

Aplicação de Inteligência Artificial para a Classificação Multiclasses de Câncer De Estômago

Alex Se Yoon Koo ¹
Leandro Augusto da Silva ^{1, 2}

¹Faculdade de Computação de Informática (FCI)
Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

²Programa de Pós-graduação em Engenharia Elétrica/Computação Aplicada

<alexksyn@gmail.com>
<leandroaugusto.silva@mackenzie.br>

2025

Resumo

O câncer de estômago apresenta alta incidência e mortalidade, sendo frequentemente diagnosticado em estágios avançados pela ausência de sintomas precoces. As limitações dos métodos diagnósticos atuais e a escassez de patologistas motivam o uso de Inteligência Artificial para aumentar a precisão. Este projeto, desenvolvido em parceria com o Hospital da Santa Casa de Misericórdia de São Paulo, visa criar uma solução inovadora para este contexto. O objetivo principal é treinar um modelo de Rede Neural Convolucional (CNN) para classificar lesões de câncer gástrico a partir de imagens do órgão retirado, capturadas por celular durante a cirurgia. A metodologia inclui a curadoria de um banco de dados com imagens anotadas por especialistas, uma revisão da literatura para selecionar as arquiteturas de CNN mais eficientes e o treinamento do modelo. Como justificativa, a pesquisa oferece uma ferramenta de auxílio diagnóstico que pode funcionar como uma "segunda opinião" em tempo real para equipes cirúrgicas, especialmente em hospitais com recursos limitados. O resultado esperado é um algoritmo validado para a classificação de lesões, que servirá de base para um futuro aplicativo, buscando melhorar a precisão diagnóstica em um momento crítico do tratamento.

Palavras-chave: rede neural convolucional; aprendizagem profunda; câncer.

Abstract

Stomach cancer has a high incidence and mortality rate, often diagnosed in advanced stages due to the absence of early symptoms. The limitations of current diagnostic methods and the shortage of pathologists motivate the use of Artificial Intelligence to increase accuracy. This project, developed in partnership with the Santa Casa de Misericórdia Hospital of São Paulo, aims to create an innovative solution for this context. The main objective is to train a Convolutional Neural Network (CNN) model to classify gastric cancer lesions from images of the removed organ, captured by a cell phone during surgery. The methodology includes curating a database of images annotated by specialists, a literature review to select the most efficient CNN architectures, and training the model. As a justification, the research offers a diagnostic support tool that can act as a real-time "second opinion" for surgical teams, especially in hospitals with limited resources. The expected outcome is a validated algorithm for lesion classification, which will serve as the basis for a future application, aiming to improve diagnostic accuracy at a critical moment of treatment.

Keywords: convolutional neural networks; deep learning; câncer.

1 Introdução

O câncer de estômago, também conhecido como câncer gástrico é o terceiro tipo mais frequente de pessoas do sexo masculino e o quinto do sexto feminino. Estima-se que em 2022 foram diagnosticados 21230 novos casos, sendo que 63% foram acometidos em homens e 27% em mulheres. Já o número de mortes alcançou um total de 14260 pessoas, mantendo a maior estatística para o sexo masculino. Em relação a idade, é um tipo de doença que atinge pessoas na idade adulta (acima dos 60 anos) e raramente apresentam sinais, dificultando o diagnóstico precoce que poderia aumentar a chance de cura (BRASIL. Ministério da Saúde. Instituto Nacional de Câncer, 2025).

Pela falta de sintomas e por ser uma doença que se desenvolve lentamente, quando o diagnóstico é feito, geralmente o tumor já se encontra em estágios avançados. Os tipos mais comuns de exames para o diagnóstico são endoscopia, quadros patológicos, imagens patológicas, exames por imagens etc. A endoscopia é tida como o tipo de exame mais preciso para o diagnóstico. No entanto, estudos apontam que até 10% dos cânceres gastrointestinais superiores não são detectados por este tipo de exame, mesmo quando a imagem é analisada por mais de um especialista e a razão para isto é a necessidade de acúmulo de anos de experiência. Por conseguinte, o uso de imagens histopatológicas que é o padrão ouro deste tipo de diagnóstico tem como agravante a escassez de patologistas, levando a alta carga de trabalho para os patologistas e consequentes erros no diagnóstico (ZHAO et al., 2022).

Neste cenário de alta demanda e dificuldade de um diagnóstico mais preciso que emerge a aplicação de Inteligência Artificial na classificação, detecção e segmentação de tumores, especialmente em caso de câncer de estômago (ZHAO et al., 2022; IIZUKA et al., 2020; KHAN et al., 2020; KHAN et al., 2021; KUNTZ et al., 2021).

Dentre os algoritmos de Aprendizagem de Máquina, os quais são treinados com dados e tornam viáveis a aplicação da Inteligência Artificial, as Redes Neurais Convolucionais (CNN, do inglês Convolution Neural Network) é uma das arquiteturas mais proeminentes para processamento de imagens em aprendizagem profunda (ZHAO et al., 2022; IIZUKA et al., 2020; KHAN et al., 2020; KHAN et al., 2021; KUNTZ et al., 2021).

Em uma revisão da literatura que envolveu a comparação de diferentes arquiteturas de CNN como AlexNet, ResNet, VGG, Inception, DenseNet e Deeplab, etc, a precisão dos estudos na detecção de câncer de estômago foi entre 77,3 – 98,7% de acurácia (ZHAO et al., 2022). Estes resultados motivam e justificam diversos outros estudos com aplicação de Inteligência Artificial em câncer de estômago (IIZUKA et al., 2020; KHAN et al., 2020; KHAN et al., 2021; KUNTZ et al., 2021).

1.1 Problema de Pesquisa

Este trabalho será parte de um projeto maior para o desenvolvimento de um aplicativo que auxilia a identificação e classificação de lesões de câncer de estômago com imagens do estômago no momento da operação.

O projeto está sendo feito em parceria com especialistas no assunto de câncer de estômago do Hospital da Santa Casa de Misericórdia de São Paulo. Este grupo tem um conjunto de casos de imagens documentadas e anotadas que serão organizadas em um conjunto de dados para os estudos desta pesquisa. Cabe destacar que já há um grande conjunto de imagens para iniciar os estudos e o que precisa é uma organização para compartilhamento e uso no treinamento dos algoritmos.

O problema a ser explorado nesta pesquisa consiste em revisar a literatura em busca dos algoritmos de CNN com melhores precisões de reconhecimento de tumores ou de outros tipos de padrões em imagens para ser adaptado em diagnósticos complexos que exigem a retirada do estômago para o diagnóstico.

Portanto trata-se de um contexto diferente aos estudos da literatura, pois trata-se de imagens coletadas com o próprio celular.

1.2 Justificativa

A principal justificativa para o estudo que se trata de uma aplicação inovadora na medicina especializada em câncer de estômago. Embora o estudo atual não tenha foco no diagnóstico de imagem como exames de endoscopia e etc, como contextualizado antes é uma doença silenciosa no sentido de sintomas ao paciente e o número de casos é extremamente significativo, sendo o terceiro tipo de câncer mais frequente na população brasileira.

Significa dizer que muitos casos são identificados apenas com procedimentos cirúrgicos, exigindo a retirada do estômago e a abertura dele, uma vez que a lesão se apresenta na parte interna do estômago.

Como também contextualizado antes a relação da precisão no diagnóstico depende da experiência da equipe médica. E nesta lacuna que o aplicativo tem como maior contribuição da pesquisa, permitindo hospitais com baixo recurso ter uma segunda opinião sobre o caso clínico analisado em tempo de operação sobre o tipo de lesão que está sendo diagnosticado.

1.3 Objetivo

O objetivo geral deste trabalho é treinar um modelo CNN para o reconhecimento de tipos de câncer de estômago em imagens coletadas por celular em momento da cirurgia. De forma específica haverá uma organização dos dados para que se tenha uma metodologia de referência para outros casos sejam armazenados em um grande repositório de dados, também será feita uma revisão da literatura em vista a seleção dos algoritmos com melhor desempenho na detecção de lesões em imagens médicas e, ainda, a análise de viabilidade de adicionar dados pessoais do paciente e de hábitos como informações complementares visando analisar a melhoria da precisão.

Cabe pontuar assim que neste projeto a entrega será o algoritmo de reconhecimento das lesões e classificação delas para uso no aplicativo que será desenvolvido no projeto geral de pesquisa.

