

TESTE DE MODELOS DE INTELIGÊNCIA ARTIFICIAL PARA ANÁLISE DE EXAMES DE PUNÇÃO ASPIRATIVA COM AGULHA FINA.

Felipe Bortoluzzi Dela Vedova¹, André Faria Ruaro²

Resumo: Este trabalho tem como foco a análise de exames de Punção Aspirativa por Agulha Fina (PAAF) da tireoide, utilizando técnicas de inteligência artificial para auxiliar na identificação de nódulos malignos. A principal motivação da pesquisa está na necessidade de métodos mais eficientes e confiáveis para auxiliar no diagnóstico citopatológico, uma vez que esse processo depende da interpretação humana, podendo estar sujeito a erros. Foram utilizados modelos de aprendizado de máquina para classificar imagens citológicas da tireoide, com o objetivo de avaliar seu desempenho diante de um conjunto limitado de dados. A base de dados utilizada foi submetida a técnicas de tratamento e aumento de dados, como equalização de contraste (CLAHE), rotações, zoom e recortes aleatórios, a fim de aumentar a qualidade das informações fornecidas. Três abordagens foram aplicadas: uma rede neural convolucional (CNN), uma Máquina de Vetores de Suporte (SVM) e uma SVM com *MobileNetV2* como ferramenta para extração de padrões visuais. Os resultados mostraram que mesmo com uma base reduzida, os modelos foram capazes de realizar a tarefa de classificação com desempenho satisfatório. A CNN apresentou maior precisão e taxa de acerto geral, enquanto o SVM com *MobileNetV2* destacou-se por sua alta sensibilidade, sendo mais eficaz na identificação correta de casos malignos. A pesquisa demonstra o potencial do uso de inteligência artificial como ferramenta de apoio em análises citológicas, contribuindo para a redução de falhas humanas e a melhoria do processo diagnóstico.

Palavras-chave: inteligência artificial; câncer de tireoide; citologia; aprendizado de máquina; redes neurais convolucionais; máquina de vetor de suporte

¹Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), felipe_vedova@unesc.net

²Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), andre.ruaro@unesc.net

ABSTRACT: This work focuses on the analysis of Fine Needle Aspiration (FNA) exams of the thyroid, using artificial intelligence techniques to assist in the identification of malignant nodules. The main motivation of the research is the need for more efficient and reliable methods to assist in the cytopathological diagnosis, since this process depends on human interpretation and may be subject to errors. Machine learning models were used to classify cytological images of the thyroid, with the aim of evaluating their performance in the face of a limited data set. The database used was subjected to data processing and augmentation techniques, such as contrast equalization (CLAHE), rotations, zoom and random cropping, in order to increase the quality of the information provided. Three approaches were applied: a convolutional neural network (CNN), a Support Vector Machine (SVM) and an SVM with *MobileNetV2* as a tool for extracting visual patterns. The results showed that even with a reduced base, the models were able to perform the classification task with satisfactory performance. CNN showed higher accuracy and overall hit rate, while SVM with *MobileNetV2* stood out for its high sensitivity, being more effective in correctly identifying malignant cases. The research demonstrates the potential of using artificial intelligence as a support tool in cytological analyses, contributing to the reduction of human errors and the improvement of the diagnostic process.

Keywords: artificial intelligence; thyroid cancer; cytology; machine learning; convolutional neural networks; support vector machine

1 INTRODUÇÃO

O câncer é uma das principais causas de morte no mundo e atinge milhões de pessoas todos os anos (Santos et al., 2016). Segundo o Instituto Nacional de Câncer José Alencar Gomes da Silva (2020), ele é caracterizado pela divisão celular descontrolada e pela habilidade de invasão a outros tecidos do organismo.

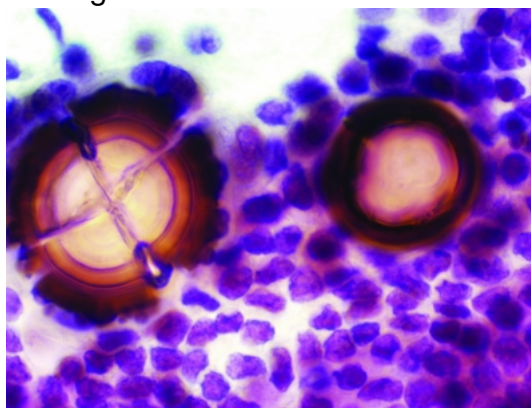
As células do corpo humano possuem um padrão de vida contínuo e natural, onde, em normalidade, as mesmas crescem, multiplicam-se e morrem de maneira ordenada, o que leva a uma proliferação celular para atender as necessidades do corpo sem apresentar algum tipo de malignidade. Diferente do crescimento normal, as células com características cancerosas continuam crescendo de forma incontrolada ao invés de morrerem, levando a novas células anormais (Instituto Nacional de Câncer José Alencar Gomes da Silva, 2020).

