



NÚCLEO DE VISÃO COMPUTACIONAL E ENGENHARIA

RELATÓRIO DE TESTE TÉCNICO

IMPLEMENTAÇÃO E ANÁLISE DE REDE NEURAL CONVOLUCIONAL

**AVALIAÇÃO DA EFICÁCIA E EFICIÊNCIA DA REDE NA CLASSIFICAÇÃO DE
CÉLULAS**

Relatório de teste técnico para o núcleo de visão computacional e engenharia.

Marcelo de Araújo

Junho / 2024

RESUMO

Este estudo apresenta um processo detalhado para a classificação de imagens de células utilizando redes neurais convolucionais, com foco na distinção entre células normais e anormais em imagens digitais de exames de Papanicolaou. O objetivo é avaliar a eficácia e eficiência do modelo proposto na tarefa de classificação de imagens celulares, diferenciando células normais de anormais com base em imagens digitais de microscopia, provenientes de exames de Papanicolaou, visando aprimorar a precisão dos diagnósticos médicos e reduzir a carga de trabalho dos citologistas. O principal propósito é avaliar a eficácia e eficiência do modelo proposto na tarefa de classificação de tais imagens.

As imagens de células provenientes de exames de Papanicolaou foram categorizadas em duas classes principais: **Normal** e **Anormal**, sendo as imagens organizadas meteticulosamente em um **DataFrame** contendo as colunas **filename** e **class**, facilitando assim a referência e manipulação durante o processamento. Para garantir a integridade dos dados, a detecção e eliminação de arquivos duplicados foram realizadas de forma eficiente utilizando **hashes MD5**. No processo de carregamento e pré-processamento das imagens, cada imagem foi redimensionada e posteriormente normalizada. Essa etapa foi crucial para preparar as imagens de forma adequada para a entrada no modelo de redes neurais convolucionais. Além disso, os rótulos associados a cada imagem foram convertidos para o formato one-hot encoding, uma técnica essencial que simplifica o treinamento e a classificação. O modelo foi então construído utilizando a biblioteca TensorFlow em conjunto com Keras, seguindo uma arquitetura sequencial. A configuração incluiu várias camadas convolucionais para capturar características significativas das imagens, camadas de max pooling para redução de dimensionalidade, e camadas densamente conectadas para processamento final. A função de ativação ReLU6 foi escolhida para promover a não linearidade na extração de características, enquanto camadas de dropout foram inseridas para mitigar o overfitting durante o treinamento. Para otimização do modelo, a função de perda de entropia cruzada categórica foi utilizada, e o algoritmo Adam foi empregado como otimizador. O modelo foi treinado extensivamente com os dados de treino e validado com os dados de teste para garantir robustez e precisão na classificação das imagens de células. Essa abordagem integrada e meteticulosa demonstra o potencial das redes neurais convolucionais na análise de imagens médicas, especificamente no contexto do diagnóstico precoce de anomalias celulares identificadas em exames de Papanicolaou.

1 INTRODUÇÃO

Segundo (SOLOMON *et al.*, 2002), "o exame de Papanicolaou, também conhecido como citologia cervical ou esfregaço de Papanicolaou, é um teste de rastreamento amplamente utilizado para a detecção precoce de câncer cervical e lesões pré-cancerosas, sendo essencial para a saúde pública, pois permite a identificação de alterações celulares que podem evoluir

para o câncer, possibilitando intervenções precoces e reduzindo significativamente a mortalidade associada ao câncer cervical". Dentro da área de visão computacional, as redes neurais convolucionais são uma ferramenta poderosa para a extração e classificação de atributos em imagens, incluindo as obtidas no exame de Papanicolaou, sendo sua aplicação extremamente útil no auxílio ao diagnóstico, uma vez que essas redes são capazes de analisar grandes volumes de imagens de forma rápida e precisa. As redes neurais convolucionais podem extrair automaticamente características relevantes das imagens de células cervicais, como anomalias na textura, forma e padrão celular, que são indicativos de possíveis lesões pré-cancerosas ou câncer cervical.

