

Aroldo de Moura Santos Júnior
Orientador: Flávio Henrique Duarte de Araújo

Uso de Transferência de Aprendizado para Detecção Automática de Melanoma

Picos - PI
Junho de 2019

Aroldo de Moura Santos Júnior
Orientador: Flávio Henrique Duarte de Araújo

Uso de Transferência de Aprendizado para Detecção Automática de Melanoma

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
Junho de 2019

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

S237u Santos Júnior, Aroldo de Moura.
Uso de transferência de aprendizado para detecção automática de melanoma. / Aroldo de Moura Santos Júnior. – Picos, PI, 2019.
36 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. PhD. Flávio Henrique Duarte de Araújo

1. Processamento de Imagem. 2. Redes Neurais Convolucionais. 3. ResNet50. I. Título.

CDD 005.1

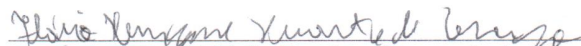
USO DE TRANSFERÊNCIA DE APRENDIZADO PARA DETECÇÃO AUTOMÁTICA DE MELANOMA


AROLDO DE MOURA SANTOS JUNIOR

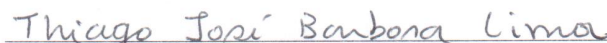
Monografia aprovada como exigência parcial para obtenção do grau de Bacharel em Sistemas de Informação.

Data de Aprovação:

Picos – PI, 12 de junho de 2019


Prof. Flávio Henrique Duarte de Araújo


Prof. Antônio Oseas de Carvalho Filho


Prof. Thiago José Barbosa Lima

Agradecimentos

Agradeço, inicialmente e acima de todos, a Jeová, meu Deus, por ter me concebido a vida, permitindo que eu chegasse até aqui e me fornecido muitas oportunidades e auxílios.

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

A minha primeira orientadora, Professora Ma. Patrícia Medyna, por aceitar me orientar, mesmo que por pouco tempo.

Ao meu pai Aroldo de Moura Santos e minha mãe Maria Vanusa de Sousa Santos 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.

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

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

Jesus lhe disse: “Esta expressão: ‘Se puder’! Ora, tudo é possível para quem tem fé.”
Marcos 9:23

Resumo

Atualmente, o câncer de pele é visto como uma das formas mais perigosas dos cânceres encontrados em seres humanos. Ele pode ser classificado em vários tipos, como melanoma, carcinoma de células basais ou escamosas e outros, onde o primeiro é o mais imprevisível. Quanto mais precoce for o diagnóstico do melanoma maior as chances de cura, por isso foram desenvolvidas técnicas de computação baseadas em aprendizagem profunda que podem desempenhar um papel importante no diagnóstico de imagens médicas. Porém, elas necessitam de uma grande quantidade de imagens o que acaba se tornando um desafio. Neste trabalho utilizamos a transferência de aprendizado, pois não foi possível treinar uma arquitetura própria devido a pequena quantidade de imagens da base de dados coletadas do arquivo ISIC e limitações de poder computacional de hardware. Foram testadas diferentes arquiteturas de Redes Neurais Convolucionais (CNN) utilizando a transferência de aprendizado e também comparamos seus desempenhos com outros classificadores utilizando descritores clássicos da literatura como o HOG, LBP e o Dayse. O modelo que obteve o melhor resultado foi a ResNet50 que atingiu uma acurácia de 88,5%, sensibilidade de 89,2%, especificidade de 87,7% e uma área sob a curva ROC de 88 ultrapassando os descritores clássicos utilizados nesse trabalho.

Palavras-chaves: Processamento de Imagens, Aprendizagem Profunda, Redes Neurais Convolucionais.

Abstract

Currently, skin cancer as one of the most dangerous forms of cancer found in humans. It can be classified into several types such as melanoma, basal and squamous cell carcinoma, where the first is the most preventable. The earlier the diagnosis of melanoma, the greater the chances of cure, so many computational techniques have been developed that may play an important role in the diagnosis of medical images. Among these techniques, those based on deep learning are the ones that produce better results, however, they require a great amount of images, which turns out to be a challenge. In this work we use transfer learning, because it was not possible to train our own architecture due to the small amount of images in the database. Different convolutional neural network (CNN) architectures were tested and we also compared their performances with classic descriptors of the literature such as HOG, LBP and Dayse. The ResNet50 achieved the best results, with 88.5% of accuracy, 89.2% sensitivity, 87.7% specificity and an area under the ROC curve of 88, performing better than classic descriptors used in this study.

Lista de ilustrações

Figura 1 – Melanoma (A) e nevos (B).	14
Figura 2 – Ilustração de uma ANN.	19
Figura 3 – Ilustração de uma rede de aprendizagem profunda.	20
Figura 4 – Ilustração de uma CNN formada por camadas convolucionais, de <i>pooling</i> e uma camada totalmente conectada.	21
Figura 5 – Exemplo da transferência de aprendizado utilizada neste trabalho.	22
Figura 6 – Ilustração de um bloco residual na arquitetura da rede ResNet.	23
Figura 7 – Ilustração da arquitetura da VGG.	24
Figura 8 – Ilustração de uma arquitetura InceptionV3.	25
Figura 9 – Ilustração de uma camada InceptionResNetV2.	26
Figura 10 –Arquiteturas InceptionV3 (esquerda) e Xception (direita).	26
Figura 11 –Esquema da metodologia utilizada neste trabalho.	31

Lista de tabelas

Tabela 1 – Tabela Comparativa de Trabalhos Relacionados.	15
Tabela 2 – Nível de exatidão de uma classificação, conforme o valor de índice Kappa.	29
Tabela 3 – Tabela com a quantidade de características extraídas pelas redes pré-treinadas e descritores.	30
Tabela 4 – Tabela de especificações da máquina utilizada neste trabalho.	31
Tabela 5 – Tabela de resultados das arquiteturas com o método de divisão da base.	32
Tabela 6 – Tabela de resultados das arquiteturas com o método de <i>cross validation</i>	32
Tabela 7 – Tabela de resultados dos descritores com o método de divisão da base.	33
Tabela 8 – Tabela de resultados dos descritores com o método de <i>cross validation</i>	33
Tabela 9 – Tabela comparativa dos resultados dos trabalhos da área com este trabalho.	33

Lista de abreviaturas e siglas

ANN	<i>Artificial Neural Network</i>
CAD	<i>Computer-aided Diagnosis</i>
CNN	Rede Neural Convolutacional (<i>Convolutional Neural Network</i>)
EDL	<i>Edinburgh Dermofit Library</i>
EDL	<i>Edinburgh Dermofit Library</i>
FCRN	<i>Fully Convolutional Residual Network</i>
HOG	<i>Histogram of Oriented Gradients</i>
INCA	Instituto Nacional de Câncer
ISIB	<i>International Symposium on Biomedical Imaging</i>
ISIC	<i>International Skin Imaging Collaboration</i>
LBP	<i>Local binary patterns</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SVM	<i>Support Vector Machine</i>
UV	Ultravioleta

Lista de símbolos

Σ	Somatório
θ	Theta

Sumário

1	Introdução	13
1.1	Objetivos	14
1.1.1	Objetivos Específicos	14
1.2	Disposição do Trabalho	14
2	Trabalhos Relacionados	15
3	Materiais e Métodos	18
3.1	Base de Dados	18
3.2	<i>Artificial Neural Network</i>	18
3.2.1	<i>Deep Learning</i>	19
3.3	<i>Convolutional Neural Network</i>	20
3.4	Transferência de Aprendizado	21
3.5	Modelo das Arquiteturas Comparadas	23
3.5.1	ResNet	23
3.5.2	VGG	23
3.5.3	InceptionV3	24
3.5.4	Inception-ResNet	25
3.5.5	Xception	26
3.6	Descritores Clássicos Comparados	27
3.6.1	Histogram of Oriented Gradients - HOG	27
3.6.2	Local Binary Patterns - LBP	27
3.6.3	Dense Daisy	27
3.7	Métricas Utilizadas	28
3.7.1	Acurácia	28
3.7.2	Sensibilidade	28
3.7.3	Especificidade	28
3.7.4	Kappa	28
3.7.5	Área sob a curva ROC - AUC	29
4	Discussões e Resultados	30
4.1	Testes Avaliativos	31
5	Conclusão	34
	Referências	35

1 Introdução

O melanoma e o câncer de pele não-melanoma (carcinoma basocelular e de células escamosas) são exemplos de alguns dos tipos de cânceres de pele diagnosticados atualmente. Embora o segundo tipo tenha uma maior incidência, o primeiro tipo possui uma taxa de fatalidade bem maior. De acordo com [Tuccitto et al. \(2019\)](#) o melanoma representa apenas 4% dos cânceres de pele, mas é responsável por 80% das mortes relacionadas a lesões de pele.

