

Renesio Joaquim de Oliveira Costa
Orientador: Flávio Henrique Duarte de Araújo

Redes Neurais Convolucionais para Identificação de Metástase em Imagens Histopatológicas.

Picos - PI
Dezembro 2019

Renesio Joaquim de Oliveira Costa
Orientador: Flávio Henrique Duarte de Araújo

Redes Neurais Convolucionais para Identificação de Metástase em Imagens Histopatológicas.

Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Universidade Federal do Piauí
Campus Senador Helvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
Dezembro 2019

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

C837r Costa, Renesio Joaquim de Oliveira.
Redes neurais convolucionais para identificação de metástase em imagens histopatológicas.. / Renesio Joaquim de Oliveira Costa. -- Picos,PI, 2019.
34 f.
CD-ROM: 4 ¾ pol.

Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação). – Universidade Federal do Piauí, Picos, 2019.
“Orientador(A): Prof. Flávio Henrique Duarte de Araújo.”

1. PatchCamelyon. 2. PCam. 3. Histopatologia – Imagens Digitais. I. Título.

CDD 005

Elaborada por Rafael Gomes de Sousa CRB 3/1163

REDES NEURAIS CONVOLUCIONAIS PARA IDENTIFICAÇÃO DE METÁSTASE EM
IMAGENS HISTOPATOLÓGICAS

RENESIO JOAQUIM DE OLIVEIRA COSTA

Monografia aprovada como exigência parcial para obtenção do
grau de Bacharel em Sistemas de Informação.

Data de Aprovação

Picos – PI, 03 de dezembro de 2019

Flávio Henrique Duarte de Araújo

Prof(a). Flávio Henrique Duarte de Araújo

Thiago José Barbosa Lima

Prof(a). Thiago José Barbosa Lima

Daniel de Sousa Luz

Prof(a). Daniel de Sousa Luz

Agradecimentos

Quero agradecer primeiramente ao Senhor Deus todo poderoso a ele toda honra e glória, para todo o sempre. Amém.

Aos meus pais Joaquim Genésio da Costa e Maria dos Remédios Feitosa de Oliveira Costa quero agradecer por sempre me darem apoio na realização das minhas conquistas e por sempre estarem comigo em todos os momentos. O maior dos agradecimentos a minha namorada Hayla Daphne Albuquerque de Oliveira Leal, por me inspirar a cada dia a ser melhor do que eu era, e me incentivar a sempre superar a mim mesmo. Agradeço também as minhas irmãs Rayllanny Maria, Iryslanny Maria e Renyelly Maria, minha avós paterna e materna quero aqui agradecer por tudo que as senhoras já fizeram por mim, pelos cuidados por ter me educado e sempre ter zelado por meu bem estar. A todos, meus singelos agradecimentos.

Quero agradecer também aos meus amigos que estavam aqui comigo, que ajudaram nesta jornada, agradecer principalmente a Jonycássio almondes, ele que enfrentou comigo essa jornada árdua e tortuosa. Quero agradecer a cada um dos amigo do grupo "*Xetelas*", Aroldo Júnior pelos conhecimentos e momentos de descontração, Matheus Feitosa pelas parcerias de estudos e também de hospedagem, meu amigo Eduardo Rocha um cara que sempre esteve a disposição, pelas conversas e desabafos, meu amigo Rafael Pedro da Silva, que compartilhou comigo tantas histórias, conhecimentos, e alegrias, meu amigo Marcos Feitosa um amigo para toda a vida. Muito obrigado "Seus cara". Amo vocês.

Agradeço ao meu orientador Flávio Henrique Duarte de Araújo, por em todos os momentos ter me guiado no caminho certo no decorrer desta pesquisa até a escrita dessa monografia. Muito obrigado.

Agradeço a equipe do DTI principalmente a Rafel Luz, Nonato Sales e Jonysson Lima, por tudo que me ensinaram, pelas orientações e por sempre me ajudarem. Agradeço também a meus amigos Matheus Lima, Deusimar Diniz, Francisco das Chagas Júnior, Daniel Veloso, Tomaz Bisneto, Walef Alves, Samuel Campos, Meteus Pinto, Estevão Viana, Milton Júnior, o grande Davi Luís e o senhor Angel Pereira quem sempre zelou por minha segurança.

Agradeço a todos os professores pelo conhecimento e pela dedicação passados, que levarei para toda a vida.

Por fim, agradeço a todos que, de forma direta ou indireta, contribuíram à realização e conclusão desta jornada.

Glória a Deus!

*Quem confia em si mesmo é tolo, mas quem segue os ensinamentos dos sábios terá
segurança.*

Provérbios 28:26

Resumo

Metástase é um tumor formado a partir de células cancerígenas que, se separaram de um primeiro tumor e migraram através dos vasos linfáticos ou vasos sanguíneos para outra parte do corpo, onde se estabeleceram. Atualmente, a detecção de metástases é realizada por patologistas revisando grandes extensões de tecidos biológicos. Este processo de detecção é tido na literatura como trabalhoso e propenso a erros. Com base na problemática de detecção, este trabalho apresenta três estratégias de *deep learning* para identificação automática, capaz de localizar tecido tumoral em imagens de microscopia da base PCam do desafio *Camelyon16*. A PCam é uma base com 327.680 imagens coloridas, destas 262.144 são dedicadas para treinamento, 32.768 são destinadas para teste de aprendizado. Com uma base dessa extensão é possível criar um modelo de classificação automática que poderá auxiliar na decisão do patologista. Para a construção desse método propomos uma arquitetura de Rede Neural Convolutiva com 9 camadas, como também utilizar outros modelos já prontos de redes neurais pré-treinadas, e também técnicas de aperfeiçoamento de aprendizagem como *fine tuning*, para assim determinarmos a melhor técnica entre essas estratégias. Obteve-se um resultado promissor com um AUC de 96,00% utilizando a técnica de *fine-tuning*, mostrando que modelos de *deep learning*, podem ser usados, para tarefas de classificação de imagens.

Palavras-chaves: *PatchCamelyon*. Histopatologia. Redes. Neurais Convolutivas. Aprendizado Profundo. Transferência de Aprendizado. *fine tuning*. PCam.

Abstract

Metastasis is a tumor formed from cancer cells that separated from a first tumor and migrated through lymphatic vessels or blood vessels to another part of the body where they settled. Currently, metastasis detection is performed by pathologists reviewing large extensions of biological tissues. This detection process is considered in the literature as laborious and error-prone. Based on the detection problem, this paper presents three deep learning strategies for automatic identification, capable of locating tumor tissue in PCam-based microscopy images of the Camelyon challenge16. PCam is a base with 327,680 color images, of which 262,144 are dedicated to training, 32,768 are intended for learning testing. With a base of this extension it is possible to create an automatic classification model that may be helping the pathologist's decision. For the construction of this method we propose a 9 layer Convolutional Neural Network architecture, as well as use other ready-made models of pre-trained neural networks, as well as learning improvement techniques such as textit fine tuning, in order to determine the best one. between these strategies. A promising result with a 96.00 % AUC was obtained using the fine-tuning technique, showing that deep learning models can be used for image classification tasks.

Keywords: PatchCamelyon. Histopathology. Convolutional Neural Networks. Deep Learning. Learning Transfer. textit fine tuning. Pcam.