No contexto das redes neurais convolucionais, a sua utilização pode melhorar a precisão dos diagnósticos, reduzindo a taxa de falsos positivos e falsos negativos, e permitindo uma triagem mais eficiente e consistente. Além disso, a automação do processo de análise de imagens com redes neurais convolucionais pode aliviar a carga de trabalho dos citologistas, permitindo que eles se concentrem em casos mais complexos e que requerem uma avaliação mais detalhada. Esse processo de análise é realizado automaticamente através das camadas convolucionais das, onde filtros são aplicados às imagens para extrair características de maneira hierárquica e progressivamente mais complexa. O processo é realizado automaticamente através das camadas convolucionais, onde filtros são aplicados à imagem para extrair essas características de maneira hierárquica e progressivamente mais complexa. (FORSYTH; PONCE, 2003) destacam a importância da extração de atributos em seu livro seminal sobre visão computacional, descrevendo-a como "o processo de transformar uma imagem em uma representação mais compacta que é mais fácil de analisar". As redes neurais convolucionais realizam essa transformação de maneira eficiente, criando uma representação compacta composta por atributos relevantes, o que é essencial para reduzir a complexidade computacional e facilitar a interpretação automática de imagens.

As camadas densamente conectadas que recebem como entrada a representação compacta extraída pelas camadas convolucionais. Esses classificadores atribuem rótulos ou categorias a objetos, ou regiões de interesse na imagem, analisando as características extraídas para determinar a classe à qual cada objeto pertence. (SZELISKI, 2011) afirma que "os classificadores são uma ferramenta essencial em visão computacional, permitindo a automação de tarefas como reconhecimento de objetos e detecção de padrões em imagens." No caso das redes neurais convolucionais, esses classificadores são integrados no final da rede neural, tornando-a uma solução completa e eficiente para a análise de grandes conjuntos de dados visuais, automatizando tarefas de reconhecimento e categorização.

2 METODOLOGIA

A metodologia adotada fundamenta-se na utilização e aplicação de algoritmos componentizados de uma arquitetura de rede neural convolucional. A construção do estudo foi organizada em múltiplas fases metodológicas, uma prática consolidada e bem fundamentada na

pesquisa científica, especialmente nas áreas de aprendizado de máquina e análise de dados. Esta abordagem sistemática assegura uma organização clara e sequencial do processo de pesquisa, garantindo que cada fase seja meticulosamente planejada e executada. Conforme discutido por (CRESWELL, 2014), uma estrutura metodológica bem definida permite maior rigor e reprodutibilidade na pesquisa. Adicionalmente, (YIN, 2018) enfatiza que a segmentação do estudo em etapas distintas facilita a análise e a interpretação dos resultados. Além disso, tal estrutura facilita a replicação do estudo e a verificação dos resultados por parte de outros pesquisadores, contribuindo para a robustez e a transparência científica.

2.1 Pré análise dos dados

O estudo iniciou-se com o carregamento e pré-processamento dos dados envolvendo a organização, análise e validação de um conjunto de dados de imagens médicas para classificação, abrangendo desde a leitura dos arquivos de imagem até a verificação de duplicatas e integridade dos dados.

Primeiramente, as imagens fornecidas foram descompactadas e o diretório principal que contém os dados das imagens foi definido. As classes de interesse foram estabelecidas como **Normal** e **Anormal**, com exemplos específicos identificados para cada categoria. Em seguida, foi realizado um percurso recursivo pelo diretório de dados para localizar arquivos de imagens com extensões específicas (.png, .jpg, .jpeg, .tif). Cada imagem foi classificada como **Normal** ou **Anormal** com base na pasta onde estava localizada. Após coletar os nomes dos arquivos e suas respectivas classes, os dados foram organizados em um **DataFrame** e exportados para um arquivo denominado `cell_images_classification.csv`. Após a geração do arquivo onde consta o nome da imagem e seu respectivo rótulo, verifico se cada imagem listada no arquivo possuía uma classe associada, caso contrário, uma mensagem foi exibida indicando a falta de associação. Para assegurar a integridade dos dados e evitar duplicatas, cada imagem foi processada para calcular seu **hash MD5**, sendo que duplicatas foram identificadas com base neste hash.