Entre os diversos tipos registrados, o câncer de tireoide vem se destacando pelo crescimento significativo na incidência de casos (Santos et al., 2016). A tireoide é uma glândula endócrina, localizada na parte anterior do pescoço e sua principal função é produzir e secretar hormônios que regulam o metabolismo do corpo, impactando diversas funções (Maciel, 2001).

De acordo o estudo de Santos et al. (2016), os casos desse tipo de câncer vêm aumentando em 1% ao ano, gerando a atenção dos especialistas, uma vez que, a tireoide pode sofrer alterações anatômicas e fisiológicas e podendo ser decorrente da geração de um carcinoma, havendo a possibilidade de ser um do tipo papilífero, medular, folicular ou anaplásico e, embora venha-se a ter um aumento nos casos desse tipo de câncer, os números de óbitos tem diminuído em decorrência dos diagnósticos precoces.

A fim de se adquirir um diagnóstico e identificar se o respectivo tumor é maligno ou benigno, o paciente passará por um exame de punção aspirativa por agulha fina, ou PAAF, que consiste na retirada de uma pequena quantidade de células para análise laboratorial, bem como a verificação da quantidade de achado, o formato e o núcleo celular (Yokozawa, 1998). A Figura 1 exhibe células de tireoide.

Figura 1 - Células da Tireoide.



Fonte: (Cytopathology, 2024).

O projeto proposto não ressalta apenas a importância no avanço da área da saúde, mas também no âmbito computacional, ao desenvolver protótipos de algoritmos de inteligência artificial (IA) com o intuito de auxiliar nas análises de exames citológicos de tireoide a fim de reduzir o erro humano e aumentar a confiabilidade, produtividade e eficiência.

A análise das amostras do exame apresenta desafios devido à

complexidade dos dados, motivando a exploração de técnicas de algoritmos de treinamento em uma IA focada em aprendizado supervisionado. Busca-se assim, o algoritmo mais eficiente, com a menor taxa de erro e otimização dos parâmetros de busca e classificação, enfrentando os obstáculos da complexidade dos dados.

Kezlarian e Lin (2021) examinaram os avanços recentes na aplicação de inteligência artificial (IA) na citopatologia da tireoide, com foco especial na biópsia aspirativa por agulha fina (PAAF), revisando a evolução nas aplicações da IA na citopatologia e comparando as tecnologias empregadas.

Os autores Kezlarian e Lin (2021) discutem a evolução das técnicas de análise, desde métodos morfométricos até a aplicação de redes neurais convolucionais, que se destacam na análise de imagens ao manter informações espaciais através das camadas da rede, permitindo classificações mais precisas, onde, apesar de alguns estudos anteriores, a maioria da literatura se concentra na distinção entre carcinomas papilares de tireoide e outros tipos de lesões.

A metodologia dos estudos revisados inclui a análise de imagens digitais de lâminas de biópsia. Esses dados são utilizados para treinar algoritmos de aprendizado de máquina, como redes neurais artificiais, máquinas de vetores de suporte e *K-vizinhos* mais próximos (Kezlarian; Lin, 2021).

Um exemplo é o estudo de Zoulias et al. (2011), que analisou 1.886 esfregaços benignos e 150 malignos, examinando 67 características microscópicas. Os resultados foram usados para treinar um classificador de votação majoritária, que atingiu uma sensibilidade de 89,1% e uma especificidade significativa.

Já Sanyal et al. (2018), desenvolveram uma CNN para classificar nódulos da tireoide onde foram utilizadas 186 imagens de carcinomas papilares e 184 imagens de lesões benignas. Ao analisar em diferentes ampliações, resultaram em uma sensibilidade de 90,48% e uma especificidade de 83,33% para lesões malignas.

Por sua vez, Guan et al. (2019) conduziu um estudo que utilizou um modelo de CNNs, demonstrando a eficácia na análise de biópsias. A revisão também ressalta que a área ainda enfrenta desafios, como a necessidade de um maior volume de amostras e a robustez dos algoritmos para aplicação clínica.

A diferenciação entre lesões foliculares benignas e malignas con-

tinua a ser complexa e várias abordagens de IA têm sido testadas para aprimorar a precisão diagnóstica. A pesquisa é apoiada por instituições de saúde e não apresenta conflitos de interesse por parte dos autores (Kezlarian; Lin, 2021).