Lista de ilustrações

Figura 1 – Exemplo de imagens da base PCam, que são divididas em classe 0, que não apresenta tecido metastático e classe 1 que apresenta tecido metastático	17
Figura 2 – Ilustração da arquitetura de uma CNN com duas camadas convolucionais, duas de <i>pooling</i> , uma totalmente conectada e a de saída	19
Figura 3 – Ilustração de um bloco residual na arquitetura da rede ResNet	20
Figura 4 – Exemplo de Transferência de Aprendizado.	22

Lista de tabelas

Tabela 1 – Tabela contendo a descrição dos trabalhos relacionados área de classificação de imagens histológicas. Acurácia (Acc) e Área sob a curva ROC (AUC).	15
Tabela 2 – Nível de exatidão de uma classificação, conforme o valor de índice <i>Kappa</i>	24
Tabela 3 – Tabela de testes da arquitetura treinada do zero	26
Tabela 4 – Tabela de resultados da arquitetura pré-treinada VGG16	29
Tabela 5 – Tabela de resultados da arquitetura pré-treinada VGG19	29
Tabela 6 – Tabela de resultados da arquitetura pré-treinada ResNet50	29
Tabela 7 – Tabela de resultados da aplicação da técnica de <i>fine-tuning</i>	30
Tabela 8 – Tabela do melhores resultados das estratégias testadas.	30
Tabela 9 – Tabela de comparação dos resultados deste trabalho com os resultados do desafio <i>fine-tuning</i>	30

Lista de abreviaturas e siglas

ANN	Rede Neural Artificial
CAD	<i>Computer-Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
PCam	<i>Patch Camelyon (Artificial Neural Networks) (Scale Invariant Feature Transform)</i>
DL	<i>Deep Learning</i>
SVM	<i>Support Vector Machine</i>
ISI	Imagem de Slide Inteira
VGG	<i>Visual Geometry Group</i>
SIFT	<i>Scale-Invariant Feature Transform</i>

Lista de símbolos

\leq	Menor que
\geq	Maior que
θ	Letra grega Theta

Sumário

1	Introdução	13
1.1	Objetivo Geral	14
1.2	Organização do Trabalho	14
2	Trabalhos Relacionados	15
2.1	Descrição dos Trabalhos	15
3	Materiais e Métodos	17
3.1	Base de Dados	17
3.2	Aprendizagem Profunda	17
3.2.1	Redes Neurais Convolucionais	18
3.2.1.1	Arquitetura das Redes Avaliadas	19
3.3	Transferência de Aprendizado	21
3.4	Métricas de Avaliação	23
4	Modelos Propostos e Resultados	26
4.1	Rede treinada do zero	26
4.2	Redes Pré-treinadas	28
4.3	<i>Fine-tuning</i>	29
5	Conclusão	31
	Referências	32

1 Introdução

De acordo com o Instituto Nacional do Câncer (INCA), a metástase é um tumor formado a partir de células cancerígenas que se separaram de um primeiro tumor e migraram através dos vasos linfáticos ou vasos sanguíneos para outra parte do corpo onde se estabeleceram. Há muitos anos, o sistema de estadiamento clínico de tumores TNM, tem sido o padrão de classificação adotado para caracterizar os tumores através de imagens histopatológicas, propondo a terapia mais adequada e, assim, estimar a sobrevida dos pacientes.

Tendo em vista a importância do exame histopatológico do tumor, é necessário uma observação detalhada na imagem da região afetada. Com auxílio desse tipo de exame onde se analisa imagens a respeito da região afetada, métodos computacionais de classificação com base em Inteligência Artificial (IA) se tornam essenciais para gerar resultados, que auxiliam no diagnóstico das imagens histopatológicas.

O Aprendizado de máquina (ML) é uma área de estudo da Inteligência Artificial (IA) que foca no aprendizado a partir de dados, estudos nessa linha de pesquisa tornaram-se cada vez mais onipresentes no estudo de dados de domínio médico. No campo da visão computacional o uso de estruturas de Redes Neurais Convolucionais (CNN) e outros modelos de Aprendizagem Profunda (DL) são utilizadas para tarefas relacionadas com classificação e segmentação de imagens médicas. Por exemplo, o uso de tais estruturas para detectar metástases em manchas extraídas de imagens histopatológicas levam a resultados que muitas vezes ultrapassam o de patologistas humanos (VEELING et al., 2018),(LIU et al., 2017).

A importância de tais resultados é aumentada pelo fato de que especialistas em patologia ao realizarem a mesma tarefa, podem discordar em mais de 20% dos casos (BEJNORDI et al., 2017). Essa tarefa também é descrita na literatura como tediosa e demorada, o que torna mais propensa a erros humanos.

A existência de sistemas computacionais que podem executar esta tarefa de classificação em níveis muito altos de precisão, mostra ser uma maneira de superar este problema. Como a demanda por patologistas é extremamente alta em todo o mundo, esses sistemas de IA poderiam ajudar muito no balanceamento de sua carga de trabalho e permitir um diagnóstico mais preciso e consistente. Esta questão, no entanto, se estende a muitas outras áreas onde as decisões apoiadas pela IA são utilizadas como ferramentas de segunda opinião a respeito de um determinado problema.

A principal contribuição do presente trabalho é abordar técnicas de classificação com base em Aprendizagem Profunda assim como um conjunto de métodos e procedimentos de Transferência de Aprendizado e o uso da tática de *fine-tuning* para melhores resultados, aplicando técnicas e estratégias computacionais para o problema da classificação de me-

tástases em Imagem de Slide Inteira (ISI). O conjunto de dados relacionado a essa tarefa é o benchmark Patch Camelyon (PCam) lançado em março de 2019 como um desafio de classificação pela patchcamelyon.grand-challenge.org (VEELING et al., 2018).

1.1 Objetivo Geral

Tendo como base o contexto elucidado, este trabalho tem como objetivo geral, avaliação de técnicas computacionais aplicadas na detecção de metástase em imagens histopatológicas extraídas de exames de microscopia de seções linfonodais.

Para tornar possível a realização do objetivo geral, tem-se como bases os seguintes objetivos específicos:

- Avaliar o uso de uma arquitetura própria treinada a partir das imagens da base PCam.
- Avaliar o uso de transferência de aprendizado por meio do uso de redes pré-treinadas e de técnicas de *fine-tuning*.
- Comparar os resultados obtidos com outros métodos clássicos da literatura.

1.2 Organização do Trabalho

Este trabalho está organizado em cinco capítulos. O presente capítulo onde são detalhados o contexto e proposta do trabalho. No Capítulo 2 são detalhados os principais trabalhos relacionados. No Capítulo 3 são detalhados os materiais e métodos, iniciando pela aquisição das imagens e embasamentos teóricos a respeito das técnicas e métricas utilizadas nesse projeto. No Capítulo 4 são apresentados e discutidos os resultados obtidos com a metodologia proposta. E por fim, no Capítulo 5 são apresentadas as conclusões e trabalhos futuros.

2 Trabalhos Relacionados

No que diz respeito aos trabalhos voltados a base escolhida, pode se notar que por se tratar de uma base bem recente, poucos utilizaram esses dados para análise, contudo há uma tendência de crescimento em estudos baseadas nesta linha de pesquisa. Foram analisados trabalhos com o mesmo intuito e objetivos semelhantes a este projeto, ou seja, uso de métodos computacionais para detecção de metástase, entretanto esses trabalhos utilizam outras bases de dados, ou uma versão mais antiga da base utilizada nesse projeto. Pôde-se observar que há uma tendência de crescimento no uso de técnicas computacionais, tanto para extração das características, quanto para classificação de imagens histopatológicas.

2.1 Descrição dos Trabalhos

A Tabela 1 resume os principais trabalhos encontrados na literatura e relacionados com o objetivo dessa pesquisa. Na Tabela 1 também são mostradas as bases de imagens utilizadas por cada autor, como também, as arquiteturas e os resultados de acurácia(Acc.) ou Área sob a curva ROC (AUC).

Tabela 1 – Tabela contendo a descrição dos trabalhos relacionados área de classificação de imagens histológicas. Acurácia (Acc) e Área sob a curva ROC (AUC).

Referências	Base de Imagens	Arquitetura	Resultados
(WANG et al., 2016)	Camelyon16	GoogleNet	Acc 98,4%
(LIU et al., 2017)	Camelyon16	Inception V3	AUC 97,5%
(VEELING et al., 2018)	Camelyon16 , Pcam e BreakHi	CNN Tradicional	Acc 96,1%
(LINMANS et al., 2018)	PCam	P4M U-Net, Unet	Acc 83,7%
(JAISWAL et al., 2019)	PCam	DenseNet201	Acc 97,94%

Fonte: Elaborada pelo autor.