2 Referencial Teórico

2.1 Redes Convolucionais

Redes Convolucionais ou CNN tem excelente desempenho de acurácia em muitas áreas do conhecimento, como visão computacional e natural processamento de linguagem. Especialmente no

campo da visão computacional, CNN é o modelo de estado da arte para classificação de imagens, recuperação de imagens, detecção de objetos e segmentação de imagens.

O uso de CNN no auxílio ao diagnóstico para identificar o câncer de estômago (gástrico) tem se destacado por proporcionar mais precisão e em menos tempo no diagnóstico. Os resultados se destacam em redução de omissões e descaracterizações das alterações de tumores, ajudando assim a resolver as limitações da falta de experiência dos médicos atuais em exames.

A arquitetura básica de uma rede CNN está ilustrada na Figura 1. Note que em uma mesma arquitetura se tem a extração de características da imagem e o treinamento para classificação da imagem. Na etapa de extração de características existe uma camada de convolução, na qual são realçadas características de forma, contorno, textura e etc por meio de elementos estruturantes que são matrizes de menores dimensões, chamadas de kernel, arranjadas em valores que possibilitam fazer tais destaques. Precedendo a convolução, se tem a etapa de pooling na qual a imagem resultante da convolução tem a sua dimensionalidade reduzida. Por último há uma camada de treinamento e classificação, que no exemplo tem uma rede neural FC (Fully Connected). A rede FC é formada por camadas de neurônio, cuja configuração pode ser ajustada como meta-parâmetros da rede. A rede FC pode também ser substituída por outros algoritmos de Aprendizagem de Máquina.

As variações de arquiteturas consistem na diferenciação e/ou composição das etapas de convolução e pooling. Clássicas estruturas de CNN incluem AlexNet, ResNet, VGG, Inception e etc. A profundidade da rede refere-se ao número de camadas que precisam ser atualizadas, parâmetros por meio de treinamento, como camadas convolucionais, camadas totalmente conectadas, etc.

A estrutura básica da rede AlexNet inclui 8 camadas no total. Existem 5 camadas convolucionais, seguidas por 3 camadas totalmente conectadas (Fully Connected). Na literatura de imagens médicas, a precisão da rede AlexNet é superior aos métodos tradicionais. O tamanho do kernel das primeiras 4 camadas convolucionais é 11×11 , 5×5 , 3×3 , 3×3 e 3×3 . Na camada de Fully Connected as duas primeiras camadas totalmente conectadas têm 4.096 neurônios respectivamente, e a saída final tem 1.000 neurônios.

A rede VGG foi proposta com base na rede AlexNet. Comparado com AlexNet, VGG usa o kernel de convolução com 3×3 e o kernel de pooling com 2×2 . A principal diferença foi feita na camada da Fully Connected. A arquitetura mais utilizada é a rede VGG-16, que tem 13 camadas convolucionais e 3 camadas totalmente conectadas.

GoogLeNet é uma rede de Inceptions proposta pelo Google cujo módulo de iniciação substitui as operações tradicionais de convolução e funções de ativação. Inception é caracterizado pelo uso de convolução 1×1 para redução de dimensionalidade, enquanto a convolução e a reagregação podem ser realizadas em múltiplas escalas. Nesta rede todas as camadas totalmente conectadas são substituídas por um pooling de média global simples, que reduz o número total de parâmetros do modelo. Posteriormente, o modelo foi melhorado com as versões de Inception-v2, Inception-v3, Inception-v4 e outras versões foram desenvolvidas.

ResNet foi projetado como um bloco residual para permitir treinar redes mais profundas e aumentar desempenho do modelo. Existem 5 estruturas de profundidade diferentes que são 18, 34, 50, 101 e 152 camadas. A rede mais comum é ResNet-50. Tem um total de 50 camadas de estrutura, incluindo 49 camadas convolucionais e uma camada totalmente conectada. O tamanho do kernel de convolução inclui três tipos: 1×1 , 3×3 e 7×7 .

DenseNet é uma rede neural convolucional com conexão densa. A proposta da rede DenseNet eliminar a necessidade de configuração do número de camadas da rede ResNet e ampliando a estrutura da rede Inception para melhorar o desempenho da rede.

2.2 Coleta de Dados

As imagens são oriundas dos especialistas da Santa Casa, assim como os dados clínicos e os diagnósticos. As imagens apresentam uma inconsistência nos seus tamanhos e pela sua distribuição desbalanceada por categoria Borrmann apresentada pela figura 1.

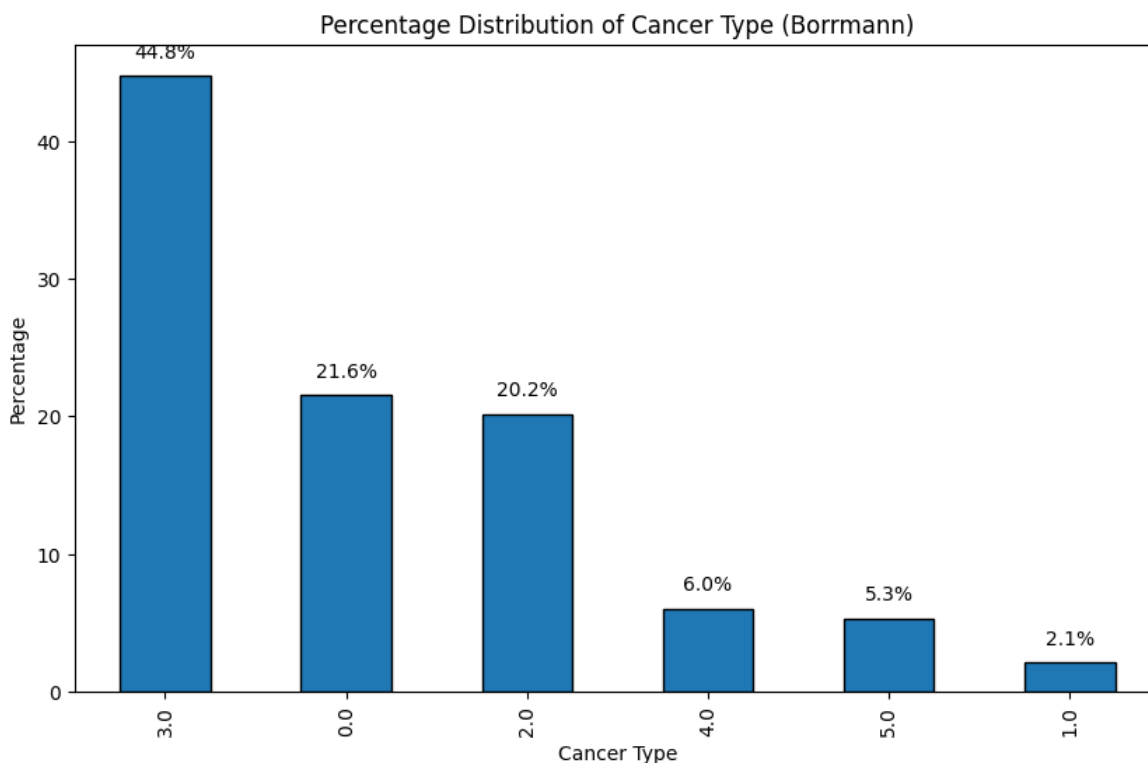


Figura 1 – Gráfico de distribuição dos tipos de Borrmann

2.3 Pré processamento das imagens

Como parte da etapa de pré processamento das imagens realizamos os seguintes procedimentos:

2.3.1 Redimensionamento das imagens

Para redimensionamento usamos um tamanho padrão para todas as imagens para 450 x 450 com um channel de 3 camadas RGB (*Red, Green, Blue*).

2.4 Técnicas de Augmentation

Para superar a barreira de desbalanceamento entre os tipos de borrmann usamos a técnica de oversampling nos tipos que apresentam quantidade mais baixos aplicando as filtragens abaixo de maneira estocástica.

2.4.1 Brilho

O brilho apresenta estar na faixa entre -40 (mais escuro) até 60 (mais claro).

