UNIVERSIDADE FEDERAL DE OURO PRETO INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS DEPARTAMENTO DE COMPUTAÇÃO

LUCAS ANDRADE FREITAS Orientadora: Prof. Dra. Andrea G. Campos Bianchi

MODELO DE DETECÇÃO E CLASSIFICAÇÃO CELULAR BASEADA EM REDES NEURAIS CONVOLUCIONAIS

Ouro Preto, MG 2022

UNIVERSIDADE FEDERAL DE OURO PRETO INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS DEPARTAMENTO DE COMPUTAÇÃO

LUCAS ANDRADE FREITAS

MODELO DE DETECÇÃO E CLASSIFICAÇÃO CELULAR BASEADA EM REDES NEURAIS CONVOLUCIONAIS

Monografia apresentada ao Curso de Ciência da Computação da Universidade Federal de Ouro Preto como parte dos requisitos necessários para a obtenção do grau de Bacharel em Ciência da Computação.

Orientadora: Prof. Dra. Andrea G. Campos Bianchi

Ouro Preto, MG 2022

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

F866m Freitas, Lucas Andrade. Modelo de detecção e classificação celular baseada em redes neurais convolucionais. [manuscrito] / Lucas Andrade Freitas 2022. 28 f.: il.: color., tab
Orientadora: Profa. Dra. Andrea Gomes Campos Bianchi. Monografia (Bacharelado). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Graduação em Ciência da Computação .
 Classificação. 2. Aprendizado de Máquina. 3. Rede Neural Convolucional. I. Bianchi, Andrea Gomes Campos. II. Universidade Federal de Ouro Preto. III. Título.
CDU 573

Bibliotecário(a) Responsável: Luciana De Oliveira - SIAPE: 1.937.800



MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE FEDERAL DE OURO PRETO REITORIA INSTITUTO DE CIENCIAS EXATAS E BIOLOGICAS DEPARTAMENTO DE COMPUTACAO



FOLHA DE APROVAÇÃO

Lucas Andrade Freitas

Modelo de detecção e classificação celular baseada em redes neurais convolucionais

Monografia apresentada ao Curso de Ciência da Computação da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Ciência da Computação

Aprovada em 26 de Outubro de 2022.

Membros da banca

Andrea Gomes Campos Bianchi (Orientadora) - Doutora - Universidade Federal de Ouro Preto Pedro Henrique Lopes Silva (Examinador) - Doutor - Universidade Federal de Ouro Preto Débora Nasser (Examinadora) - Mestre - IBM Marcone J. Freitas Souza (Examinador) - Doutor - Universidade Federal de Ouro Preto

Andrea Gomes Campos Bianchi, Orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 26/10/2022.



Documento assinado eletronicamente por Andrea Gomes Campos Bianchi, PROFESSOR DE MAGISTERIO SUPERIOR, em 03/11/2022, às 15:24, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



A autenticidade deste documento pode ser conferida no site <u>http://sei.ufop.br/sei/controlador_externo.php?</u> acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0416022** e o código CRC **BB8E194F**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.014507/2022-51

Agradecimentos

Gostaria de agradecer primeiro lugar a Deus por sempre ter me guiado, dando saúde e foco para todos os desafios. Agradeço toda minha família, por sempre me apoiar em minhas decisões e me auxiliar sempre que necessitei. Em especial meu pai Carlos, minha mãe Adlene, meu irmãos Gabriell e Rafael, meu irmão Gabriell nunca esquecerei de tudo que você fez por mim, e em quanto eu estiver respirando irei honrar o teu nome, muito obrigado por todos os conselhos. A minha namorada Isabelle Christina, que teve em todos esses 4 anos ao meu lado, me auxiliando, aconselhando e guiando para as minhas conquistas, sempre a me lembrar que nunca estive sozinho nessa caminhada. A professora Andreia que sempre me auxiliou, com todo carinho, dedicação e humildade, sempre compartilhando o seu conhecimento, tornando esse trabalho realidade com a sua participação. Agradeço também, aos meu amigos, Gabriel Bicalho, Vinícius de Paula, Halliday Gauss, Carlos Magalhães, Willian Gomes, Emanuel Jesus, Marcos Pontes e Vinícius Targa, pelos momentos vividos juntos, pelas noite em claros estudando, e por toda colaboração e companheirismo que sempre nos ajudou em toda caminhada na UFOP. Serei eternamente grato a todos. À Universidade Federal de Ouro Preto (UFOP), à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelos financiamentos concedidos.

Resumo

Esta monografia visa analisar amostras de células cervicais de exame de Papanicolaou através da metodologia baseada no aprendizado profundo. É apresentado um modelo para classificação de células cervicais sem a necessidade de segmentação da célula (núcleo ou citoplasma) para realizar o treinamento do modelo. Essa abordagem usa no treinamento a informação da localização do núcleo disponibilizada pelo banco de dados e faz recortes ao redor do núcleo usando dimensões variadas. Para a realização do trabalho propostos os desenvolvimentos de diversos modelos de CNN (*convolutional neural networks* de classificação. Para o treinamento do modelo foram utilizadas 400 imagens reais do exame de Papanicolaou extraídas de uma base de dados de células cervicais. Espera-se que os resultados possam auxiliar os citopatologistas a melhorar a taxa de classificação correta dos exames realizados em laboratórios, proporcionando uma melhoria na qualidade dos resultados.

Palavras-chave: Aprendizado de Máquina. Classificação. Rede Neural Convolucional.

Abstract

This monograph aims at examining cervical cells through a methodology based on deep learning. It is a classification model for the offer model or classification model to carry out the training to carry out the training. This approach uses the location information provided by the database in training and cuts around the core using varying dimensions. To carry out the proposed developments of several CNN models (*convolutional* classification. For the training of the neural network model, 400 real images of the Pap smear from a cervical cell database were used. - The results can help improve a correct classification rate of exams performed in laboratories, providing an improvement in the quality of the results.

Keywords: Machine Learning. Classification. Convolutional Neural Network.

Lista de Ilustrações

Figura 2.1 – Exemplo de rede neural convolucional e suas diferentes camadas (VARGAS; PAES;	
VASCONCELOS, 2016)	4
Figura 2.2 – Exemplo de um processo de convolução (MARIED et al., 2017)	6
Figura 2.3 – Exemplo de uma utilização do <i>Max Polling</i> de dimensão 2x2 e com salto 2 sobre o	
mapa de características (SHARMA; MANGAT, 2016)	7
Figura 3.1 – Imagem retirada da base de dados CRIC Cervix.	10
Figura 3.2 – Classificações geradas da base de dados CRIC Cervix	10
Figura 3.3 – Exemplo da geração de imagens com o núcleo centralizado e caixas de 90x90	12
Figura 3.4 – Diferentes tamanhos de caixa para a mesma célula	12
Figura 3.5 – Processo de geração das imagens descentralizadas.	13
Figura 3.6 – Exemplo visual do uso da técnica da fronteira.	13
Figura 3.7 – Fluxograma da arquitetura interna da <i>Inception V3</i> (SZEGEDY et al., 2016)	16
Figura 3.8 – Fluxograma da arquitetura interna da <i>EfficientNet</i> .(AHMED; SABAB, 2020)	17
Figura 3.9 – (A) Arquitetura da <i>MobileNet</i> (B) Explicação da camada de <i>Depthwise Separable</i>	
<i>Convolution</i> . (SINGH et al., 2015)	18
Figura 4.1 – Matriz de confusão do modelo <i>EfficientNet</i> para o tamanho de caixa 30x30	20
Figura 4.2 – Matriz de confusão do modelo <i>EfficientNet</i> para o tamanho de caixa 50x50	21
Figura 4.3 – Matriz de confusão do modelo <i>EfficientNet</i> para o tamanho de caixa 70x70	22
Figura 4.4 – Matriz de confusão do modelo <i>XceptionNet</i> para o tamanho de caixa 90x90	23
Figura 4.5 – Exemplo de uma imagem da base da dados CRIC com sua verdadeira classificação.	24
Figura 4.6 – Predição e distribuição das classes do sistema proposto.	25

