

Rafael Pedro da Silva Neto  
Orientador: Antônio Oséas de Cavalho Filho

# **Metodologia automática para classificação de lesões de mama utilizando Transfer Learning**

Picos - PI  
Dezembro de 2019

Rafael Pedro da Silva Neto  
Orientador: Antônio Oséas de Cavalho Filho

## **Metodologia automática para classificação de lesões de mama utilizando Transfer Learning**

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 de 2019

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

**S586m** Silva Neto, Rafael Pedro da.  
Metodologia automática para classificação de lesões de mama utilizando Transfer Learning. / Rafael Pedro da Silva Neto. -- Picos,PI, 2019.  
37 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. Antônio Oséas de Cavalho Filho.”

1. Processamento Digital de Imagens. 2. Lesões de Mama. 3. Transfer Learning. I. Título.

**CDD 004.54**

*Elaborada por Rafael Gomes de Sousa CRB 3/1163*

METODOLOGIA AUTOMÁTICA PARA CLASSIFICAÇÃO DE LESÕES DE MAMA  
UTILIZANDO TRANSFER LEARNING

RAFAEL PEDRO DA SILVA NETO

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

  
\_\_\_\_\_  
Prof. Antônio Oseas de Carvalho Filho

Thiago José Barbosa Lima  
\_\_\_\_\_  
Prof. Thiago José Barbosa Lima

Jonnison Lima Ferreira  
\_\_\_\_\_  
Prof. Jonnison Lima Ferreira

# Agradecimentos

Agradeço, primeiramente e acima de todos, a meu Deus, digno de toda Honra e toda Glória, por ter me concebido a vida, permitindo que eu chegasse até aqui e me fornecido muitas oportunidades e auxílios.

Ao meu orientador, Professor Dr. Antônio Oséas, por reservar parte de seu precioso tempo para revisar, corrigir e auxiliar de outras diversas maneiras a conclusão deste trabalho.

Ao meu pai José Eurípedes de Paula Vasconcelos e minha mãe Marli Marlene da Silva Vasconcelos por terem me educado conforme os valores e tradições da moral cristã. Aos meus avós, tios, tias e todos os membros de minha família que me deram apoio.

Aos meus amigos Mateus, Renésio, Jonycassio, Aroldo Júnior, Eduardo e muitos outros, os quais me incentivaram em diversos momentos difíceis desta jornada árdua, compartilhando noites em claro, várias dificuldades e vitórias.

À minha namorada Ingrid Maressa, por estar comigo em cada momento difícil me incentivando e fortalecendo.

A todos aqueles que, direta ou indiretamente, fizeram parte da minha formação.

*Peçam e vocês receberão; procurem e vocês acharão; batam, e a porta será aberta para  
vocês.*

*Mateus 7:7*

# Resumo

As lesões de Mama é um dos tipos de lesões mais comuns entre as mulheres no Brasil e no mundo, respondendo por cerca de 28% dos novos casos a cada ano. Estatísticas indicam aumento da sua incidência tanto nos países desenvolvidos quanto nos em desenvolvimento. Quando os pacientes são diagnosticados com a lesão, precisam saber se o mesmo mostra-se Benigno ou Maligno. Tal informação pode ser obtida por meio de exames de biopsia, que é um exame de diagnóstico no qual o médico retira um pedaço de tecido do interior da mama, geralmente de um nódulo, para avaliar no laboratório e observar se existem células lesionadas. Nesse trabalho, foi desenvolvido um método computacional para classificação de imagens diferenciando lesões malignas e benignas de mama. Para tal método, foram testadas diferentes arquiteturas de Redes Neurais Convolucionais (CNN). O Transfer Learning foi empregada para lidar com a limitação da pequena quantidade de imagens na base de dados. O conjunto de imagens utilizado foi o Digital Database for Screening Mammography (DDSM). Com base nos testes realizados, a ResNet foi a arquitetura que obteve o melhor resultado utilizando do classificador Suport Vector Machine (SVM) com kernel Polinomial, atingindo uma precisão de 81,73%, uma sensibilidade de 85,66%, uma especificidade de 78,40%, Kappa de 0,63 e curva ROC de 0,82. Finalmente, acreditamos que nosso método pode integrar uma ferramenta de CAD atuando como triagem de pacientes ou fornecendo uma segunda opinião ao especialista.

**Palavras-chaves:** Lesões de mama, Processamento digital de imagens, Transfer Learning.

# Abstract

Breast lesions are one of the most common types of injuries among women in Brazil and worldwide, accounting for about 28% of new cases each year. Statistics indicate increased incidence in both developed and developing countries. When patients are diagnosed with the lesion, they need to know if it is Benign or Malignant. This information can be obtained through biopsy tests, which is a diagnostic test in which the doctor removes a piece of tissue from inside the breast, usually from a lump, to evaluate in the laboratory and see if there are any injured cells. In this work, we developed a computational method for image classification differentiating malignant and benign breast lesions. For such method, different Convolutional Neural Networks (CNN) architectures were tested. Transfer Learning was employed to deal with limiting the small amount of images in the database. The set of images used was the Digital Database for Screening Mammography (DDSM). Based on the tests performed, ResNet was the architecture that obtained the best result using the Support Vector Machine (SVM) classifier with Polynomial kernel, reaching an accuracy of 81.73 %, a sensitivity of 85.66 %, a specificity 78.40%, Kappa 0.63 and ROC curve 0.82. Finally, we believe that our method can integrate a CAD tool acting as patient screening or providing a second opinion to the specialist.

**Keywords:** Breast Lesions, Digital Image Processing, Transfer Learning.



# Lista de ilustrações

Figura 1 – Metodologia proposta. . . . .	17
Figura 2 – Benigno (A) e Maligno (B). . . . .	18
Figura 3 – Ilustração de uma ANN. . . . .	19
Figura 4 – Ilustração de uma rede de <i>Deep Learning</i> . . . . .	20
Figura 5 – Ilustração da arquitetura de uma CNN. . . . .	22
Figura 6 – Ilustração do <i>Transfer Learning</i> . . . . .	23
Figura 7 – Ilustração da arquitetura da rede ResNet. . . . .	24
Figura 8 – Ilustração da arquitetura da rede VGG. . . . .	25
Figura 9 – Ilustração da arquitetura da rede Xception. . . . .	26

# Lista de tabelas

Tabela 1 – Comparação entre os trabalhos relacionados. . . . .	15
Tabela 2 – Nível de exatidão de uma classificação, conforme o valor de índice <i>Kappa</i>	29
Tabela 3 – Descrição do Ambiente de Desenvolvimento . . . . .	30
Tabela 4 – Resultados da rede Resnet. . . . .	31
Tabela 5 – Resultados da rede Xception. . . . .	31
Tabela 6 – Resultados da rede VGG16. . . . .	31
Tabela 7 – Resultados da rede VGG19. . . . .	32
Tabela 8 – Resultados de todas as features concatenadas. . . . .	32
Tabela 9 – Tabela de comparação dos resultados . . . . .	33

# Lista de abreviaturas e siglas

ANN	<i>Artificial Neural Network</i>
BCDR	<i>Breast Cancer Digital Repository</i>
CAD	<i>Computer-Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
DDSM	<i>Database for Screening Mammography</i>
HGD	<i>Histogram of the Gradient Divergence</i>
HOG	<i>Histogram of Oriented Gradients</i>
ICIAR	<i>International Conference on Image Analysis and Recognition</i>
INCA	Instituto Nacional de Câncer
LBP	<i>Local Binary Patterns</i>
RBF	<i>Radial Basis Function</i>
ROI	<i>Regions of Interest</i>
SVM	<i>Support Vector Machine</i>
VGG	<i>Visual Geometry Group</i>
XGBoost	<i>eXtreme Gradient Boost</i>