O termo melanoma se aplica a um tipo de câncer de pele grave que pode ocorrer em qualquer parte do corpo que possua melanócitos. A principal causa do melanoma é a excessiva exposição à radiação ultravioleta (UV), que pode acabar prejudicando as células da pele afetando o seu crescimento. As pessoas mais afetadas por esse e tipo de câncer são as de pele clara, devido ao baixo teor de melanina.

Estima-se que, em anos anteriores, 105.000 pessoas foram diagnosticadas com melanoma e que 33.000 morreram da doença em todo o mundo ([SHEHA; MABROUK; SHARAWY, 2012](#)). De acordo com o Instituto Nacional de Câncer (INCA) ¹, no ano de 2018, estima-se que foram diagnosticados 6.260 novos casos de câncer de pele melanoma no Brasil.

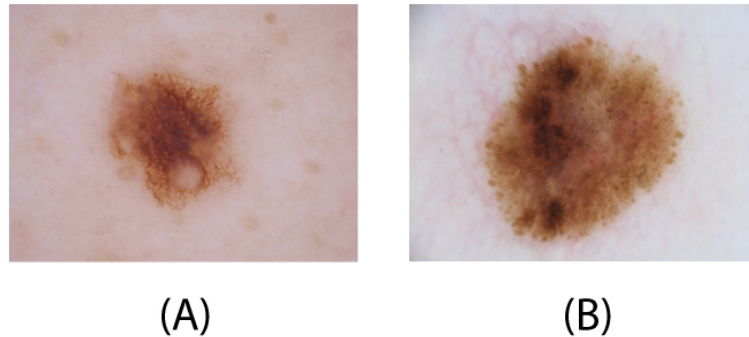
Segundo [Ercal et al. \(1994\)](#), o diagnóstico de melanoma maligno é um verdadeiro desafio, pois possui características bem semelhantes a outros tipos de lesões na pele. Em muitos casos, os dermatologistas devem realizar uma biópsia (um procedimento médico de laboratório) para analisar de forma detalhada se um tumor é maligno ou benigno. Como este é um procedimento que envolve um certo nível de despesa e morbidade, particularmente em pacientes com múltiplos sinais atípicos, técnicas alternativas de detecção precoce estão sendo procuradas para uma rápida e conveniente detecção do câncer de pele.

Os médicos geralmente confiam na sua experiência profissional e avaliam as lesões de cada paciente, caso a caso, levando em conta os padrões de lesão local do paciente em comparação com as demais existentes no corpo. No entanto, as diferenças visuais entre melanoma e lesões cutâneas benignas podem ser muito sutis, como podemos observar na Figura 1, tornando difícil distinguir os dois casos, mesmo para médicos especialistas ([LOPEZ et al., 2017](#)). A Figura 1 mostra dois exemplos de imagens dermatoscópicas da base de dados da Colaboração Internacional de Imagem de Pele (*International Skin Imaging Collaboration – ISIC*)²: uma de melanoma (A) e outra de nevos (B), onde pode-se perceber que ambos possuem características semelhantes, como a coloração e a textura.

¹ Disponível em <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-pele-melanoma>> acesso em 29/04/19

² O ISIC é uma parceria acadêmica e industrial projetada para facilitar a aplicação de imagens digitais da pele para ajudar a reduzir a mortalidade por melanoma. Disponível em <<https://www.isic-archive.com/>> acesso em 30/04/19

Figura 1 – Melanoma (A) e nevos (B).



Fonte: ISIC.

Os sistemas de detecção auxiliada por computador (*Computer-Aided Design - CAD*) foram desenvolvidos para reduzir os gastos e melhorar a capacidade de inúmeros profissionais que lidam com imagens. Estudos recentes, como por exemplo [Elsheikh, Mustafa e Ibraheem \(2018\)](#), comprovaram que os sistemas CAD são capazes de revelar lesões de pele, como o melanoma. Embora não sejam utilizados para emitir o diagnóstico final, eles servem como uma segunda opinião para os especialistas.

1.1 Objetivos

Com base no exposto, o objetivo geral desse trabalho é desenvolver um método computacional para a detecção automática de melanoma em imagens médicas de pele.

1.1.1 Objetivos Específicos

1. Analisar a performance de diferentes arquiteturas de redes e também descritores clássicos aliado a transferência de aprendizado para diagnosticar casos de melanoma.
2. Comparar o desempenho do método proposto com outros trabalhos de detecção de melanoma presentes na literatura.

1.2 Disposição 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. Já no Capítulo 3 são apresentados conceitos importantes ao desenvolvimento e compreensão do trabalho bem como a base de dados e técnicas utilizadas. Logo em seguida, o Capítulo 4 aborda as principais características deste trabalho, tais como ambiente e método de desenvolvimento, resultados alcançados e outras informações úteis a respeito da metodologia aqui implementada. Por fim, no Capítulo 5 são apresentadas as conclusões deste trabalho, bem como possíveis trabalhos futuros.

2 Trabalhos Relacionados

No que concerne aos trabalhos desta linha de pesquisa, pode-se observar uma tendência de crescimento no uso de técnicas de aprendizado profundo. A Tabela 1 mostra as principais informações dos trabalhos relacionados e em seguida há uma breve descrição de cada um deles.

Tabela 1 – Tabela Comparativa de Trabalhos Relacionados.

Referências	Base de Imagens	Arquiteturas ou Descritores	Resultados
Mahbod et al. (2019)	ISIC 2016 ISIC 2017 Total: 2.037 imagens	AlexNet VGG16 ResNet	AUC de 83,83%
Menegola et al. (2017)	Atlas ISIC 2016 Total: 900 imagens	VGG16	Acurácia de 83,5%
Chang (2017)	ISIC Total: 2000 imagens	InceptionV3	Acurácia de 80%
Yu et al. (2017)	ISIC 2016 Total: 900 imagens	VGG16 GoogleNet FCRN	Acurácia de 94%
Lopez et al. (2017)	ISIC <i>Dermofit Image Library</i> Dermnet Total: 26.000 imagens	VGG	Sensibilidade de 79%
Esteva et al. (2017)	ISIC EDL <i>Stanford Hospital</i> Total: 900 imagens	InceptionV3	Acurácia de 72%.
Majtner, Yildirim-Yayilgan e Hardeberg (2016)	ISIC 2016 Total: 900 imagens	RSurf LBP	Acurácia de 82%.

Fonte: Elaborada pelo autor.

[Mahbod et al. \(2019\)](#) abordaram um método computacional totalmente automático para classificação de lesões de pele, que emprega características profundas otimizadas. Eles utilizaram três modelos profundos pré-treinados, o AlexNet, o VGG16 e o ResNet, como geradores de características profundas que são usadas para treinar o classificador *Support Vector Machine (SVM)*. Eles usaram as imagens do desafio ISIC 2016 e ISIC 2017, totalizando 2037 imagens dermatoscópicas, incluindo 411 melanomas malignos, 254 ceratose seborreica e 1372 nevos benignos. As imagens são de vários tamanhos, ângulos fotográficos e condições de iluminação. O resultado obtido foi de uma área sob a curva ROC de 83,83% para classificação de melanoma.

Menegola et al. (2017) investigam o uso de transferência de aprendizado em redes de aprendizagem profunda. Os conjuntos de dados empregados nesse trabalho foram o *Atlas of Dermoscopy (Atlas)* e o *ISBI Challenge 2016 / ISIC* que contém 900 imagens para treinamento (sendo 273 melanomas) e 379 para testes (sendo 115 melanomas). Os autores aplicaram a transferência de aprendizado na arquitetura VGG16 com *fine-tuning* pré-treinada com pesos da ImageNet e também o SVM para classificar melanoma e lesões benignas, alcançando uma acurácia de 83.5%.