Lista de Tabelas

Tabela 3.1 – Distribuição de núcleos por classe CRIC Cervix.	11
Tabela 3.2 – Distribuição das células centralizadas.	12
Tabela 3.3 – Distribuição de classes descentralizadas para tamanho de caixa 30x30. .	14
Tabela 3.4 – Distribuição de classes descentralizadas para tamanho de caixa 50X50. .	14
Tabela 3.5 – Distribuição de classes descentralizadas para tamanho de caixa 70X70.	14
Tabela 3.6 – Distribuição de classes descentralizadas para tamanho de caixa 90X90. . .	14
Tabela 3.7 – Distribuição final das classes para tamanho de caixa 30x30.	15
Tabela 3.8 – Distribuição final das classes para tamanho de caixa 50x50.	15
Tabela 3.9 – Distribuição final das classes para tamanho de caixa 70x70.	15
Tabela 3.10–Distribuição final das classes para tamanho de caixa 90x90.	15
Tabela 4.1 – Especificação do Hardware utilizado.	19
Tabela 4.2 – Métricas do modelo EfficientNet para o tamanho de caixa 30X30.	20
Tabela 4.3 – Métricas dos modelos para o tamanho de caixa 50X50. <	20
Tabela 4.4 – Métricas dos modelos para o tamanho de caixa 70x70. <	21
Tabela 4.5 – Métricas dos modelos para o tamanho de caixa 90x90.	22

Lista de Abreviaturas e Siglas

- ABNT Associação Brasileira de Normas Técnicas
- DECOM Departamento de Computação
- UFOP Universidade Federal de Ouro Preto
- CRIC Center for Recognition and Inspection of Cells
- CNN Convolutional Neural Network
- UX User Experience
- LSIL Lesão Intraepitelial Escamosa de Baixo Grau
- HSIL Lesão Intraepitelial Escamosa de Alto Grau
- ASC-H Células Escamosas Atípicas Não Podendo Excluir HSIL
- ASC-US Células Escamosas Atípicas de Significado Indeterminado
- NILM Negativo para Lesão Intraepitelial ou Malignidade
- SCC Carcinoma de Células Escamosas
- SVM Support Vector Machine
- KNN K Nearest Neighbor
- JSON JavaScript Object Notation

Lista de Símbolos

- Γ Letra grega Gama
- Λ Lambda
- ζ Letra grega minúscula zeta
- ∈ Pertence

Sumário

1	Intr	odução
	1.1	Motivação
	1.2	Objetivos
		1.2.1 Objetivos específicos
	1.3	Organização do Trabalho
2	Rev	isão Bibliográfica
	2.1	Trabalhos Relacionados
	2.2	Fundamentação Teórica
		2.2.1 Rede Neural Convolucional
		2.2.1.1 Camada de Convolução
		2.2.1.2 Processo de <i>Pooling</i>
		2.2.2 Métricas de Avaliação
		2.2.2.1 F1-Score
		2.2.2.2 Acurácia
		2.2.2.3 Revocação
		2.2.2.4 Precisão
		2.2.2.5 Especifidade
3	Des	envolvimento
	3.1	Base de Dados Cric Cervix 9
	3.2	Pré-processamento
		3.2.1 Núcleo Centralizado
		3.2.2 Núcleo Descentralizado
	3.3	Arquitetura das Redes Neurais Convolucionais
		3.3.1 Inception V3 e XceptionNet 15
		3.3.2 <i>EfficientNet</i>
		3.3.3 <i>MobileNet</i>
4	Res	ultados
	4.1	Descrição dos Experimentos
	4.2	Tamanho de Caixa 30x30 19
	4.3	Tamanho de Caixa 50x50 20
	4.4	Tamanho de Caixa 70x70 21
	4.5	Tamanho de Caixa 90x90 22
	4.6	Discussão dos Resultados
	4.7	Predição do Sistema
5	Con	siderações Finais
	5.1	Conclusão

5.2	Traball	10s Fi	ituros	•	 •	 •	•	 •	 •	•	•••	•	•	• •	•	•	 •	•	 •	•	 26
Referênc	cias		••••	• •	 •	 •	•		 •				•		••	•	 •	•	 •		 27

1 Introdução

A organização mundial de saúde (World Health Organization, 2018)) estipulou em 2018 que o câncer do colo do útero vem registrando uma média de 530.000 novos casos da doença, sendo mais frequente na população feminina, representando o terceiro tumor mais frequente nessa população, sendo responsável por 265 mil óbitos anuais, chegando a ser a quarta causa mais frequente de morte por câncer em mulheres. No entanto, um fato positivo é que essa neoplasia tem desenvolvimento lento, elevando assim consideravelmente a taxa de cura quando essas lesões são diagnosticadas precocemente pelo exame citopatológico ou Papanicolaou. Esse exame foi introduzido pelo médico George Nicholas Papanicolaou. Com o passar dos anos, o exame de Papanicolau se tornou a principal forma de diagnóstico, principalmente pelo fato dos outros métodos de exame serem extremamente caros nos países em desenvolvimento. Essa exame é capaz de identificar alterações nas células descamadas do colo do útero com base em vários estudos citomorfológicos. Os critérios são baseados em características nucleares, sendo um exemplo o aumento nuclear, irregularidade da membrana, hipercromasia e relação do tamanho do núcleo e citoplasma.

Normalmente numa rotina um citopatologista costuma avaliar até 300.000 células cervicais em um único esfregaço. Além disso, dependendo da sua nacionalidade pode chegar a avaliar 100 esfregaços por dia. Por exemplo, no Brasil a recomendação de análises diárias é de no máximo 70 esfregaços por dia, mas nem sempre esse limite é respeitado, o que pode ocasionar erros na hora do resultado do exame. Uma investigação conduzida desde antes de 1990 mostra que as taxas de falsos-negativos, que representa os casos de células anormais que são classificados como normais, podem variar entre 2% a 62% (SUNG et al., 2021) nos exames de Papanicolaou, o que é considerado um problema grave.