# Lista de símbolos

$\theta$  Theta

# Sumário

<b>1</b>	<b>Introdução</b>	<b>13</b>
1.1	Objetivo Geral	14
1.2	Publicações	14
1.3	Organização do Trabalho	14
<b>2</b>	<b>Trabalhos Relacionados</b>	<b>15</b>
<b>3</b>	<b>Materiais e Métodos</b>	<b>17</b>
3.1	Base de Dados	18
<b>4</b>	<b>Referencial Teórico</b>	<b>19</b>
4.1	<i>Artificial Neural Network</i>	19
4.2	<i>Deep Learning</i>	20
4.2.1	<i>Convolutional Neural Network</i>	21
4.3	<i>Transfer Learning</i>	22
4.4	Arquitetura das Redes Avaliadas	23
4.4.1	Modelo 1 - ResNet	24
4.4.2	Modelo 2 - VGG	24
4.4.3	Modelo 3 - Xception	25
4.5	Classificação	26
4.5.1	<i>Support Vector Machine</i>	26
4.5.2	<i>Random Forest</i>	27
4.5.3	<i>eXtreme Gradient Boost</i>	27
4.6	Métricas de Avaliação	27
4.6.1	Acurácia	27
4.6.2	Sensibilidade	28
4.6.3	Especificidade	28
4.6.4	<i>Kappa</i>	28
4.6.5	Área Sob a Curva ROC	29
<b>5</b>	<b>Resultados e Discussões</b>	<b>30</b>
5.1	Treinamento e Avaliação	30
5.2	Testes Avaliativos	30
<b>6</b>	<b>Conclusão</b>	<b>34</b>
	<b>Referências</b>	<b>35</b>

# 1 Introdução

Atualmente, o câncer de mama é uma das principais ocorrências de câncer em mulheres com mais de 40 anos de idade em todo o mundo. Um diagnóstico precoce é muito importante, e um método comum de diagnóstico é o uso da mamografia, que é basicamente um raio-X da região da mama tirada em uma condição especial. A partir da imagem, um médico treinado o examina buscando microcalcificações e massas de várias formas. Se encontrados, esses artefatos na imagem podem ser um sinal da presença de um tumor benigno ou maligno (FERREIRA; BORGES, 2003).

Segundo Pontes (2013), a diferença entre benigno e maligno é definida pela aparência e estrutura das células atacadas pelo tumor. Os tumores benignos são constituídos por células bem semelhantes às que os originaram e não possuem a capacidade de provocar metástases. Já os malignos são agressivos e possuem a capacidade de infiltrar outros órgãos. A multiplicação celular é “controlada” pelos chamados oncógenes ativadores, que regulam o crescimento e a morte da célula. No tumor benigno há uma mutação na estrutura genética dos oncógenes, mas nada capaz de “descontrolá-la”. O mesmo não ocorre com os tumores malignos, que crescem sem controle por conta da alteração genética.

Ainda não é possível prevenir o câncer de mama em função da multiplicidade de fatores relacionados ao surgimento da doença, desta forma, o diagnóstico precoce é a melhor forma de combater a doença dentre os métodos para detecção e classificação de câncer de mama. Tramonte et al. (2016) aponta que o diagnóstico do câncer de mama, em países com recursos limitados, é frequentemente feito em um estágio já avançado.

Especificamente, durante o procedimento de diagnóstico, os especialistas avaliam a organização do tecido por meio de imagens microscópicas. No entanto, a grande quantidade de dados e a complexidade das imagens tornam essa tarefa custosa e não trivial. O método tradicional de inspeção microscópica é importantíssimo em analisar a estrutura do tecido. No entanto, esta técnica de inspeção manual é altamente demorada e propensa a erros manuais, principalmente no processo de classificação em lesões malignas ou benignas, pois, é nesse diagnóstico que é fornecido a informação de espalhamento de células cancerosas para os tecidos circundantes a partir de seu ponto de ocorrência (INCA, 2018).

Os sistemas de detecção auxiliada por computador (*Computer-Aided Diagnosis* - CAD) foram desenvolvidos para reduzir os gastos e melhorar a capacidade de inúmeros profissionais que lidam com imagens. Pesquisas efetuadas mostraram que o uso de tais sistemas estão se tornando primordiais na resolução de diagnóstico médico por meio de imagens Elsheikh, Mustafa e Ibraheem (2018). Um problema relacionado à eles é seu alto custo computacional, além de ser preciso um longo tempo para trazer um resultado e necessitar de uma grande quantidade de dados para processamento.

## 1.1 Objetivo Geral

Em vista do contexto apresentado, este trabalho tem como objetivo geral desenvolver um método computacional de baixo custo computacional e eficiente para a classificação automática de imagens de lesões de mama em benignas e malignas em resolução ao problema de uma limitada base de imagens, reduzindo resultados falso positivo e falso negativo.

Para a concretização do objetivo geral, tem-se os seguintes objetivos específicos:

- Avaliar a viabilidade do uso das CNNs por meio do teste de diferentes arquiteturas de redes e o uso do Transfer Learning nas arquiteturas aplicadas ao problema descrito.
- Utilizar métodos de avaliação das lesões de mama utilizado pelos especialistas com intuito de empregar na etapa de validação dos resultados obtidos pelo método proposto;
- Construir uma ferramenta capaz de quantificar automaticamente regiões de lesões de mama.

## 1.2 Publicações

- NETO, R.; CARVALHO, A. O. Transfer learning aplicado a classificação de lesões mamárias. Simpósio de Sistemas de Informação, 2019.
- NETO, R.; CARVALHO, A. O. Automatic classification of breast lesions using transfer learning. IEEE Latin America Transactions, 2019. **(aceito para publicação)**

## 1.3 Organização do Trabalho

Os próximos capítulos estão dispostos da seguinte forma: o Capítulo 2 contém, de forma breve, descrições dos principais trabalhos relacionados encontrados na literatura. No Capítulo 3 apresenta os materiais e métodos abordados no trabalho, onde são explicados as etapas de desenvolvimento da pesquisa, iniciando pela aquisição das imagens, métricas de avaliação, CNNs e a descrição das arquiteturas e descritores comparados. O Capítulo 4 trata de referenciais teóricos para o entendimento do trabalho desenvolvido. O Capítulo 5 discute os resultados obtidos com a metodologia proposta. E por fim, no Capítulo 6 são apresentadas as conclusões e trabalhos futuros.

## 2 Trabalhos Relacionados

Estudos conhecidos sobre o problema da classificação de imagens vem se tornando relevantes com o passar dos anos e se popularizando no meio científico. Com isso, a Tabela 1 mostra um resumo das principais informações dos trabalhos relacionados para a classificação de lesões de mama, em seguida, apresentamos um resumo dos trabalhos relacionados.

Tabela 1 – Comparação entre os trabalhos relacionados.

Trabalho	Base de Imagens	Algoritmos	Acurácia
( <a href="#">RAKHLIN et al., 2018</a> )	ICIAR 2018	ResNet, VGG e Inception	87,2%
( <a href="#">GUO et al., 2018</a> )	ICIAR 2018	GoogleNet	80,0%
( <a href="#">SILVA, 2016</a> )	DDSM	LBP	90,18%
( <a href="#">AREVALO et al., 2016</a> )	BCDR-FM	CNN, HOG e HGD	–

Em ([RAKHLIN et al., 2018](#)), foi proposto uma abordagem computacional baseada em redes neurais profundas, onde são utilizadas algumas arquiteturas de redes (ResNet, VGG e Inception) pré-treinadas na base ImageNet para a classificação de imagens histológicas de câncer de mama da base *International Conference on Image Analysis and Recognition* (ICIAR) 2018.