Essa etapa é extremamente essencial para o pré-processamento de dados de imagens médicas, desde a classificação inicial até a verificação de integridade e duplicatas tal abordagem é fundamental para manipulação, análise e processamento dos dados, demonstrando um processo estruturado para garantir a qualidade e a consistência dos dados utilizados em tarefas de aprendizado de máquina.

2.2 Preparação do conjunto de dados

O objetivo nessa etapa é carregar, pré-processar e preparar o conjunto de dados de imagens médicas para classificação utilizando TensorFlow e bibliotecas relacionadas. Neste trabalho, inicia-se carregando a tabela que contém os nomes dos arquivos de imagem e suas

respectivas classes. Este arquivo é essencial para associar cada imagem ao seu rótulo correto durante o treinamento.

Logo em seguida, cada classe de imagem é mapeada para um número inteiro único, o que facilita o treinamento de modelos de aprendizado de máquina. Esse mapeamento é criado a partir das classes únicas encontradas nessa tabela. Sequencialmente, percorro o diretório principal e suas subpastas em busca de arquivos de imagem no formato **.tif**, onde cada imagem é carregada e redimensionada para um tamanho padrão e convertida em um array NumPy. Por fim, carregam-se todas as imagens e associam-se corretamente suas classes, os dados são convertidos para arrays NumPy, facilitando a manipulação e uso subsequente nos modelos de aprendizado de máquina.

Esta etapa é um processo sistemático para preparar dados das imagens médicas para tarefas de classificação. Cada etapa, desde o carregamento das imagens até a preparação dos dados, é cuidadosamente implementada para garantir a consistência e a eficácia do modelo final, sendo essencial para a construção de modelos robustos e precisos.

2.3 Treino e Teste

Primeiramente, destaco que a divisão entre treino e teste visa avaliar fundamentalmente a capacidade de generalização do modelo, isto é, sua habilidade de realizar previsões precisas em dados não observados anteriormente. Nesse sentido, é crucial seguir as orientações de (GONZALEZ; WOODS; EDDINS, 2008) para uma avaliação apropriada dos algoritmos de visão computacional, que enfatizam a necessidade de dividir os dados em conjuntos de treinamento e teste, assegurando a capacidade do modelo de generalizar para novos exemplos. Para isso, adotei a prática comum de particionar o conjunto de dados em dois subconjuntos mutuamente exclusivos: o conjunto de treinamento, utilizado para treinar o modelo, e o conjunto de teste, empregado para avaliar o desempenho do modelo em dados não observados.

No contexto deste estudo, os dados foram divididos em conjuntos de treino e teste utilizando a função **train_test_split** do scikit-learn. Essa divisão é essencial para evitar que o modelo se ajuste demais aos dados de treinamento, permitindo uma avaliação confiável do desempenho em dados não vistos. Optou-se por alocar 80% das imagens para treinamento e 20% para teste. Além disso, o parâmetro **random_state = 42** garante a reprodutibilidade dos resultados. Isso significa que, em diferentes execuções do código, a mesma divisão entre os conjuntos de treino e teste será obtida, facilitando a comparação consistente entre diferentes modelos e abordagens.

Finalmente, as intensidades dos pixels das imagens são normalizadas para o intervalo $[0, 1]$, garantindo que as características das imagens tenham uma escala comum e as labels de classe são convertidas para o formato one-hot encoding usando a função **to_categorical**, tais estratégias são benéficas para o treinamento eficaz do modelo e classificação multiclasse.

2.4 Modelo

Conforme mencionado por (GEETHA; SENDHILKUMAR, 2023), a construção de um modelo é uma arte delicada, onde cada variável e sua interação devem ser consideradas com cuidado, como peças em um quebra-cabeça, para formar um todo coeso e preditivo. Assim, a análise comparativa desses modelos em diferentes cenários se torna crucial para determinar a estratégia mais eficaz para a aplicação em questão.

O modelo de rede neural é configurado utilizando a API Sequential do TensorFlow Keras, conhecida por permitir a construção de modelos camada a camada de maneira sequencial. Este tipo de abordagem é especialmente útil para a criação de arquiteturas robustas e eficazes, como a definida neste contexto. A camada de entrada é definida por **Input()**, onde shape é definido como 50 representando a dimensionalidade das imagens de entrada e 3 denotando a quantidade de canais de cor.