Neste trabalho será desenvolvido um modelo de classificação de células, esse sistema poderá as classificar em 3 classes distintas, sendo elas (com lesão, sem lesão e não célula). Em Diniz et al. (2021a) é usado para a elaboração do modelo de treinamento, imagens que possuem o núcleo já segmentado. Esse autores utilizaram um trabalho manual para efetuar a segmentação, sendo este processo extremamente custoso, delicado e trabalhoso. Esta monografia é uma proposta de expansão baseado no trabalho Antunes (2021) que usa metodologias tradicionais, já nosso sistema será baseado no uso de CNN (*Convolutional Neural network*). na fase de treinamento são realizados cortes de diferentes dimensões ao redor do núcleo, não tendo a necessidade de possuir as imagens já segmentadas, com este procedimento é possível vários testes, seguindo as mesmas métricas de avaliação, a fim de conseguir identificar o quanto a segmentação impacta no resultado final. Além disso, esta pesquisa contribui com uma abordagem diferente, na qual não há a necessidade de segmentar imagens na etapa de treinamento. Os testes realizados mostraram que os resultados obtidos foram tão satisfatórios quanto os obtidos em Diniz et al. (2021a)

1.1 Motivação

A motivação para realização deste trabalho é a criação de uma ferramenta que possa auxiliar os profissionais na área da saúde, mais especificamente os citopatologistas e médicos, em suas tarefas de diagnóstico de lesões celulares. Espera-se, com essa ferramenta, uma diminuição da sobrecarga de trabalho desses profissionais na realização de suas tarefas.

Por outro lado, quando pensamos nos algoritmos de inteligência artificial, também foi fonte de motivação o desenvolvimento de uma ferramenta que dispensa a segmentação de regiões de interesse em uma fase anterior à da classificação, evitando um esforço desnecessário e, ao mesmo tempo, economizando tempo na conclusão da tarefa de classificação. Dessa maneira, resultando numa melhora na taxa de sobrecarregamento do serviço dos profissionais, os ajudando no dia a dia na classificação das células.

1.2 Objetivos

O objetivo dessa monografia é desenvolver e validar um modelo para a classificação de imagens reais de citologia convencional do colo do útero (exames de Papanicolau), tendo como sua principal aplicação, ajudar na tomada de decisão do profissional quanto ao diagnóstico, visto que o diagnóstico incorreto é um problema bastante pertinente na área dos citopatologistas.

1.2.1 Objetivos específicos

1. Desenvolvimento de um modelo que diferencia as regiões da imagem em três classes, as diferenciando em células com lesão, sem lesão e não célula;

2. Proposta de classificação de imagens de Papanicolaou sem o uso de informações de segmentação de núcleo ou citoplasma

3. Investigação dos resultados encontrados, através do uso das imagens com tamanho de caixa variado;

4. Investigação das métricas obtidas em 4 tipos de modelos de redes neurais convolucionais diferentes.

1.3 Organização do Trabalho

Neste Capítulo é descrita a organização desse trabalho. No Capítulo 2 aborda conceitos importantes e trabalhos com temas relacionados. No Capítulo 3 é apresentada a metodologia que será desenvolvida. No Capítulo 4 são relatados as análises e os resultados do trabalho desenvolvido. No Capítulo são apresentadas conclusões sobre o trabalho.

2 Revisão Bibliográfica

Ao decorrer do tempo, diferentes tipos de abordagens para a classificação de células foram desenvolvidas, nessa seção passaremos por algumas delas, pois elas foram de suma importância para o entendimento do problema e usadas como referência para o desenvolvimento do trabalho. As revisões focam em explorar as metodologias adotadas, na explicação dos algoritmos escolhidos e resultados obtidos.

2.1 Trabalhos Relacionados

Em Diniz et al. (2021b), é proposto um algoritmo de redução da carga de trabalho para citopatologistas que analisam exames de Papanicolaou. A detecção é baseada na análise das características dos núcleos das células das imagens do esfregaço de Papanicolaou. No artigo citado são utilizados diversos algoritmos tradicionais de aprendizado de máquina para realizar uma classificação hierárquica, em até 6 classes. Dentre os algoritmos utilizados está presente a Árvore de decisão, k-NN (*K-Nearest Neighbors*), *Random Forest*, Métodos de Ridge. No artigo citado, foi concluído que o melhor resultado, se obtém ao usar o Random Forest como classificador.

Gómez, Sánchez-DelaCruz e Mata (2017) fizeram uma comparação de vários algoritmos para classificação do câncer de colo de útero separando as células em duas classes: normais e anormais. Eles seguiam o seguinte fluxo, aquisição de dados, pré-processamento, treinamento e seleção do modelo, o banco de dados utilizado era fornecido pelo hospital da Herley University, que tinha disponível 917 imagens, onde 242 eram células normais e 675 anormais. Extraíram das imagens 20 características morfológicas cada uma com seu peso e encontraram que as combinações de algoritmos Bagging + MultilayerPerceptron e AdaBoostM1 + LMT foram os melhores cenários analisados por eles.

Mariarputham e Stephen (2015) relatam que a tarefa de classificação precisa das imagens do esfregaço de Papanicolaou, podia ser melhorada através da seleção de características específicas bem definidas e da seleção do melhor classificador. O sistema apresentado, envolvia a preparação dos esfregaços de Papanicolaou, através de imagens capturadas por microscópio óptico, que são passadas para um pré-processamento onde ocorre a segmentação do núcleo do citoplasma. Logo após, são extraídas características sobre a textura das imagens segmentadas. Após a realização destas etapas, é realizado a classificação em 7 possíveis classes, utilizando o SVM (*Support Vector Machine*). É ressaltado que não existe um conjunto exclusivo de características capazes de fornecer diagnósticos eficientes para todas e quaisquer classes de um sistema de classificação proposto, no entanto, o SVM foi demonstrado como uma solução eficiente para o problema proposto.

Antunes (2021) propõe uma classificação automática de recortes de células cervicais com base em característica não geométricas, são utilizado métodos tradicionais na sua arquitetura interna, seguindo o seguinte fluxo, as imagens são retiradas do banco de dados CRIC, logo após essa imagens são recortadas em sub-imagens com dimensão 50x50, pelo fato de utilizar algoritmos tradicionais, a imagem precisa passar por uma extração de características, é aplicado diversas metodologias para conseguir extrair o máximo de informação daquela imagem, logo em seguida é feito o treinamento utilizando o SVM, depois são realizados testes aplicando-se a validação cruzada, sendo 20% do total das imagens do banco de dados, utilizadas somente para teste, por fim é extraída métricas sobre a eficiência do sistema, onde se pode concluir que a abordagem teve ótimos resultados.

2.2 Fundamentação Teórica

2.2.1 Rede Neural Convolucional

Rede neural convolucional é um algoritmo de aprendizado profundo, que comparado aos demais obtém uma grande vantagem pelas particularidades de ser capaz de construir filtros que se baseiam nas características da imagem de entrada dispensando a necessidade de préprocessamento. Ademais na rede neural convolucional as imagens percorrem diversas camadas de convolução, cada camada será responsável pela extração de um conjunto de características das amostras da base de dados. Dessa maneira o seu desempenho no reconhecimento de imagens é uma das principais razões da eficácia comprovada do aprendizado profundo. A Figura 2.1 é um exemplo de uma rede neural convolucional.