([GUO et al., 2018](#)) propôs uma estratégia que utiliza CNN híbrida que é projetada com base na arquitetura GoogleNet para mesclar mais informações importantes na decisão, onde o *Transfer Learning* e aumento de dados foram empregados para lidar com as limitações da pequena quantidade de imagens na base ICIAR 2018.

Em ([SILVA, 2016](#)), utilizando a base *Digital Database for Screening Mammography* (DDSM), foram combinadas abordagens estrutural e estatística para a análise de textura. A abordagem estrutural foi realizada por meio da técnica de Local Binary Patterns (LBP) e as abordagens estatísticas, através das técnicas da função K de Ripley e de índice de diversidade.

O método proposto por ([AREVALO et al., 2016](#)), apresenta uma metodologia para classificação automática das lesões de imagem da mama utilizando representação do aprendizado. O método é avaliado de dados de benchmarking comprovado por biópsia 736 casos. O método desenvolvido compreende dois estágios principais: (i) pré-processamento para aprimorar os detalhes da imagem e (ii) treinamento supervisionado para aprender os recursos e a extração de características de lesões de imagem da mama usando *Histogram of Oriented Gradients* (HOG) e *Histogram of the Gradient Divergence* (HGD). Os resultados apresentam desempenho de 0,787 a 0,822 em termos da área sob a curva ROC.



A análise dos trabalhos encontrados contribuiu para identificar que as técnicas de aprendizagem profunda estão sendo bastante utilizadas em aplicações de imagens médicas. Também observou-se o uso de hardwares com grande capacidade de processamento para atingir os objetivos propostos por cada trabalho, por isso, foi investigado nesse trabalho o uso dessas técnicas com um hardware de baixo poder de processamento. Para contornar o problema da baixa quantidade de imagens e da limitação do hardware utilizado neste trabalho, foi utilizado o Transfer Learning que será abordada mais a frente. Com as pesquisas realizadas, viu-se necessário a utilização de classificadores externos para a validação. Foram então utilizados os classificadores Support Vector Machine (SVM), Random Forest e eXtreme Gradient Boost (XGBoost). Em seguida, veio a etapa de desenvolvimento e testes computacionais.

### 3 Materiais e Métodos

Nessa seção são apresentados a base de imagens, conceitos importantes para uma melhor compreensão deste trabalho, as principais arquiteturas de CNNs utilizadas para o processamento das imagens. Por fim são apresentadas as métricas de avaliação dos resultados obtidos.

A primeira etapa foi a coleta das imagens de duas classes de lesões de mama, malignas e benignas. Para compor a base de dados utilizada nesse trabalho, foram selecionadas imagens do arquivo da DDSM, por ser de fácil acesso e aberta ao público.

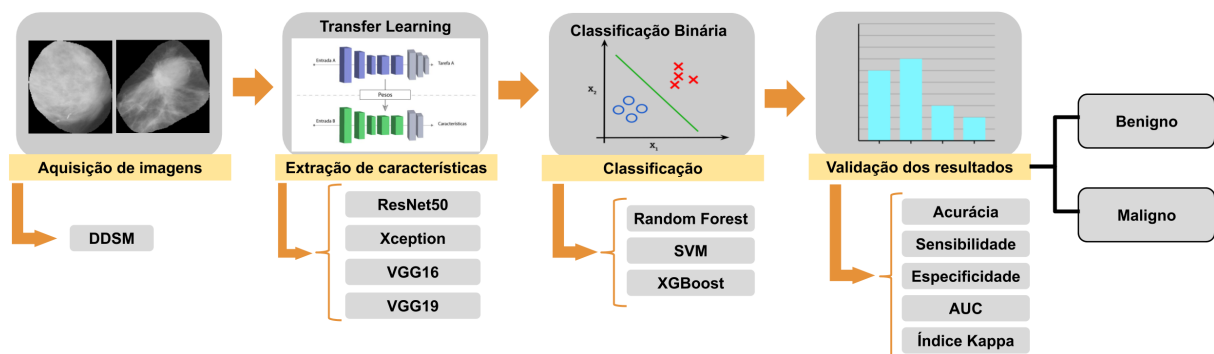
A segunda etapa consistiu em extrair características dessas imagens utilizando as redes pré-treinadas com pesos da ImageNet, disponíveis na biblioteca Keras. Houve um processamento da base para alterar o tamanho de entrada dos dados para o formato exigido por cada uma das redes e, devido o *Transfer Learning*, as suas camadas de classificação foram removidas.

Na terceira etapa, foram utilizados os classificadores: SVM, *Random Forest* e o XGBoost, com as características extraídas das arquiteturas para a validação. Nesta etapa, foi utilizado o método *Cross Validation*. Foram testados ao todo três kernels diferentes do SVM: Linear, Polynomial e RBF, numa tentativa de alcançar maiores resultados.

Por fim, na quarta e última etapa, para a avaliação do método proposto nesse trabalho foram utilizadas as métricas: Sensibilidade (S), Especificidade (E), Acurácia (A), Área Sob a Curva ROC (AUC) e o índice *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).

A Figura 1 descreve um resumo das etapas do método utilizado.

Figura 1 – Metodologia proposta.



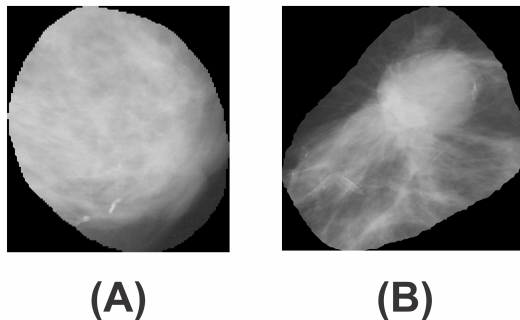
Fonte: Elaborada pelo autor.

### 3.1 Base de Dados

A DDSM (HEATH et al., 2000) é uma base de imagens públicas contendo 2.620 estudos/exames de mamografia digitais. Contém casos normais, benignos e malignos com informações patológicas confirmadas. A DDSM está disponível para uso da comunidade de pesquisa de análises mamográficas de imagens. O apoio primário para este projeto foi uma doação do Programa de Pesquisa do Câncer de Mama do Comando de Pesquisa Médica e Material do Exército dos EUA. A DDSM é mantida na Universidade do Sul da Flórida com o objetivo de mantê-lo acessível na web.

Todas os exames possuem marcações indicando os contornos das lesões. Estas marcações são realizadas por um conjunto de especialistas. Com isso, em nosso trabalho, utilizamos as marcações disponíveis para segmentar as *Region of Interest* (ROI) e com isso aplicar o método proposto. A Figura 2 mostra dois exemplos de imagens da base de dados da DDSM, uma de lesão benigna (A) e outra de lesão maligna (B) após segmentação e extração da ROI.

Figura 2 – Benigno (A) e Maligno (B).



Fonte: Banco de Dados da DDSM

Tradicionalmente as mais variadas tecnologias são utilizadas para oferecer um ótimo desempenho no processamento de imagens e uma dessas está rapidamente se tornando um dos mais estudados e procurados campos dentro da ciência da computação moderna, onde ela oferece facilidades e poder computacional. Uma dessas tecnologias que foi abordada nesse trabalho é denominada *Artificial Neural Network* (ANN) e ela será apresentada a seguir.