Wang et al. (2016) propuseram um sistema de aprendizagem profunda para a detecção automática de câncer metastático de slide inteiro em imagens de linfonodos. Os autores avaliaram o desempenho de quatro arquiteturas de redes de aprendizagem profunda conhecidas, para esta tarefa de classificação: GoogLeNet, AlexNet, VGG16 e FaceNet na base de imagens Camelyon16. Eles conseguiram o melhor resultado com o GoogLeNet, atingindo uma precisão de 98.4%.

Liu et al. (2017) desenvolveram uma abordagem para classificação de imagens histológicas, onde os autores utilizam a arquitetura de rede pré-treinada *Inception* (V3) para extração das características, usando o classificador *Random forest*. O conjunto de imagens

é o *Camelyon16*, do desafio de detecção de tumor em nível de lesão. O método proposto obteve 97.5% de AUC.

Veeling et al. (2018) desenvolveram um novo modelo de classificação treinando o modelo com a base de imagens histopatológicas *Camelyon16* e com foco na estabilidade das previsões. Para avaliar o modelo proposto os autores utilizaram a base Pcam e testaram a arquitetura com a base de dados *BreakHis*. O modelo proposto alcançou a precisão de 96.1%.

Linmans et al. (2018) propuseram um modelo de segmentação semântica que explora simetrias de rotação, onde são utilizadas duas arquiteturas de redes (Unet e P4M U-Net), além disso o aumento de dados é empregado. O conjunto de imagens utilizadas nesse projeto é o *PCam* fornecido como parte do desafio de classificação de imagens histológicas. Eles conseguiram com a rede P4M U-Net 83,7% de precisão.

Jaiswal et al. (2019) propuseram um modelo baseado em CNN, validado no conjunto de dados da PCam, onde usaram uma abordagem de aprendizagem semi-supervisionada em diferentes modelos pré-treinados como VGG16, *InceptionResNetV2*, *InceptionV3*, *Xception*, *ResNet101* e *DenseNet201* no sentido de transferência de aprendizagem fazendo a diferenciação entre classes de baixa densidade. O *DenseNet201* obteve uma acurácia de 97%.

Pesquisas efetuadas mostraram que, as técnicas de aprendizagem profunda estão se tornando primordiais, na resolução de diagnóstico médico por meio de imagens, por esta razão foi determinado o uso de técnicas envolvendo arquiteturas de CNN.

A principal contribuição desse trabalho é a avaliação de três técnicas baseadas em aprendizagem profunda. A primeira técnica é uma CNN própria treinada com a inicialização aleatória de pesos. Como segunda técnica foram utilizadas as arquiteturas pré-treinadas VGG16, VGG19, e *ResNet50* para extrair características e com ajuda de classificadores, classificar os dados. Como terceira estratégia, foi utilizada a técnica de *fine-tuning*. Além disso os modelos foram treinados e testados com a versão mais atual da base *PCAM* que foi disponibilizada em março de 2019, como desafio de classificação do *Camelyon16 Challenge*. Os resultados melhores resultados obtidos nesses testes serão comparados com os resultados presentes no desafio, visto que ainda não há trabalhos publicados que utilizaram a versão mais atual da base PCam.

3 Materiais e Métodos

Nesta seção são apresentados a base de imagens, conceitos importantes para uma melhor compreensão deste trabalho, e as principais arquiteturas e técnicas de CNNs utilizadas para o processamento das imagens histológicas.

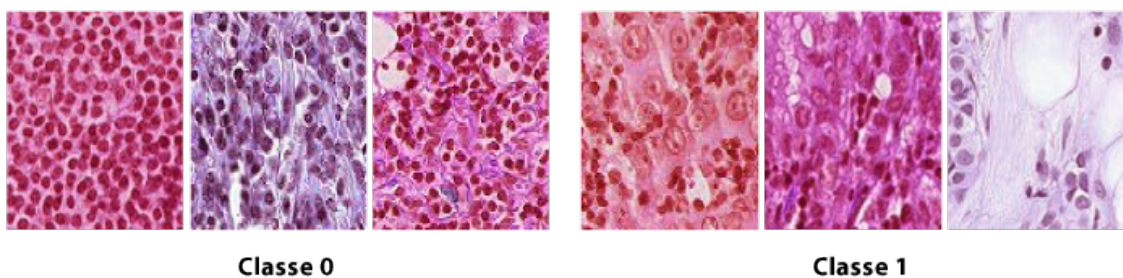
3.1 Base de Dados

O conjunto de imagens utilizadas nesse trabalho consiste em uma base de dados lançada em março de 2019, a base conta com 327.680 imagens coloridas com dimensão (96 x 96px) extraídas de exames histopatológicos de seções de linfonodos. Cada imagem é anotada com um rótulo binário indicando a presença ou não tecido metastático.

O conjunto de imagens é dividido em treinamento com 262.144(2^{18}) exemplos e um teste de 32.768 (2^{15}) exemplos. Não há sobreposição em ISIs entre as divisões e todas as divisões têm um equilíbrio de 50/50 entre exemplos positivos e negativos.

A PCam é derivada do Camelyon16 Challenge, que contém WSIs com 400 H & E de seções de linfonodo sentinela. As lâminas foram adquiridas e digitalizadas em dois centros diferentes usando uma objetiva de 40x (resolução de pixels resultante de 0,243 microns). A Figura 1 apresenta exemplos de imagens das duas classes, classe 0 representa tecido não metastático e classe 1 tecido metastático.

Figura 1 – Exemplo de imagens da base PCam, que são divididas em classe 0, que não apresenta tecido metastático e classe 1 que apresenta tecido metastático



Fonte: Banco de dados PCam (PCam) 2019

1

3.2 Aprendizagem Profunda

A Aprendizagem Profunda é uma sub área do aprendizado de máquina que trata de modelos vagamente inspirados no cérebro humano. Tecnicamente, toda Rede Neural

¹ Disponível em <https://github.com/basveeling/pcam>

com mais de uma camada escondida é chamada de Rede Neural Profunda e tudo que ela aprende é por meio do Aprendizado Profundo. Para [LeCun, Bengio e Hinton \(2015\)](#) a Aprendizagem Profunda é uma área que contém uma pilha de multicamadas simples, todas (ou a maioria) estão sujeitas à aprendizagem, e muitos dos quais calcula mapeamentos de entrada-saída não lineares.

Tradicionalmente, a qualidade dos algoritmos depende muito da representação dos dados em certas características (as chamadas *features*). Assim, a criação de recursos válidos e pré-processamento consomem grande parte dos esforços dos especialistas. Já os algoritmos do tipo Aprendizagem Profunda têm uma abordagem inovadora, pois dispensam grande parte desse pré-processamento e geram automaticamente propriedades invariantes em suas camadas hierárquicas de representação ([HELABS, 2019](#)).

[Brigade \(2016\)](#) afirma que atualmente o reconhecimento de imagens por máquinas treinadas através de Aprendizagem Profunda em alguns cenários possuem uma taxa de acerto maior que a de humanos. Várias arquiteturas de Aprendizagem Profunda, tais como CNN têm sido aplicadas em áreas como visão computacional, reconhecimento automático de fala, processamento de linguagem natural, reconhecimento de áudio e bioinformática, onde elas têm se mostrado capazes de produzir resultados significantes.

