



UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

LUCAS GARCIA BATISTA

**ANÁLISE E CLASSIFICAÇÃO DE LESÕES DE PELE POR MEIO DE  
ESTRATÉGIAS DE APRENDIZADO ATIVO**

DISSERTAÇÃO DE MESTRADO

CORNÉLIO PROCÓPIO  
2022

LUCAS GARCIA BATISTA

## **ANÁLISE E CLASSIFICAÇÃO DE LESÕES DE PELE POR MEIO DE ESTRATÉGIAS DE APRENDIZADO ATIVO**

## **ANALYSIS AND CLASSIFICATION OF SKIN LESION THROUGH ACTIVE LEARNING STRATEGIES**

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito parcial para obtenção do título de Mestre em Informática.

Orientadora: Profa. Dra. Priscila Tiemi Maeda Saito

Co-orientador: Prof. Dr. Pedro Henrique Bugatti

CORNÉLIO PROCÓPIO  
2022



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/)

Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es) e que licenciem as novas criações sob termos idênticos. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.





LUCAS GARCIA BATISTA

**ANÁLISE E CLASSIFICAÇÃO DE LESÕES DE PELE POR MEIO DE ESTRATÉGIAS DE APRENDIZADO ATIVO**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Informática da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Computação Aplicada.

Data de aprovação: 24 de Fevereiro de 2022

Prof.a Priscila Tiemi Maeda Saito, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Alexandre Rossi Paschoal, Doutorado - Universidade Tecnológica Federal do Paraná

Prof Marcelo Ponciano Da Silva, Doutorado - Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 26/02/2022.

Dedico este trabalho a Deus. Sem ele nada seria possível.

## **AGRADECIMENTOS**

À professora Priscila pela sua atenção dedicada ao longo de todo o projeto no qual o fez com paciência e sabedoria.

## RESUMO

BATISTA, Lucas Garcia. ANÁLISE E CLASSIFICAÇÃO DE LESÕES DE PELE POR MEIO DE ESTRATÉGIAS DE APRENDIZADO ATIVO. 63 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

Segundo o Instituto Nacional de Câncer, dentre todos os tumores malignos, o câncer de pele não-melanoma e melanoma são os mais frequentes no Brasil. Apesar de apresentar menor incidência, em relação ao tipo não-melanoma, o tipo melanoma apresenta crescimento acelerado e maior letalidade. O diagnóstico precoce, ajuda no tratamento e na cura dos pacientes. Vários estudos têm sido realizados nos últimos anos na área de visão computacional para auxiliar no diagnóstico precoce do câncer de pele. Abordagens de aprendizado profundo, utilizando arquiteturas de redes neurais convolucionais por meio de transferência de aprendizado, têm sido largamente adotadas. Apesar de apresentarem resultados significativos, é necessária uma grande quantidade de dados de treinamento anotados. No entanto, a obtenção de dados rotulados por um especialista torna-se inviável em se tratando de grandes conjuntos de dados. Abordagens de aprendizado ativo possibilitam selecionar um pequeno conjunto de dados mais informativos para o treinamento do classificador. Portanto, o presente projeto de pesquisa explora abordagens de aprendizado ativo, incluindo a proposta e a validação de estratégias de seleção de amostras mais informativas para classificação de lesões de pele.

**Palavras-chave:** aprendizado ativo, aprendizado profundo, câncer, lesão de pele, redes neurais convolucionais, classificação de imagens.

## ABSTRACT

BATISTA, Lucas Garcia. ANALYSIS AND CLASSIFICATION OF SKIN LESION THROUGH ACTIVE LEARNING STRATEGIES. 63 f. Dissertação – Programa de Pós-Graduação em Informática, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2022.