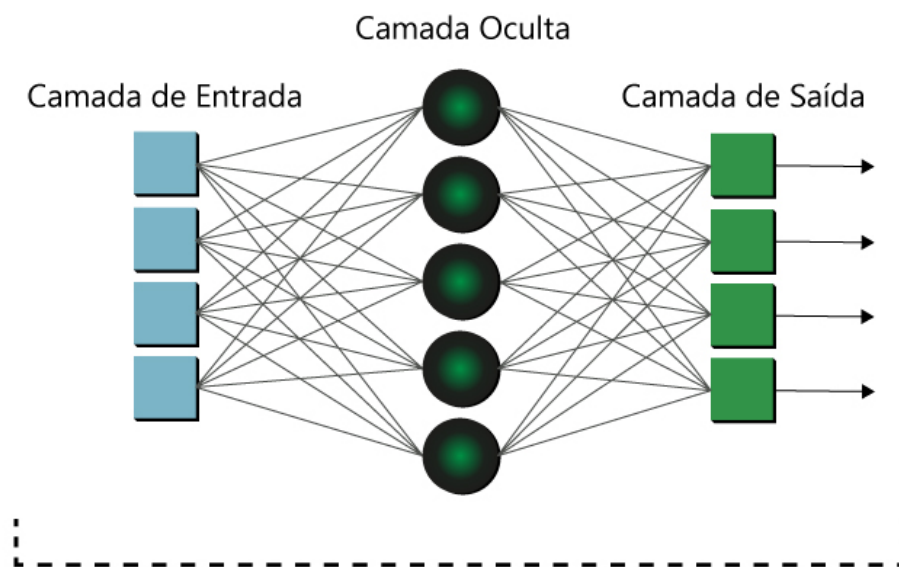
## 4 Referencial Teórico

Este capítulo aborda tópicos relevantes ao contexto deste trabalho: conceitos sobre ANN, seus métodos e arquiteturas avaliadas; conceitos sobre os classificadores utilizados e, por fim; as etapas de validação de resultados.

### 4.1 *Artificial Neural Network*

As ANN's são redes computacionais que tentam, a grosso modo, simular o processo de decisão de uma rede de célula nervosa do sistema nervoso central biológico humano ou animal (DANIEL, 2013). São capazes de aprender à partir de um conjunto de dados, trabalhando de forma especializada com reconhecimento de padrões. Elas, semelhante ao sistema nervoso, possuem componentes lógicos (Neurônios Artificiais), organizados em um espécie de rede, que recebem entradas (características) e passam por uma camada oculta que decide quais as saídas mais adequadas a cada uma dessas entradas. A Figura 3 traz a ilustração de uma ANN.

Figura 3 – Ilustração de uma ANN.



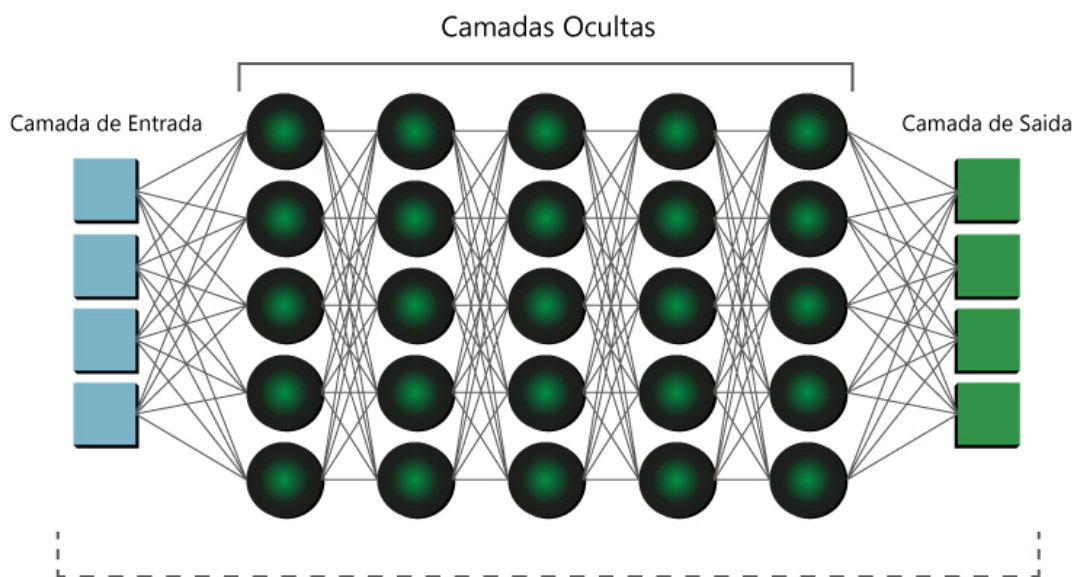
Fonte: Adaptação de Vázquez (2017).

## 4.2 Deep Learning

A Aprendizagem Profunda (*Deep Learning* em inglês) é uma sub área do aprendizado de máquina que trata de modelos vagamente inspirados no cérebro humano. Tecnicamente, toda Rede Neural 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 \(2015\)](#), o *Deep Learning* é uma área que contém uma pilha de multicamadas simples, todos (ou a maioria) estão sujeitos à aprendizagem, e muitos dos quais calcula mapeamentos de entrada-saída não lineares.

A profundidade do modelo é determinada pela quantidade de camadas ocultas. Os modelos de *Deep Learning* atuais podem conter dezenas ou até centenas de camadas ocultas e todas são aprendidas automaticamente necessitando apenas da inserção dos dados de treinamento. A Figura 4 nos mostra uma ilustração de uma rede de *Deep Learning*.

Figura 4 – Ilustração de uma rede de *Deep Learning*.



Fonte: Adaptação de [Vázquez \(2017\)](#)

Tradicionalmente, a qualidade dos algoritmos depende muito da representação dos dados em características. Assim, a criação recursos válidos e pré-processamento consomem grande parte dos esforços dos especialistas. Já os algoritmos do tipo Aprendizado Profundo 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 *Deep Learning* em alguns cenários possuem uma taxa de acerto

maior que a de humanos. Várias de suas arquiteturas 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.

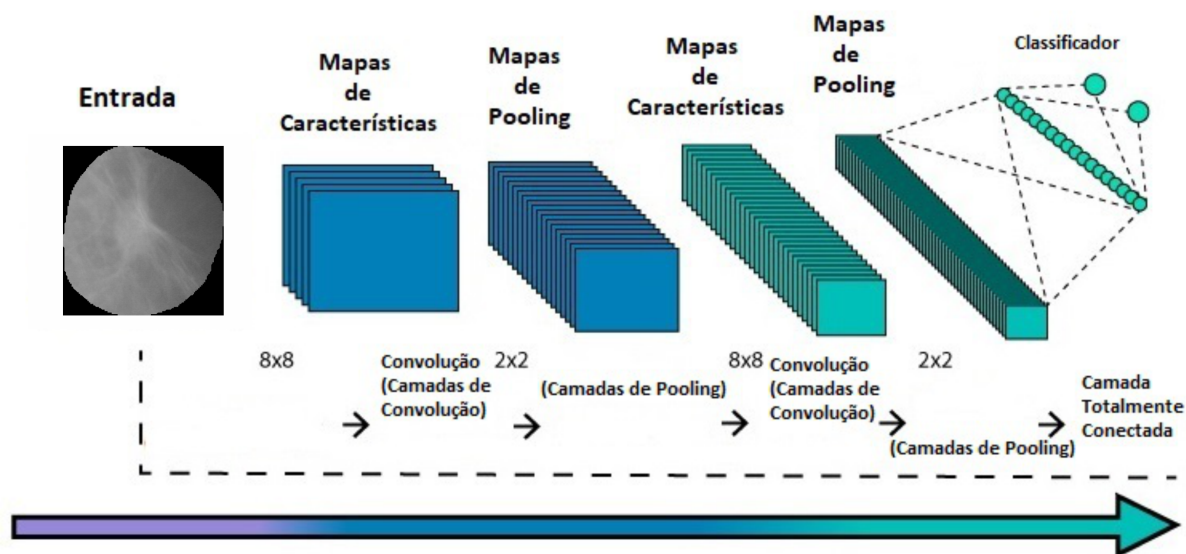
Ao passar dos anos surgiu uma categoria de ANN baseada em *Deep Learning* que inovaram por utilizar convolução em algumas de suas camadas. Essa categoria que passou a ser o novo padrão de visão computacional, é chamada de Convolutional Neural Network (CNN), que será descrita a seguir.