Chang (2017) relata uma técnica de aprendizagem profunda utilizando a transferência de aprendizado. Como base de dados foram utilizadas 2000 imagens dermatoscópicas coletadas no arquivo da ISIC sendo 374 imagens de melanoma, 1372 imagens de nevos e 254 imagens de verruga seborreica. Ele aplicou a transferência de aprendizado em uma rede neural Inceptionv3 atingindo uma acurácia de 80%.

Yu et al. (2017) propuseram um novo método para o reconhecimento do melanoma, com redes neurais convolucionais profundas. Os autores utilizaram o conjunto de dados do desafio ISIC 2016 que possui imagens da base de dados da ISIC e é dividido em 900 imagens para treino e 350 imagens para teste. Eles aplicaram diferentes arquiteturas com diferentes profundidades, dentre elas a VGG 16, a GoogleNet com 22 camadas de profundidade e *Fully Convolutional Residual Network - FCRN*. A que atingiu os melhores resultados foi a FCRN de 50 camadas de profundidade obtendo uma acurácia de 94%, uma sensibilidade de 91% e uma especificidade de 95%.

Lopez et al. (2017) propuseram uma abordagem baseada em transferência de aprendizado para treinamento com a arquitetura VGG usando pesos pré-treinados da ImageNet. As imagens utilizadas nesse artigo são do conjunto de imagens do desafio ISIC 2016, da *Dermofit Image Library* e Dermnet, onde os autores obtiveram um valor de sensibilidade de 78% e uma precisão de 79% para classificação de lesões benignas ou malignas.

Esteva et al. (2017) propuseram um método para classificação de lesões cutâneas usando uma única CNN (arquitetura InceptionV3) treinada no banco de dados ImageNet, removendo a camada de classificação da rede e atualizando com sua própria base de dados. Nesse trabalho foi utilizado um conjunto de dados que vem de uma combinação de repositórios de dermatologia de acesso aberto, do ISIC, da *Edinburgh Dermofit Library - EDL* e dos dados do *Stanford Hospital*. O método proposto alcançou acurácia de 72%.

Majtner, Yildirim-Yayilgan e Hardeberg (2016) abordaram um sistema automatizado de reconhecimento de melanoma. Os autores coletaram as imagens na base de dados da ISIC separando 900 imagens (727 imagens de lesões benignas e 173 imagens de melanoma) para treinamento e 379 imagens (304 imagens de lesões benignas e 75 imagens de melanoma) para teste. Todo o conjunto de dados consiste em imagens de várias fontes, simulando o uso real das soluções propostas. Eles adotam um método de aprendizagem profunda combinado com os descritores RSurf e *Local Binary Patterns - LBP* onde conseguiram atingir uma acurácia de 82%, uma sensibilidade de 53%, uma especificidade de 89% e uma área sobre curva de 78%.

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 resolvemos investigar 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, resolvemos usar a transferência de aprendizado que será abordada mais a frente. Também decidimos testar outros descritores clássicos para confirmar se as CNNs realmente apresentavam um melhor desempenho com uma quantidade de imagens reduzidas. Com as pesquisas realizadas, também vimos que seria necessário utilizar classificadores externos para a validação, por isso utilizamos dois: o SVM e o RandomForest. Em seguida, veio a etapa de desenvolvimento e testes computacionais.

3 Materiais e Métodos

Neste capítulo são apresentados a base de imagens e as principais arquiteturas de CNNs e descritores clássicos comparados. Por fim são apresentadas as métricas de avaliação dos resultados obtidos.

3.1 Base de Dados

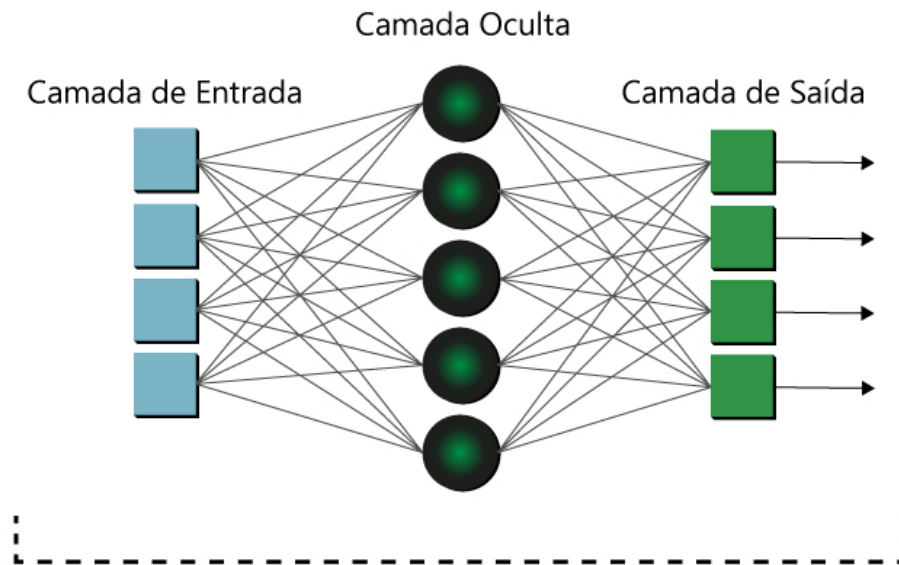
A base de imagens utilizada nesse trabalho foi coletada da International Skin Imaging Collaboration - ISIC) que contém atualmente uma grande coleção de imagens dermatoscópicas de livre acesso ao público, que foram coletadas dos principais centros clínicos do mundo todo e adquiridas a partir de uma variedade de dispositivos dentro de cada centro. Todas as imagens recebidas no arquivo da ISIC são rastreadas quanto à privacidade e garantia de qualidade. A maioria das imagens tem metadados clínicos associados, avaliados por especialistas reconhecidos em melanoma.

Foram escolhidas duas categorias (classes) para compor a base de imagens que, de acordo com [Codella et al. \(2015\)](#), são as mais difíceis de serem distinguidas pelos especialistas. Ela é composta por 800 imagens sendo, 400 de melanoma e 400 de nevos atípicos, que não são melanoma mas, possuem características bem semelhantes como a cor, o formato e tamanho. Elas possuem dimensões variadas que vão desde 722x542, até 6708x4419 pixels e foram escolhidas individualmente, descartando aquelas que possuíam a presença de artefatos, naturais (cabelos, veias) ou artificiais (bolhas de ar, marcas de régua, gráficos de calibração de cores) pois, segundo [Yu et al. \(2017\)](#), tais artefatos podem embaçar ou obstruir as lesões da pele dificultando ainda mais o processamento dessas imagens. A seguir serão apresentados os conceitos das técnicas utilizadas nesse trabalho, começando com Redes Neurais Artificiais que é a base dos estudos desta área.

3.2 *Artificial Neural Network*

As Redes Neurais artificiais (*Artificial Neural Network* - ANN) 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 2](#) traz a ilustração de uma ANN.

Figura 2 – Ilustração de uma ANN.



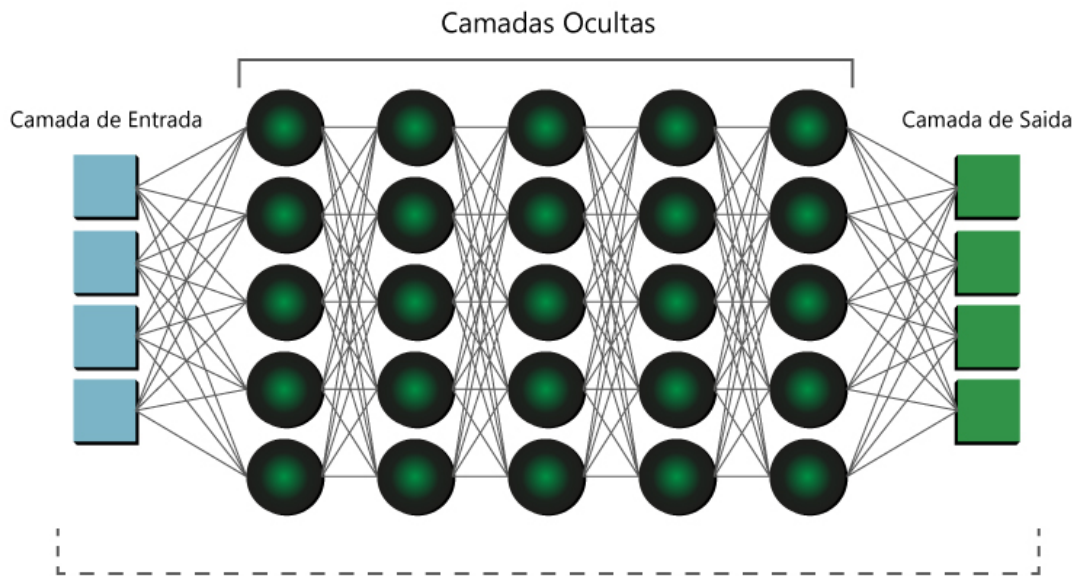
Fonte: Adaptação de [Vázquez \(2017\)](#).