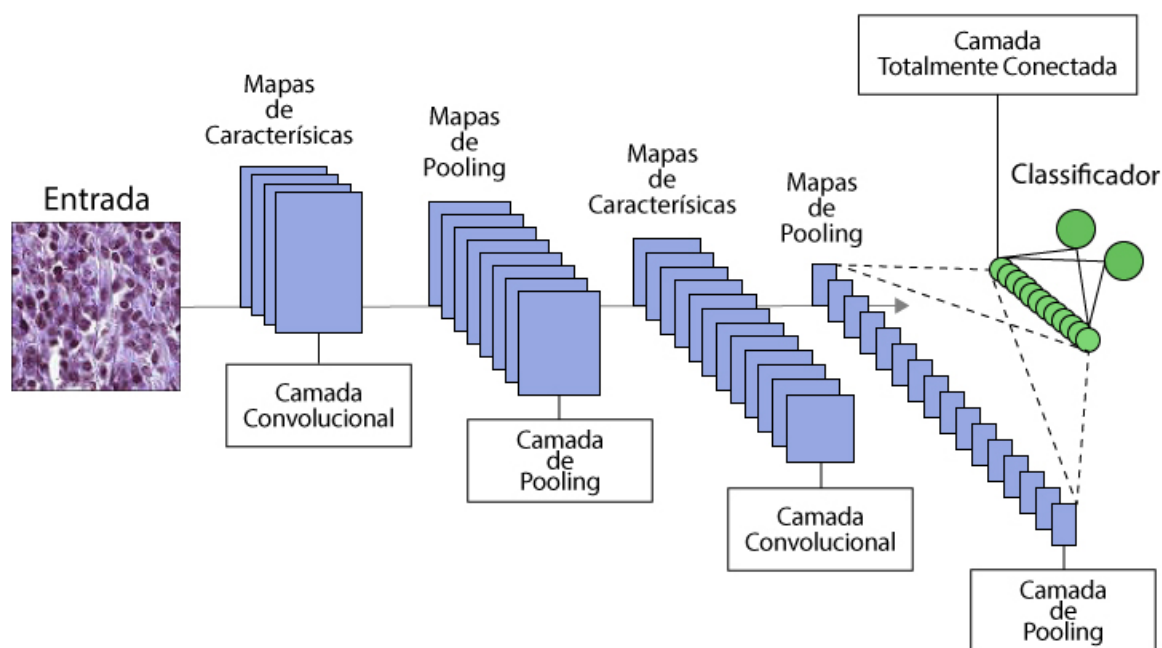
A seguir são abordadas as fundamentações sobre as Redes Neurais Convolucionais (CNN) utilizadas nesse projeto.

3.2.1 Redes Neurais Convolucionais

CNN ou *ConvNets* são uma categoria de redes neurais usadas no reconhecimento e processamento de imagens que é especificamente projetada para processar dados de pixel. [Ponti Moacir Antonelli \(2018\)](#) versa que sua característica marcante é ser composta basicamente de camadas convolucionais, que processa as entradas considerando campos locais. Essas redes constituem uma escolha de topologia ou arquitetura projetadas para reduzir o número de parâmetros a serem aprendidos otimizando o tempo de treinamento.

As CNNs são projetadas para processar dados que vêm na forma de múltiplos arrays, por exemplo, uma imagem colorida composta por três matrizes contendo intensidades de pixel nos três canais de cores ([LECUN; BENGIO; HINTON, 2015](#)). Elas são capazes de capturar com sucesso as dependências espaciais e temporais em uma imagem através da aplicação de filtros. [Kovaleski \(2018\)](#) afirma que a mesma se utiliza desses filtros para analisar pequenos grupos de dados, extraindo suas características mais relevantes e relata que apenas estas características são passadas adiante para as próximas camadas, como apresentado na Figura 2.

Figura 2 – Ilustração da arquitetura de uma CNN com duas camadas convolucionais, duas de *pooling*, uma totalmente conectada e a de saída



Fonte: Adaptação de [Saha \(2018\)](#).

Ao contrário de uma Rede Neural regular, as camadas de uma CNN possuem neurônios dispostos em 3 dimensões: largura, altura e profundidade. Além disso, os neurônios em uma camada não se conectam a todos os neurônios da camada seguinte, mas apenas a uma pequena região dela. Por fim, a saída final será reduzida a um único vetor de pontuações de probabilidade, organizadas ao longo da dimensão de profundidade.

A arquitetura executa um melhor ajuste ao conjunto de dados da imagem devido à redução no número de parâmetros envolvidos e à capacidade de reutilização dos pesos. A grande vantagem da aplicação das CNNs para o Processamento Digital de Imagens (PDI) é sua grande capacidade de extrair características, por trabalhar com convoluções, ou seja, criando “sub-quadros de bits” de forma que seja possível realçar contornos e áreas de interesse de cada imagem ([ROSEBROCK, 2017](#)).

A seguir será apresentada as fundamentações e principais características das mais diversas arquiteturas de redes convolucionais utilizados nesse trabalho para a extração de características das imagens.

3.2.1.1 Arquitetura das Redes Avaliadas

Após analisar a bibliografia que resultou da pesquisa a respeito do tema, foram encontrados diversos modelos de arquitetura compatíveis com a proposta desse trabalho. Com isso, foi possível selecionar dois dos principais modelos, visto que essas duas arquiteturas

são as mais comuns na literatura. Para um estudo mais profundo e consequente avaliação da melhor performance, os modelos utilizados são descritos à seguir.

- **Modelo 1 - *Visual Geometry Group* - VGG**

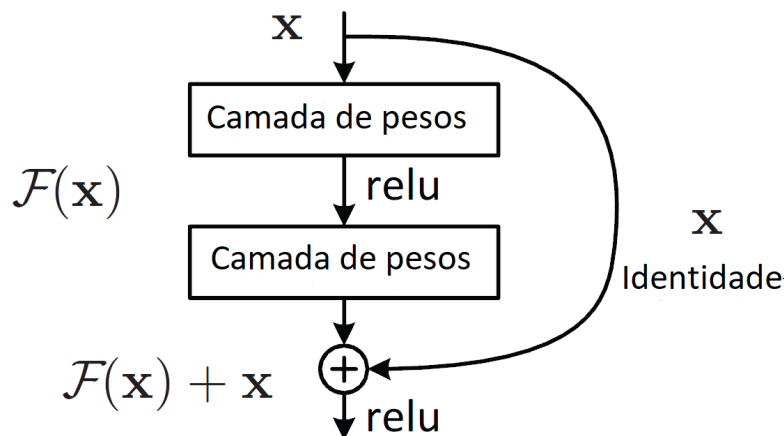
A VGG é um modelo de CNN inicialmente proposto por [Simonyan e Zisserman \(2014\)](#). A mesma garantiu a primeira e segunda posições nas tarefas de localização e classificação, respectivamente, no desafio ImageNet 2014. Ela faz a melhoria sobre o AlexNet construída por [Krizhevsky, Sutskever e Hinton \(2012\)](#) substituindo grandes filtros do kernel por múltiplos filtros do tamanho 3x3 um após o outro. Nela as profundidades mais comuns são as de 16 e 19 camadas de profundidade.

Com a introdução da VGG, elas trouxeram alguns padrões: sugeriu-se que todos os filtros tivessem tamanho de 3x3, pools máximos deveriam ser colocados após cada 2 convoluções e o número de filtros deveria ser dobrado após cada *max-pooling*. A VGGNet é considerada a rede preferida pela comunidade para Transferência de Aprendizado, pois sua arquitetura uniforme é boa para extrair características de imagens. Assim, a VGG tem sido utilizada como modelo na construção de novas redes, como Unet, TernausNet e várias outras redes.

- **Modelo 2 - ResNet**

A ResNet é uma arquitetura que tem como principal característica suas conexões residuais. Na aprendizagem residual, em vez de tentar aprender algumas características, tal arquitetura é voltada para aprender alguns resíduos. Residual pode ser simplesmente entendido como subtração de características aprendidas da entrada dessa camada e a ResNet faz isso usando conexões de atalho, que são conexões em que o gradiente é diretamente retropropagado para camadas anteriores. Na Figura 3 é demonstrado o funcionamento de um bloco residual.

Figura 3 – Ilustração de um bloco residual na arquitetura da rede ResNet



Fonte: Adaptação de [He et al. \(2016\)](#).