Figura 2.1 – Exemplo de rede neural convolucional e suas diferentes camadas (VARGAS; PAES; VASCONCELOS, 2016).

Para Vargas, Paes e Vasconcelos (2016) uma rede neural convolucional consiste em múltiplas camadas com funções diferentes. Inicialmente é comum aplicar sobre o dado de entrada camadas ditas de convolução. Uma camada de convolução é composta por diversos neurônios, cada um é responsável por aplicar um filtro em um pedaço específico da imagem. Podemos

imaginar cada neurônio sendo conectado a um conjunto de *pixels* da camada anterior e que a cada uma dessas conexões se atribui um peso. A combinação das entradas de um neurônio, utilizando os pesos respectivos de cada uma de suas conexões, produz uma saída passada para a camada seguinte. Os pesos atribuídos às conexões de um neurônio podem ser interpretados como uma matriz que representa o filtro de uma convolução de imagens no domínio especial (conhecido também como kernel ou máscara).

As redes neurais convolucionais foram propostas em 1998 (LECUN et al., 2015), quando ele propôs a arquitetura *LeNet*. Focado no problema de identificação de códigos postais, sendo o trabalho pioneiro nesse ramo. Contudo como o método foi proposto em 1998, e existindo uma grande limitação de *hardware* da época, mais especialmente na parte das unidades gráficas de processamento (*Graphics Processing Unit - GPU*)) a *LeNet* e as redes neurais convolucionais foram deixadas em *stand-by* devido a sua inviabilidade de implementação. Sendo novamente inseridas somente em 2012 com o novo método de implementação *AlexNet* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) foi considerada um momento histórico para a Inteligência Artificial em geral, quando conseguiu conquistar o primeiro lugar no campeonato de reconhecimento visual em grande escala realizada pela *ImageNet*, a arquitetura é estruturada com 8 camadas de convolução e 60 milhões de parâmetros, além disso ela conta com a Unidade Linear Retificada (*(Rectified Linear Unit - ReLU)*) omo função de ativação, e para conter o sobre-ajuste é usado o *Dropout*. Como o passar dos tempos várias outras arquiteturas continuam sendo propostas, entre elas vale a pena ressaltar *VGGNet*(SIMONYAN; ZISSERMAN, 2014) e *ResNet*(HE et al., 2016) que são arquiteturas globalmente conhecidas e diariamente utilizadas, e também se destacaram devido a performance delas no mesmo campeonato de reconhecimento visual em grande escala da *ImageNet*.

2.2.1.1 Camada de Convolução

A imagem é uma matriz que possui propriedades como largura, altura e profundidade, onde cada elemento representa uma intensidade dos pixels, que pode variar dentro do intervalo de 0 a 255. A profundidade irá determinar a quantidade de canais de cores disponíveis, caso a imagem estiver em escalas de cinza ela terá apenas um canal, já nos outros casos se tratando de RGB (*Red, Green, Blue*) terá a sua disposição 3 canais uma para cada faixa de cor. Para ser possível a extração de características de uma imagem se faz necessário a aplicação de *kernels* ou filtros, através da convolução. A camada de convolução tem como função aplicar máscaras nas matrizes de entrada tendo considerado uma vizinhança já pré-definida, assim será produzido um mapa de características que armazenará os pesos das conexões entre os neurônios. Na Figura 2.2 é representado um exemplo do cálculo da convolução, sendo considerado uma matriz de dimensão 6x6 e sua máscara de 3x3. A convolução além desses elementos pode incluir um parâmetro chamado de *Stride*, que tem como função indicar quantos passos a máscara é movida ao longo

do processo de entrada. Outro parâmetro é o *padding* que indica o número que será inserido caso a entrada não possua elementos suficientes, no exemplo demonstrado o *Stride* equivalente é igual a 1, e não se faz necessário o valor de *padding*.



Figura 2.2 – Exemplo de um processo de convolução (MARIED et al., 2017)

2.2.1.2 Processo de Pooling

A camada de Subamostragem ou simplesmente processo de *polling*, é responsável por representar a redução da dimensionalidade do mapa de características, para assim diminuir o número de parâmetros e consequentemente também a complexidade do modelo (LOY, 2019). Nas Redes Neurais Convolucionais é comum utilizar a função *Max Pool*, que tem como objetivo capturar os pixels mais significativos da imagem, sendo que ele selecionará o maior valor de uma vizinhança que já foi pré-definida. A Figura 2.3 representa o antes e o depois da operação *Max Pool* sobre o mapa de características.



Figura 2.3 – Exemplo de uma utilização do *Max Polling* de dimensão 2x2 e com salto 2 sobre o mapa de características (SHARMA; MANGAT, 2016)

2.2.2 Métricas de Avaliação

Para melhor análise dos resultados e comparação com algoritmos desenvolvidos por outra abordagem (DINIZ et al., 2021a), os resultados passam por diferentes métricas de avaliação, cujo o propósito é conseguir saber o quão assertivo será o algoritmo, para isso ele é submetido a 5 métricas de avaliação.

Para calcular essas 5 métricas, é preciso mensurar 4 variáveis, sendo elas: o Verdadeiro Positivo (*True positive - TP*), que se refere ao número de núcleos alterados corretamente classificado como anormal, o Verdadeiro Negativo (*True negative - TN*), que se refere aos núcleos normais corretamente classificado como normal, o Falso Positivo (*False positive - FP*), que se refere aos núcleos que são normais porém foram classificados como anormal, e por fim o Falso Negativo (*False Negative - FN*), que por sua vez se refere aos núcleos que possuem alteração e foram classificados como normais.

2.2.2.1 F1-Score

F1-Score é a relação dada entre a precisão e revocação, no qual tem o intuito de fornecer uma medição única e balanceada para o algoritmo. A Equação está descrita em 2.1.

$$F1 = 2 * \frac{Precisao * Revocacao}{Precisao + Revocacao}$$
(2.1)

2.2.2.2 Acurácia

O intuito da acurácia é apontar a performance geral do algoritmo, o seu resultado indicará a porcentagem de classificações verdadeiramente negativas e positivas dentre toda a classificação dos conjuntos de dados. A Equação está descrita em 2.2.

$$Acuracia = \frac{TP + TN}{TP + TN + FP + FN}$$
(2.2)

2.2.2.3 Revocação

A revocação é a porcentagem de dados positiva que foram corretamente classificados como positivos. Ela é mais indicada nas situações em que os falsos negativos (FN) são mais danosos ao sistema que os falsos positivos (FT). A Equação está descrita em 2.3:

$$Revocacao = \frac{TP}{TP + FN}$$
(2.3)

2.2.2.4 Precisão

O intuito da precisão é expor dentre as classificações positivas concluídas do nosso algoritmo, quais estão corretas, sendo assim uma opção aconselhável quando os falsos positivos (FP) são mais prejudiciais do que os falsos negativos (FN). A Equação está descrita em 2.4:

$$Precisao = \frac{TP}{TP + FP}$$
(2.4)

2.2.2.5 Especifidade

Por fim, a última métrica extraída é a especifidade, que tem como intuito colher as porcentagem dos dados normais (sem lesão) que são corretamente classificados como verdadeiro negativo (TN). A Equação está descrita em 2.5:

$$Especifidade = \frac{TN}{TN + FP}$$
(2.5)

3 Desenvolvimento

Neste capítulo são descritas as fases do desenvolvimento desta monografia. Na Seção 3.1 é especificada a base de dados. Na Seção 3.2 é mostrado passo a passo como foi realizada a etapa do pré-processamento das imagens. E por fim, na Seção 3.3 são evidenciadas as arquiteturas dos modelos escolhidos para representar a rede neural convolucional de classificação e suas particularidades.