3.2.1 *Deep Learning*

A aprendizagem profunda (*Deep Learning* em inglês) é um subcampo do aprendizado de máquina que tem como foco principal desenvolver meios para resolver problemas de forma semelhante ao processo de tomada de decisão humano. O termo profunda não se refere a qualquer tipo de compreensão alcançada pelo método, ele se refere à ideia de múltiplas camadas de representação sucessivas ([CHOLLET, 2018](#)).

A profundidade do modelo é determinada pela quantidade de camadas ocultas. Os modelos de aprendizagem profunda 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 3 nos mostra uma ilustração de uma rede de aprendizagem profunda.

Figura 3 – Ilustração de uma rede de aprendizagem profunda.



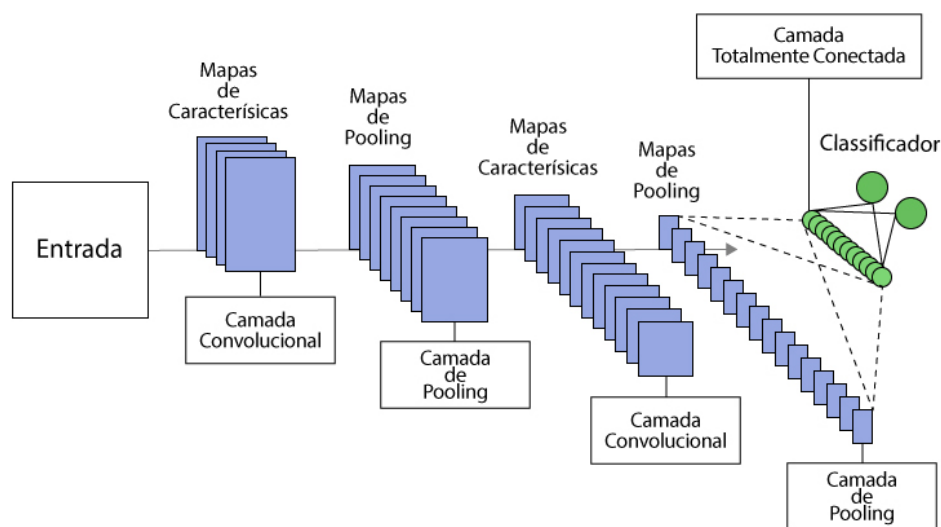
Fonte: Adaptação de [Vázquez \(2017\)](#)

Ao passar dos anos surgiu uma categoria de redes neurais artificiais baseada em aprendizagem profunda 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 rede neural convolucional que será descrita a seguir.

3.3 Convolutional Neural Network

As redes neurais convolucionais (CNNs) se tornaram o novo padrão em visão computacional e são fáceis de treinar quando existe grande quantidade de amostras rotuladas que representam as diferentes classes-alvo. Algumas das vantagens consistem em: (a) capacidade de extrair características relevantes através de aprendizado de transformações (kernels) e (b) depender de menor número de parâmetros de ajustes do que redes totalmente conectadas com o mesmo número de camadas ocultas. Como cada unidade de uma camada não é conectada com todas as unidades da camada seguinte, há menos pesos para serem atualizados, facilitando assim o treinamento ([ARAÚJO et al., 2017](#)). A Figura 4 nos mostra uma ilustração de uma CNN com duas camadas convolucionais, duas de *pooling* e uma totalmente conectada.

Figura 4 – Ilustração de uma CNN formada por camadas convolucionais, de *pooling* e uma camada totalmente conectada.



Fonte: Adaptação de Saha (2018)

Como observado na Figura 4 a rede neural convolucional possui três principais camadas: convolucionais, de *pooling* e totalmente conectadas. De acordo com Araújo et al. (2017) as camadas convolucionais são responsáveis por extrair atributos dos volumes de entradas. As camadas de *pooling* são responsáveis por reduzir a dimensionalidade do volume resultante após as camadas convolucionais e ajudam a tornar a representação invariante a pequenas translações na entrada. As camadas totalmente conectadas são responsáveis pela propagação do sinal por meio da multiplicação ponto a ponto e o uso de uma função de ativação.

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 desenvolvida uma técnica chamada de transferência de aprendizado, que será descrita a seguir.

3.4 Transferência de Aprendizado

O termo transferência de aprendizado se refere ao processo de transferir o conhecimento da tarefa de origem para uma tarefa de destino permitindo um progresso rápido ou duradouro ao modelar a segunda tarefa. Na área de aprendizagem de máquina essa técnica é utilizada quando não é possível obter uma base de dados grande o suficiente para resolver uma tarefa.

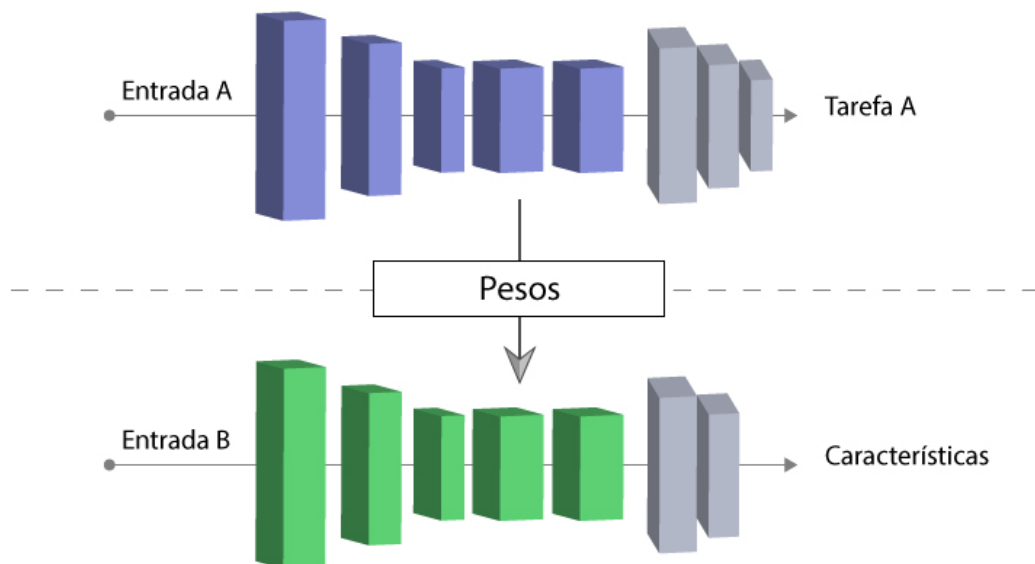
Transferência de aprendizado e adaptação de domínio referem-se à situação em que o que foi aprendido em um ambiente... é explorado para melhorar a generalização em outro ambiente (GOODFELLOW; BENGIO; COURVILLE, 2016).

Existem três medidas comuns pelas quais a transferência pode melhorar o aprendizado. O primeiro é o desempenho inicial alcançável na tarefa alvo usando apenas o conhecimento transferido, antes que qualquer outra aprendizagem seja feita, comparada ao desempenho inicial de um agente ignorante. A segunda é a quantidade de tempo que leva para aprender completamente a tarefa-alvo, dado o conhecimento transferido em comparação com a quantidade de tempo para aprendê-la do zero. O terceiro é o nível de desempenho final alcançável na tarefa alvo, comparado ao nível final sem transferência (TORREY; SHAVLIK, 2010).