O artigo, "Análise de Desempenho de Algoritmos de Aprendizado de Máquina na Classificação de Exames de Câncer do Colo do Útero", de Sobrinho (2018), tem como objetivo analisar o desempenho de diferentes algoritmos de aprendizado de máquina na classificação de células normais e anormais de colo do útero através de exames Papanicolau.

O autor do estudo descrito anteriormente utilizou a ferramenta *Waikato Environment for Knowledge Analysis* (WEKA) para realizar as análises. A metodologia envolveu a seleção de uma base de dados pré-existente, que continha descrições de células normais e anormais obtidas através do exame de *Papanicolaou*. Os algoritmos de aprendizado de máquina escolhidos para os experimentos foram *J48*, *Random Forest*, *Naive Bayes* e *Multilayer Perceptron*. Para isso, foi utilizada uma base de dados do Hospital da Universidade de Herlev, na Dinamarca, contendo 917 registros, com 20 atributos descritivos de células. A base original tinha 7 classes, mas uma versão secundária foi criada com apenas 2 classes: normal e anormal. O desempenho dos algoritmos foi avaliado com métricas específicas, a seleção de atributos foi aplicada para identificar os mais relevantes (Sobrinho, 2018).

A dissertação de Igor Cavalcante Guedes, "Avaliação Quantitativa da Citopatologia de Boca: Comparação entre a Análise Humana e Inteligência Artificial", investiga a detecção precoce do câncer bucal, com foco no carcinoma espinocelular (CECB), uma condição que apresenta alta mortalidade devido à sua detecção tardia. O estudo compara a eficácia da análise citopatológica realizada por humanos e por um sistema de inteligência artificial (IA) em esfregaços celulares bucais, utilizando a técnica de Papanicolau. Os resultados indicam que o uso do programa *Papanicolaou Slide - Image Examiner (PSIE)*, reduziu o tempo de análise em 16,6 vezes em relação à análise humana. A concordância entre os resultados obtidos por pesquisadores humanos e o sistema de IA foi notável, com um índice de concordância de 0,9 entre os especialistas e entre 0,544 e 0,548 com a IA, sugerindo que a inteligência artificial pode ser uma ferramenta viável para análises rotineiras (Guedes, 2024).

O autor (Guedes, 2024) propõe no mesmo artigo a ampliação da base de dados de imagens, visando melhorar a concordância entre os

resultados. Além disso, sugere a realização de estudos longitudinais para acompanhar pacientes ao longo do tempo, o desenvolvimento de modelos multicategóricos para a estratificação do risco de CECB e a investigação do uso contínuo da IA para o monitoramento de indivíduos expostos a fatores de risco. Essas direções têm como objetivo aprimorar a eficácia da detecção precoce do câncer bucal e integrar a IA de maneira mais abrangente na prática clínica.

As Redes Neurais Convolucionais (CNN) analisam diferentes regiões dos contornos da borda aplicando filtros convolucionais, extraindo bordas, texturas e formas e utilizando pesos para essas informações. Cada filtro é uma matriz que aplica operações de convolução gerando um mapa de características. Para reduzir o custo computacional as CNNs possuem a camada de *Pooling* que reduz a dimensão dos mapas de características. Por fim, a camada densa interliga todos os neurônios da última camada com todos os neurônios da camada anterior, combinando as características e realizando a classificação (Silva, 2018).

A Máquina de Vetores de Suporte (SVM) é amplamente utilizada na área de visão computacional, devido à sua eficiência em lidar com grandes quantidades de detalhes nos dados e com conjuntos de dados menores. Seu funcionamento se baseia na utilização de algoritmos de classificação e regressão que tentam maximizar a margem entre classes distintas ao encontrar o hiperplano que melhor separa as classes (Bellini, 2024). Essa abordagem garante maior capacidade de generalização e menor propensão ao erro em novas previsões.

O artigo inicia-se com a introdução que apresenta a problemática e justificativa do projeto, apresentando o objetivo proposto, a fundamentação teórica e relatando os trabalhos correlatos. A próxima seção “Materiais e Métodos” contextualiza o ambiente de desenvolvimento e como foi desenvolvido, desde a base de dados até o desenvolvimento do projeto. Por fim, serão apresentados e discutidos os resultados obtidos e conclusão em seções separadas, finalizando com as referências.