Um dos problemas que a ResNet resolve é o famoso desaparecimento de gradiente. Isso ocorre porque, quando a rede é muito profunda, os gradientes de onde a função de perda é calculada diminuem facilmente para zero. Este resultado nos pesos nunca atualiza seus valores e, portanto, nenhum aprendizado será realizado.

Com tal arquitetura, os gradientes podem fluir diretamente através das conexões de salto. Para [Chen et al. \(2017\)](#) a ResNet fornece uma estrutura de treinamento para facilitar formação de redes substancialmente mais profundas que as utilizadas anteriormente.

3.3 Transferência de Aprendizado

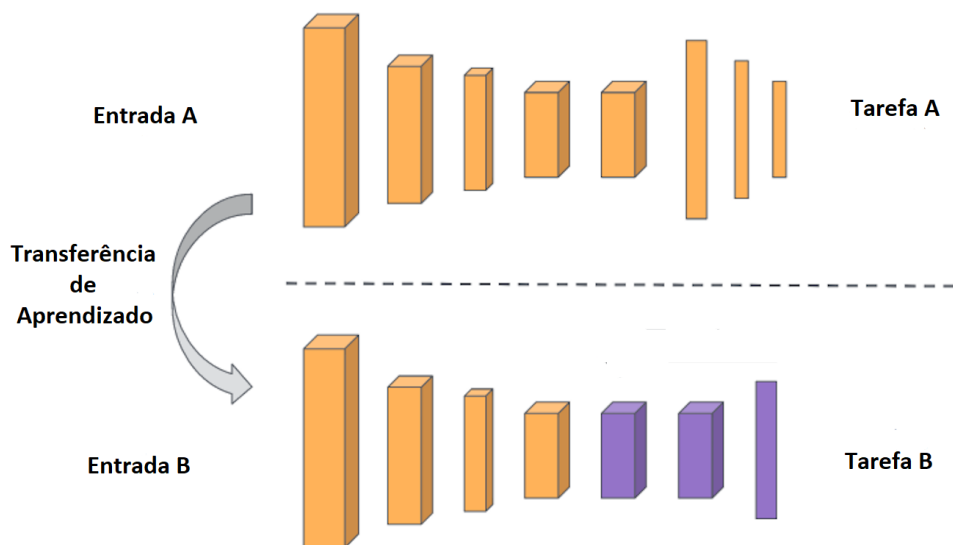
A Transferência de Aprendizado é um método de aprendizado de máquina, em que, um modelo desenvolvido para uma tarefa é reutilizado como ponto de partida para um modelo em uma segunda tarefa ([PONTI MOACIR ANTONELLI, 2018](#)). É uma abordagem popular em Aprendizado Profundo na qual os modelos pré-treinados são usados como fase inicial em tarefas de visão computacional.

Normalmente, precisa-se de muitos dados para treinar uma Rede Neural, mas nem sempre é possível ter acesso a dados suficientes. É aí que a Transferência de Aprendizado pode ser utilizada, porque com ela pode-se construir um modelo de aprendizado de máquina sólida com comparativamente poucos dados de treinamento, já que o modelo já é pré-treinado ([TORREY; SHAVLIK, 2010](#)).

Embora as redes profundas tenham mostrado-se promissoras na representação de informações, tais modelos profundos supervisionados, exigem grandes quantidades de dados rotulados e recursos de computação, que são caros ([DU et al., 2017](#)). O processo de Transferência de Aprendizado é demonstrado pela Figura 4

Na Transferência de Aprendizado, os modelos pré-treinados são usados para extrair características de imagens, sendo estes modelos treinados em milhões de imagens pertencentes a milhares de classes, em várias GPUs de alta potência por vários dias e todos esses recursos já aprendidos são submetidos para resolver problemas relacionados.

Figura 4 – Exemplo de Transferência de Aprendizado.



Fonte: Adaptação de [Subodh \(2018\)](#).

[Hu et al. \(2015\)](#) corrobora em afirmar que as vantagens da Transferência de Aprendizado incluem aceleração do tempo de treinamento, redução parâmetros e uma arquitetura de redes neurais profundas que busca manter alto desempenho de classificação, especialmente em um conjunto limitado de imagens. Existem duas abordagens principais para implementar na Transferência de Aprendizado; são elas: inicialização de peso e extração de características ([BROWNLEE, 2019](#)). No caso da primeira abordagem os pesos nas camadas reutilizadas podem ser usados como ponto de partida para o processo de treinamento e adaptados em resposta ao novo problema, essa técnica é conhecida como *fine-tuning*.

A estratégia de *fine-tuning* consiste em dar continuidade ao treinamento de uma rede pré-treinada utilizando o algoritmo *Backpropagation* e uma nova base de imagens. Em outras palavras, os pesos de todas as camadas de uma rede pré-treinada, com exceção da última camada, são utilizados para a inicialização de uma nova CNN ([ARAÚJO et al., 2017](#)). É possível fazer o *fine-tuning* de todas as camadas de uma CNN, ou somente das últimas camadas. Isso é motivado pelo fato que as primeiras camadas da rede contém extratores mais genéricos que podem ser utilizados para diferentes tarefas, como detectores de bordas e de cores, porém, as camadas mais profundas possuem detalhes específicos da base com a qual a rede foi originalmente treinada ([ARAÚJO et al., 2017](#)).

Na segunda abordagem pode-se remover a camada de saída (aquela que fornece as probabilidades de estar em cada uma das 1000 classes) e, em seguida, usar toda a rede como um extrator de características fixo para o novo conjunto de imagens. Depois disso um classificador externo é usado para classificação das imagens. Nesse trabalho utilizamos

a segunda abordagem envolvendo os classificadores (*Random Forest* e SVM), por causa que os mesmos eram os mais abordados nos trabalhos pesquisados.

Portanto, ela também economiza muito tempo de treinamento, porque às vezes pode levar dias ou até semanas para treinar uma Rede Neural Profunda do zero em uma tarefa complexa. Além desse benefício a Transferência de Aprendizado resulta em menor erro de generalização (GOODFELLOW; BENGIO; COURVILLE, 2016).

Na próxima Seção são apresentadas as fundamentações dos descritores utilizados nesse projeto.

3.4 Métricas de Avaliação

Para a avaliação do método proposto nesse trabalho foram utilizadas as seguintes métricas: Sensibilidade (S), Especificidade (E), Acurácia (A), Área Sob a Curva ROC (AUC) e *Kappa* (K). Essas métricas fazem uso da matriz de confusão, que indica a classificação correta ou incorreta das classes em uso, agrupando os resultados em função de quatro valores, sendo eles: Falso Negativo (FN), Falso Positivo (FP), Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN).

1. Acurácia

Esta é a métrica mais simples. Consiste basicamente o número de acertos (positivos) dividido pelo número total de exemplos. Essa métrica é mais apropriada para bases com a mesma proporção de exemplos para cada classe, e quando as penalidades de acerto e erro para cada classe forem as mesmas (MISHRA, 2018). Em problemas com classes desproporcionais, ela causa uma falsa impressão de bom desempenho. A acurácia pode ser calculada conforme a Equação 3.1.

$$A = \frac{VP + VN}{VP + VN + FP + FN}. \quad (3.1)$$

2. Sensibilidade

A sensibilidade é uma medida da proporção de casos positivos reais que foram previstos como positivos (ou positivos verdadeiros). Em outras palavras, um teste altamente sensível é aquele que identifica corretamente pacientes com uma doença, o mesmo pode ser útil para descartar uma doença se uma pessoa tiver um resultado negativo (KUMAR, 2018). A sensibilidade pode ser calculada conforme a Equação 3.2

$$S = \frac{VP}{VP + FN}. \quad (3.2)$$

3. Especificidade

A especificidade de um teste (também chamado de Taxa Negativa Verdadeira) é a proporção de pessoas sem a doença que terão um resultado negativo. Em outras palavras, a especificidade de um teste se refere a quão bem um teste identifica pacientes que não têm uma doença (KUMAR, 2018). A especificidade pode ser calculada com a Equação 3.3

$$E = \frac{VN}{VN + FP}. \quad (3.3)$$

4. *Kappa*

O coeficiente *Kappa* é uma medida de concordância estatística formulada por Jacob Cohen, tal métrica mede concordância entre dois avaliadores que classificam cada N itens em C categorias mutuamente exclusivas. O *Kappa* de Cohen é forte dependente de vários fatores, incluindo, por exemplo, número de classes e número de exemplos a serem classificados. O *Kappa* pode ser calculado conforme a Equação 3.4

$$K = \frac{\theta_1 - \theta_2}{1 - \theta_2}, \quad (3.4)$$

onde

$$\theta_1 = \frac{VP + VN}{VP + VN + FP + FN}, \quad (3.5)$$

e

$$\theta_2 = \frac{[(VP + FN)(VP + FP)] + [(VN + FN)(VN + FP)]}{(VP + VN + FP + FN)^2}. \quad (3.6)$$

Sobre a avaliação dos classificadores, Landis e Koch (1977) particionaram o nível de exatidão do Índice *Kappa* conforme a Tabela 2.