Segundo Karpathy et al. (2016), basicamente há três formas de se utilizar a transferência de aprendizado com CNNs. A primeira consiste em usar a CNN como extrator de características. Este é o caso das CNNs pré-treinadas da ImageNet, que remove a última camada totalmente conectada, transformando a CNN em extrator de características. A segunda forma é o *fine-tuning*, onde os pesos da rede pré-treinada são ajustados continuando a retropropagação. E a terceira forma são os modelos pré-treinados. Como as CNNs modernas levam de 2 a 3 semanas para treinar em várias GPUs no ImageNet, é comum ver pessoas liberando seus últimos pontos de verificação da CNN para o benefício de outras pessoas que podem usar as redes para o *fine-tuning*.

A Figura 5 nos mostra um esquema de transferência de aprendizado aplicada neste trabalho onde utilizamos os pesos de CNNs treinadas para a ImageNet e removemos a camada de classificação, assim o resultado da rede são as características que descrevem a imagem de entrada. Depois disso utilizamos um outro classificador para a classificação.

Figura 5 – Exemplo da transferência de aprendizado utilizada neste trabalho.



Fonte: Elaborada pelo autor.

A seguir serão abordados as arquiteturas e os descritores onde utilizamos essa técnica e que posteriormente foram avaliados e comparados neste trabalho.

3.5 Modelo das Arquiteturas Comparadas

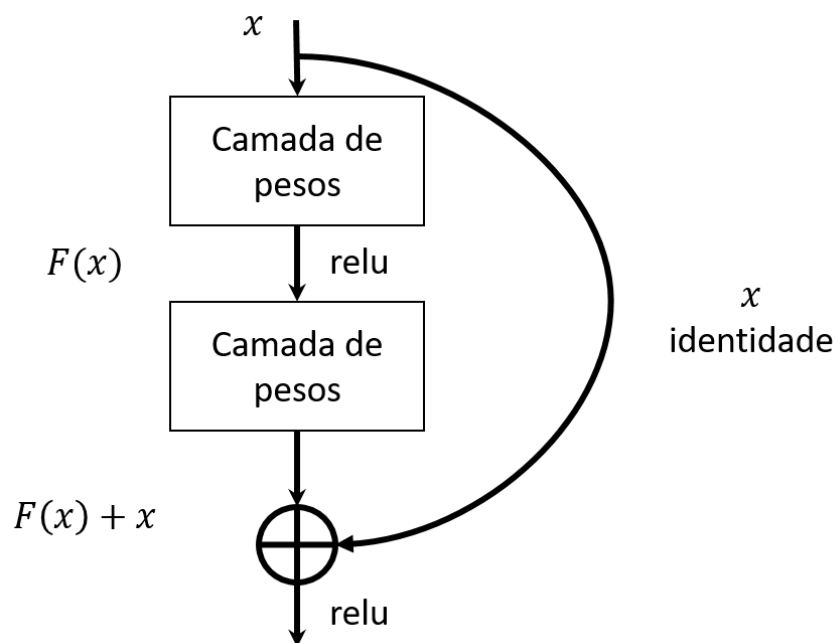
Após analisar a bibliografia que resultou da pesquisa a respeito do tema, foram encontrados diversos modelos de arquitetura compatíveis com a proposta desse trabalho. Foram realizadas comparações e avaliações para determinar qual a melhor performance dentro do objetivo do trabalho. Com isso, foi possível selecionar nove principais modelos para estudo mais profundo e consequente avaliação da melhor performance. Tais modelos são descritos à seguir.

3.5.1 ResNet

A ResNet é uma arquitetura proposta por [He et al. \(2016\)](#) que tem como principal característica suas conexões residuais. Na aprendizagem residual, o foco não é aprender características, mas sim aprender alguns resíduos. Resíduo pode ser entendido como a subtração da característica aprendida na camada de entrada.

Um dos problemas que as ResNets resolvem 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 várias aplicações da regra de cadeia. Com tal arquitetura, os gradientes podem fluir diretamente através das conexões de salto. A Resnet fornece uma estrutura de treinamento para facilitar formação de redes substancialmente mais profundas. A Figura 6 mostra uma ilustração de um bloco residual, característico da ResNet.

Figura 6 – Ilustração de um bloco residual na arquitetura da rede ResNet.



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

3.5.2 VGG

A rede VGG proposta por [Simonyan e Zisserman \(2014\)](#) foi a primeira a utilizar filtros pequenos (3x3) em cada camada convolucional. Ele faz a melhoria sobre o AlexNet,

construída por [Krizhevsky, Sutskever e Hinton \(2012\)](#), onde filtros grandes (9x9 e 11x11) eram usados para capturar características similares na imagem. A grande contribuição da VGG foi a ideia de que múltiplas convoluções 3x3 em sequência podiam substituir efeitos de filtros de máscaras maiores (5x5 e 7x7), e que resultavam em maior custo computacional ([ARAÚJO et al., 2017](#)). A proposta original da rede VGG era muito mais profunda que a da AlexNet. Nela as profundidades mais comuns são as de 16 e 19 camadas. A Figura 7 nos mostra uma ilustração da arquitetura utilizada na VGG.

Figura 7 – Ilustração da arquitetura da 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 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	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 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Fonte: [Simonyan e Zisserman \(2014\)](#).

3.5.3 InceptionV3

Baseado no artigo de [Szegedy et al. \(2016\)](#), o InceptionV3 é uma versão melhorada das anteriores (GoogleNet e InceptionV2), que teve como uma de suas principais mudanças a alteração da camada de convolução tradicional de 7x7 para 3 camadas de 3x3. Na sua parte *inception* ela possui 3 módulos de criação 35 x 35 com 288 filtros cada (A). Isso é reduzido para 17 x 17 com 768 filtros (B) e logo em seguida há 5 instâncias de módulos fatoriais de criação onde ocorre uma nova redução para 8 x 8 x 1280 (C). Ao todo ela possui 42 camadas. A Figura 8 nos mostra uma ilustração da arquitetura InceptionV3.

Figura 8 – Ilustração de uma arquitetura InceptionV3.

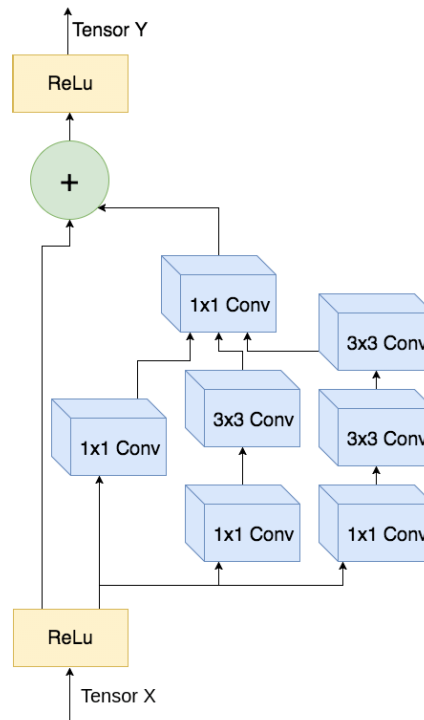
Tipo	Tamanho do Fragmento	Tamanho da Entrada
conv	$3 \times 3 / 2$	$299 \times 299 \times 3$
conv	$3 \times 3 / 1$	$149 \times 149 \times 32$
conv padded	$3 \times 3 / 1$	$147 \times 147 \times 32$
pool	$3 \times 3 / 2$	$147 \times 147 \times 64$
conv	$3 \times 3 / 1$	$73 \times 73 \times 64$
conv	$3 \times 3 / 2$	$71 \times 71 \times 80$
conv	$3 \times 3 / 1$	$35 \times 35 \times 192$
3×Inception	(A)	$35 \times 35 \times 288$
5×Inception	(B)	$17 \times 17 \times 768$
2×Inception	(C)	$8 \times 8 \times 1280$
pool	8×8	$8 \times 8 \times 2048$
linear	logits	$1 \times 1 \times 2048$
softmax	classifier	$1 \times 1 \times 1000$

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

3.5.4 Inception-ResNet

É uma combinação proposta por [Szegedy et al. \(2017\)](#) que une as conexões residuais (ResNet) e a arquitetura Inception. De acordo com [He et al. \(2016\)](#), as conexões residuais são essencialmente importantes para o treinamento de arquiteturas muito profundas. Como as redes do Inception tendem a ser muito profundas, foi proposto a substituição do estágio da concatenação do filtro da arquitetura Inception por conexões residuais. Isso permitiria o Inception colher os benefícios da abordagem residual, mantendo sua eficiência computacional. A Figura 9 mostra um exemplo de uma camada InceptionResnetV2.