2 MATERIAIS E MÉTODOS

Tendo em vista o objetivo de testar modelos de inteligência artificial para análises de imagens clínicas, definindo os métodos de desenvolvimento e testes, assim como a coleta de dados, esta etapa foi essencial para garantir a organização do experimento, a reprodutibilidade dos resultados e a escolha adequada das ferramentas e técnicas utilizadas ao longo

da pesquisa dando-se através dos seguintes passos.

2.1 AMBIENTE DE TESTE

Para a implementação e testes dos sistemas desenvolvidos neste estudo, foi necessário configurar um ambiente de desenvolvimento compatível com os requisitos necessários das bibliotecas de *machine learning*.

Para utilização da rede neural convolucional, foi utilizada a biblioteca do *TensorFlow Keras*, no entanto, a versão atual do TensorFlow 2.19.0 não está mais disponível para *Windows*. Foi instalado o WSL com Ubuntu 22.04.5 e instalado o *Python* versão 3.11.11, última versão da linguagem suportada pelo *TensorFlow*. Os protótipos foram testados utilizando um processador i5-10300H com 24GB de RAM.

2.2 BASE DE DADOS

As amostras utilizadas foram retiradas do fórum Cytopathology (2024), uma instituição internacional voltada para estudo e inovações na área da citopatologia e patologia, com cerca de 300 imagens livres para usos institucionais. Contudo, o álbum possuía classes como “indeterminado” e “suspeito” que não se enquadram no presente estudo, sendo assim, foram feitos os downloads apenas das classes maligna e benigna.

2.3 TRATAMENTO DAS IMAGENS

Com o objetivo de melhorar o desempenho dos classificadores de inteligência artificial mesmo com poucas imagens disponíveis, foram usadas algumas estratégias para aumentar a quantidade e a qualidade da base de dados. Essas técnicas buscavam não só gerar mais exemplos para o treino, mas também reduzir ruídos e destacar melhor as características importantes.

Diante disso, foi realizada a divisão da base de dados original em 70% para treino, 15% para validação durante o treino e 15% para o teste final, a base de dados foram separadas antes de realizar qualquer processamento das mesmas. Devido à quantidade limitada de dados, foram aplicadas técnicas de aumento de dados. Para contornar esta limitação, foi realizado o pré-processamento manual, gerando cópias das imagens com rotações de 45°, 60° e 90°, zoom 1.2x e recortes aleatórios de 10%, ampliando o conjunto de dados.

Outra técnica utilizada foi a implementação de uma classe do *Keras ImageDataGenerator*, que aplica transformações de rotação, zoom e

cortes em tempo real para melhorar a generalização. Esta classe também aplica a função de pré-processamento das amostras.

A função de processamento tem o objetivo de evitar ruídos e realçar os detalhes, aplicando a técnica de *CLAHE*, que diferentemente de outros métodos que clareiam a imagem por inteiro perdendo qualidade em certas áreas, ao usar o *CLAHE*, a imagem é segmentada em regiões equalizando o contraste separadamente, evitando perda de informação.

2.4 MODELO DE CNN

O projeto utiliza a arquitetura de rede neural convolucional *MobileNetV2*, implementada no *TensorFlow Keras*, como *backbone*. A arquitetura contém pesos pré-treinados, o que facilita a extração de características, como bordas e texturas, além de otimizar o desenvolvimento e reduzir o custo computacional ao processar a base de dados com dimensões de entrada de 240x240 em etapas usando convoluções separáveis em profundidade, reduzindo gradualmente as dimensões ao aplicar global *average pooling*.

Duas das principais características do *MobileNetV2* são os blocos residuais invertidos (*inverted residuals*), expandindo o número de canais de convolução, aumentando a capacidade de captar características robustas em imagens médicas, e a técnica de gargalos lineares, que preserva informações em espaços de baixa dimensionalidade, evitando perdas de dados ao manter valores negativos.

As camadas adicionadas no topo, ao aplicar técnicas como *globalAveragePooling2D*, *Dense*, *BatchNormalization*, *Dropout* e ativação *softmax*, visam evitar que o modelo memorize a base de dados, priorizando a sensibilidade da classe “maligna” e evitando *overfitting*. Na segunda etapa, foi aplicado o *fine-tuning*, descongelando as últimas camadas do *MobileNetV2* e o treinamento novamente com 10 épocas ajustando os pesos das camadas superiores.