3.1 Base de Dados Cric Cervix

A base de dados selecionada para a extração das imagens dessa monografia, juntamente dos núcleos celulares é a CRIC (REZENDE; BIANCHI; CARNEIRO, 2021), essa coleção de imagens de células cervicais, foram segmentadas e classificadas por citopatologistas independentes. A base de dados contém 11.534 núcleos e citoplasmas que foram obtidos através de 400 imagens de exame do Papanicolaou. Na Figura 3.1 temos um exemplo de imagem extraida da base de dados. Os dados da base da dados são imagens completas de um exame, mapeando a posição x e y de cada célula existente. As células dessa coleção são classificadas em seis classes, que estão representadas na Figura 3.2 sendo elas:

- 1. Negativo para lesão intraepitelial ou malignidade (NILM);
- Células escamosas atípicas de significado indeterminado, possivelmente não neoplásicas (ASC-US);
- 3. Lesão intraepitelial escamosa de baixo grau (LSIL);
- 4. Células escamosas atípicas, não pode excluir lesão de alto grau (ASC-H)
- 5. Lesão intraepitelial escamosa de alto grau (HSIL);
- 6. Carcinoma de células escamosas (SCC)



Figura 3.1 – Imagem retirada da base de dados CRIC Cervix.



Figura 3.2 – Classificações geradas da base de dados CRIC Cervix.

Como explicado anteriormente a base da dados possui seis classes de classificação diferentes, porém neste trabalho as classes serão aglomeradas em apenas duas classes, células sem lesão ou normal (NILM) e células com lesão ou anormal (LSIL, HSIL, ASC-H, ASC-US, NILM). A base é composta por imagens contendo um número variável de núcleos os que nos leva a ter mais núcleos de certas classificações, tendo bem mais núcleos normais do que anormais,

deixando assim a base desbalanceada, a distribuição do numero de lesão celulares em cada classe está disponível na Tabela 3.1.

Classes	Categoria	Tipo de célula	Total de lesões
1	Anormal	LSIL	1.360
2	Anormal	HSIL	1.703
3	Anormal	ASC-H	925
4	Anormal	ASC-US	606
5	Anormal	SCC	161
6	Normal	NILM	6.779

Tabela 3.1 – Distribuição de núcleos por classe CRIC Cervix.

3.2 Pré-processamento

No pré-processamento foram adotadas duas estratégias diferentes em cima das imagens para geração do conjunto de treino e teste, gerando dois tipos de imagens, sendo elas imagens com o núcleo centralizado, e imagens com o núcleo descentralizado.

3.2.1 Núcleo Centralizado

Na base de dados escolhida para este trabalho, para cada imagem de exame do Papanicolaou, é mapeada a posição x e y de cada célula presente na imagem, tendo o núcleo como ponto de referência. Para a geração das imagens com o núcleo centralizado, foram criadas caixas de quatro tamanhos distintos (30x30, 50x50, 70x70 e 90x90), a partir do ponto x e y fornecido pela base de dados, dessa maneira se garante que no centro da imagem sempre esteja o núcleo.

A Figura 3.3 exemplifica o processo de geração das imagens, utilizando um tamanho de caixa 90x90, na qual os quadrados verdes são as células classificadas como sem lesão e os quadrados vermelhos são as células classificadas como com lesão. Já a Figura 3.4 representa um exemplo de uma mesma célula em diferentes tamanhos de caixa.



Figura 3.3 – Exemplo da geração de imagens com o núcleo centralizado e caixas de 90x90.



Figura 3.4 – Diferentes tamanhos de caixa para a mesma célula.

Após a criação das imagens, para cada tamanho de imagem 30x30, 50x50, 70x70 e 90x90 foram gerados os mesmos números de imagens, na qual está representada na Tabela 3.2.1. Para a elaboração do modelo final adotamos a estratégia 70-30, que consiste em separar 70% das imagens para treino e 30% das imagens para teste.

Tabela 3.2 – Distribuição das células centralizada									
	Conjunto	Com Lesão	Sem Lesão						

Conjunto	Com Lesão	Sem Lesão
Treinamento	3.433	4.745
Teste	1.322	2.034
Total	4.755	6.779

3.2.2 Núcleo Descentralizado

Para as imagens com o núcleo descentralizado, foi adotada uma abordagem diferente, pois necessitamos ter imagens na qual o núcleo não estivesse no centro da imagem, para isso geramos várias caixas em sequência na imagem do exame de Papanicolaou, e para cada caixa era consultado se existia alguma célula com ou sem lesão ali dentro, caso existisse essa caixa era salva e a imagem faria parte do conjunto de treinamento ou teste, as caixas que não possuíam célula dentro delas eram salvas com o rótulo de não célula, pois a classe não célula é uma das classes que foram incluídas no modelo. Além disso, em alguns casos o núcleo da célula se encontrava no limite do tamanho da caixa, para que isso não acontecesse adotamos uma estratégia de fronteira, na qual a caixa só é salva se o núcleo estiver a pelo menos 10 *pixels* de distância das suas bordas.

A Figura 3.5 representa o processo de geração das imagens descentralizadas, as caixas amarelas representam as imagens classificadas como não célula, as vermelhas como com lesão e as verdes como sem lesão. Na Figura 3.6 é mostrada a estratégia, na imagem mais à esquerda sem o uso da fronteira, mais caixas são criadas nas imagens, a da direita, considerando a política da fronteira, várias caixas deixam de ser criadas por não estar respeitando o limite de 10 *pixels* estipulado, vale ressaltar que no exemplo não foi mapeado a classe não célula apenas para facilitar a visualização, dessa maneira apenas mapeamos as classes com lesão e sem lesão.



Figura 3.5 – Processo de geração das imagens descentralizadas.



Figura 3.6 – Exemplo visual do uso da técnica da fronteira.

Diferentemente da geração das imagens centralizadas, nas imagens descentralizadas houve variações na quantidade de caixas contendo células, isso se deve por causa da estratégia da fronteira, já que cada caixa varia o tamanho, podendo haver tamanhos de caixas que geraram mais imagens, e tamanhos de caixa que geraram menos imagens, dessa maneira será separado uma tabela para cada uma das quatro tamanhos de caixa para saber como ficou a distribuição de classes.

Conjunto	Com Lesão	Sem Lesão	Não Célula
Treinamento	1.614	2.248	6.567
Teste	692	964	2.814
Total	2.306	3.212	9.381

Tabela 3.3 – Distribuição de classes descentralizadas para tamanho de caixa 30x30.

Tabela 3.4 – Distribuição de classes descentralizadas para tamanho de caixa 50X50.