Figura 9 – Ilustração de uma camada InceptionResNetV2.

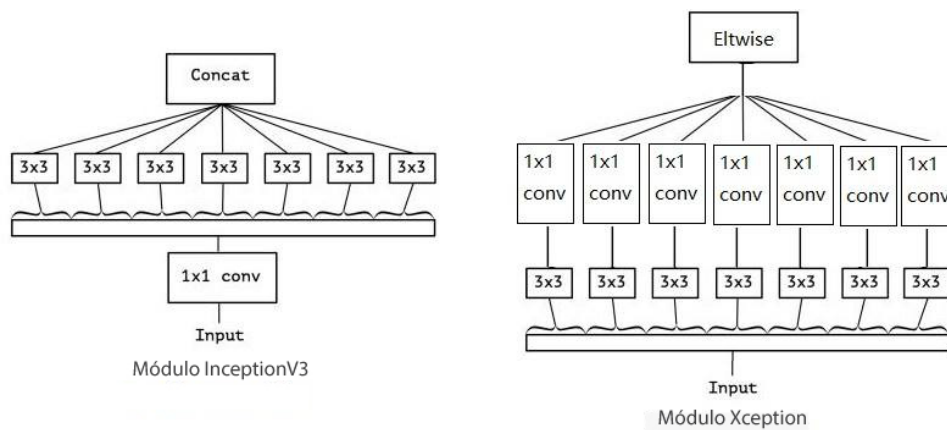


Fonte: [Nguyen-Meidine et al. \(2017\)](#).

3.5.5 Xception

Também conhecida como *Extreme Inception*, é uma variação mais forte da *Inception* que possui como uma de suas características principais a substituição dos módulos de iniciação padrões por convoluções separadas em profundidade. A arquitetura Xception possui 36 camadas convolucionais formando a base de extração de características da rede ([CHOLLET, 2017](#)). A Figura 10 nos mostra uma comparação de um módulo da InceptionV3 e um da Xception.

Figura 10 – Arquiteturas InceptionV3 (esquerda) e Xception (direita).



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

3.6 Descritores Clássicos Comparados

Com a pesquisa realizada também foi possível selecionar 3 descritores clássicos para comparar sua performance com as arquiteturas citadas acima. Esses descritores trabalham com imagens não segmentadas, por isso foi decidido utiliza-los, pois as imagens não foram segmentadas.

3.6.1 Histogram of Oriented Gradients - HOG

Ao utilizar esse descritor é necessário realizar a divisão de uma determinada imagem em pequenas regiões conectadas, chamadas de células, e, em seguida, para cada uma dessas partes, o cálculo de um histograma de gradientes orientados para os *pixels*. É possível então acumular esses histogramas através de múltiplas células para obter como resultado o vetor de características.

No descritor de característica HOG, a distribuição (histogramas) de direções de gradientes (gradientes orientados) é usada como características. Os gradientes (derivadas x e y) de uma imagem são úteis devido ao fato da magnitude dos gradientes ser grande em torno das bordas e cantos (regiões de mudanças bruscas de intensidade) e sabemos que bordas e cantos contêm muito mais informações sobre a forma do objeto do que as regiões planas.

O descritor HOG consiste então em um vetor concatenado dos componentes dos histogramas de células normalizadas de todas as regiões do bloco. Esses blocos normalmente se sobrepõem, o que significa que cada célula contribui mais de uma vez para o descritor final. Existem duas geometrias de blocos principais: blocos R-HOG retangulares e blocos C-HOG circulares.

3.6.2 Local Binary Patterns - LBP

O LBP é um tipo de descritor visual usado para classificação em visão computacional. O LBP calcula uma representação local da textura e essa representação local é construída comparando cada pixel com sua vizinhança de pixels. [Topi et al. \(2000\)](#) descreve que o LBP contém a densidade de cada recurso em uma região e que, para cada pixel, apenas a saída de melhor correspondência de detector de recurso é utilizada.

O uso de uma vizinhança circular e a interpolação bilinear de valores em coordenadas de pixel não inteiro permitem qualquer raio e número de pixels na vizinhança. Outra extensão importante é a definição dos chamados padrões uniformes, que pode ser usado para reduzir o comprimento do vetor de recursos e implementar um descritor simples de rotação invariante. Um padrão binário local é chamado uniforme se o padrão binário contiver no máximo duas transições bit a bit de 0 a 1 ou vice-versa

3.6.3 Dense Daisy

O Daisy é um descritor de imagem baseado em histogramas de orientação de gradiente semelhantes ao descritor SIFT (*Scale Invariant Feature Transform*). Ele é formulado de uma maneira que permite a extração rápida e densa.

[Zhu, Bichot e Chen \(2011\)](#) afirma que o Daisy usa uma configuração circular de vizinhança em vez de uma retangular, além do mais a parte do descritor é feito pela

concatenação de todos os vetores dos círculos, depois que eles são normalizados para a norma da unidade. O DAISY usa o histograma de gradiente filtrado pelo filtro de Gauss como uma descrição de recurso, para que o desfoque da imagem tenha um efeito mais proeminente (LI; ZIC, 2014). Além disso, eles são invariantes a rotação sem depender de uma orientação de referência, melhorando ainda mais sua robustez.

3.7 Métricas Utilizadas

Para avaliar as arquiteturas e descritores apresentados, foram utilizadas cinco métricas consideravelmente difundidas, que serão descritas a seguir. As quatro primeiras fazem uso da matriz de confusão, que indica um ranking correto ou incorreto das classes em uso, agrupando os resultados em quatro categorias, que são: Verdadeiro Positivo (VP), Verdadeiro Negativo, (VN) Falso Negativo (FN) e Falso Positivo (FP).

3.7.1 Acurácia

Consiste no número de acertos (positivos) dividido pelo número total de exemplos. É recomendado utilizar esse tipo de métrica em bases com a mesma proporção de exemplos para cada classe, caso contrário haverá uma falsa impressão de bom desempenho. A acurácia pode ser calculada com a Equação 3.1.

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

3.7.2 Sensibilidade

A sensibilidade é uma medida da proporção de casos reais que foram previstos como positivos (positivos verdadeiros). Em outras palavras, um teste altamente sensível é aquele que identifica corretamente pacientes com uma doença. A sensibilidade pode ser calculada com a Equação 3.2.

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

3.7.3 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. Ela mostra o quão bem o teste identifica pacientes que não têm uma doença. A especificidade pode ser calculada com a Equação 3.3.

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

3.7.4 Kappa

O índice Kappa é um coeficiente de concordância para escalas nominais que mede o relacionamento entre a concordância, além da casualidade, e a discordância esperada (ROSENFELD; FITZPATRICK-LINS, 1986). Este índice vem sendo recomendado como

uma medida apropriada da exatidão por representar inteiramente a matriz de confusão, e não apenas os elementos que se situam na diagonal principal, o que ocorre quando se calcula a exatidão global da classificação (ROSENFELD; FITZPATRICK-LINS, 1986). O Kappa pode ser encontrado com a Equação 3.4.

$$Kappa = \frac{\Theta_1 - \Theta_2}{1 - \Theta_2} \quad (3.4)$$

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 Kappa (K)	Qualidade
$K < 0.2$	Ruim
$0.2 \leq K < 0.4$	Razoável
$0.4 \leq K < 0.6$	Bom
$0.6 \leq K < 0.8$	Muito Bom
$K \geq 0.8$	Excelente

Fonte: Araujo (2014)

3.7.5 Área sob a curva ROC - AUC

Segundo Braga (2001), a área sob a curva ROC (AUC) é um método gráfico simples e robusto que permite estudar a variação da sensibilidade e especificidade, para diferentes valores de corte. Este método 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.

O próximo capítulo aborda a metodologia utilizada neste trabalho bem como os resultados alcançados na classificação do melanoma.

4 Discussões e Resultados

A primeira etapa (1) foi a coleta das imagens de duas classes de lesões cutâneas, o melanoma e nevos atípicos, pois o nevos é um tipo de lesão de pele que apresenta características muito semelhantes ao melanoma, como a cor, o formato e a textura. Para compor a base de dados utilizada nesse trabalho, foram selecionadas imagens do arquivo da ISIC, por ser de fácil acesso e aberta ao público. A segunda etapa (2) 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 a transferência de aprendizado, as suas camadas de classificação foram removidas. Com isso as arquiteturas e descritores produziram a quantidade de características mostradas na Tabela 3.