Para otimizar o treinamento, foram adicionadas quatro *callbacks* no treinamento do algoritmo. A *OversampleToBalance* foi utilizada para balancear as classes, aumentando as amostras da classe minoritária. A *Early Stopping* interrompe o treinamento caso o *recall* não melhore após 3 épocas, economizando recursos. A *ReduceLROnPlateau* reduz a taxa de aprendizado caso o *loss* não apresente melhora, permitindo ajustes mais finos nos pesos. Já a *ModelCheckpoint* salva o modelo sempre que houver uma melhora no *recall* da classe “maligna”. Após o treinamento, o modelo

salvo é carregado novamente.

Para o treinamento, as imagens foram redimensionadas para 240x240 *pixels*, tratadas com a função *CLAHE* e, por fim, processadas pela função *preprocess_input*, que converte a escala de *pixels* de 0 a 255 para o intervalo de -1 a 1, conforme exigido pelo *MobileNetV2*.

2.5 MODELO SVM

Utilizando o *framework scikit-learn* para implementar um modelo de Máquina de Vetor de Suporte (SVM), as amostras foram tratadas com a função *CLAHE* com a adição da técnica HOG (Histogram of Oriented Gradients), que foca no aprendizado das formas geradas pelos contornos. Essas formas são convertidas em números que posteriormente são padronizados, porém, para reduzir a dimensionalidade, foi aplicada a técnica de PCA (Análise de Componentes Principais) a fim de manter somente as partes essenciais.

Como o acerto de Maligno tem maior importância, foram ajustados os pesos informando que a classe maligna é mais grave, alterados parâmetros dos hiperparâmetros com o intuito de melhorar a distinção das classes usando o *kernel RBF* e *GridSearchCV* para validação cruzada.

Além disso, o desempenho do SVM depende de parâmetros importantes, como o "C", que regula o equilíbrio entre a margem e os erros de classificação, o *Gamma*, que define a influência dos pontos e o *Degree*, que ajusta a complexidade do kernel polinomial. A escolha do *kernel* (RBF, polinomial, sigmoide, etc.) também é essencial, pois determina como os dados são mapeados para espaços de maior dimensão e influencia diretamente os resultados obtidos. Ferramentas como a biblioteca *Scikit-learn*, em *Python*, oferecem suporte a SVMs e facilitam sua aplicação prática em diferentes cenários de aprendizado de máquina (Santos et al., 2016).

2.5.1 Modelo SVM com MobileNetV2

Outra técnica aplicada, ao invés de utilizar o HOG, foi implementado o *backbone* de rede neural convolucional apenas como extrator de características ao desabilitar a camada de classificação de características ao utilizar o *MobileNetV2*, capturando padrões visuais complexos de bordas. As características são extraídas criando um modelo intermediário que retorna um vetor de 1280 elementos por imagem da camada de *pooling*, que serão usados como entrada para treinar o SVM.

Neste, foi alterado a função de pré-processamento para incluir o

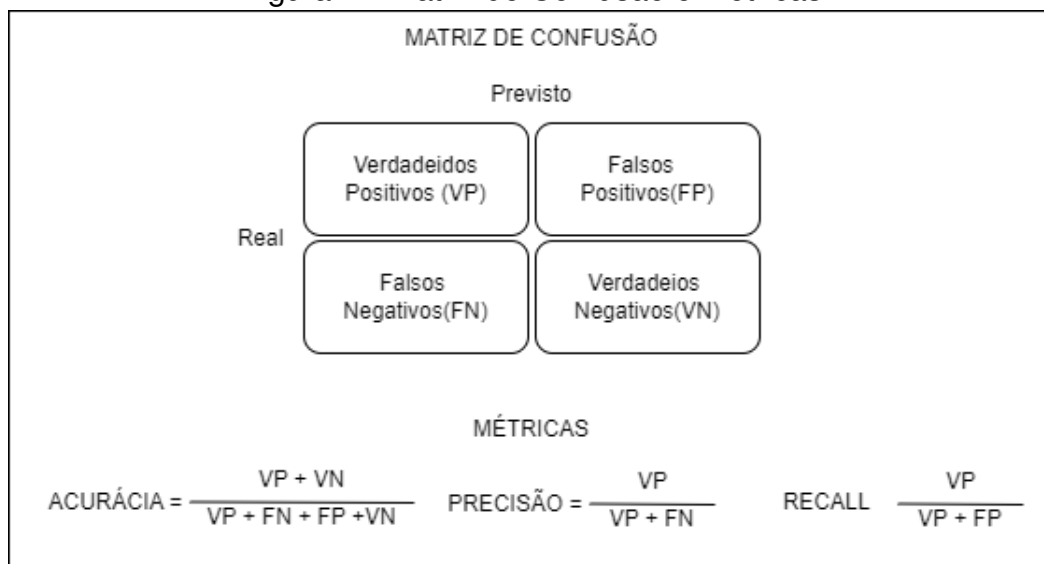
Keras preprocess_input que converte a imagem para o padrão de entrada do *MobileNetV2*. As configurações implementadas aqui seguem a mesma implementação do SVM, utilizando o *kernel RBF* e maximizando os hiperparâmetros.