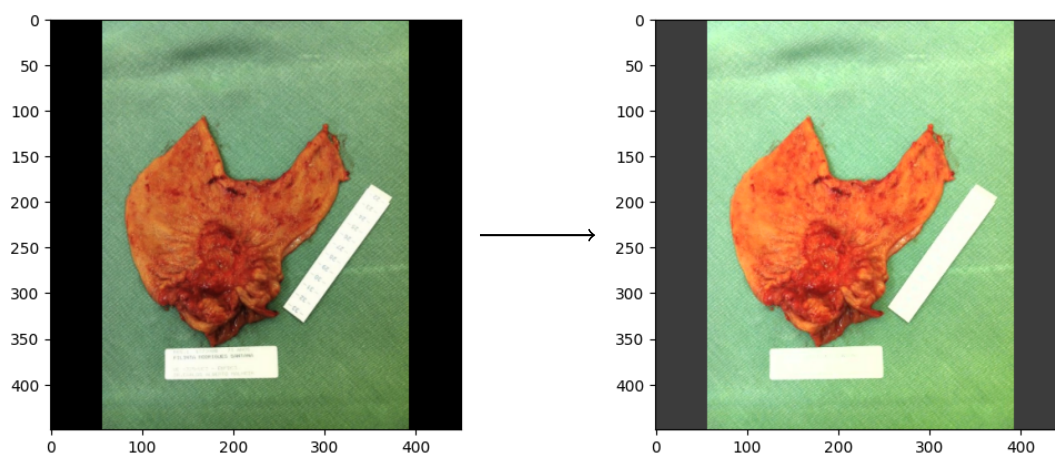


Figura 2 – Exemplo de uma transformação de brilho de uma figura (esquerda) para outra (direita).

2.4.2 Zoom

O *zoom* foi feita apenas o zoom in em um intervalo de valor entre 0.5 até 1, sendo esse o tamanho da imagem.

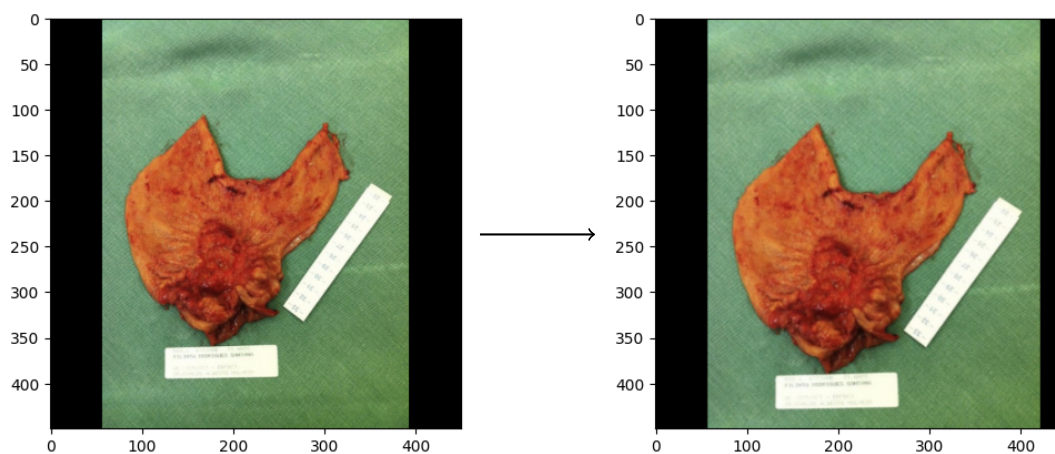


Figura 3 – Exemplo de uma transformação de Aumento de uma figura (esquerda) para outra (direita).

2.4.3 Inversão Horizontal

A inversão horizontal é o processo da imagem apresentar o seu inverso pelo eixo y.

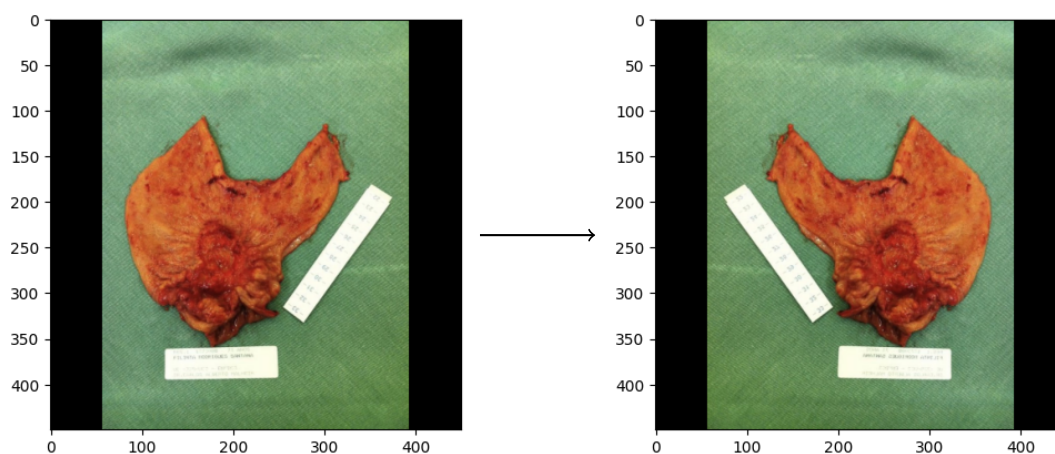


Figura 4 – Exemplo de uma transformação de inversão horizontal de uma figura (esquerda) para outra (direita).

2.4.4 Inversão Vertical

A inversão vertical é o processo da imagem apresentar o seu inverso pelo eixo x.

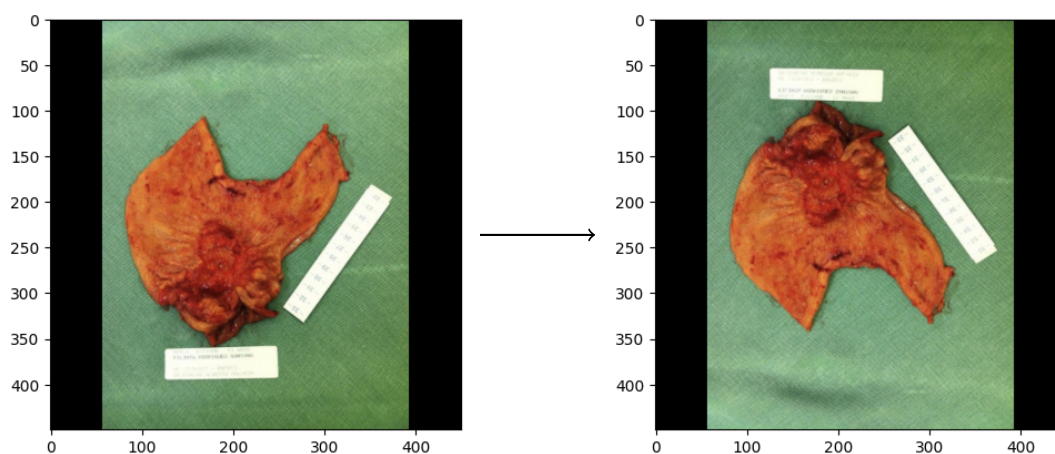


Figura 5 – Exemplo de uma transformação de inversão vertical de uma figura (esquerda) para outra (direita).

2.4.5 Rotação

A rotação da figura foi feita no intervalo de -30 a 30.

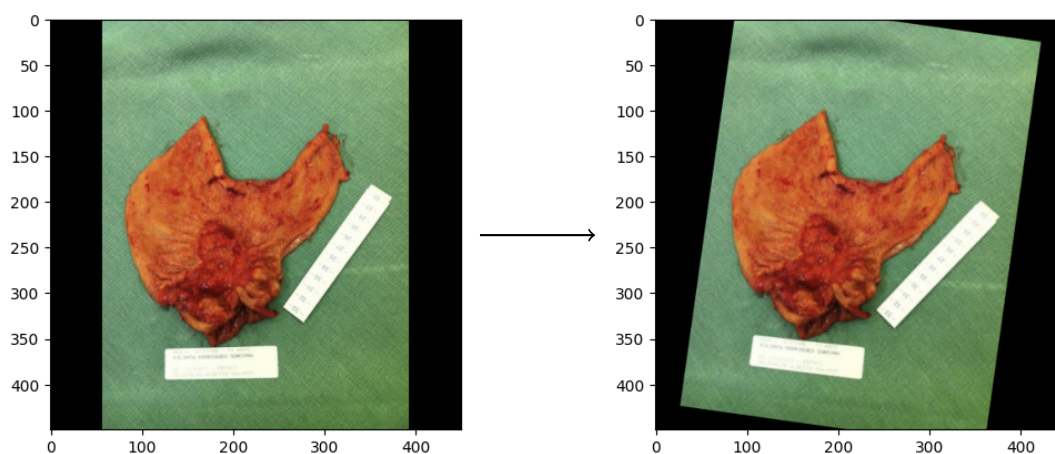


Figura 6 – Exemplo de uma transformação de rotação de uma figura (esquerda) para outra (direita).

2.4.6 Ruído de Sal e Pimenta

O ruído de sal e pimenta é um tipo de ruído onde as imagens apresentam pontos pretos e brancos.