O modelo inclui um total de cinco camadas convolucionais, seguidas por camadas de **max-pooling** para redução de dimensionalidade, camadas de **dropout** para regularização e camadas totalmente conectadas para a extração de características das imagens e classificação subsequente. A função de ativação **relu6** é especificada para todas as camadas convolucionais e totalmente conectadas. Esta função é uma variante da função de ativação ReLU (Rectified Linear Unit), limitada superiormente a 6 para mitigar problemas de explosão de gradientes, frequentemente observados em redes profundas.

Cada camada **Conv2D** é configurada com um determinado número de filtros e um tamanho de kernel específico, aplicando a função de ativação para introduzir não-linearidades na rede. As camadas **MaxPooling2D** são empregadas após as camadas convolucionais para realizar subamostragem, preservando as características mais proeminentes das imagens. As camadas de **dropout** são inseridas com taxas de 25% e 50%, respectivamente, após as camadas convolucionais e densas. Essas camadas ajudam a prevenir o overfitting ao desativar aleatoriamente uma fração das unidades de saída durante o treinamento.

A camada **Flatten()** é utilizada para transformar a saída das camadas convolucionais em um vetor unidimensional, preparando os dados para serem processados por camadas densas subsequentes. A camada **Dense()** representa a camada de saída da rede, com número de neurônios correspondentes ao número de classes no problema de classificação e a função de ativação softmax gerando as probabilidades para cada classe, garantindo que a soma das saídas seja igual a 1 e facilitando a interpretação da saída como probabilidades de pertencimento a cada classe.

3 RESULTADOS E DISCUSSÕES

Nesta seção, apresento os resultados e discussões sobre a utilização do modelo proposto na classificação de imagens digitais de exames de Papanicolaou. A base de dados consiste em imagens de 100 x 100 píxeis que foram redimensionadas para 50 x 50 píxeis. O objetivo do

modelo é distinguir entre células normais e anormais em imagens histopatológicas, explorando a eficácia da arquitetura desenvolvida, avaliando a precisão do modelo e discutir a aplicabilidade dos resultados para diagnósticos.

Após o treinamento, foi avaliado o desempenho do modelo nos conjuntos de treinamento e teste. Os resultados foram muito encorajadores: o modelo alcançou uma acurácia de aproximadamente 86.62% nos dados de teste e 88.86% nos dados de treinamento. A perda média foi de 31.04% nos dados de teste e 26.85% nos dados de treinamento, indicando uma boa capacidade de generalização.