Conjunto	Com Lesão	Sem Lesão	Não Célula
Treinamento	1.225	1.732	6.560
Teste	532	743	2.812
Total	1.757	2.475	9.372

Tabela 3.5 – Distribuição de classes descentralizadas para tamanho de caixa 70X70.

Conjunto	Com Lesão	Sem Lesão	Não Célula
Treinamento	1.539	2.106	6.502
Teste	659	902	2787
Total	2.198	3.008	9.289

Tabela 3.6 – Distribuição de classes descentralizadas para tamanho de caixa 90X90.

Conjunto	Com Lesão	Sem Lesão	Não Célula
Treinamento	1.582	2.372	6.582
Teste	678	1.017	3.206
Total	2.260	3.389	9.788

Depois de gerados as caixas de informações das células descentralizadas para todos tamanhos de caixa, foi preciso balancear o banco de dados para prosseguir com o treinamento, já que cada classe possuía um número diferente de instâncias, sendo a classe com lesão a limitadora, dessa forma foi somado a classe com lesão centralizada e com lesão descentralizada gerando o número total da classe, esse número foi usado para limitar as outras classes, sendo assim algumas instancias foram deletadas a fim de aproximar a quantidade de imagens das classes deixando o banco de dados balanceado. As tabelas a seguir representam o número total de instâncias por classe depois de balanceada para cada tamanho de caixa.

Set	Com Lesão	Sem Lesão	Não Célula
Treinamento	4943	4970	4956
Teste	2118	2130	2124
Total	7061	7100	7080

Tabela 3.7 – Distribuição final das classes para tamanho de caixa 30x30.

Tabela 3.8 – Distribuição final das classes para tamanho de caixa 50x50.

Set	Com Lesão	Sem Lesão	Não Célula
Treinamento	4558	4578	4571
Teste	1954	1962	1959
Total	6512	6540	6530

Tabela 3.9 – Distribuição final das classes para tamanho de caixa 70x70.

Set	Com Lesão	Sem Lesão	Não Célula
Treinamento	4867	4886	4893
Teste	2086	2094	2097
Total	6953	6980	6890

Tabela 3.10 – Distribuição final das classes para tamanho de caixa 90x90.

Set	Com Lesão	Sem Lesão	Não Célula
Treinamento	4924	4935	4942
Teste	2111	2115	2118
Total	7035	7050	7060

3.3 Arquitetura das Redes Neurais Convolucionais

Com a evolução da inteligência artificial, diferentes arquiteturas de rede neural convolucional têm sido utilizadas para classificar imagens. Como esse trabalho tem como objetivo a investigação do uso de diferentes modelos para o processamento das imagens, foram utilizados quatro modelos diferentes sendo eles (*EfficientNet*, *MobileNet*, *XceptionNet* e *Inception V3*) a razão pelo qual eles foram escolhidos, pois eles são o mesmos utilizados em (DINIZ et al., 2021a) e temos o interesse em comparar os resultados obtidos, já que serão utilizados os mesmos modelos variando-se apenas o pré-processamento.

3.3.1 Inception V3 e XceptionNet

A *Inception V-3* é uma arquitetura focada na resolução de classificação de imagens que foi treinada no conjunto de dados *ImageNet* (SZEGEDY et al., 2016). Esta rede foi apresentada em 2015, tendo bastante relevância pelo seu bom desempenho atrelado ao baixo custo computacional,

isso se deve ao fato dela ter reduzido os parâmetros estimados pela rede, fazendo com que a rede possua um melhor desempenho computacional quando comparado às redes VGG durante seu treinamento.

Umas das principais características da arquitetura *Inception*, representada na Figura 3.9, é atuar como um extrator de características em vários níveis de convolução 1x1, 3x3 e 5x5 dentro do mesmo módulo da rede. A razão de ter sido a escolhida para o desenvolvimento desta monografia está relacionada a sua ótima taxa de acurácia e principalmente ao baixo custo computacional.



Figura 3.7 – Fluxograma da arquitetura interna da Inception V3 (SZEGEDY et al., 2016).

Já a arquitetura *XceptionNet* desenvolvida por (CHOLLET, 2017) é uma variação da *Inception*. Possuindo o mesmo número de parâmetros, porém com um maior ganho de desempenho, que se deve ao uso mais eficiente dos parâmetros do modelo.

3.3.2 EfficientNet

A arquitetura *EfficientNet* desenvolvida por (TAN; LE, 2019) é considerada uma família de modelos de redes neurais projetadas por meio de uma pesquisa de arquitetura neural, na qual o bloco de construção básico é o *Mobile Inverted Bottleneck Conv Block*, sendo ampliada em 3 dimensões: profundidade, largura e resolução, a arquitetura possui 8 variações, *EfficientNetBO* até *EfficientNetB8*, sendo a B0 a versão base da implementação, as outras versões aumentam os parâmetros de entrada gradualmente a cada aplicação. As *EfficientNetB7* que atinge uma precisão de 84,4% para a base de dados *ImageNet*, sendo 8,4x menor e 6,1x mais rápido na inferência do que a melhor *ConvNet* existente em (HUANG et al., 2019).



Figura 3.8 – Fluxograma da arquitetura interna da *EfficientNet*.(AHMED; SABAB, 2020)

3.3.3 MobileNet

A arquitetura *MobileNet* desenvolvida por (HOWARD et al., 2017) é também considerada uma família de modelos de redes neurais que tem como principal característica a aplicação de modelos para computação móvel e embarcada. As *MobileNets* são baseadas em uma arquitetura simplificada que usa convoluções separáveis em profundidade para construir redes neurais profundas leves. Tendo como seu principal diferencial a camada de *Depthwise Separable Convolution* que é uma camada onde é aplicado uma sequência de normalização e ReLU. Sendo a particularidade responsável de tornar o modelo mais leve computacionalmente.



Figura 3.9 – (A) Arquitetura da *MobileNet* (B) Explicação da camada de *Depthwise Separable Convolution*. (SINGH et al., 2015)

4 Resultados

Neste Capítulo são apresentados os resultados que foram obtidos no desenvolvimento deste trabalho. A Seção 4.1 mostra a descrição dos experimentos. A Seção 4.2 mostra os resultados obtidos através do tamanho de caixa 30x30. A Seção 4.3 por sua vez irá mostrar os resultados obtidos pelo tamanho de caixa 50x50. A Seção 4.4 os resultados do tamanho de caixa 70x70. A Seção 4.5 os resultados do tamanho de caixa 90x90. A Seção 4.6 abordará uma discussão sobre os resultados das caixas. E por final, a Seção 4.7 irá mostrar os resultados na hora da predição e a distribuição de classes de uma imagem desconhecida.

4.1 Descrição dos Experimentos

Para o desenvolvimento deste trabalho os experimentos foram executados pela máquina virtual do *Google Colaboratory*. As particularidades do *Hardware* estão presentes na Tabela 4.1. Por sua vez, a implementação dos modelos das redes neurais convolucionais foram desenvolvidas utilizando a biblioteca *TensorflowKeras*.

Parâmetros	Especificação
GPU	Nvidia K80
Mémoria da GPU	12 GB
Performance	4.1TFLOPS
Número de núcleos da GPU	2
Memória RAM	12GB
Memória em Disco	358 GB

Tabela 4.1 – Especificação do *Hardware* utilizado.