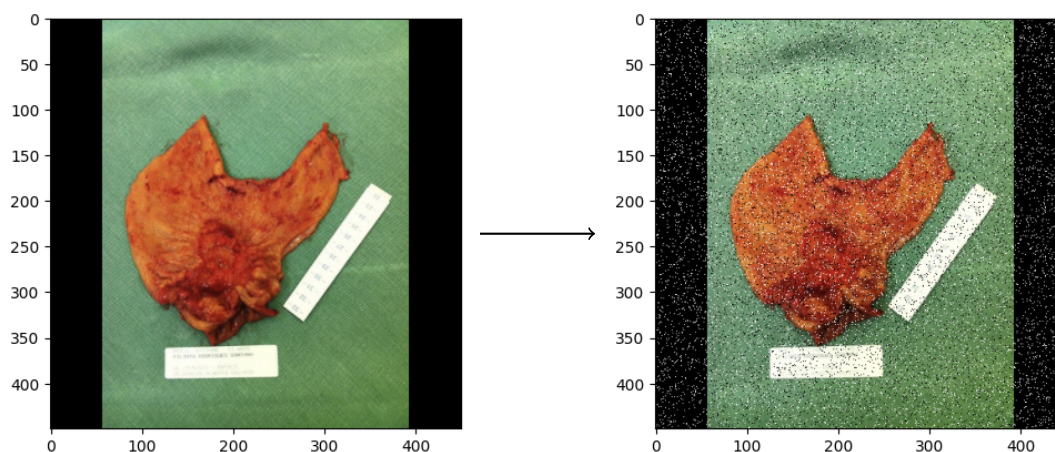


Figura 7 – Exemplo de uma aplicação de ruído sal e pimenta de uma figura (esquerda) para outra (direita).

2.4.7 Normalização das imagens

A normalização apresenta ser uma técnica de padronizar todas as imagens dentro do intervalo de 0 a 1 respeitando o seu valor e portando todas as imagens tanto no treino quanto no teste foram normalizadas divididas por 255.

2.5 Estudo dos algoritmos

3 Rede Neural

Nesta seção explicamos os algoritmos estudados e utilizados no decorrer da pesquisa.

Para o desenvolvimento da nossa pesquisa utilizamos um dos mais conhecidos de Deep Learning chamada de Redes Neurais, onde a arquitetura é dividida em camadas onde cada uma possui vários neurônios que são ativados pela sua função de ativação seja ReLU (*Rectified Linear Activation*), Logistic (*Sigmoid*) ou Tangente Hiperbólico (*Tanh*).

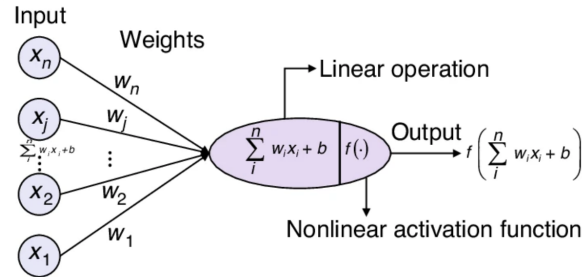


Figura 8 – Exemplo de um neurônio com a sua função de ativação. Fonte: *Optical neural networks: progress and challenges*, Nature Reviews Physics.

O nosso modelo é do tipo *Feed Forward Neural Network*, onde a rede neural é densa em todas as suas camadas, ou seja, a primeira camada está totalmente conectada com a segunda camada e essa com a terceira e assim sucessivamente, também chamada de *Fully Connected Layer* (Camada totalmente conectada ou Densa), conjuntamente com as camadas de convolução e pooling como mostra na figura 8.

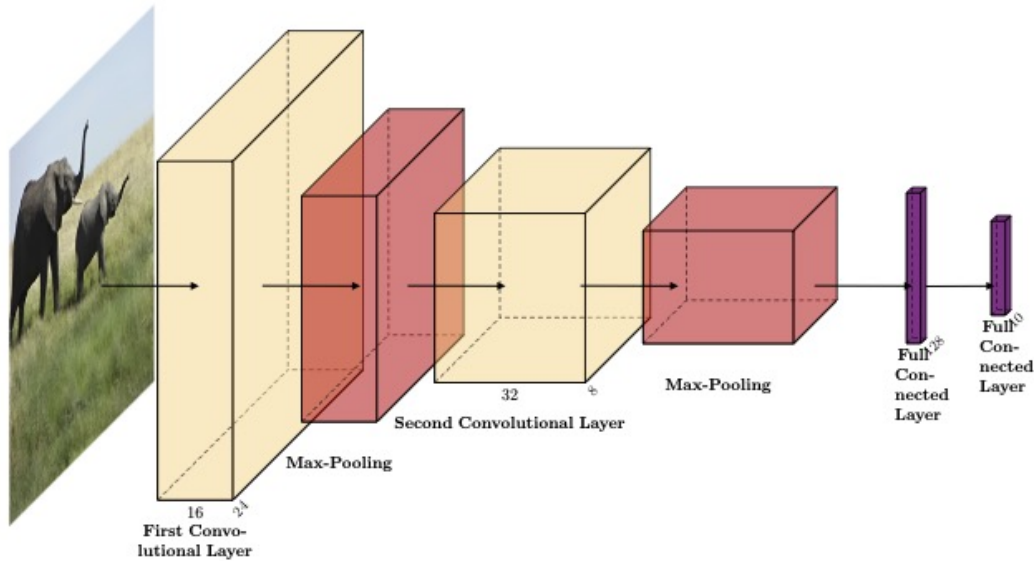


Figura 9 – Exemplo de uma rede neural convolucional.

4 Arquiteturas

Usamos duas arquiteturas de redes neurais convolucionais: ResNet50 e VGG16. A principal diferença entre essas duas arquiteturas se dá pela sua complexidade e profundidade sendo a ResNet50 mais recomendável para problemas que exigem mais complexidade.

4.0.1 ResNet50

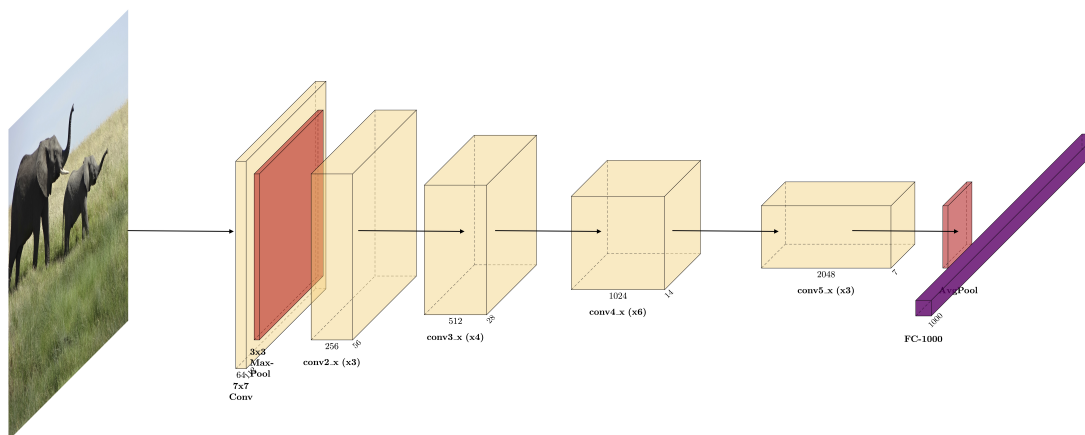


Figura 10 – Exemplo de uma rede neural convolucional de arquitetura ResNet50.