Tabela 2 – Nível de exatidão de uma classificação, conforme o valor de índice *Kappa*

Índice <i>Kappa</i> (K)	Qualidade
$K < 0.2$	Ruim
$0.4 \leq K < 0.6$	Bom
$0.6 \leq K < 0.8$	Muito Boa
≥ 0.8	Excelente

5. Área Sob a Curva ROC (*Area Under ROC*)

É uma métrica para classificação binária, é provavelmente uma das mais populares. Braga (2001) afirma que a curva ROC é um método gráfico simples e robusto, permite estudar a variação da sensibilidade e especificidade, para diferentes valores de corte. AUC é uma abreviação da área sob a curva. É usado na análise de

classificação para determinar qual dos modelos utilizados prediz melhor as classes. Quanto mais próximo o AUC de um modelo chegar a 1, melhor ele será (GIANNINI, 2012). Assim, os modelos com AUCs mais altos são preferidos em relação àqueles com AUCs mais baixos.

Após apresentar fundamentações, métodos e técnicas ao longo desse trabalho, a seguir são descritos os resultados obtidos ao fim das etapas de treinamento e testes avaliativos, além disso será abordado uma discussão sobre o processo proposto para resolver o problema abordado.

4 Modelos Propostos e Resultados

Neste trabalho é proposto um método automático para a classificação de imagens extraídas de exames histopatológicos de seções de linfonodos. Para determinar a arquitetura e/ou técnica que produzisse os melhores resultados, testamos três estratégias: uma CNN elaborada do zero para classificar as imagens que são divididas em duas classes, a que não apresenta e a que apresenta tecido metastático; Redes Neurais pré-treinadas usadas para extrair características das imagens, com ajuda de classificadores para determinar a classe das imagens. Foram utilizados os classificadores SVM e seus diferentes *kernels* e o *Random forest*; Como terceira estratégia utilizamos o *fine-tuning*, utilizando a arquitetura clássica inicializada com os pesos do modelo treinado para a base ImageNet, e este treinamento foi continuado utilizando a base utilizada nesse projeto, juntamente com as camadas da CNN proposta. Abaixo são detalhadas as estratégias, juntamente com os resultados obtidos nos teste.

4.1 Rede treinada do zero

Abaixo são mostrados os resultados obtidos de acordo com os testes feitos durante a implementação do zero da Rede Neural Convolutacional, logo em seguida é discutida a implementação em cada modelo. A Tabela 3 apresenta o modelo implementado (Modelo), com a quantidade de imagens utilizadas para o treinamento (Quant. Imagens). Estão também listadas as seguinte métricas: Acurácia(Acc.), Especificidade (Esp.), Sensibilidade(Sens.), Área sob a curva ROC (AUC), e o coeficiente *Kappa*(*Kappa*). Vale ressaltar que após cada camada de convolução havia uma camada de *pooling* com filtro de tamanho 2X2. Todos os modelos foram elaborados utilizando a API *Keras*, uma API de alto nível para redes neurais usando a biblioteca *tensorflow* como *backend*.

Tabela 3 – Tabela de testes da arquitetura treinada do zero

Modelo	Quant. Imagens	Acc.%	Esp.%	Sens.%	AUC%	<i>Kappa</i>
Modelo 1	10,000	49,63	0,0	0,0	50,00	0,19
Modelo 2	10.000	73,68	51,08	96,26	82,00	0,37
Modelo 3	10.000	79,81	64,77	94,84	84,00	0,59
Modelo 4	100.000	82,26	68,15	93,85	90,00	0,63
Modelo 5	262.144	83,63	72,49	94,76	93,00	0,67

O Modelo 1 foi uma rede criada sem camadas de convolução ou *pooling*, apenas com uma camada de entrada com 32 neurônios, uma camada oculta com 16 neurônios, e 2 neurônios na camada de saída. Para o treinamento usamos um total de 10.000 (dez mil) imagens, com um tamanho de lote igual a 32, isso combinado com um treinamento de

300 épocas. Com a base de testes no modelo 1, atingimos uma acurácia de 49,63%, com uma especificidade de 00,00% e sensibilidade de 00,00%, conseguindo um AUC 50,00%. Ou seja com essa implementação não foi possível aprender o problema e a rede jogou todos os exemplos de teste para a mesma classe.

No modelo 2 foram feitas modificações na rede do modelo 1, adicionamos 2 camadas de convolução e 2 de *pooling*, todas com filtros de tamanho 3x3 e 2x2 respectivamente. Para treinamento foram separadas 10.000 (dez mil) imagens. O modelo continha uma camada de entrada com 32 neurônios, uma camada oculta com 16 neurônios e uma camada de saída com 2 neurônios, um tamanho de lote igual a 32, com um total de 300 épocas. Na base de testes, no modelo 2, atingiu-se uma acurácia de 73,68%, com uma especificidade de 51,08%, e sensibilidade de 96,26%, conseguindo um AUC 82,00%. Com isso determinamos que ao adicionar as camadas de convolução a rede começava a aprender o problema.

No modelo 3 foi adicionado 3 camadas de convolução e 3 de *pooling*, todas com filtros de tamanho 3x3 e 2x2 respectivamente. Para treinamento foram separadas 10.000 (dez mil) imagens. O modelo continha uma camada de entrada com 64 neurônios, uma camada oculta com 32 neurônios e uma camada de saída com 2 neurônios, um tamanho de lote igual a 32, com um total de 500 épocas, tendo cada época um número de passos por época igual ao número de imagens para treinamento dividido pelo tamanho do lote. Na base de testes, no modelo 3 atingiu-se uma acurácia de 79,68%, com uma especificidade de 64,77% e sensibilidade de 94,26%, conseguindo um AUC 84,00%. Ao aumentarmos o número de épocas e a quantidade de neurônios, notamos um aumento na Acurácia como também nas demais métricas.

Após a análise dos resultados, de acordo com as implementações, foi encontrado um modelo que se saia melhor em aprender o problema, e que gerava baixo custo computacional. Após acharmos a melhor arquitetura decidiu-se aumentar a quantidade de imagens de treinamento, como mostrados os resultados do modelo 4 e 5.

No modelo 4 foram utilizadas 3 camadas de convolução e 3 de *pooling*, todas com filtros de tamanho 3x3 e 2x2 respectivamente, porém com uma diferença do modelo anterior foi adicionado uma camada de normalização entre cada camada de *pooling*. Para treinamento foram separadas 100.000 (cem mil) imagens. O modelo continha uma camada de entrada com 128 neurônios, uma camada oculta com 64 neurônios e uma camada de saída com 2 neurônios, diminuimos o tamanho de lote pela metade do modelo 3 ficando igual a 16, com um total de 2500 épocas, tendo cada época um número de passos por época igual 1000. Na base de testes, no modelo 4, atingimos uma acurácia de 82,26%, com uma especificidade de 67,33% e sensibilidade de 93,85%, conseguindo um AUC 90,00%. Concluiu-se que a normalização entra cada camada fez com que a rede mostrasse uma performance de classificação mais acurada.