2.6 MÉTRICAS

Uma das formas de avaliação dos algoritmos de classificação é a matriz de confusão, uma tabela 2x2 auxiliando no cálculo de outras métricas. A matriz auxilia a visualizar os resultados mostrando a quantidade de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos (Fontana, 2020).

Como demonstrado na Figura 2, com os valores da matriz de confusão, obtém-se o cálculo a acurácia, que mede a precisão de acertos ao somar o total de verdadeiros positivos e negativos com o total de dados de teste. Com uma acurácia alta podemos dizer que os valores estão com uma maior chance de acerto. (Fontana, 2020).

Figura 2 - Matriz de Confusão e Métricas.



Fonte: Elaborado pelo autor.

Diferente da acurácia que mede a quantidade de valores positivos e negativos, a precisão verifica a proporção de previsões positivas corretas (TP) em relação ao total de previsões positivas. A precisão mede a proporção das previsões positivas que realmente são positivas (Strauss; Júnior; Ferreira, 2022).

De acordo com Strauss, Júnior e Ferreira (2022), Júnior e Ferreira (2022), *Recall* ou sensibilidade, é a quantidade de previsões de verda-

deiros positivos que o modelo fez e que foram previstas corretamente. É a proporção de instâncias realmente positivas que foram corretamente identificadas, sendo calculada pela razão entre o número de verdadeiros positivos pela soma entre os verdadeiros positivos e os falsos negativos. Essa métrica tem uma importância maior, mesmo que reduza a precisão ou gere mais falsos positivos, pois é preferível marcar uma imagem benigna como maligna do que classificar um maligno como benigno.

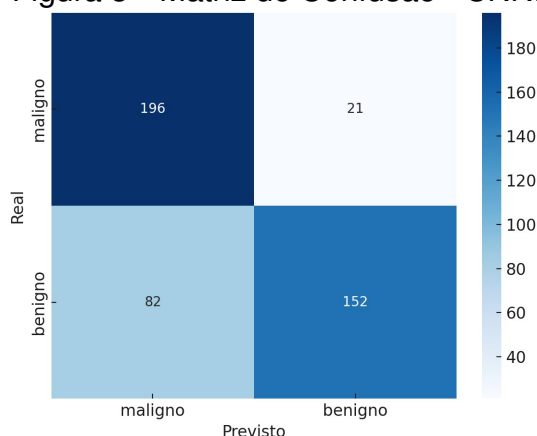
3 DISCUSSÃO E RESULTADOS

Um dos maiores obstáculos encontrado neste projeto foi a escassez de imagens disponíveis. Embora tenha-se encontrado diversos sites com amostras citológicas, o material era composto de tipos celulares que não eram de células da tireoide e por fim, a maioria das bases eram de exames de ultrassom, desviando da temática do estudo.

Por se tratar de um exame médico, o foco dos modelos é elevar a sensibilidade da classe maligna, tentando manter o melhor nível de precisão, devido à gravidade de falsos negativos.

Com o CNN, evidenciado na na Figura 3, alcançou-se uma acurácia de 77% com precisão de 70% da classe maligna ao separar corretamente, conforme a Figura 3, das 165 das 251 imagens previstas. Porém, a sensibilidade do modelo alcançou 90%, isso significa que dos 217 casos cancerígenos deixou de classificar 21. Estes resultados foram semelhantes ao de Sanyal et al. (2018), que, com base reduzida, alcançaram 90,48% de sensibilidade e 83,33% de especificidade.

Figura 3 - Matriz de Confusão - CNN.

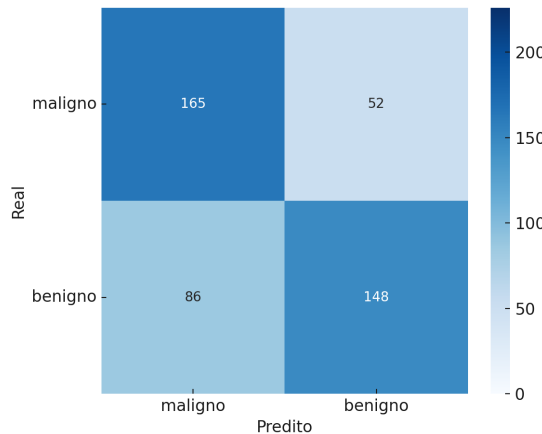


Fonte: Elaborado pelo autor.