4.0.2 VGG16

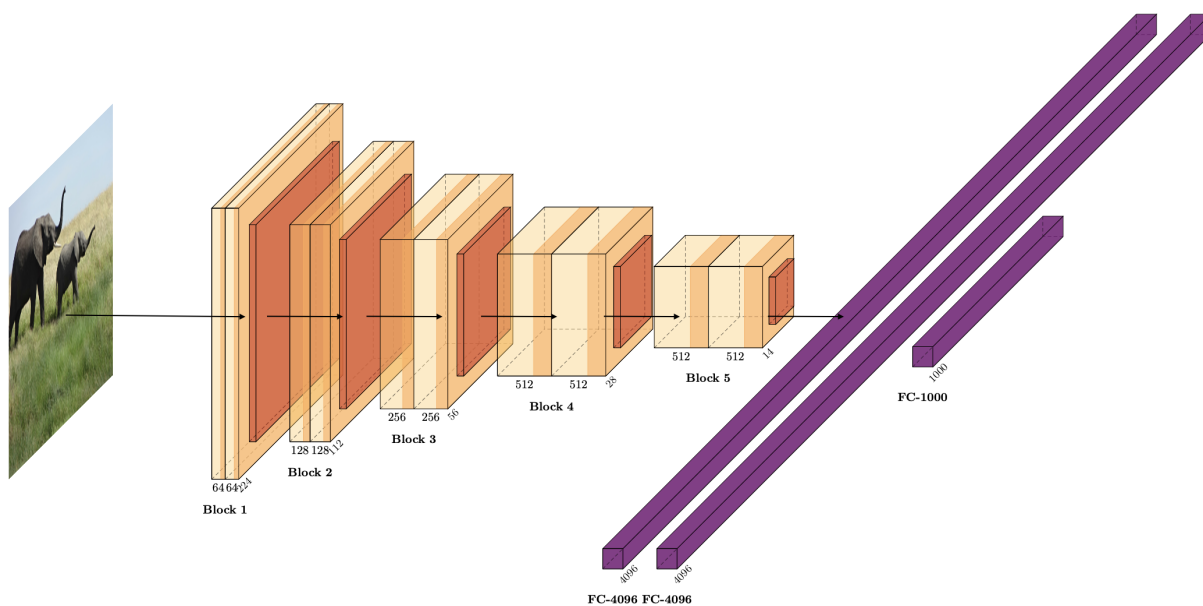


Figura 11 – Exemplo de uma rede neural convolucional de arquitetura VGG16.

4.1 Parametrização dos algoritmos

Nesta seção, iremos detalhar como foi a implementação do nosso modelo, principalmente para o seu treino e como a pesquisa foi conduzida em detalhes. Os experimentos foram feitos tanto para a arquitetura VGG16 e ResNet50.

4.1.1 Experimento 1

Nesse experimento pegamos os pesos dos neurônios do modelo treinado com *learning rate* de 0.001, modificar a última camada convolucional e retrainar o modelo.

4.1.2 Experimento 2

Nesse experimento liberamos os dois últimos blocos convolucional para treinamento (*fine-tuning*) que possuem a mesma dimensão, além de aumentar o batch size para 32.

4.1.3 Experimento 3

Nesse experimento diferente da anterior, ao invés de realizar apenas um treinamento de fine tuning diretamente, primeiro fizemos um treinamento com os pesos das arquiteturas com 10 épocas com *learning rate* de 0.001 e depois realizamos um treinamento com *learning rate* de 0.0001 por 100 épocas.

4.1.4 Experimento 4

Nesse experimento adicionamos o filtro de ruído de sal e pimenta, além de realizar o treinamento somente com as classes 0 e 2 e outra por 0 e 3. Para balanceamento do dataset utilizamos técnicas de *augmentation* e *subsampling*.

4.1.5 Experimento 5

No experimento 5, modificamos a camada de saída para dois neurônios, e alterar os ruídos para que seja aplicada de maneira estocástica, além de aumentar a quantidade de imagens usando técnicas de *augmentation* e realizar treinamentos para diferentes classes como: 0 e 2, 0 e 3 e 2 e 3. Neste mesmo experimento não utilizamos a rede neural de 2 saídas na arquitetura, pois a arquitetura VGG16 não apresentava um resultado melhor quando comparada com Resnet50.

5 Resultados Experimentais e Discussão

5.1 Experimento 1

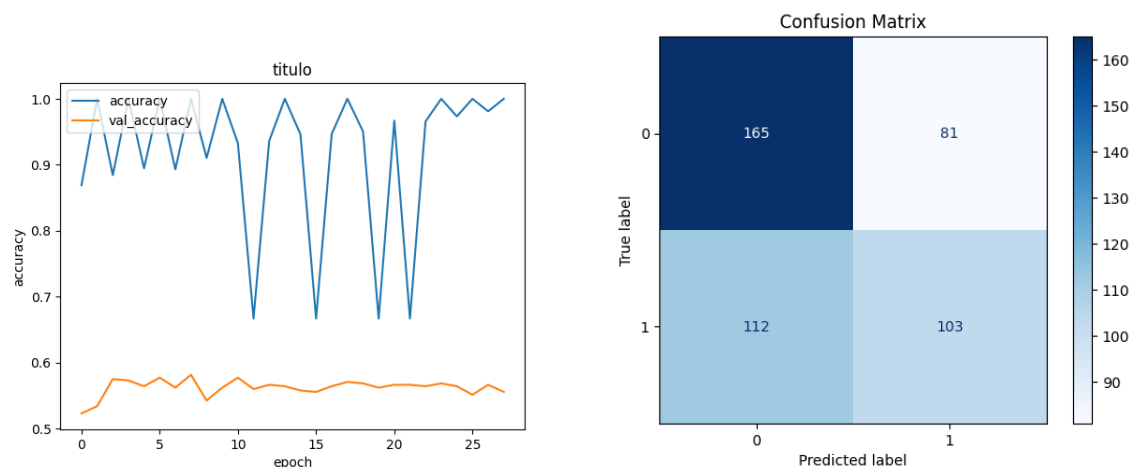


Figura 12 – Modelo ResNet com o seu treinamento e matriz de confusão.

Neste experimento podemos observar que a curva de aprendizagem não apresenta uma estabilidade durante o seu treinamento o que justifica a sua matriz de confusão das predições realizadas no seu dados de teste.

5.2 Experimento 2

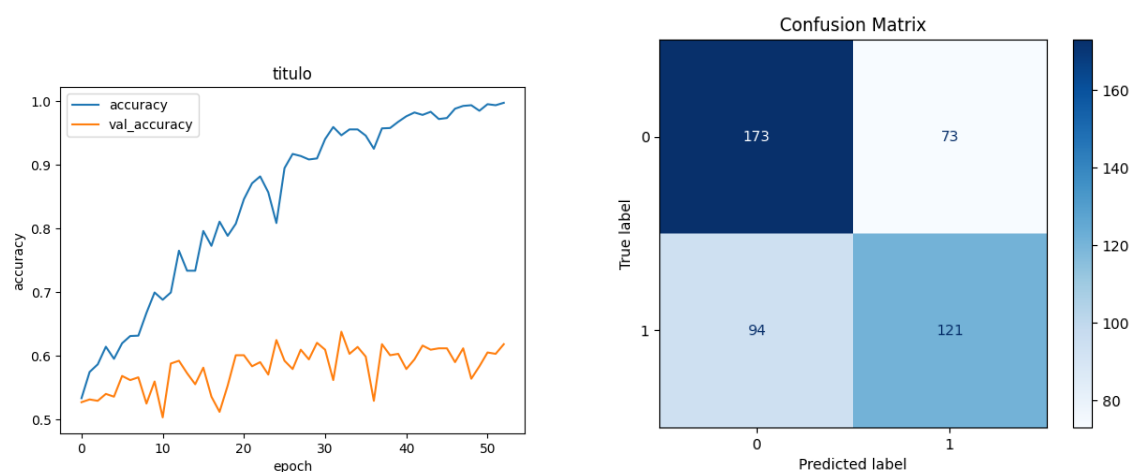


Figura 13 – Modelo ResNet com o seu treinamento e matriz de confusão.

Neste experimento podemos observar que as curvas de aprendizagem apresentam uma certa estabilidade quando comparada com o experimento 1, porém com os dados de validação ela apresenta uma certa instabilidade.

5.3 Experimento 3

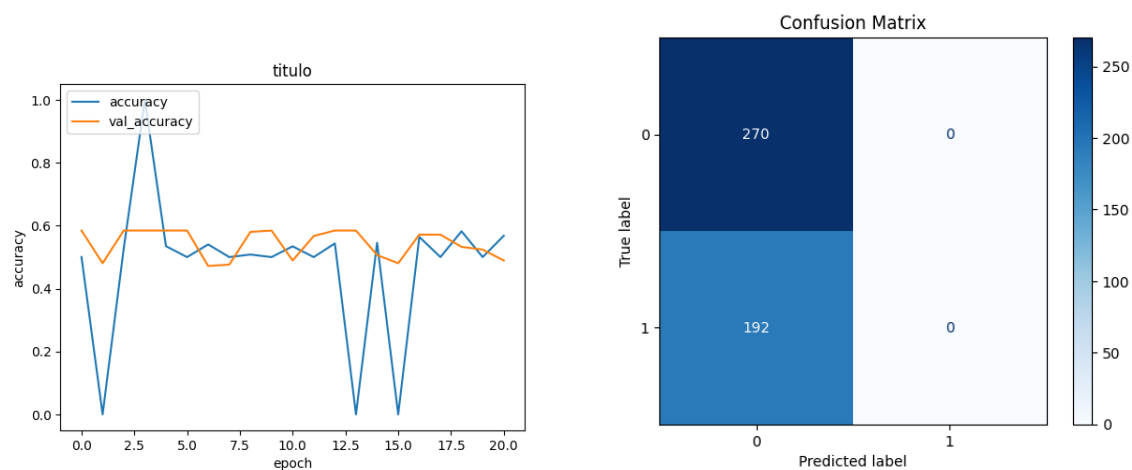


Figura 14 – Modelo ResNet com o seu treinamento e matriz de confusão.