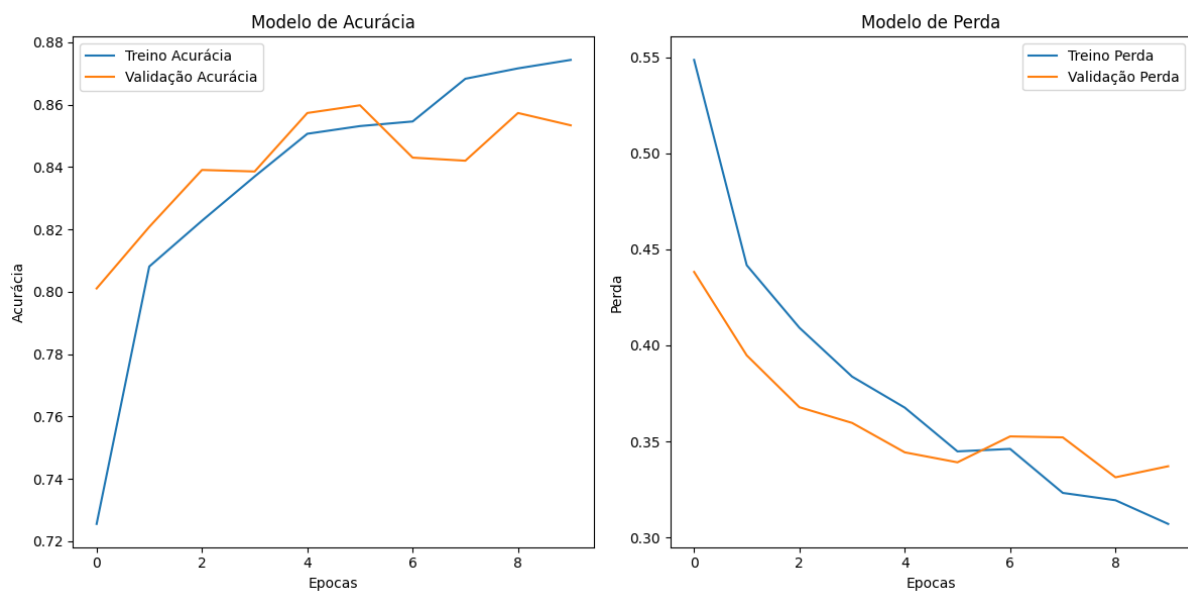


Figura 1 – Modelo de Acurácia e Perda.

Além da acurácia e da perda, foi calculadas a precisão e a revocação do modelo. A precisão foi medida em 86.63%, o que indica a proporção de instâncias positivas (células anormais) corretamente classificadas pelo modelo. A revocação também foi de 86.62%, representando a capacidade do modelo de identificar todas as instâncias positivas.



Figura 2 – Revocação.

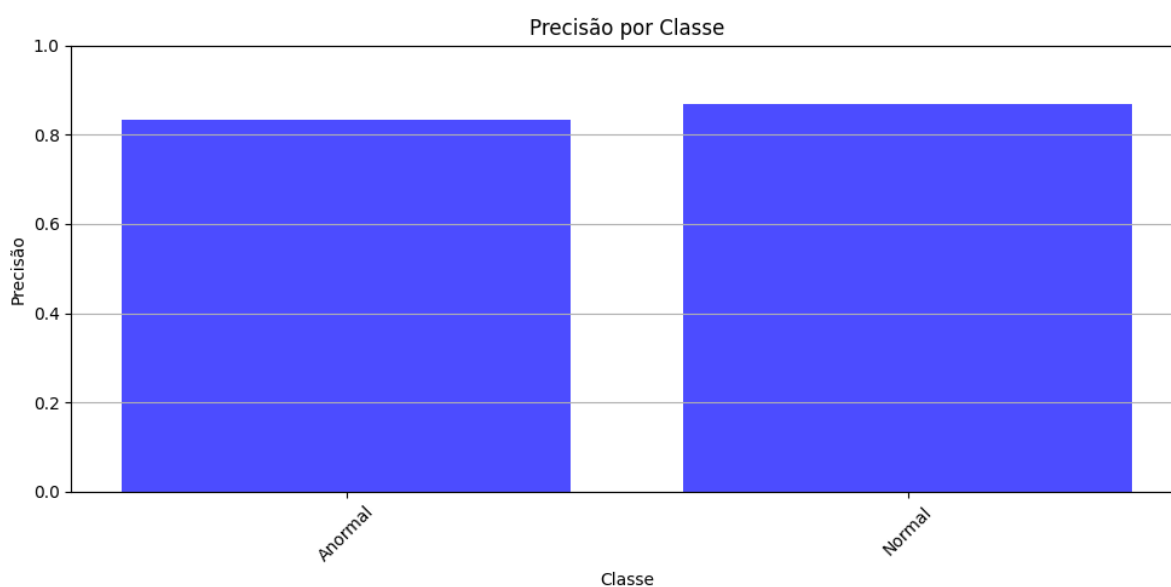


Figura 3 – Precisão.

A matriz de confusão é uma ferramenta fundamental para a avaliação do desempenho de modelos de classificação, oferecendo uma visão detalhada sobre a eficácia do modelo em termos de suas previsões. No contexto deste estudo, onde foi desenvolvido um modelo de rede neural convolucional para classificar imagens de células cervicais como normais ou anormais, a matriz de confusão é apresentada da seguinte forma:

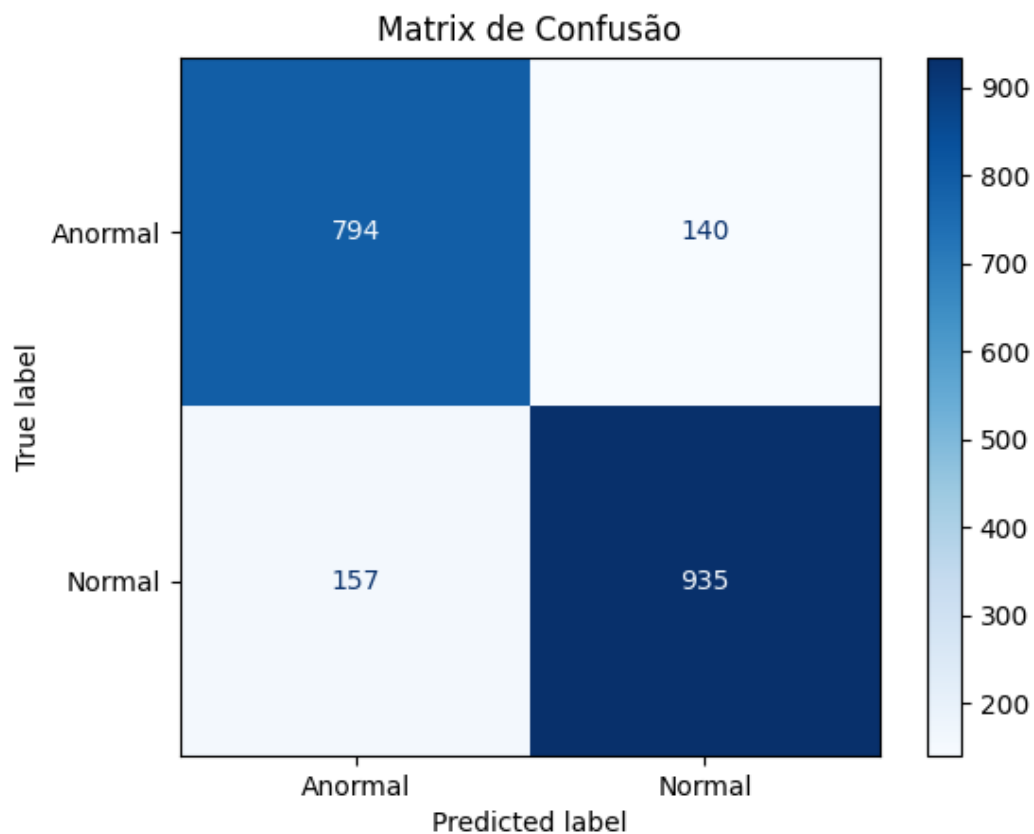


Figura 4 – Matriz de Confusão.

Na matriz, podemos observar que o modelo obteve um bom desempenho geral, no entanto, o modelo teve algumas dificuldades em classificar corretamente as instâncias anormais. A matriz de confusão indica que o modelo é altamente eficaz na classificação das células cervicais, com alta precisão e revocação. As métricas calculadas sugerem que o modelo tem um bom equilíbrio entre a capacidade de identificar corretamente as amostras positivas e negativas.

A especificidade alta sugere uma forte habilidade do modelo em identificar corretamente as amostras negativas, crucial para minimizar falsos positivos em diagnósticos clínicos. O F1-Score consistente indica um equilíbrio robusto entre precisão e revocação, enquanto a variabilidade na especificidade pode sugerir áreas para otimização adicional do modelo.