#### 4.2.1 *Convolutional Neural Network*

AS CNN's 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 projetados 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 5.

Figura 5 – Ilustração da arquitetura de uma CNN.



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](#)).

Quanto maior a quantidade de amostras (base de dados) que representam as diferentes classes, mais fácil será o treinamento de uma CNN. Mas, nem sempre é possível utilizar grandes quantidades de imagens devido a empecilhos como barreiras de acesso a essas bases ou até mesmo limitação do hardware, o que atrapalha o uso dessa ferramenta. Para contornar esse problema, foi desenvolvido uma técnica chamada de *Transfer Learning*, que será descrita a seguir.

### 4.3 *Transfer Learning*

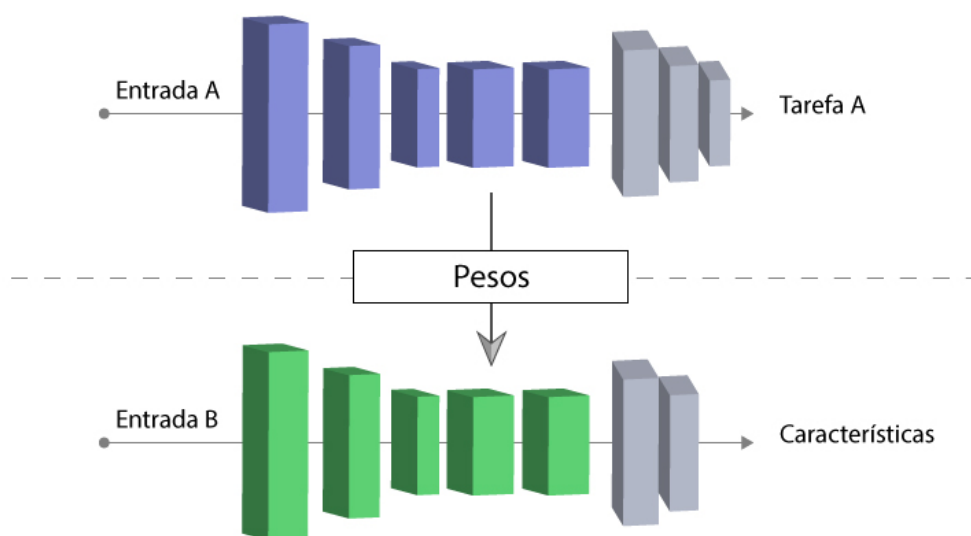
O *Transfer Learning* é um método de aprendizado de máquina em que um modelo desenvolvido para uma tarefa é reutilizado como ponto de partida para um modelo em

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 o *Transfer Learning* entra em jogo, porque com ela pode-se construir um modelo de aprendizado de máquina sólida com comparativamente poucos dados de treinamento, porque 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 *Transfer Learning* é demonstrado pela Figura 6.

Figura 6 – Ilustração do *Transfer Learning*.



Fonte: Adaptação de Subodh (2018).

Portanto, o *Transfer Learning* também economiza muito tempo de treinamento, pois pode-se levar dias ou até semanas para treinar uma Rede Neural profunda do zero em uma tarefa complexa e, além desse benefício, resulta em menor erro de generalização (GOODFELLOW; BENGIO; COURVILLE, 2016).

A seguir serão abordados as arquiteturas utilizadas que foram avaliados neste trabalho.

#### 4.4 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.

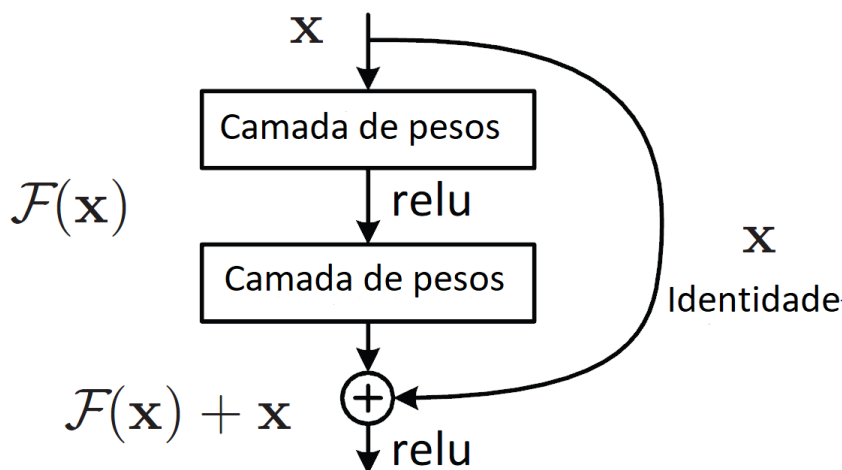


Com isso, foi possível selecionar três principais modelos para estudo mais profundo e consequente avaliação da melhor performance. Tais modelos são descritos à seguir.

#### 4.4.1 Modelo 1 - 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 retro propagado para camadas anteriores. Na Figura 7 é demonstrado o funcionamento de um bloco residual.

Figura 7 – Ilustração da 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 após a aplicação da regra da cadeia por várias vezes. 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.

#### 4.4.2 Modelo 2 - 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. A VGGNet é considerada a rede preferida pela comunidade para *Transfer Learning*, pois sua arquitetura uniforme é boa para extrair características de imagens. A Figura 8 nos mostra uma ilustração da arquitetura utilizada na VGG.

Figura 8 – Ilustração da arquitetura da rede VGG.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

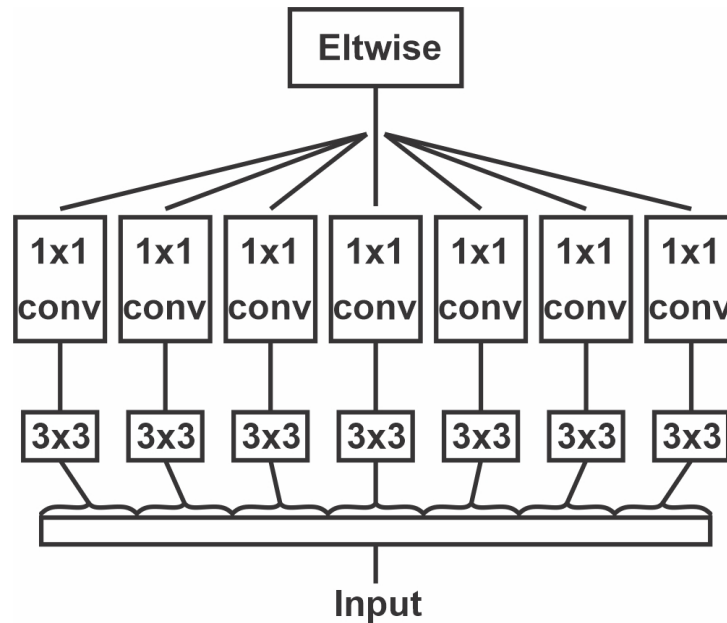
Fonte: Simonyan e Zisserman (2014).