Neste experimento é nítido que não foi feita uma aprendizagem robusta apresentando resultados menores comparada com os experimentos anteriores.

5.4 Experimento 4

5.4.1 Classe 0 e 2

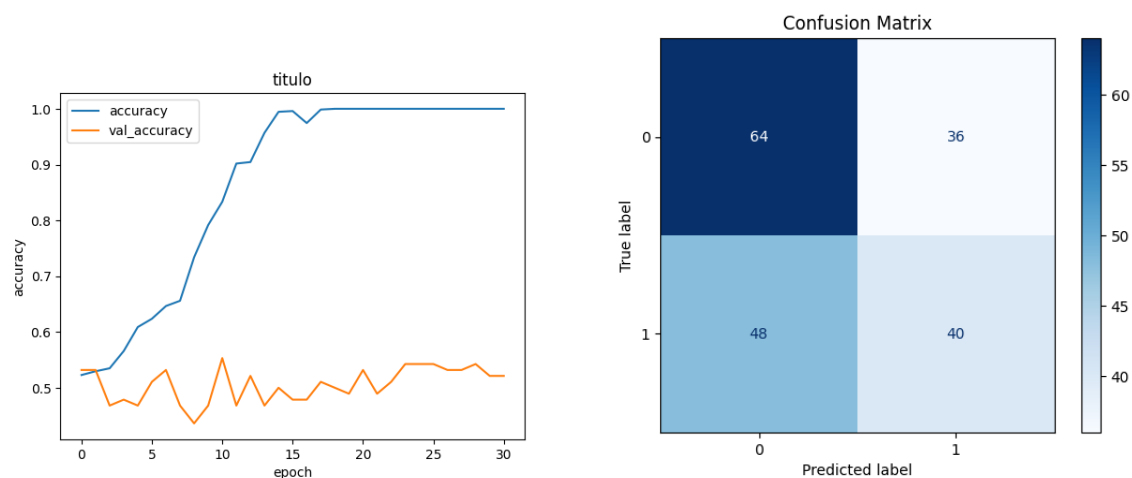


Figura 15 – Modelo ResNet com o seu treinamento e matriz de confusão.

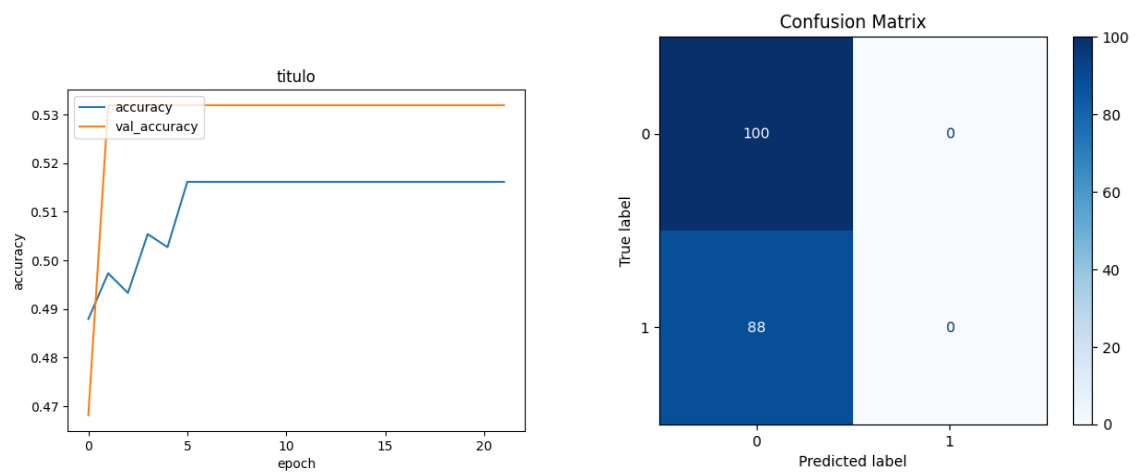


Figura 16 – Modelo VGG com o seu treinamento e matriz de confusão.

5.4.2 Classe 0 e 3

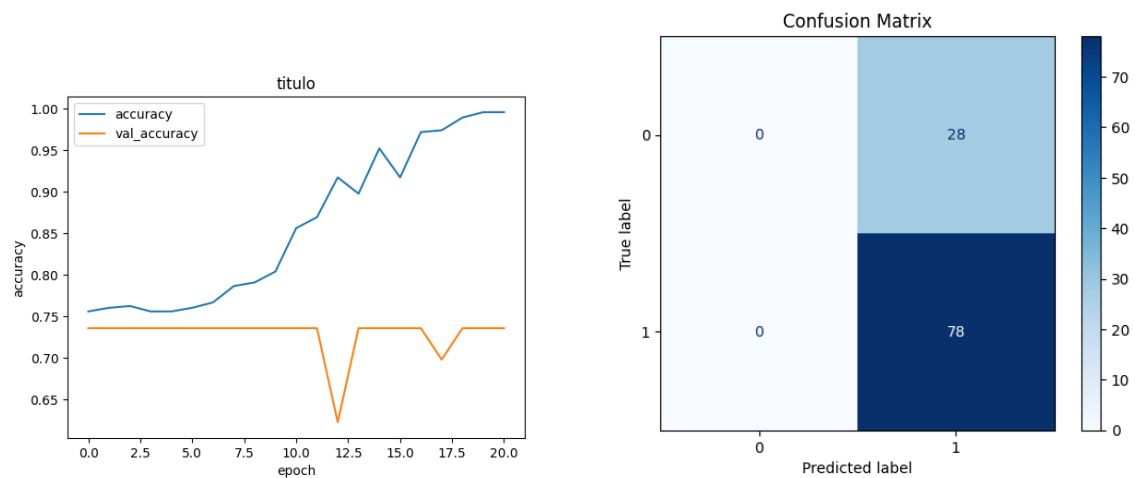


Figura 17 – Modelo ResNet com o seu treinamento e matriz de confusão.

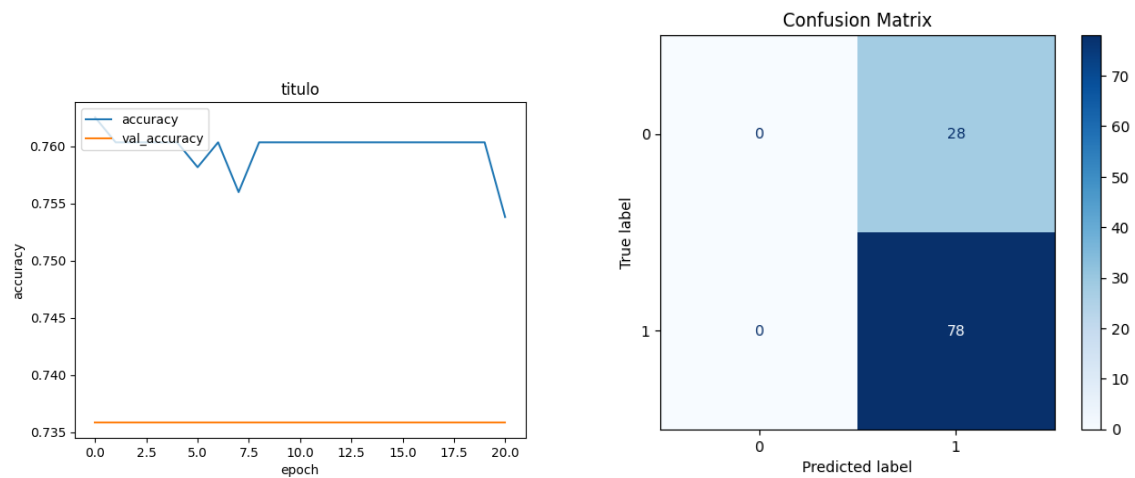


Figura 18 – Modelo VGG com o seu treinamento e matriz de confusão.