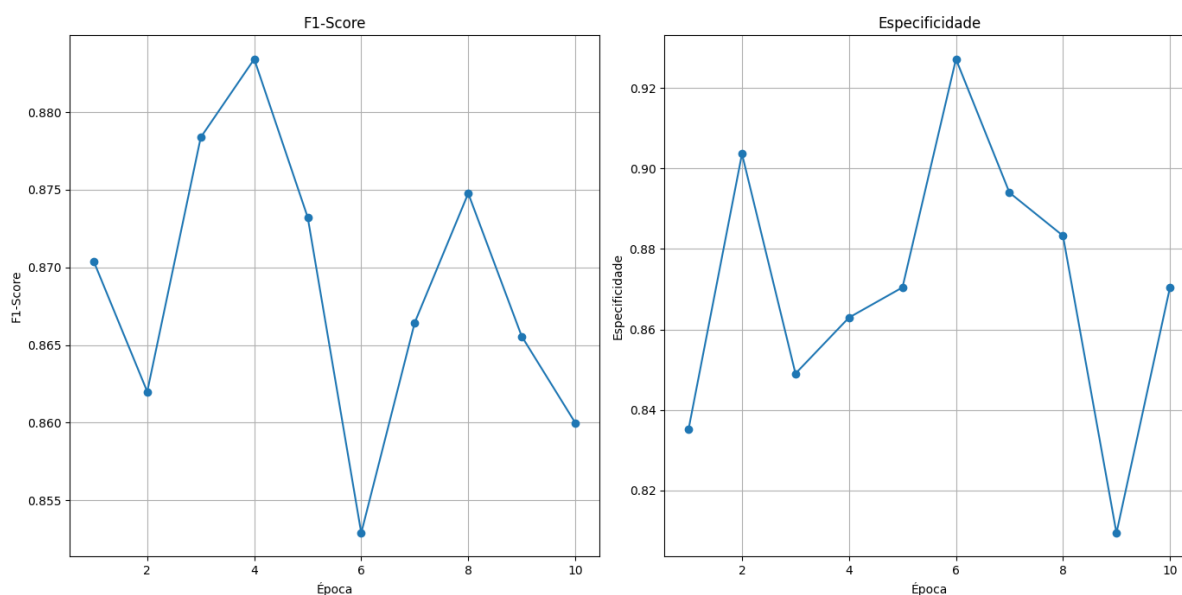


Figura 5 – F1-SCORE e Especificidade.

4 CONCLUSÕES

A conclusão deste estudo evidencia a eficácia e a robustez do modelo da rede neural convolucional desenvolvida na classificação de imagens de células citológicas, apresentando alto desempenho em termos de acurácia e outras métricas-chave. Estes resultados são promissores para a aplicação clínica na detecção precoce de anormalidades celulares. Futuros trabalhos podem explorar arquiteturas de modelos mais complexas e conjuntos de dados maiores para melhorar ainda mais a precisão e a confiabilidade do sistema. O presente estudo investigou a eficácia do modelo de rede neural convolucional com a função de ativação ReLU6 para a classificação de imagens de células citológicas, diferenciando entre células normais e anormais. A análise dos resultados demonstrou um desempenho robusto do modelo, com métricas de avaliação que indicam uma alta capacidade de generalização e precisão.

Os indicadores mostrados no tópico acima afirmam que o modelo possui uma alta taxa de acertos tanto para a classe normal quanto para a classe anormal, com um número relativamente baixo de erros de classificação. Os resultados deste estudo são encorajadores e demonstram o potencial das redes neurais convolucionais com ReLU6 para a classificação automática de imagens citológicas. As implicações clínicas são significativas, pois um sistema automatizado eficiente pode contribuir para a detecção precoce de anormalidades celulares, melhorando as taxas de sucesso no tratamento de câncer cervical. Para trabalhos futuros, recomenda-se explorar arquiteturas de redes neurais mais complexas e conjuntos de dados maiores para aumentar ainda mais a precisão e a robustez do modelo. Além disso, a implementação de técnicas de data augmentation e ajuste de hiper parâmetros pode ajudar a melhorar a variabilidade, a especificidade e o desempenho geral do sistema.

5 REFERÊNCIAS

CRESWELL, J. W. **Research design: Qualitative, quantitative, and mixed methods approaches**. 4th. ed. [S.l.]: Sage Publications, 2014.

FORSYTH, D. A.; PONCE, J. **Computer Vision: A Modern Approach**. [S.l.]: Prentice Hall, 2003.

GEETHA, T.; SENDHILKUMAR, S. **Machine Learning: Concepts, Techniques and Applications**. Chapman & Hall/CRC Press, 2023. ISBN 9781003290100. Disponível em: <<https://books.google.com.br/books?id=hnCezwEACAAJ>>.

GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. **Digital image processing using MATLAB**. [S.l.]: Gatesmark Publishing Knoxville, 2008. ISBN 978-0982085400.

SOLOMON, D. *et al.* The 2001 bethesda system: Terminology for reporting results of cervical cytology. **JAMA : the journal of the American Medical Association**, v. 287, p. 2114–9, 05 2002.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. [S.l.]: Springer Science & Business Media, 2011.

YIN, R. K. **Case study research and applications: Design and methods**. 6th. ed. [S.l.]: Sage Publications, 2018.