#### 4.4.3 Modelo 3 - Xception

A arquitetura Xception foi proposta por (CHOLLET, 2017) e significa *Extreme version of Inception* sendo uma versão da arquitetura Inception, onde essa arquitetura tem 36 camadas convolucionais formando a base de extração de características da rede. A sua principal característica é que ela substitui o módulo Inception por convoluções separáveis

em profundidade, isso faz com que o número de conexões sejam menores e o modelo se torne mais leve. A Figura 9 demonstra um módulo da rede Xception.

Figura 9 – Ilustração da arquitetura da rede Xception.



Fonte: Adaptação de [ProgrammerSought \(2017\)](#).

A Figura 9 mostra que a arquitetura Xception em vez de particionar dados de entrada em vários blocos compactados, a mesma leva isso um passo adiante, ela mapeia as correlações espaciais para cada canal de saída separadamente e, em seguida, executa uma convolução de 1x1 em profundidade para capturar a correlação entre canais.

## 4.5 Classificação

A classificação consiste em reconhecer novos objetos de um conjunto de recursos extraídos de um grupo de objetos de interesse e, em seguida, treinar um sistema para reconhecer novos objetos desses recursos ([AZEVEDO; CONCI, 2008](#)). Para esta classificação, foram utilizados: SVM, Random Forest e Xgboost.

### 4.5.1 *Support Vector Machine*

É um algoritmo de classificação de aprendizado de máquina supervisionado proposto por [Vapnik \(2013\)](#), [Vapnik \(1998\)](#). Neste trabalho, usamos o SVM com o kernel da função base radial (RBF), parâmetro  $\gamma = 10$  e  $C = 10000$ . O valor de  $C$  foi escolhido alto para que uma pequena margem fosse aceita pela função de decisão na classificação correta de todos os pontos. O parâmetro  $\gamma$  foi escolhido baixo, de modo que um único exemplo teve pouca influência no treinamento de dados.

### 4.5.2 *Random Forest*

É uma combinação de múltiplas previsões de árvores em que cada árvore depende dos valores de um vetor independente, sorteado aleatoriamente e com a mesma distribuição para todas as árvores da floresta (BREIMAN, 2001). Para este trabalho, o *Random Forest* foi usada com os seguintes parâmetros: bag Size Percent = 100, batch Size = 100, número de execuções slots = 1, max Depth = 0 (ilimitado), número de atributos escolhidos aleatoriamente = 0, número de iterações a ser realizado = 100, número mínimo de instâncias por folha = 1,0, variação mínima para divisão = 0,001 e o número aleatório semente a ser usado = 1.

### 4.5.3 *eXtreme Gradient Boost*

É uma biblioteca proposta por Chen et al. (2015) com base na estrutura de aumento de gradiente desenvolvida por Friedman (2001), Friedman, Hastie e Tibshirani (2000). O Xgboost pode ser usado para várias funções objetivas, como regressão e classificação (SONG et al., 2016). O Xgboost é extensível e seus parâmetros podem ser facilmente alterados. Os parâmetros utilizados no XGBoost foram os seguintes: profundidade máxima = 7, taxa de aprendizado = 0,1, ite = 1000, gama = 0, passo delta máximo = 1 e objetivo = "multi:softmax".

## 4.6 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).

### 4.6.1 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. Em problemas com classes desproporcionais, ela pode causar uma falsa impressão de bom desempenho. A acurácia pode ser calculada conforme a Equação 4.1.

$$A = \frac{VP + VN}{VP + VN + FP + FN} \quad (4.1)$$

### 4.6.2 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. A sensibilidade pode ser calculada conforme a Equação 4.2

$$S = \frac{VP}{VP + FN} \quad (4.2)$$

### 4.6.3 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. A especificidade pode ser calculada com a Equação 4.3

$$E = \frac{VN}{VN + FP} \quad (4.3)$$

### 4.6.4 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 4.4

$$K = \frac{\theta_1 - \theta_2}{1 - \theta_2}, \quad (4.4)$$

onde

$$\theta_1 = \frac{VP + VN}{VP + VN + FP + FN}, \quad (4.5)$$

e

$$\theta_2 = \frac{[(VP + FN)(VP + FP)] + [(VN + FN)(VN + FP)]}{(VP + VN + FP + FN)^2} \quad (4.6)$$

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*

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

Fonte: [Araujo \(2014\)](#)

#### 4.6.5 Área Sob a Curva 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. 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 utilizado para resolver o problema proposto.

## 5 Resultados e Discussões

Este capítulo mostra e discute os resultados obtidos com a metodologia proposta por esse trabalho, obtendo assim, a classificação das regiões de interesse de imagens de lesões de mama. Além disso, ele apresenta o comparativo das arquiteturas de redes convolucionais apresentadas na seção 4.4.

### 5.1 Treinamento e Avaliação

Nesse trabalho, propomos um método automático para classificação de imagens de lesões de mama em duas classes (benigno e maligno) que foram coletadas da base DDSM usando abordagens de *Transfer Learning*, sendo 1155 ROIs, compostas de 530 lesões benignas e 625 malignas.

Foram utilizadas redes pré-treinadas com pesos da ImageNet (DENG et al., 2009) para extração das características dessas imagens. No processo final de classificação é utilizado os classificadores SVM com kernels (Linear, RBF e Polynomial), Random Forest e o XGBoost, conforme descrito na Seção 4.5.

O método de avaliação empregado nesse trabalho é o *Cross Validation*, onde consiste em dividir a base de imagens em  $k$  partes, usando  $k-1$  partes para treino e a parte remanescente para teste, fazendo isso  $k$  vezes. Também foi feito a concatenação de todas as características extraídas pelos descritores para classificação, e por fim, fazemos um último teste que conta com a seleção de atributos utilizando o *Greedy StepWise* (CARUANA; FREITAG, 1994).

### 5.2 Testes Avaliativos

Para a realização dos experimentos aqui descritos, utilizou-se um computador, cujas especificações são descritas na Tabela 3.

Tabela 3 – Descrição do Ambiente de Desenvolvimento

Características	Descrição
Modelo	Acer Aspire VX5-591G-54PG
RAM	8,00 GB
Processador	Intel Core i5-7300HQ 2.5 GHz
Sistema Operacional	Windows 10

Nesse trabalho foram empregados dois cenários de testes. No primeiro cenário, temos os resultados com todas as características dos modelos. Já no segundo cenário, temos os resultados com a seleção de atributos feitas com *Greedy StepWise*.

A seguir, os resultados obtidos com a metodologia proposta serão apresentados e discutidos. As métricas nas tabelas são: Acurácia (A), Sensibilidade (S), Especificidade (E), *Kappa* (K) e área sob a curva ROC (AUC). Os melhores resultados obtidos em cada método estão destacados em negrito.

Tabela 4 – Resultados da rede Resnet.