Neste experimento podemos observar que os modelos usando a arquitetura ResNet50 apresenta um resultado melhor quando comparada com a arquitetura VGG16, seja na sua curva de aprendizagem durante o seu treinamento quanto nos dados de validação.

5.5 Experimento 5

5.5.1 Classe 0 e 2

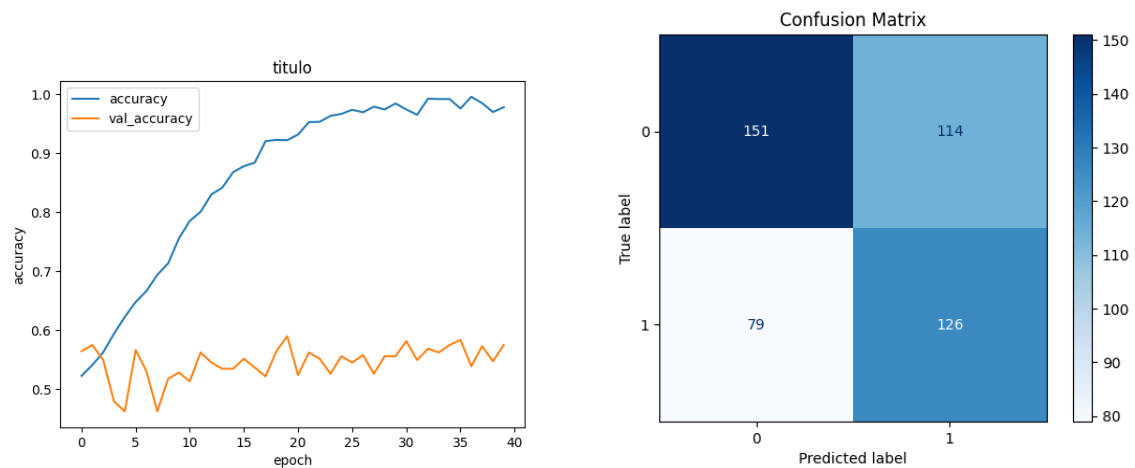


Figura 19 – Modelo ResNet com o seu treinamento e matriz de confusão.

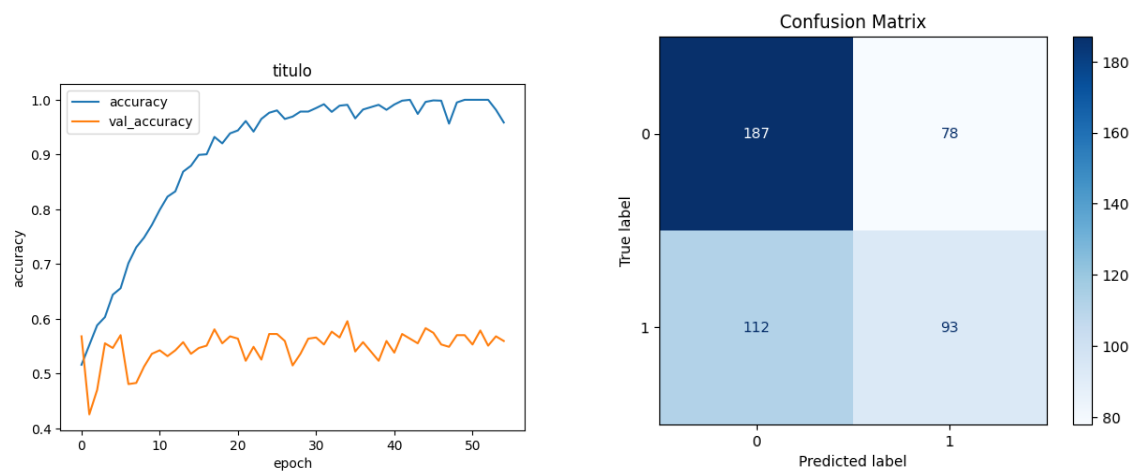


Figura 20 – Modelo ResNet com o seu treinamento e matriz de confusão.

5.5.2 Classe 2 e 3

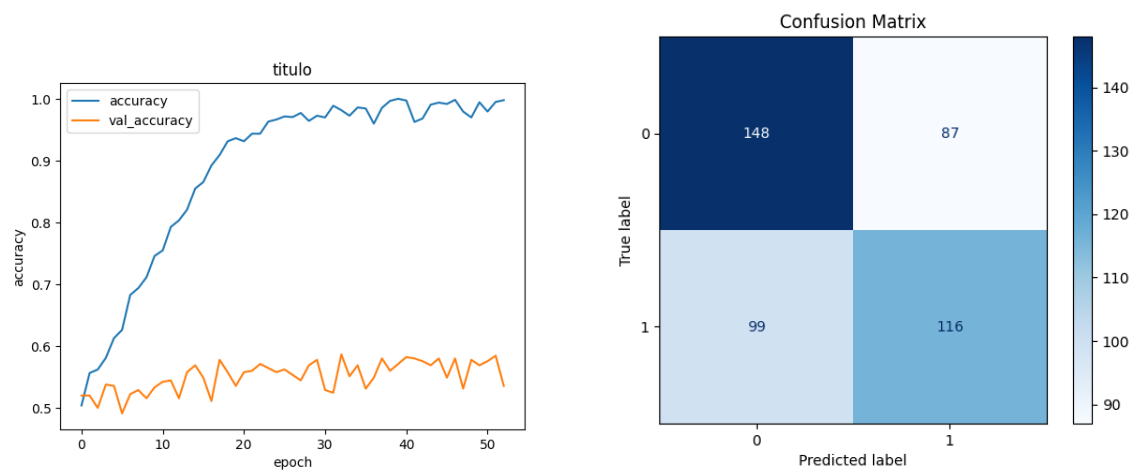


Figura 21 – Modelo ResNet com o seu treinamento e matriz de confusão.

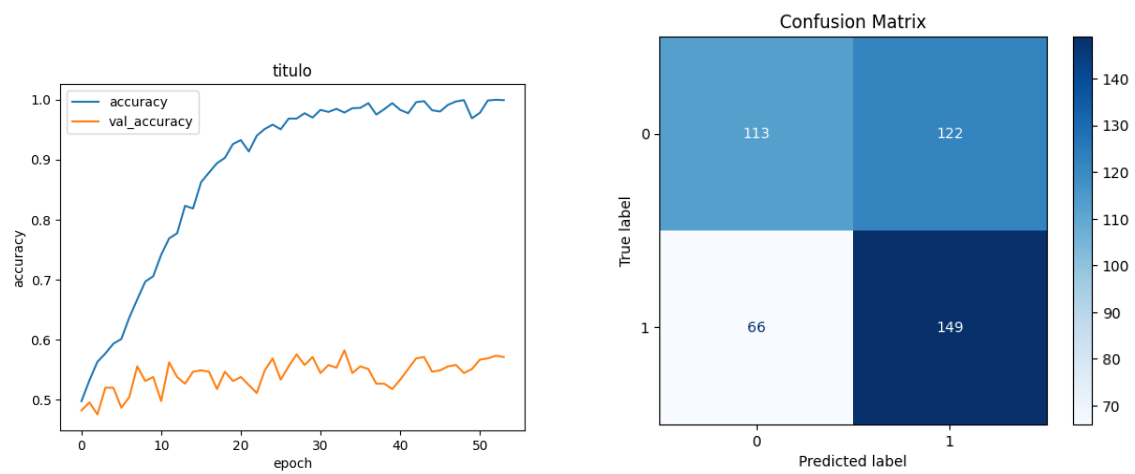


Figura 22 – Modelo ResNet com o seu treinamento e matriz de confusão.

5.5.3 Classe 0 e 3

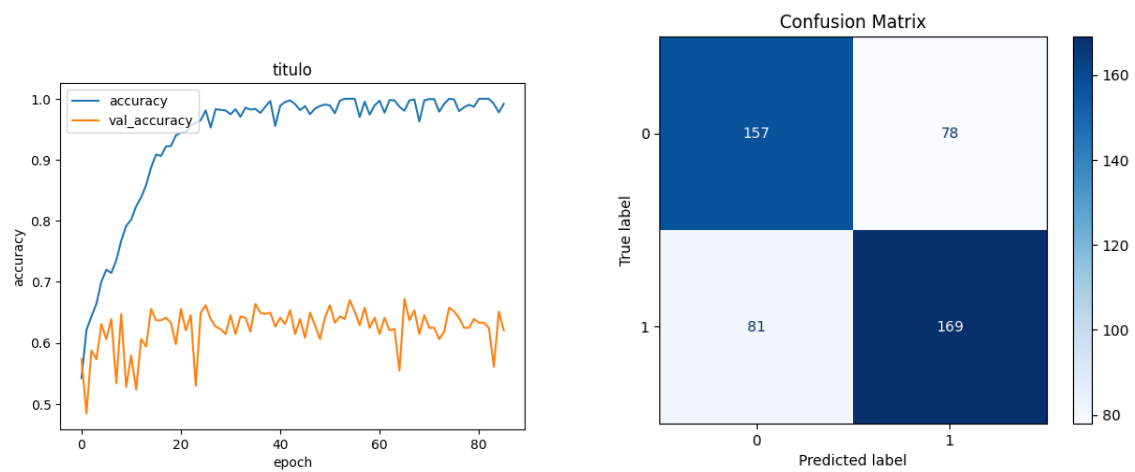


Figura 23 – Modelo ResNet com o seu treinamento e matriz de confusão.