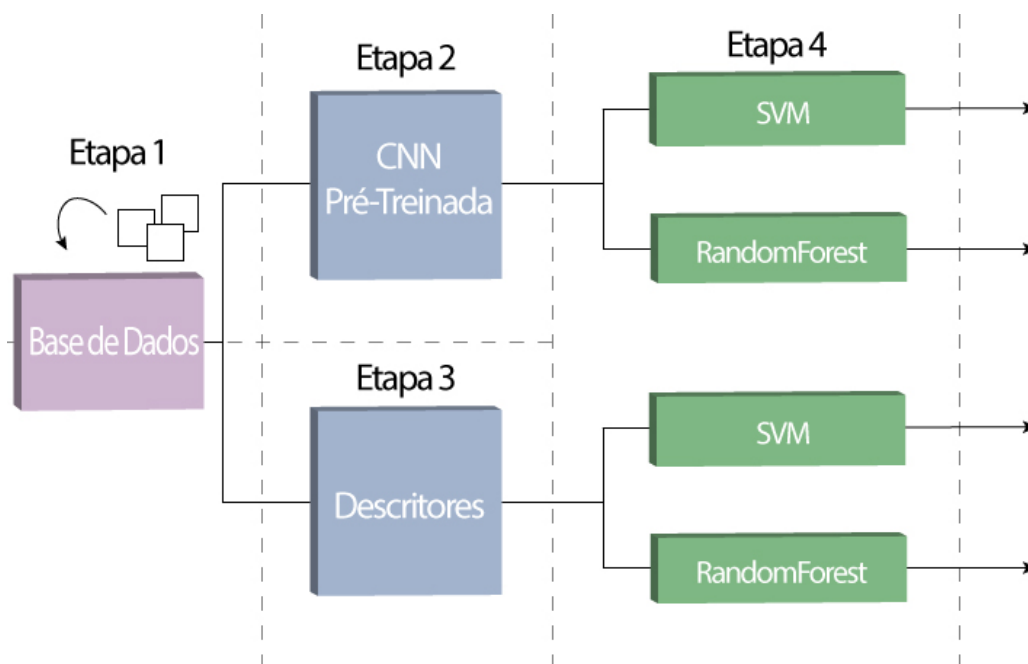
Tabela 3 – Tabela com a quantidade de características extraídas pelas redes pré-treinadas e descritores.

Algoritmos	Quantidade
Xception	204800
InceptionV3	131072
ResNet50	100352
VGG16	25088
VGG19	25088
InceptionResNetV2	1536
HOG	576
Daisy	104
LBP	18

Fonte: Elaborada pelo autor.

Na terceira etapa (3) utilizamos os descritores clássicos para a extração de características da base. A quarta e última etapa (4) foi o uso dos classificadores, *Support Vector Machine (SVM)* e o *RandomForest*, com as características extraídas das arquiteturas e dos descritores para a validação. Nesta última etapa, utilizamos o método de divisão de base de dados onde foi dividido 75% para treino e 25% para teste e depois usamos o método *cross validation*. Nos dois, o número de árvores estimadas do RandomForest foi alterado (entre 10 e 300 árvores) e foram testados ao todo quatro *kernels* diferentes do SVM: Linear, Polynomial, RBF e Sigmoid, numa tentativa de alcançar maiores resultados. A Figura 11 mostra um esquema da metodologia aqui descrita.

Figura 11 – Esquema da metodologia utilizada neste trabalho.



Fonte: Elaborada pelo autor.

4.1 Testes Avaliativos

Para a realização dos testes, utilizou-se um computador com acesso à internet, cujas especificações são descritas na Tabela 4.

Tabela 4 – Tabela de especificações da máquina utilizada neste trabalho.

Característica	Descrição
Modelo	X450LC
RAM	6GB
Memoria de Vídeo	2GB
Placa de Vídeo	Nvidia GeForce GT 720M
Processador	Intel I5 4200 @ 1.60GHz
Sistema Operacional	Windows 10

Fonte: Elaborada pelo autor.

A seguir, os resultados obtidos com a metodologia proposta serão apresentados e discutidos. As métricas nas tabelas estão simbolizadas por suas letras iniciais: Acurácia (A), Sensibilidade (S), Especificidade (E), Kappa (K), Área Sob a Curva ROC (AUC), Precisão (P), *Recall* (R). As Tabelas 5 e 6 mostram os resultados obtidos com as arquiteturas mencionadas no Capítulo 3 utilizando o método de divisão da base de dados (75% para treino e 25% para teste) e *cross validation*, respectivamente. Já as Tabelas 7 e 8 mostram os resultados obtidos com os descritores utilizando novamente o método de divisão da base de dados (75% para treino e 25% para teste) e *cross validation*, respectivamente.

Tabela 5 – Tabela de resultados das arquiteturas com o método de divisão da base.

SVM					
Arquiteturas	A (%)	S (%)	E (%)	K	AUC
ResNet50	0.88	0.89	0.87	0.76	0.88
InceptionResNetV2	0.81	0.80	0.82	0.63	0.82
VGG19	0.81	0.82	0.80	0.63	0.82
InceptionV3	0.80	0.78	0.82	0.60	0.80
VGG16	0.76	0.76	0.77	0.53	0.77
Xception	0.76	0.76	0.76	0.52	0.76
RandomForest					
Arquiteturas	A (%)	S (%)	E (%)	K	AUC
VGG19	0.82	0.83	0.81	0.64	0.82
ResNet50	0.81	0.82	0.80	0.62	0.81
VGG16	0.81	0.81	0.80	0.62	0.81
InceptionV3	0.80	0.83	0.77	0.60	0.80
Xception	0.77	0.79	0.75	0.54	0.77
InceptionResNetV2	0.75	0.79	0.70	0.49	0.75

Tabela 6 – Tabela de resultados das arquiteturas com o método de *cross validation*.

SVM				
Arquiteturas	A (%)	AUC	P (%)	R
ResNet50	0.80	0.87	0.82	0.77
InceptionV3	0.76	0.83	0.74	0.79
VGG16	0.75	0.82	0.78	0.74
VGG19	0.74	0.81	0.76	0.73
InceptionResNetV2	0.73	0.81	0.72	0.77
Xception	0.73	0.81	0.75	0.70
RandomForest				
Arquiteturas	A (%)	AUC	P (%)	R
VGG19	0.78	0.86	0.83	0.74
Xception	0.78	0.86	0.83	0.74
ResNet50	0.77	0.85	0.78	0.75
VGG16	0.76	0.84	0.79	0.74
InceptionV3	0.75	0.82	0.75	0.74
InceptionResNetV3	0.73	0.81	0.73	0.69

Como pode ser observado, ResNet50 obteve o melhor resultado nos dois métodos, *cross validation* e divisão da base, sendo o último o melhor, atingindo uma acurácia de 88,5%, sensibilidade de 89,2%, especificidade de 87,7% e uma área sob a curva ROC de 88 com o classificador SVM (*kernel* Linear). Já a VGG19 se saiu melhor dentre as redes e os descritores com o *RandomForest* (200 árvores estimadas), porém não ultrapassou o SVM. As Tabelas 7 e 8 nos mostram os resultados obtidos com os descritores.

Tabela 7 – Tabela de resultados dos descritores com o método de divisão da base.

SVM					
Descritores	A (%)	S (%)	E (%)	K	AUC
HOG	0.70	0.83	0.81	0.64	0.82
LBP	0.49	0.0	1.0	0.0	0.50
Daisy	0.49	0.0	1.0	0.0	0.50
RandomForest					
Descritores	A (%)	S (%)	E (%)	K	AUC
HOG	0.73	0.81	0.72	0.74	0.82
LBP	0.63	0.66	0.60	0.26	0.63
Daisy	0.61	0.64	0.62	0.54	0.50

Tabela 8 – Tabela de resultados dos descritores com o método de *cross validation*.