Para todos os modelos com seus diferentes tamanho de caixas, foram definidas algumas variáveis padrão, todos os experimentos executaram 40 épocas, tendo como função de ativação o *Softmax*. Já como otimizador foi utilizado a função de Adam, como já mencionado a base de dados foi dividida em 70% para treino e 30% para teste. Por fim o tamanho do lote foi fixado em 32 para todos os experimentos.

4.2 Tamanho de Caixa 30x30

No tamanho de caixa 30x30, nem todos os modelos estavam aptos a receber imagens desse tamanho, dessa maneira apenas o modelo *EfficientNet* estava apto a processar essas imagens, na Tabela 4.2 estão representadas as métricas encontradas pelo modelo. A Figura 4.1 representa a matriz de confusão do modelo, que representa a distribuição das classes que o modelo inferiu.

Arquitetura	Precisão	Acurácia	F1-Score	Revocação	Especifidade
EfficientNet	0,87	0,91	0,87	0,87	0,93

Tabela 4.2 – Métricas do modelo *EfficientNet* para o tamanho de caixa 30X30.





Figura 4.1 – Matriz de confusão do modelo *EfficientNet* para o tamanho de caixa 30x30.

4.3 Tamanho de Caixa 50x50

Para o tamanho de caixa 50x50, dois modelos estão aptos a processar as imagens, sendo eles o modelo EfficientNet e MobileNet. na Tabela 4.3 estão representadas as métricas encontradas pelo modelos. A Figura 4.2 representa a matriz de confusão do modelo EfficientNet, que obteve resultados ligeiramente melhores que o modelo MobileNet, tendo quase todas as métricas 0.1% maior e com o mesmo valor de acurácia.

Tabela 4.3 – Métricas dos modelos para o tamanho de caixa 50X50.

Arquitetura	Precisão	Acurácia	F1-Score	Revocação	Especifidade
EfficientNet	0,9	0,93	0,9	0,9	0,95
MobileNet	0,89	0,93	0,89	0,89	0,94



Figura 4.2 – Matriz de confusão do modelo *EfficientNet* para o tamanho de caixa 50x50.

4.4 Tamanho de Caixa 70x70

No tamanho de caixa 70x70, todos os modelos estão aptos a processar as imagens, na Tabela 4.4 estão representadas as métricas encontradas pelo modelos. A Figura 4.4 representa a matriz de confusão do modelo *EfficientNet*, que obteve novamente o melhor resultados, nota-se dessa vez que todos os modelos tiveram valores muitos próximos variando no máximo 0,3% entre o melhor e o pior modelo.

Arquitetura	Precisão	Acurácia	F1-Score	Revocação	Especifidade
EfficientNet	0,9	0,94	0,9	0,9	0,95
MobileNet	0,89	0,92	0,89	0,89	0,94
InceptionV3	0,88	0,92	0,88	0,88	0,94
XceptionNet	0,9	0,93	0,9	0,9	0,95

Tabela 4.4 – Métricas dos modelos para o tamanho de caixa 70x70.



Figura 4.3 – Matriz de confusão do modelo *EfficientNet* para o tamanho de caixa 70x70.

4.5 Tamanho de Caixa 90x90

Por fim, para o último tamanho de caixa, todos modelos também estão aptos a processar imagens. Na Tabela 4.5 estão representadas as métricas encontradas pelo modelos. A Figura 4.4 representa a matriz de confusão do modelo *XceptionNet*, que nesse tamanho de caixa obteve os melhores resultados, consegue-se perceber que nesse tamanho de caixa os modelos variaram mais entre si, tendo uma diferença de 0,8% do melhor modelo para o pior.

Arquitetura	Precisão	Acurácia	F1-Score	Revocação	Especifidade
EfficientNet	0,86	0,91	0,86	0,86	0,93
MobileNet	0,87	0,92	0,87	0,87	0,94
InceptionV3	0,81	0,87	0,81	0,81	0,9
XceptionNet	0,89	0,93	0,89	0,89	0,95

Tabela 4.5 – Métricas dos modelos para o tamanho de caixa 90x90.



Figura 4.4 – Matriz de confusão do modelo *XceptionNet* para o tamanho de caixa 90x90.

4.6 Discussão dos Resultados

Nesta Seção é feita uma breve discussão sobre os resultados encontrados pelos diferentes modelos e diferentes tamanhos de caixa. Nota-se que os melhores resultados foram encontrados com o **tamanho de caixa 70x70** com a utilização do modelo *EfficientNet*, os modelos estavam adaptando melhor a tamanhos de caixa maiores, pois foi notável que houve um salto de melhora do tamanho de caixa 30x30 para o tamanho de caixa 50x50 e um outro salto de melhora no tamanho de caixa 70x70. Porém ao chegar ao tamanho de caixa 90x90, os modelos se diferenciam mais entre si, encontrando resultados variados e em média todos os modelos obtiveram piores resultados em comparação ao tamanho de caixa 70x70. Dessa maneira para o desenvolvimento do sistema final, foi feito o uso do modelo *EfficientNet* e o tamanho de caixa 70x70.

4.7 Predição do Sistema

Nesta Seção será evidenciado como o modelo se comportou na hora de fazer predições em imagens desconhecidas, evidenciando como ele classificou as diferentes caixas dentro das imagens. Para isso utilizamos a mesma lógica utilizada para geração das imagens descentralizadas. sendo assim, a primeira coisa a se fazer foi redimensionar a imagem original para um tamanho que seja possível realizar caixas de 70x70 por toda a imagem, sem deixar de cobrir nenhuma parte da imagem.

Depois disso, cada caixa gerada foi passada para a inferência do modelo. Caso ele conclua como com lesão, foi desenhado um quadrado vermelho em torno da caixa, um quadrado verde para não lesão, seguindo o mesmo padrão adotado pela base de dados CRIC, e para os casos que foi concluído como não célula não foi desenhado nenhum quadrado, a fim de facilitar a visualização das células e não poluir a imagem com muita informação. A Figura 4.5 representa a verdadeira classificação de uma imagem presente na base da dados CRIC. Já a Figura 4.6 representa a mesma imagem da Figura anterior e como ela foi classificada pelo modelo proposto.



Figura 4.5 – Exemplo de uma imagem da base da dados CRIC com sua verdadeira classificação.



Figura 4.6 – Predição e distribuição das classes do sistema proposto.

Ao comparar as duas figuras conseguimos perceber que o modelo proposto conseguiu identificar muito bem as células, deixando de mapear apenas duas células, obtendo um excelente desempenho em diferenciar célula de não célula, porém de todas as quatro células que ele reconheceu ele concluiu todas como com lesão, sendo que somente uma era realmente classificada como com lesão, e as outras três não lesão. Dessa forma, conseguimos perceber que o modelo se confundiu na hora de concluir o tipo de lesão da célula.

Uma possível causa para isso seria o modelo ter sido treinado com essas três classes juntas, sendo elas com lesão, não lesão e não célula, para uma possível melhora poderia ter sido desenvolvidos dois modelos distintos, um responsável para identificar células tendo somente duas classes, sendo elas célula e não célula, e das caixas que esse modelo concluir como célula seria passado para o segundo modelo que seria treinado para diferenciar células com lesão de células sem lesão.