Rede	Classificador	A(%)	S(%)	E(%)	K	AUC
Resnet50 Sem seleção 2048 features	RF	79,48	81,88	77,44	0,58	0,85
	SVM LINEAR	78,70	77,54	79,68	0,57	0,78
	SVM RBF	81,38	85,47	77,76	0,62	0,81
	SVM POLY	<b>81,73</b>	<b>85,66</b>	<b>78,40</b>	<b>0,63</b>	<b>0,82</b>
	XGBoost	77,84	80,84	75,84	0,56	0,78
ResNet50 Com seleção 137 features	RF	78,00	81,13	75,36	0,56	0,85
	SVM LINEAR	<b>79,22</b>	<b>83,20</b>	<b>75,84</b>	<b>0,58</b>	<b>0,79</b>
	SVM RBF	76,71	79,62	74,24	0,53	0,76
	SVM POLY	78,96	85,47	73,44	0,58	0,79
	XGBoost	75,06	78,11	72,48	0,50	0,75

Tabela 5 – Resultados da rede Xception.

Rede	Classificador	A(%)	S(%)	E(%)	K	AUC
Xception Sem seleção 2048 features	RF	78,09	80,18	76,32	0,56	0,85
	SVM LINEAR	76,36	76,41	76,32	0,52	0,76
	SVM RBF	79,91	83,58	76,80	0,59	0,80
	SVM POLY	<b>80,00</b>	<b>83,96</b>	<b>76,64</b>	<b>0,60</b>	<b>0,80</b>
	XGBoost	76,28	78,11	74,72	0,53	0,76
Xception Com seleção 135 features	RF	77,66	80,37	76,96	0,55	0,85
	SVM LINEAR	78,44	82,26	75,20	0,56	0,78
	SVM RBF	76,96	85,47	69,76	0,54	0,77
	SVM POLY	<b>79,22</b>	<b>84,52</b>	<b>74,72</b>	<b>0,58</b>	<b>0,79</b>
	XGBoost	75,50	77,74	73,60	0,55	0,76

Tabela 6 – Resultados da rede VGG16.

Rede	Classificador	A(%)	S(%)	E(%)	K	AUC
VGG16 Sem seleção 4096 features	RF	78,44	85,09	72,80	0,57	0,85
	SVM LINEAR	78,70	79,05	77,92	0,57	0,78
	SVM RBF	79,65	86,98	73,44	0,59	0,80
	SVM POLY	<b>79,91</b>	<b>88,49</b>	<b>72,64</b>	<b>0,60</b>	<b>0,80</b>
	XGBoost	74,81	78,30	71,84	0,50	0,75
VGG16 Com seleção 120 features	RF	77,57	83,96	72,16	0,55	0,84
	SVM LINEAR	78,78	87,73	71,20	0,57	0,79
	SVM RBF	76,27	89,24	65,28	0,53	0,77
	SVM POLY	<b>78,87</b>	<b>90,00</b>	<b>69,44</b>	<b>0,58</b>	<b>0,79</b>
	XGBoost	74,72	80,38	69,92	0,50	0,75



Tabela 7 – Resultados da rede VGG19.

Rede	Classificador	A(%)	S(%)	E(%)	K	AUC
VGG19 Sem seleção 4096 features	RF	78,18	83,39	73,76	0,56	0,85
	SVM LINEAR	75,75	75,09	76,32	0,51	0,75
	SVM RBF	<b>79,48</b>	<b>88,22</b>	<b>73,76</b>	<b>0,59</b>	<b>0,80</b>
	SVM POLY	79,13	86,79	72,64	0,58	0,79
	XGBoost	76,10	79,06	73,60	0,52	0,76
VGG19 Com seleção 113 features	RF	<b>79,30</b>	<b>85,47</b>	<b>74,08</b>	<b>0,58</b>	<b>0,84</b>
	SVM LINEAR	77,57	83,96	72,16	0,55	0,78
	SVM RBF	76,36	83,01	70,72	0,53	0,76
	SVM POLY	78,00	86,98	70,40	0,56	0,78
	XGBoost	74,89	79,25	71,20	0,50	0,75

Tabela 8 – Resultados de todas as features concatenadas.

Rede	Classificador	A(%)	S(%)	E(%)	K	AUC
Concatenadas Sem seleção 12288 features	RF	78,78	81,88	76,16	0,57	0,85
	SVM LINEAR	77,48	76,41	78,40	0,54	0,77
	SVM RBF	80,80	80,62	79,42	0,60	0,80
	SVM POLY	<b>81,29</b>	<b>85,47</b>	<b>77,76</b>	<b>0,62</b>	<b>0,81</b>
	XGBoost	80,87	82,83	79,20	0,62	0,81
Concatenadas Com seleção 244 features	RF	80,51	84,15	77,44	0,61	0,79
	SVM LINEAR	80,51	84,71	76,96	0,61	0,80
	SVM RBF	80,34	86,60	75,04	0,60	0,80
	SVM POLY	<b>81,38</b>	<b>86,60</b>	<b>76,96</b>	<b>0,62</b>	<b>0,81</b>
	XGBoost	81,13	82,83	79,68	0,62	0,81

No primeiro cenário, foi obtido o melhor resultado com a ResNet50 com classificador SVM Polynomial, com acurácia de 81.73%, sensibilidade de 85.66%, especificidade de 78.40%, Kappa de 0.63 e curva ROC de 0,82. No segundo cenário, com seleção de atributos, não houveram melhoras nos resultados, sendo o melhor com a VGG19 com classificador *Randon Forest*, com com acurácia de 79.30%, sensibilidade de 85.47%, especificidade de 74.08%, Kappa de 0.58 e curva ROC de 0,84.

Tendo em vista todos os testes realizados, nota-se que o modelo ResNet50 obteve melhores resultados na maioria dos casos, assim como o classificador SVM com kernel Polynomial, que obteve melhores resultados que os demais classificadores. Um ponto de destaque é o valor do *Kappa*, pois com base na Tabela 2 o algoritmo utilizado tem um nível muito bom de exatidão na classificação. Os demais resultados obtidos foram satisfatórios de acordo com as métricas especificadas, indicando uma boa capacidade discriminativa e classificatória do algoritmo para as classes definidas.

A comparação com outros trabalhos dentro da literatura é uma tarefa um pouco complicada, já que há uma algumas diferenças nas maneiras de abordagem do problema

e também nas ferramentas e técnicas utilizadas para solucioná-los. A Tabela 9 apresenta um resumo quantitativo em relação aos trabalhos apresentados na Seção 2.

Tabela 9 – Tabela de comparação dos resultados

<b>Trabalho</b>	<b>A(%)</b>	<b>S(%)</b>	<b>E(%)</b>	<b>K</b>	<b>AUC</b>
( <a href="#">RAKHLIN et al., 2018</a> )	93,80	96,50	88,00	–	97,3
( <a href="#">GUO et al., 2018</a> )	80,00	–	–	–	–
( <a href="#">SILVA, 2016</a> )	90,18	–	–	–	–
( <a href="#">AREVALO et al., 2016</a> )	–	–	–	–	0,82
Esse trabalho	81,73	85,66	78,40	0,63	0,82

Em comparação com os demais trabalhos, este trabalho se mostra eficiente no uso do Transfer Learning na extração de características da base DDSM, visto que o resultado ultrapassou em acurácia o trabalho de [Guo et al. \(2018\)](#) e se igualou na métrica de AUC conseguido no trabalho de [Arevalo et al. \(2016\)](#).

Por fim, o método se mostra promissor, apesar de não alcançar resultados que excedam o estado da arte, foram aplicadas diferentes estratégias e arquiteturas que possibilitam estudos futuros com base nas experiências apresentadas em neste trabalho.

## 6 Conclusão

Neste trabalho foi avaliado e comparado o desempenho de quatro arquiteturas de redes utilizando *Transfer Learning* para a extração de características e três classificadores para a classificação automática de lesões de mama em imagens médicas.