No modelo 5 foram utilizadas 3 camadas de convolução e 3 de *pooling*, todas com filtros de tamanho 3x3 e 2x2 respectivamente, tendo uma camada de normalização entre cada

camada de *pooling*. Para treinamento utilizamos todas as imagens disponíveis na base para treinamento um total de 262.144 (duzentos e sessenta e dois mil e cento e quarenta e quatro) imagens. O modelo continha uma camada de entrada com 192 neurônios, uma camada oculta com 128 neurônios e uma camada de saída com 2 neurônios, com um total de 3000 épocas. Na base de testes, no modelo 5, atingiu-se uma acurácia de 83,63%, com uma especificidade de 72,49% e sensibilidade de 94,76%, conseguindo um AUC 93,00% . Esse método se mostrou melhor em todas as métricas depois do aumento da base de treino, e da quantidade de épocas, com auxílio das camadas de convolução e *pooling* juntamente com a camada de normalização.

4.2 Redes Pré-treinadas

Baseado nos resultados das redes implementadas anteriormente, foram feitos testes com arquiteturas pré-treinadas, ou seja, que já tinham um modelo e pesos definidos, arquiteturas estas que já passaram por um treinamento como dito na seção de referencial teórico.

Para a obtenção dos resultados foram usadas apenas 10.000 (dez mil) imagens da base de testes, para a extração de características, visto que, um número maior de imagens elevava muito o custo computacional. Foram usadas as redes neurais VGG16, VGG19 e a *ResNet50* para a obtenção de características da imagem. Extraídas as características das imagens, utilizamos os classificadores SVM e *Random Forest*. O classificador *Random Forest* foi implementado com um total de 1000 (hum mil) árvores de estimativas, um número tido na literatura como ideal. O classificador SVM foi usado, em conjunto com cada uma de suas funções chamadas *kernel*, onde testou-se cada uma dessas funções matemáticas para obter os resultados da classificação.

As Tabelas 4, 5, 6 mostram os resultados obtidos com as arquiteturas pré-treinadas. Cada tabela mostra o classificador utilizado e qual *kernel* selecionado. Vale ressaltar que em todas as arquiteturas utilizamos a mesma quantidade de imagens e que nos classificadores utilizou-se uma taxa de *cross-validation*, com K igual a 10.

Tabela 4 – Tabela de resultados da arquitetura pré-treinada VGG16

Classificador	kernel	Acc.%	Esp.%	Sens.%	AUC%	Kappa
SVM	<i>PolyKernel</i>	89,00	88,00	90,00	90,00	0,79
SVM	RBF	89,00	88,00	89,00	79,00	0,78
SVM	<i>Sigmoid</i>	77,00	77,00	77,00	78,00	0,75
SVM	Linear	84,00	85,00	83,00	84,00	0,68
<i>random forest</i>	-	87,00	86,00	87,00	87,00	0,74

Tabela 5 – Tabela de resultados da arquitetura pré-treinada VGG19

Classificador	kernel	Acc.%	Esp.%	Sens.%	AUC%	Kappa
SVM	<i>PolyKernel</i>	88,00	88,00	89,00	89,00	0,77
SVM	RBF	89,00	88,00	89,00	89,00	0,78
SVM	<i>Sigmoid</i>	81,00	81,00	79,00	81,00	0,61
SVM	Linear	83,00	84,00	82,00	90,00	0,66
<i>random forest</i>	-	86,00	86,00	86,00	87,00	0,73

Tabela 6 – Tabela de resultados da arquitetura pré-treinada ResNet50

Classificador	kernel	Acc.%	Esp.%	Sens.%	AUC%	Kappa
SVM	<i>PolyKernel</i>	90,00	89,00	91,00	90,00	0,80
SVM	RBF	90,00	89,00	91,00	90,00	0,80
SVM	<i>Sigmoid</i>	75,00	77,00	74,00	76,00	0,51
SVM	Linear	87,00	88,00	87,00	88,00	0,75
<i>random forest</i>	-	88,00	88,00	87,00	88,00	0,76

Como percebe-se, os resultados descritos na Tabela 6 foram os melhores resultados e apresentam valores de acurácia ótimos de 90,00 % na taxa média de acertos, 91,00% para sensibilidade média e 89,00% para especificidade. Sendo que o maior valor alcançado dentre todos os experimentos foi de acurácia 90,0%, também apresentando os bons resultados de 90,00% para AUC e 0,8 para o coeficiente *kappa*. utilizando a arquitetura pré-treinada *ResNet50* e fazendo uso do classificador SVM juntamente com o *Kernel* RBF ou *polyKernel*. Com base nos resultados obtidos, constata-se que a arquitetura e o classificador configurado com umas das duas funções conseguiu os resultados mais promissores e equilibrados em comparação as outras arquiteturas. Já a arquitetura VGG16 apresentou no geral menores resultados em comparação as outras arquiteturas .

4.3 Fine-tuning

Com o intuito de melhorar os resultados obtidos, tanto na arquitetura criada do zero, como nos modelos pré-treinados utilizados, fizemos uso da técnica de *fine-tuning*.

Realizamos testes com as arquiteturas pré-treinadas, ou seja, as redes VGG 16 e 19 e também a *ResNet50* todas elas em conjunto com as camadas da rede treinada do zero.

A Tabela 7 mostra os resultados obtido com a técnica *fine-tuning*, onde estão descritos modelo, que se refere a arquitetura utilizada para ser treinada com a nova base.

Tabela 7 – Tabela de resultados da aplicação da técnica de *fine-tuning*

Modelo	Acc.%	Esp.%	Sens.%	AUC%	Kappa
VGG16	84,99	83,58	86,40	93,00	0,69
VGG19	86,12	76,83	95,41	96,00	0,72
<i>ResNet50</i>	81,93	67,62	96,24	91,00	0,63

Como percebe-se os resultados da técnica *fine-tuning*, foram bem próximos dos resultados da rede treinada a partir do zero, considerando como métricas principais a Acurácia e a AUC. A técnica usada aqui mostrou uma acurácia média, em relação com as outras duas estratégias testadas, ficando maior que o melhor resultado da arquitetura implementada do zero e menor que o melhor resultado da rede pré-treinada. A Tabela 8 mostra os melhores resultados de cada uma das estratégias testadas. Para comparação dos resultados decidiu-se adotar as seguinte métricas, com base na métrica exigida no desafio lançado em março de 2019 pela *grand-challenge.org*, que mede a precisão do modelo a partir dos resultados da métrica da AUC.

Tabela 8 – Tabela do melhores resultados das estratégias testadas.

Estratégia	Acc.%	AUC%	Kappa
Rede do zero	83,63	93,00	0,67
Pré-treinada <i>Resnet50</i>	90,00	90,00	0,80
<i>fine-tuning</i> com VGG19	86,12	96,00	0,72

Levando em consideração a métrica utilizada no desafio que é a Área Sob a Curva ROC a estratégia de *fine-tuning* conseguiu um resultado ótimo podendo se colocar como um dos 12 melhores do desafio da *grand-challenge.org*, alcançando um AUC de 96,00%. Já o resultado da arquitetura treinada do zero ficou colocada na vigésimo terceira posição do desafio. A Tabela 9 mostra os principais resultados submetidos para o desafio juntamente com o melhor resultado obtido nesta pesquisa.