Já com o modelo SVM, apresentado na Figura 4, obteve-se uma

acurácia de 69% e uma precisão de 65%. Das 251 amostras previstas como malignas, apenas 165 eram de fato positivas, enquanto 86 foram falsos positivos. Dos 217 casos confirmados, 165 foram corretamente identificados, porém 52 foram classificados como benignos, resultando em uma sensibilidade de 76%.

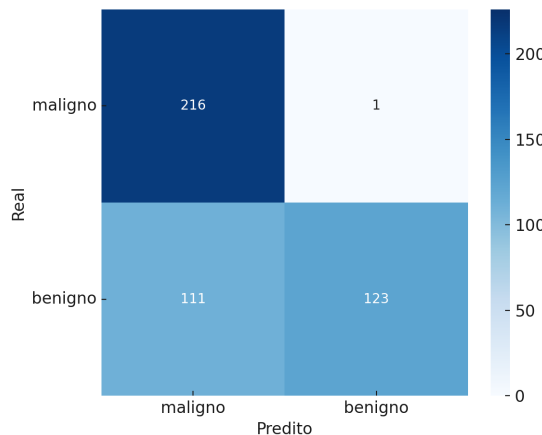
Figura 4 - Matriz de Confusão - SVM.



Fonte: Elaborado pelo autor.

Porém, ao utilizar a arquitetura *MobileNetV2* do CNN como extrator de características e treinar o SVM, atingiu-se a acurácia de 75% e 66% de precisão ao acertar 216 das 327 imagens previstas como malignas. Esses valores são poucos a baixos em relação ao CNN, entretanto, o mesmo conseguiu obter uma sensibilidade de 99% ao acertar 216 das 217 imagens, como ilustrado na Figura 5.

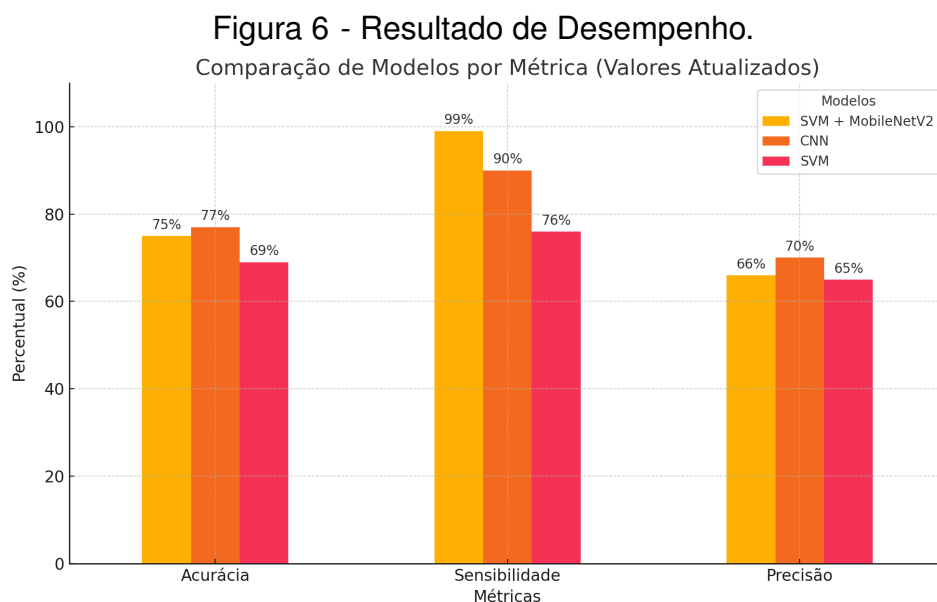
Figura 5 - Matriz de Confusão - SVM + MobilenNetV2.



Fonte: Elaborado pelo autor.

Esses resultados evidenciam diferenças relevantes entre os modelos, servindo de base para a análise complementar apresentada a seguir

com a Figura 6.



Fonte: Elaborado pelo autor.

É possível reparar que o modelo que teve o desempenho mais baixo foi o SVM puro, esperava-se um melhor desempenho deste por conseguir lidar melhor com base de dados menores, contudo não conseguiu diferenciar tão bem as classes somente com seus extratores de características, porém ao adicionar uma arquitetura pré-processada como o *MobileNetV2* somente como extrator de características, o SVM conseguiu obter a melhor sensibilidade sem se distanciar muito da precisão do CNN, que obteve a melhor acurácia e precisão. Esses resultados são valores semelhantes encontrados na pesquisa de Kezlarian e Lin (2021), que em sua revisão bibliográfica encontrou pesquisas com a sensibilidade variando entre 68-99%.