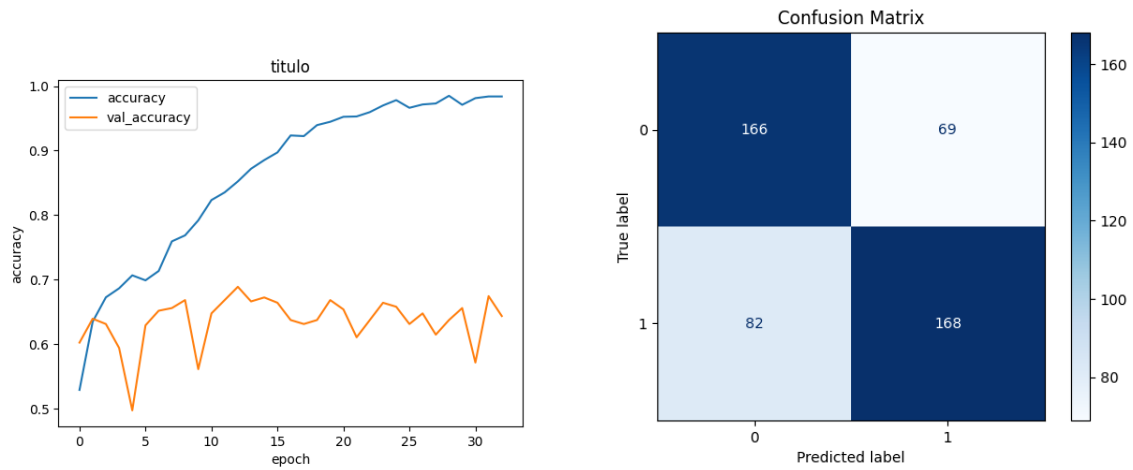


Figura 24 – Modelo ResNet com o seu treinamento e matriz de confusão.

Neste experimento fica mais fácil de notar a diferença entre os modelos de arquitetura ResNet50 e VGG16 quando comparado com os experimentos anteriores, tanto na sua curva de aprendizagem quanto na sua matriz de confusão.

Experimento	Saída Neural	Classes	Acurácia	Loss
E1	1	Todas	0.5817	1.18
E2	1	Todas	0.63	0.98
E3	1	Todas	0.5844	0.68
E4	1	0 e 2	0.55	0.89
E4	1	0 e 3	0.73	0.58
E5	1	0 e 2	0.5893	1.40
E5	1	0 e 3	0.67	1.97
E5	1	2 e 3	0.5866	1.83
E5	2	0 e 2	0.5957	1.94
E5	2	0 e 3	0.6886	0.80
E5	2	2 e 3	0.5822	1.78

Tabela 1 – Resnet50 Results

Experimento	Classes	Acurácia	Loss
E1	Todas	0.53	0.69
E2	Todas	0.64	1.77
E3	Todas	0.5844	0.69
E4	0 e 2	0.53	0.69
E4	0 e 3	0.7358	0.58
E5	0 e 2	0.56	0.9506
E5	0 e 3	0.51	1.085
E5	2 e 3	0.48	0.0

Tabela 2 – VGG16 Results

Analisando os experimentos, podemos observar que enquanto os experimentos 2 e 5 apresentam

a curva de treinamento estável enquanto que os experimentos 1, 3 e 4 apresentam uma curva instável durante o seu treinamento. Considerando as matrizes de confusão, os modelos que apresentaram uma instabilidade durante o seu treinamento houve uma maior classificação para somente uma classe o que nos diz que o modelo não aprendeu a diferenciar os tipos de Borrmann.

Além disso, podemos observar que o modelo conseguiu distinguir bem as classes quando usamos duas saídas na camada de saída da rede neural apesar da sua acurácia ser menor quando comparada com um dos modelos, o que indica que essa métrica neste problema não deve ser um fator decisivo de um modelo, mas sim como outras métricas também devem ser consideradas como a matriz de confusão, como por exemplo o experimento 4 com 5.

Por outro lado, os modelos apresentaram um desempenho maior quando treinada com as imagens do tipo câncer 0 e 3 tanto nas redes neurais com 1 saída quanto com 2 saídas, enquanto que para os tipos de câncer 0 e 2 ou 2 e 3 não apresentaram uma diferença notável de 10%.

Sobre as arquiteturas, podemos notar que entre a VGG16 e Resnet50, esta última apresenta um desempenho melhor devido a sua complexidade e profundidade do seu modelo quando comparada com a VGG16 que apresenta o seu oposto: uma arquitetura mais simples e menos profunda com menos blocos convolucionais.

6 Conclusão

Este estudo avaliou a aplicação de Redes Neurais Convolucionais para a classificação de câncer gástrico utilizando imagens capturadas por celular durante procedimentos cirúrgicos. Os experimentos demonstraram que, para este cenário de imagens não padronizadas ("in-the-wild"), a arquitetura ResNet50 apresentou desempenho superior à VGG16. A maior profundidade e complexidade da ResNet50 permitiram uma extração de características mais robusta, resultando em uma melhor generalização frente à complexidade das lesões.

Os resultados quantitativos apontaram que a estratégia de simplificação do problema para duas saídas (classificação binária ou pares de classes, especificamente classes 0 e 3) obteve os melhores índices, alcançando uma acurácia de aproximadamente 68%. Em contrapartida, notamos que a classificação multiclasse completa enfrentou dificuldades de convergência e estabilidade, evidenciadas pelas curvas de aprendizado instáveis em alguns experimentos, o que sugere que a acurácia isolada não deve ser a única métrica de avaliação, sendo necessário analisar também a matriz de confusão.

Em suma, o uso de Deep Learning para "segunda opinião" em tempo real é viável, mas dependente da qualidade do pré-processamento e da escolha da arquitetura. Para trabalhos futuros, recomendamos prioritariamente a implementação de técnicas de segmentação e remoção de fundo para que o modelo foque exclusivamente na lesão tecidual. Além disso, sugere-se a continuidade dos estudos utilizando a ResNet50 como baseline e a exploração de técnicas avançadas de balanceamento de dados para mitigar a disparidade entre as classes de Borrmann.

Referências

BRASIL. Ministério da Saúde. Instituto Nacional de Câncer. *Câncer de estômago*. 2025. Acesso em: mar. 2025. Disponível em: <<https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/estomago>>. Acesso em: 13 mar. 2025.

IIZUKA, O. et al. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Scientific Reports*, v. 10, n. 1504, 2020.

KHAN, M. A. et al. Multiclass stomach diseases classification using deep learning features optimization. *University of Leicester (Journal contribution)*, 2021. Disponível em: <<https://hdl.handle.net/2381/14423369.v1>>.

KHAN, M. A. et al. Stomachnet: Optimal deep learning features fusion for stomach abnormalities classification. *IEEE Access*, v. 8, p. 197969–197981, 2020.

KUNTZ, S. et al. Gastrointestinal cancer classification and prognostication from histology using deep learning: Systematic review. *European Journal of Cancer*, v. 155, p. 200–215, 2021.

SATO, M. *O uso de imagens de satélite na agricultura: conheça os benefícios*. Agroblog, 2020. Disponível em: <<https://agrosmart.com.br/blog/imagens-de-satelite-agricultura-beneficios>>. Acesso em: 07 de abril de 2024.

UNIVERSIDADE PRESBITERIANA MACKENZIE. *Guia do TCC: Orientações gerais para a elaboração do trabalho de conclusão dos cursos de graduação*. São Paulo: UPM, 2021.

ZHAO, Y. et al. Identification of gastric cancer with convolutional neural networks: a systematic review. *Multimedia Tools and Applications*, v. 81, p. 11717–11736, 2022.