SVM				
Descritores	A (%)	AUC	P (%)	R
HOG	0.70	0.77	0.72	0.68
Daisy	0.52	0.69	0.54	0.87
LBP	0.49	0.36	0.40	0.80
RandomForest				
Descritores	A (%)	AUC	P (%)	R
HOG	0.73	0.81	0.72	0.74
Daisy	0.67	0.74	0.70	0.66
LBP	0.61	0.64	0.62	0.54

O descritor que obteve o melhor resultado foi o HOG atingindo uma acurácia de 73% com o classificador *RandomForest* com 200 arvores estimadas, sendo inferior ao da ResNet50. A Tabela 9 apresenta uma breve comparação entre os resultados encontrados utilizando a metodologia proposta e alguns trabalhos citados no Capítulo 2. Porém, alguns dos trabalhos utilizaram metodologias e bases diferentes da usada neste trabalho. Portanto, os dados apresentados na Tabela 9, são apenas para ilustrar/comparar os resultados obtidos com os trabalhos da área, sem desmerecer nenhum dos métodos citados.

Tabela 9 – Tabela comparativa dos resultados dos trabalhos da área com este trabalho.

Trabalhos	A (%)	S (%)	E (%)	AUC
Mahbod et al. (2019)	-	-	-	0.83
Menegola et al. (2017)	0.83	-	-	-
Chang (2017)	0.80	-	-	-
Yu et al. (2017)	0.94	0.91	0.95	-
Lopez et al. (2017)	-	0.78	-	-
Esteva et al. (2017)	0.72	-	-	-
Majtner, Yildirim-Yayilgan e Hardeberg (2016)	0.82	0.53	0.89	0.78
Método Proposto	0.88	0.89	0.87	0.88

Fonte: Elaborada pelo autor.

5 Conclusão

Neste trabalho avaliamos e comparamos o desempenho de seis arquiteturas de redes profundas e três descritores clássicos utilizando transferência de aprendizado para a detecção automática de melanoma em imagens médicas. Os resultados obtidos demonstraram que as redes de aprendizado profundo obtiveram um desempenho superior aos descritores em descrever as imagens de pele.

A arquitetura ResNet50 atingiu o melhor resultado dentre os algoritmos usados neste trabalho, com valores acima de 75% em todas as métricas de avaliação, dividindo a base e usando o classificador SVM. Os bons resultados na avaliação fornecem oportunidades para propor novas metas de pesquisa na área de lesões de pele, principalmente em casos de melanoma.

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 do melanoma. 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: aumentar a quantidade de amostras utilizando mais bases de imagens e também operações de aumento de dados; desenvolver um método automático para a detecção do melanoma em meio a outros tipos de lesões de pele; testar outras técnicas de transferência de aprendizado como por exemplo, o *fine-tuning*;

Referências

- ARAÚJO, 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.
- 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 3 vezes nas páginas 20, 21 e 24.
- BRAGA, A. *Curvas ROC: aspectos funcionais e aplicações*. Tese (Doutorado), 2001. Citado na página 29.
- CHANG, H. Skin cancer reorganization and classification with deep neural network. *arXiv preprint arXiv:1703.00534*, 2017. Citado 3 vezes nas páginas 15, 16 e 33.
- 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 26.
- CHOLLET, F. *Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*. [S.l.]: MITP-Verlags GmbH & Co. KG, 2018. Citado na página 19.
- CODELLA, N. et al. Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images. In: SPRINGER. *International Workshop on Machine Learning in Medical Imaging*. [S.l.], 2015. p. 118–126. Citado na página 18.
- DANIEL, G. *Principles of artificial neural networks*. [S.l.]: World Scientific, 2013. Citado na página 18.
- ELSHEIKH, S. H.; MUSTAFA, Z. A.; IBRAHEEM, B. A. Computer aided diagnosis system of skin cancer using artificial neural network. 2018. Citado na página 14.
- ERCAL, F. et al. Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on biomedical engineering*, IEEE, v. 41, n. 9, p. 837–845, 1994. Citado na página 13.
- ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, Nature Publishing Group, v. 542, n. 7639, p. 115, 2017. Citado 3 vezes nas páginas 15, 16 e 33.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016. Citado na página 21.
- 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 2 vezes nas páginas 23 e 25.

- KARPATHY, A. et al. Cs231n convolutional neural networks for visual recognition. *Neural networks*, v. 1, 2016. Citado na página 22.
- 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 24.
- LANDIS, J. R.; KOCH, G. G. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, JSTOR, p. 363–374, 1977. Citado na página 29.
- LI, L.; ZIC, J. Image matching algorithm based on feature-point and daisy descriptor. *journal of multimedia*, v. 9, n. 6, p. 829–834, 2014. Citado na página 28.
- LOPEZ, A. R. et al. Skin lesion classification from dermoscopic images using deep learning techniques. In: IEEE. *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*. [S.l.], 2017. p. 49–54. Citado 4 vezes nas páginas 13, 15, 16 e 33.
- MAHBOD, A. et al. Skin lesion classification using hybrid deep neural networks. In: IEEE. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2019. p. 1229–1233. Citado 2 vezes nas páginas 15 e 33.
- MAJTNER, T.; YILDIRIM-YAYILGAN, S.; HARDEBERG, J. Y. Combining deep learning and hand-crafted features for skin lesion classification. In: IEEE. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. [S.l.], 2016. p. 1–6. Citado 3 vezes nas páginas 15, 16 e 33.
- MENEGOLA, A. et al. Knowledge transfer for melanoma screening with deep learning. In: IEEE. *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. [S.l.], 2017. p. 297–300. Citado 3 vezes nas páginas 15, 16 e 33.
- NGUYEN-MEIDINE, L. T. et al. A comparison of cnn-based face and head detectors for real-time video surveillance applications. In: . [S.l.: s.n.], 2017. Citado na página 26.
- PROGRAMMERSOUGHT. From inception, rexnnet to xception to mobilenets, shufflenet, mobilenetv2, shufflenetv2. URL: [http://www.programmersought.com/article/4817156943/\(visited on 04/05/2019\)](http://www.programmersought.com/article/4817156943/(visited%20on%2004/05/2019)), 2017. Citado na página 26.
- ROSENFELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric engineering and remote sensing*, v. 52, n. 2, p. 223–227, 1986. Citado 2 vezes nas páginas 28 e 29.
- 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%20on%2030/04/2019)), 2018. Citado na página 21.
- SHEHA, M. A.; MABROUK, M. S.; SHARAWY, A. Automatic detection of melanoma skin cancer using texture analysis. *International Journal of Computer Applications*, Citeseer, v. 42, n. 20, p. 22–26, 2012. Citado na página 13.

- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. Citado 2 vezes nas páginas 23 e 24.
- SZEGEDY, C. et al. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-First AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2017. Citado na página 25.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 2818–2826. Citado 2 vezes nas páginas 24 e 25.
- TOPI, M. et al. Robust texture classification by subsets of local binary patterns. In: IEEE. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. [S.l.], 2000. v. 3, p. 935–938. Citado na página 27.
- 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 22.
- TUCCITTO, A. et al. Cancer stem cell challenges in melanoma characterization and treatment. In: *Cancer Stem Cell Resistance to Targeted Therapy*. [S.l.]: Springer, 2019. p. 115–135. Citado na página 13.
- VÁZQUEZ, F. Deep learning made easy with deep cognition. URL: [https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351\(visited on 03/05/2019\)](https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351(visited%20on%2003/05/2019)), 2017. Citado 2 vezes nas páginas 19 e 20.
- YU, L. et al. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE transactions on medical imaging*, IEEE, v. 36, n. 4, p. 994–1004, 2017. Citado 4 vezes nas páginas 15, 16, 18 e 33.
- ZHU, C.; BICHOT, C.-E.; CHEN, L. Visual object recognition using daisy descriptor. In: IEEE. *2011 IEEE International Conference on Multimedia and Expo*. [S.l.], 2011. p. 1–6. Citado na página 27.



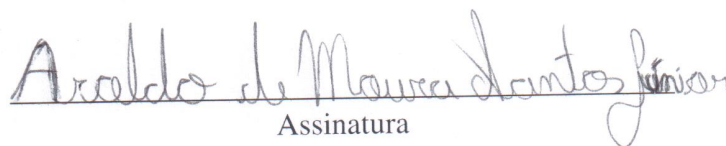
**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- Tese
- Dissertação
- Monografia
- Artigo

Eu, **Aroldo de Moura Santos Júnior**, 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 **Uso de Transferência de Aprendizado para Detecção Automática de Melanoma** 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 21 de Junho de 2019.


Assinatura