According to the National Cancer Institute, among all malignant tumors, non-melanoma skin cancer and melanoma are the most frequent in Brazil. Despite having a lower incidence, compared to the non-melanoma type, the melanoma type has accelerated growth and greater lethality. Early diagnosis helps in the treatment and cure of patients. Several studies have been carried out in recent years in the area of computer vision to assist in the early diagnosis of skin cancer. Deep learning approaches, using convolutional neural network architectures through transfer of learning, have been widely adopted. Despite presenting significant results, a large amount of annotated training data is needed. However, obtaining labeled data by a specialist is not feasible in the case of large data sets. Active learning approaches make it possible to select a small set of more informative data for classifier training. Therefore, this research project explores active learning approaches, including the proposal and validation of selection strategies of more informative samples for the classification of skin lesions.

**Keywords:** active learning, deep learning, cancer, skin lesion, convolutional neural networks, image classification.

## LISTA DE FIGURAS

FIGURA 1	– Exemplo de uma Rede Neural Convolutacional. ....	18
FIGURA 2	– Exemplo do processo de convolução. ....	19
FIGURA 3	– Exemplo do processo de <i>pooling</i> . ....	19
FIGURA 4	– Comparativo de arquiteturas CNNs ....	21
FIGURA 5	– Pipeline da abordagem de aprendizado ativo. ....	23
FIGURA 6	– Principais etapas adotadas em trabalhos para a classificação de lesões de pele. ....	27
FIGURA 7	– <i>Pipeline</i> da abordagem proposta para classificação de lesões de pele, utilizando técnicas de aprendizado profundo e ativo ....	36
FIGURA 8	– Exemplos de imagens do conjunto de dados HAM10000 a partir de cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc. ....	38
FIGURA 9	– Exemplos de imagens do conjunto de dados $D_1$ com aplicação de aumento de dados em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) vasc (com exceção da classe <i>nv</i> ). A primeira imagem de cada coluna, corresponde à imagem original ....	39
FIGURA 10	– Exemplos de imagens do conjunto de dados $D_2$ com aplicação de segmentação utilizando a arquitetura U-Net em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc. ..	40
FIGURA 11	– Exemplos de imagens do conjunto de dados $D_3$ com aplicação de segmentação com revisão manual do especialista em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc. ..	41
FIGURA 12	– Resultados considerando a abordagem de aprendizado supervisionado tradicional. Matrizes de confusão correspondentes às melhores configurações (i.e. combinação extrator-classificador) para cada conjunto de dados: (a) $D_1$ com ACC-RF, (b) $D_1$ com BIC-RF, (c) $D_2$ com BIC-RF e (d) $D_3$ com BIC-RF. ....	49
FIGURA 13	– Exemplos de imagens apresentando similaridade/dissimilaridade visual (maiores/menores confusões) entre as classes: (a) bkl, (b) mel, (c) nv e (d) vasc ....	50
FIGURA 14	– Resultados considerando a abordagem de aprendizado ativo. Acurácias a cada porcentagem de amostras rotuladas obtidas pelas estratégias de aprendizado ativo DBE, EN, LC, MS, MST-BE, RAND e RDS com o classificador RF para os conjuntos de dados e respectivas configurações extrator-classificador. (a) $D_1$ , ACC-RF, (b) $D_1$ , BIC-RF, (c) $D_2$ , BIC-RF e (d) $D_3$ , BIC-RF ....	51
FIGURA 15	– Resultados considerando a abordagem de aprendizado ativo. Acurácias e tempos de organização, seleção e classificação obtidas pelas estratégias de aprendizado ativo DBE, EN, LC, MS, MST-BE e RDS, para os conjuntos de dados e respectivas configurações	

extrator-classificador. (a) $D_1$ , ACC-RF, (b) $D_1$ , BIC-RF, (c) $D_2$ , BIC-RF e (d) $D_3$ , BIC-RF. ....	52
---	----

