, o uso dessa tecnologia pode contribuir significativamente para a melhoria da detecção e do tratamento de tumores cerebrais, promovendo benefícios para a saúde pública.

REFERENCES

- [1] BILSKY, Mark H. Visão geral dos tumores intracranianos. MSD Manuals, jul. 2024. Disponível em: <https://www.msdmanuals.com/pt/profissional/dist%C3%BArbios-neurol%C3%B3gicos/tumores-intracranianos-e-espinhais/vis%C3%A3o-geral-dos-tumores-intracranianos>. Acesso em: 3 de fevereiro de 2025.
- [2] Badža, M.M.; Barjaktarović, M.Č. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. Appl. Sci. 2020, 10, 1999. Disponível em: <https://doi.org/10.3390/app10061999>.
- [3] Nickparvar, M. "MRI Brain Tumor Dataset: Brain Tumor MRI Dataset." Kaggle. Disponível em: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>.
- [4] Krizhevsky, A; Sutskever, I; Hinton, G. "ImageNet Classification with Deep Convolutional Neural Networks." Communications of the ACM, 2012.

- [5] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. Dive into Deep Learning. Cambridge University Press.