A arquitetura ResNet50 atingiu o melhor resultado dentre os algoritmos usados neste trabalho, alcançando valores acima de 78% nas métricas de avaliação, com a utilização do classificador SVM com kernel Polynomial. Os bons resultados na avaliação fornecem oportunidades para propor novas metas de pesquisa na área de lesões de mama.

As arquiteturas de redes investigadas transferiram com sucesso o conhecimento do ImageNet codificados como características convolucionais para o problema na presença de dados limitados de treinamento.

O trabalho apresentado mostra uma boa classificação no uso da *Transfer Learning* evidenciando uma eficiente ferramenta para a análise automática de imagens, mesmo com um número limitado de imagens e recursos computacionais escassos.

Desta forma, este trabalho apresenta contribuições em duas áreas de estudo. Na área da saúde, oferecendo uma metodologia como segunda opinião a profissionais da área médica para auxiliar no diagnóstico de lesões de mama, e na área da computação, fornecendo informações importantes para a área de processamento de imagens, como forma de auxiliar pesquisas futuras e também de terceiros.

Como trabalhos futuros, pretende-se:

1. Utilizar outras arquiteturas de CNN como *Transfer Learning* para comparação de resultados;
2. Utilizar outros classificadores; e,
3. Usar métodos automáticos para estimar parâmetros para classificação, afim de melhorar os resultados obtidos.

# Referências

- ARAUJO, F. *Descoberta de conhecimento em base de dados para o aprendizado da regulação médica/odontológica em operadora de plano de saúde*. Tese (Doutorado) — Dissertação de Mestrado, 2014. Citado na página 29.
- AREVALO, J. et al. Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer Methods and Programs in Biomedicine*, v. 127, p. 248 – 257, 2016. ISSN 0169-2607. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169260715300110>>. Citado 2 vezes nas páginas 15 e 33.
- AZEVEDO, E.; CONCI, A. *Computação gráfica: teoria e prática*. [S.l.]: Elsevier, 2008. Citado na página 26.
- BRAGA, A. *Curvas ROC: aspectos funcionais e aplicações*. Tese (Doutorado) — Universidade do Minho, 2001. Citado na página 29.
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, Oct 2001. ISSN 1573-0565. Citado na página 27.
- 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%2F05%2F2019)), 2016. Citado na página 20.
- CARUANA, R.; FREITAG, D. Greedy attribute selection. In: *Machine Learning Proceedings 1994*. [S.l.]: Elsevier, 1994. p. 28–36. Citado na página 30.
- CHEN, T. et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, p. 1–4, 2015. Citado na página 27.
- 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 24.
- CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 1251–1258. Citado na página 25.
- DANIEL, G. *Principles of artificial neural networks*. [S.l.]: World Scientific, 2013. Citado na página 19.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: IEEE. *2009 IEEE conference on computer vision and pattern recognition*. [S.l.], 2009. p. 248–255. Citado na página 30.
- 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 23.
- ELSHEIKH, S. H.; MUSTAFA, Z. A.; IBRAHEEM, B. A. Computer aided diagnosis system of skin cancer using artificial neural network. 2018. Citado na página 13.

- FERREIRA, C. B. R.; BORGES, D. L. Analysis of mammogram classification using a wavelet transform decomposition. *Pattern Recognition Letters*, Elsevier, v. 24, n. 7, p. 973–982, 2003. Citado na página 13.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Additive logistic regression: A statistical view of boosting. *The Annals of Statistics*, v. 28, p. 337–407, 04 2000. Citado na página 27.
- FRIEDMAN, J. H. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 29, n. 5, p. 1189–1232, 2001. ISSN 00905364. Citado na página 27.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016. Citado na página 23.
- GUO, Y. et al. Breast cancer histology image classification based on deep neural networks. In: SPRINGER. *International Conference Image Analysis and Recognition*. [S.l.], 2018. p. 827–836. Citado 2 vezes nas páginas 15 e 33.
- 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 24.
- HEATH, M. et al. The digital database for screening mammography. In: MEDICAL PHYSICS PUBLISHING. *Proceedings of the 5th international workshop on digital mammography*. [S.l.], 2000. p. 212–218. Citado na página 18.
- 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), 2019. Citado na página 20.
- INCA. *Instituto Nacional do Câncer Brasil. Estimativa 2018 Incidência do Câncer no Brasil*. 2018. Citado na página 13.
- 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 21.
- 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 25.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977. Citado na página 28.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015. Citado 2 vezes nas páginas 20 e 21.
- PONTES, D. L. de B. *Câncer benigno e maligno*. 2013. Disponível em: <<https://www.einstein.br/noticias/noticia/cancer-benigno-maligno>>. Citado na página 13.
- PONTI MOACIR ANTONELLI, D. C. G. B. P. Como funciona o deep learning. 2018. Citado 2 vezes nas páginas 21 e 23.

- PROGRAMMERSOUGHT. From inception, rexnnet to xception to mobilenets, shufflenet, mobilenetv2, shufflenetv2. URL: <http://www.programmersought.com/article/4817156943/>(visited on 04/05/2019), 2017. Citado na página 26.
- RAKHLIN, A. et al. Deep convolutional neural networks for breast cancer histology image analysis. In: SPRINGER. *International Conference Image Analysis and Recognition*. [S.l.], 2018. p. 737–744. Citado 2 vezes nas páginas 15 e 33.
- ROSEBROCK, A. *Deep Learning for Computer Vision with Python: ImageNet Bundle*. [S.l.]: PyImageSearch, 2017. Citado na página 22.
- 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), 2018. Citado na página 22.
- SILVA, T. *Diferenciação do padrão de malignidade e benignidade de massas em mamografias utilizando características geométricas e máquina de vetor de suporte*. Tese (Doutorado) — Dissertação de Mestrado. Programa de Pós-Graduação em Ciência da Computação . . . , 2016. Citado 2 vezes nas páginas 15 e 33.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado 2 vezes nas páginas 24 e 25.
- SONG, R. et al. extreme gradient boosting for identifying individual users across different digital devices. In: CUI, B. et al. (Ed.). *Web-Age Information Management*. Cham: Springer International Publishing, 2016. p. 43–54. ISBN 978-3-319-39937-9. Citado na página 27.
- SUBODH, M. Transfer learning using tensorflow. URL: <https://medium.com/@subodh.malgonde/transfer-learning-using-tensorflow-52a4f6bcde3e/>(visited on 06/05/2019), 2018. Citado na página 23.
- 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 23.
- TRAMONTE, M. S. et al. Delay in diagnosis of breast cancer in a public oncologic hospital. *Medicina (Ribeirão Preto. Online)*, v. 49, n. 5, p. 451–462, 2016. Citado na página 13.
- VAPNIK, V. The support vector method of function estimation. In: \_\_\_\_\_. *Nonlinear Modeling: Advanced Black-Box Techniques*. Boston, MA: Springer US, 1998. p. 55–85. ISBN 978-1-4615-5703-6. Citado na página 26.
- VAPNIK, V. *The nature of statistical learning theory*. [S.l.]: Springer science & business media, 2013. Citado na página 26.
- VÁZQUEZ, F. Deep learning made easy with deep cognition. URL: <https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351>(visited on 03/05/2019), 2017. Citado 2 vezes nas páginas 19 e 20.



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

**Identificação do Tipo de Documento**

- Tese  
 Dissertação  
 Monografia  
 Artigo

Eu, **Rafael Pedro da Silva Neto**, 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 **Metodologia automática para classificação de lesões de mama utilizando Transfer Learning** 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 06 de Fevereiro de 2020.

*Rafael Pedro da Silva Neto*

Assinatura