Tabela 9 – Tabela de comparação dos resultados deste trabalho com os resultados do desafio *fine-tuning*

Equipe	Posição	AUC%
<i>zhongqiaoyong (Dr. Strange)</i>	DenseNet201	97.37
GeertLitjens	<i>DenseNet with TTA</i>	96.91
<i>ISMI19-KarmaCamelyon</i>	NASNET	96.23
<i>e.boeijenk (ISMI19-KarmaCamelyon)</i>	Ensemble convnet, nasnet, recnn	96.20
Este trabalho	<i>fine-tuning com VGG19</i>	96.00

¹ Disponível em <https://patchcamelyon.grand-challenge.org/evaluation/results/>

5 Conclusão

Neste trabalho foram apresentadas e testadas três estratégias de *deep learning* para classificação de imagens histopatológicas para identificação do câncer metastático. A estratégia de melhor resultado foi a que aplicamos a técnica de *fine-tuning* que ao ser treinada com os pesos da rede pré-treinada juntamente com as camadas da rede própria implementada do zero, obteve como melhor resultado na base de teste chegando a 96,00% na área sob a curva ROC e uma sensibilidade de 95,41%, que é um resultado melhor que os resultados obtidos por um patologista humano que conseguiu 72.00% de sensibilidade de acordo com (VEELING et al., 2018). Com base nos resultados obtidos o trabalho apresentado mostra boa eficácia no âmbito de classificação automática de imagens.

Dessa maneira este trabalho apresenta contribuições em duas áreas de estudo. Na área da saúde oferecendo uma metodologia como segunda opinião a profissionais patologistas. Esse método pode ajudar no diagnóstico de tumores metastáticos em imagens histopatológicas de seções linfonodais. A outra área é a da computação, fornecendo informações importantes de processamento de imagens como forma de auxiliar pesquisas futuras. Este trabalho pretende futuramente testar novos modelos de extração de características como também de classificação, tratando como trabalhos futuros: utilizar outras bases de pesos de redes neurais pré-treinadas, para comparação de resultados e testar outras estratégias de classificação de imagens, a fim de melhorar os resultados.

Referências

- ARAÚJO, F. H. et al. Redes neurais convolucionais com tensorflow: Teoria e prática. *SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais-Artigos e Minicursos*, Sociedade Brasileira de Computação, v. 1, p. 382–406, 2017. Citado na página 22.
- BEJNORDI, B. E. et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama*, American Medical Association, v. 318, n. 22, p. 2199–2210, 2017. Citado na página 13.
- BRAGA, A. *Curvas ROC: aspectos funcionais e aplicações*. Tese (Doutorado), 2001. Citado na página 24.
- BRIGADE, D. S. A diferença entre inteligência artificial, machine learning e deep learning. URL: [https://medium.com/data-science-brigade/a-diferen%C3%A7a-entre-intelig%C3%Aancia-artificial-machine-learning-e-deep-learning-930b5cc2aa42/\(visited on 14/05/2019\)](https://medium.com/data-science-brigade/a-diferen%C3%A7a-entre-intelig%C3%Aancia-artificial-machine-learning-e-deep-learning-930b5cc2aa42/(visited%20on%2014/05/2019)), 2016. Citado na página 18.
- BROWNLEE, J. How to improve performance with transfer learning for deep learning neural networks. URL: [https://machinelearningmastery.com/how-to-improve-performance-with-transfer-learning-for-deep-learning-neural-networks/\(visited on 14/05/2019\)](https://machinelearningmastery.com/how-to-improve-performance-with-transfer-learning-for-deep-learning-neural-networks/(visited%20on%2014/05/2019)), 2019. Citado na página 22.
- CHEN, Z. et al. Resnet and model fusion for automatic spoofing detection. In: *INTERSPEECH*. [S.l.: s.n.], 2017. p. 102–106. Citado na página 21.
- DU, X. et al. Fused dnn: A deep neural network fusion approach to fast and robust pedestrian detection. In: *IEEE. 2017 IEEE winter conference on applications of computer vision (WACV)*. [S.l.], 2017. p. 953–961. Citado na página 21.
- GIANNINI, T. C. Desafios atuais da modelagem preditiva de distribuição de espécies. *Rodriguésia-Instituto de Pesquisas Jardim Botânico do Rio de Janeiro*, v. 63, n. 3, 2012. Citado na página 25.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016. Citado na página 23.
- HE, K. et al. 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. Citado na página 20.
- HELABS. Deep learning é o motor que propulsiona a inteligência artificial. URL: [https://helabs.com/blog/deep-learning-propulsiona-a-inteligencia-artificial/\(visited on 14/05/2019\)](https://helabs.com/blog/deep-learning-propulsiona-a-inteligencia-artificial/(visited%20on%2014/05/2019)), 2019. Citado na página 18.
- HU, W. et al. Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, Hindawi, v. 2015, 2015. Citado na página 22.
- JAISWAL, A. K. et al. Semi-supervised learning for cancer detection of lymph node metastases. *arXiv preprint arXiv:1906.09587*, 2019. Citado 2 vezes nas páginas 15 e 16.

- KOVALESKI, P. de A. *Implementação de Redes Neurais Profundas para Reconhecimento de ações em Vídeo*. Tese (Doutorado) — Universidade Federal do Rio de Janeiro, 2018. Citado na página 18.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 20.
- KUMAR, A. Ml metrics – sensitivity vs specificity difference. URL: [https://vitalflux.com/ml-metrics-sensitivity-vs-specificity-difference/\(visited on 14/05/2019\)](https://vitalflux.com/ml-metrics-sensitivity-vs-specificity-difference/(visited on 14/05/2019)), 2018. Citado 2 vezes nas páginas 23 e 24.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977. Citado na página 24.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015. Citado na página 18.
- LINMANS, J. et al. Sample efficient semantic segmentation using rotation equivariant convolutional networks. *arXiv preprint arXiv:1807.00583*, 2018. Citado 2 vezes nas páginas 15 e 16.
- LIU, Y. et al. Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442*, 2017. Citado 2 vezes nas páginas 13 e 15.
- MISHRA, A. Metrics to evaluate your machine learning algorithm. URL: [https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234/\(visited on 14/05/2019\)](https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234/(visited on 14/05/2019)), 2018. Citado na página 23.
- PONTI MOACIR ANTONELLI, D. C. G. B. P. Como funciona o deep learning. 2018. Citado 2 vezes nas páginas 18 e 21.
- ROSEBROCK, A. *Deep Learning for Computer Vision with Python: ImageNet Bundle*. [S.l.]: PyImageSearch, 2017. Citado na página 19.
- SAHA, S. A comprehensive guide to convolutional neural networks?—?the eli5 way. URL: [https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53/\(visited on 30/04/2019\)](https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53/(visited on 30/04/2019)), 2018. Citado na página 19.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado na página 20.
- SUBODH, M. Transfer learning using tensorflow. URL: [https://medium.com/@subodh.malgonde/transfer-learning-using-tensorflow-52a4f6bcde3e/\(visited on 06/05/2019\)](https://medium.com/@subodh.malgonde/transfer-learning-using-tensorflow-52a4f6bcde3e/(visited on 06/05/2019)), 2018. Citado na página 22.
- TORREY, L.; SHAVLIK, J. Transfer learning. In: *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. [S.l.]: IGI Global, 2010. p. 242–264. Citado na página 21.
- VEELING, B. S. et al. Rotation equivariant cnns for digital pathology. In: SPRINGER. *International Conference on Medical image computing and computer-assisted intervention*. [S.l.], 2018. p. 210–218. Citado 5 vezes nas páginas 13, 14, 15, 16 e 31.

WANG, D. et al. Deep learning for identifying metastatic breast cancer. *arXiv preprint arXiv:1606.05718*, 2016. Citado na página [15](#).



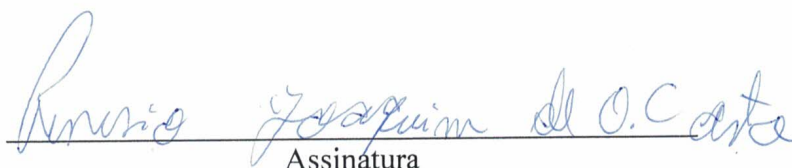
TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”

Identificação do Tipo de Documento

- () Tese
() Dissertação
(X) Monografia
() Artigo

Eu, **Renesio Joaquim de Oliveira Costa**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação **Redes Neurais Convolucionais para Identificação de Metástase em Imagens Histopatológicas**, de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 07 de FEVEREIRO de 2020.


Assinatura