5 Considerações Finais

Neste capítulo são apresentadas as considerações finais sobre esta monografia. Na Seção 5.1 é apresentado a conclusão do trabalho e na Seção 5.2 será apresentado os trabalhos futuros.

5.1 Conclusão

É de suma consciência que classificar corretamente núcleos celulares em imagens reais do exame de Papanicolau, é uma tarefa árdua e custosa, sendo uma condição necessária para fornecer melhores diagnósticos. Dessa maneira neste trabalho é proposta a elaboração de um sistema de apoio para classificação, sem a necessidade de usar imagens segmentadas para treinamento.

Os modelos desenvolvidos variando os diferentes tamanhos de caixas conseguiram ótimos resultados, levando em consideração que em Diniz et al. (2021a) são utilizadas imagens segmentadas para classificação de um sistema de duas classes, tendo o seu melhor modelo atingido 0,96 de acurácia. Já o nosso melhor modelo foi o *EfficientNet* juntamente do tamanho de caixa 70x70 atingiu uma acurácia de 0,94. Dessa maneira, conclui-se que é possível ter métricas com valores adequados sem a necessidade de utilizar imagens segmentadas.

Em relação a investigação do uso dos modelos e os tamanhos de caixa, conseguimos perceber que na maioria dos modelos todos os resultados ficaram bem próximos uns dos outros, sendo o modelo *EfficientNet* que alcançou os melhores resultados em dois tipos de tamanho de caixa. Já em relação ao tamanho das caixas, a que gerou melhores resultados foi a de tamanho 70x70.

Dessa maneira, podemos dizer que nosso objetivo principal foi concluído, que foi a elaboração de um modelo, que ajudasse na tomada de decisão dos profissionais da saúde, levando em consideração que o modelo alcançou resultados excelentes, mas manter essas métricas na hora de inferir uma imagem totalmente desconhecida ainda é um grande desafio.

5.2 Trabalhos Futuros

Como trabalhos futuros, pretende-se justamente a elaboração dos dois modelos sugeridos na Seção 4.7 a fim de investigar se haverá uma melhora na diferenciação entre células com lesão de células sem lesão. Além disso, outro fator importante seria avaliar tamanhos de caixas maiores que 90x90, para saber se poderiam obter melhores resultados.

Referências

AHMED, T.; SABAB, N. Classification and understanding of cloud structures via satellite images with efficientunet. In: [S.l.: s.n.], 2020.

ANTUNES, D. W. Classificação automática de recortes de células cervicais baseada em características não geométricas. 2021. Disponível em: http://www.monografias.ufop.br/handle/35400000/3075>.

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings* of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.: s.n.], 2017.

DINIZ, D. N.; REZENDE, M. T.; BIANCHI, A. G. C.; CARNEIRO, C. M.; LUZ, E. J. S.; MOREIRA, G. J. P.; USHIZIMA, D. M.; MEDEIROS, F. N. S. de; SOUZA, M. J. F. A deep learning ensemble method to assist cytopathologists in pap test image classification. *Journal of Imaging*, v. 7, n. 7, 2021. ISSN 2313-433X. Disponível em: https://www.mdpi.com/2313-433X/7/7/111.

DINIZ, D. N.; REZENDE, M. T.; BIANCHI, A. G. C.; CARNEIRO, C. M.; USHIZIMA, D. M.; MEDEIROS, F. N. S. de; SOUZA, M. J. F. A hierarchical feature-based methodology to perform cervical cancer classification. *Applied Sciences*, v. 11, n. 9, 2021. ISSN 2076-3417. Disponível em: https://www.mdpi.com/2076-3417/11/9/4091>.

GÓMEZ, O. H.; SÁNCHEZ-DELACRUZ, E.; MATA, A. P. de la. Classification of cervical cancer using assembled algorithms in microscopic images of papanicolaou. *Res. Comput. Sci.*, v. 139, p. 125–134, 2017.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *CoRR*, abs/1704.04861, 2017. Disponível em: http://arxiv.org/abs/1704.04861>.

HUANG, Y.; CHENG, Y.; BAPNA, A.; FIRAT, O.; CHEN, D.; CHEN, M.; LEE, H.; NGIAM, J.; LE, Q. V.; WU, Y.; CHEN, z. Gpipe: Efficient training of giant neural networks using pipeline parallelism. In: WALLACH, H.; LAROCHELLE, H.; BEYGELZIMER, A.; ALCHé-BUC, F. d'; FOX, E.; GARNETT, R. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019. v. 32. Disponível em: https://proceedings.neurips.cc/paper/2019/file/093f65e080a295f8076b1c5722a46aa2-Paper.pdf>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGES, C.; BOTTOU, L.; WEINBERGER, K. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. v. 25. Disponível em: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.

LECUN, Y. et al. Lenet-5, convolutional neural networks. URL: http://yann. lecun. com/exdb/lenet, v. 20, n. 5, p. 14, 2015.

LOY, J. Neural Network Projects with Python: The ultimate guide to using Python to explore the true power of neural networks through six projects. [S.1.]: Packt Publishing Ltd, 2019.

MARIARPUTHAM, E. J.; STEPHEN, A. Nominated texture based cervical cancer classification. *Comput Math Methods Med*, v. 2015, p. 586928, jan. 2015.

MARIED, E.; ELDALI, M.; ZIADA, O.; BABA, A. A Literature Study of Deep learning and its application in Digital Image Processing. [S.l.], 2017.

REZENDE, M. T.; BIANCHI, A. G.; CARNEIRO, C. M. Cervical cancer: Automation of pap test screening. *Diagnostic Cytopathology*, Wiley Online Library, v. 49, n. 4, p. 559–574, 2021.

SHARMA, B.; MANGAT, K. K. An improved nucleus segmentation for cervical cell images using fcm clustering and bpnn. In: IEEE. 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). [S.1.], 2016. p. 1924–1929.

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

SINGH, D.; VERMA, A.; ANEJA, M.; SINGH, B. Cervical cell classification using random forests. In: . [S.l.: s.n.], 2015.

SUNG, H.; FERLAY, J.; SIEGEL, R. L.; LAVERSANNE, M.; SOERJOMATARAM, I.; JEMAL, A.; BRAY, F. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, Wiley Online Library, v. 71, n. 3, p. 209–249, 2021.

SZEGEDY, C.; VANHOUCKE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.1.: s.n.], 2016.

TAN, M.; LE, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In: CHAUDHURI, K.; SALAKHUTDINOV, R. (Ed.). *Proceedings of the 36th International Conference on Machine Learning*. PMLR, 2019. (Proceedings of Machine Learning Research, v. 97), p. 6105–6114. Disponível em: https://proceedings.mlr.press/v97/tan19a.html.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: SN. *Proceedings of the xxix conference on graphics, patterns and images.* [S.I.], 2016. v. 1, n. 4.

World Health Organization. *Global Cancer Statistics*. 2018. http://gco.iarc.fr/. Disponível em: <http://gco.iarc.fr/>.