4 CONCLUSÃO

Este trabalho teve como objetivo comparar o desempenho de diferentes modelos de inteligência artificial na classificação de imagens citológicas da tireoide, visando auxiliar na detecção de nódulos malignos. Foram desenvolvidos e testados três algoritmos: uma rede neural convolucional (CNN), uma Máquina de Vetores de Suporte (SVM) e uma SVM combinada com o *MobileNetV2*. O objetivo foi alcançado, permitindo avaliar a eficácia de cada abordagem mesmo em um cenário com dados limitados.

Os resultados mostraram que a CNN obteve a maior taxa de acerto e precisão, enquanto a combinação da SVM com o *MobileNetV2*

se destacou pela alta sensibilidade, identificando 99% dos casos malignos, um aspecto essencial para aplicações médicas. Já o modelo SVM puro teve menor desempenho, indicando a importância de utilizar bons métodos de extração de características.

Para estudos futuros, recomenda-se o uso de bases de dados maiores e mais específicas, além da aplicação de técnicas que melhorem a interpretação dos resultados pelos profissionais da saúde. Isso pode tornar os modelos mais úteis e confiáveis no apoio ao diagnóstico clínico.

REFERÊNCIAS

BELLINI, R. **Aplicação de máquinas de suporte vetorial na classificação textual**. Universidade de Caxias do Sul. 2024.

CYTOPATHOLOGY, T. P. S. of. **Image Atlas of Gynecologic Cytology**. 2024. Acesso em 02 jun. 2025. Disponível em: <<https://www.papsociety.org/image-atlas/>>.

FONTANA, É. **Introdução aos algoritmos de aprendizagem supervisionada**. Departamento de Engenharia Química, Universidade Federal do Paraná. 2020.

GUAN, Q. et al. **Deep convolutional neural network vgg-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study**. Journal of Cancer, v. 10, n. 20, p. 4876–4882. 2019.

GUEDES, I. C. **Avaliação quantitativa da citopatologia de boca: comparação entre a análise humana e inteligência artificial**. Porto Alegre, RS, Brasil. 2024.

Instituto Nacional de Câncer José Alencar Gomes da Silva. **ABC do Câncer: Abordagens Básicas para o Controle do Câncer**. 6. ed., revista e atualizada. ed. Rio de Janeiro: INCA, 2020.

KEZLARIAN, B.; LIN, O. **Artificial intelligence in thyroid fine needle aspiration biopsies**. Acta cytologica, v. 65, n. 4, p. 324–329. 2021.

MACIEL, R. **Citologia aspirativa da tiróide: utilidade diagnóstica atual e perspectivas futuras**. [S.l.]: SciELO Brasil, 2001. 217–218 p.

SANTOS, L. M. S. d. et al. **Evolução temporal da mortalidade por câncer de tireoide no brasil no período de 2000 a 2012**. Revista Brasileira de Análises Clínicas, Salvador, v. 48, n. 2, p. 133–137. 2016.

SANYAL, P. et al. **Artificial intelligence in cytopathology: a neural network to identify papillary carcinoma on thyroid fine-needle aspiration cytology smears**. Journal of Pathology Informatics, v. 9, p. 43. 2018.

SILVA, R. E. V. d. **Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens**. 2018. 51 f. Trabalho de Conclusão de Curso (Graduação em Sistemas de Informação) – Universidade Federal do Ceará, Campus de Quixadá, Quixadá.

SOBRINHO, D. d. A. L. **Análise de desempenho de algoritmos de aprendizagem de máquina na classificação de exames de câncer no colo do útero**. Tese (Doutorado) — Universidade Estadual do Piauí (UESPI), 2018.

STRAUSS, E.; JÚNIOR, M. V. B.; FERREIRA, W. L. L. **A importância de utilizar métricas adequadas de avaliação de performance em modelos preditivos de machine learning**. Projectus, v. 7, n. 2, p. 52–62. 2022.

YOKOZAWA, T. **Câncer da tireóide detectado pela punção aspirativa por agulha fina guiada pelo ultra-som**. Arquivos Brasileiros de Endocrinologia & Metabologia, SciELO Brasil, v. 42, p. 296–298. 1998.

ZOULIAS, E. A. et al. **A decision support system for assisting fine needle aspiration diagnosis of thyroid malignancy**. Analytical and Quantitative Cytology and Histology, v. 33, n. 4, p. 215–222. 2011.