## LISTA DE TABELAS

TABELA 1	– Trabalhos que exploram automatização na classificação de lesões de pele, considerando diferentes abordagens, bases de dados, quantidades de classes, métricas de avaliação, tais como: SE (sensibilidade), ACC (acurácia), SP (especificidade), AUC ( <i>area under the ROC curve</i> )	28
TABELA 2	– Trabalhos que exploram abordagens baseadas em morfologia para classificação de lesões de pele.	29
TABELA 3	– Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando CNNs tradicionais com transferência de aprendizado, para classificação de lesões de pele.	31
TABELA 4	– Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando CNNs próprias, para classificação de lesões de pele.	31
TABELA 5	– Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando aumento de dados, para classificação de lesões de pele.	33
TABELA 6	– Descrição do conjunto de dados HAM10000, incluindo o tipo (maligno ou benigno) e a distribuição das amostras em cada categoria de lesão, considerando o conjunto de dados original (com 10.015 amostras) e o conjunto após a aplicação da estratégia de aumento de dados (com 40.308 amostras)	37
TABELA 7	– Descritores de imagens, tipos e quantidades de características extraídas.	42
TABELA 8	– Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias $\pm$ desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados $D_1$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco	45
TABELA 9	– Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias $\pm$ desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados $D_2$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco	46



TABELA 10	Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias $\pm$ desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados $D_3$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco .....	47
TABELA 11	Resultados considerando a abordagem de aprendizado supervisionado tradicional. Métricas de avaliação (acurácia, precisão, sensibilidade, F1-Score) obtidas pelas melhores configurações (i.e. combinação extrator-classificador) para cada conjunto de dados ( $D_1, D_2, D_3$ ). Os melhores resultados obtidos são destacados em negrito .....	48

## LISTA DE SIGLAS

BoF	<i>Bag-of-Features</i>
BIC	<i>Border/Interior Classification</i>
CNN	<i>Convolutional Neural Network</i>
DBE	<i>Decreasing Boundary Edges</i>
EN	<i>Entropy</i>
GHI	<i>Generalized Intersection Histogram</i>
PHOG	<i>Generates Pyramid Histograms of Oriented Gradients</i>
GAN	<i>Generative Adversarial Networks</i>
HOG	<i>Histograms of Gradients</i>
HSV	<i>Hue Saturation Value</i>
HSI	<i>Hue Saturation Intensity</i>
INCA	<i>Instituto Nacional de Câncer José Alencar da Silva</i>
CIE	<i>International Commission on Illumination</i>
<i>k</i> -NN	<i>k-Nearest Neighbor</i>
LC	<i>Least Confidence</i>
LR	<i>Logistic Regression</i>
MS	<i>Margin Sampling</i>
MLP	<i>Multilayer Perceptron</i>
MST-BE	<i>Minimum-Spanning Tree Boundary Edges</i>
NB	<i>Naive Bayes</i>
OCS	<i>Opponent Color Space</i>
PPGI	<i>Programa de Pós-graduação em Informática</i>
RBE	<i>Random Boundary Edges</i>
RF	<i>Random Forest</i>
RGB	<i>Red Green e Blue</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SVM	<i>Support Vector Machine</i>
UC	<i>Unsupervised Clustering</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
1.1	Objetivos	15
1.1.1	Objetivo geral	15
1.1.2	Objetivos específicos	16
1.2	Organização do Texto	16
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>17</b>
2.1	Rede Neural Convolutacional	17
2.1.1	Arquiteturas CNN	20
2.2	Abordagens de Aprendizado Ativo	22
2.3	Trabalhos Relacionados	26
2.3.1	Abordagens baseadas em morfologia	27
2.3.2	Abordagens baseadas em aprendizado profundo	30
<b>3</b>	<b>METODOLOGIA</b>	<b>35</b>
<b>4</b>	<b>EXPERIMENTOS E RESULTADOS</b>	<b>43</b>
4.1	Descrição dos Experimentos	43
4.2	Análise e discussão dos resultados	45
4.3	Aprendizado Ativo	48
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>53</b>
	<b>REFERÊNCIAS</b>	<b>55</b>

## 1 INTRODUÇÃO

Atualmente, o câncer é uma das doenças que mais causam morte (INSTITUTE, 2019). A mesma pode ocorrer em qualquer região do corpo, sendo a sua denominação dada geralmente a um grupo de mais de 100 doenças que têm em comum o crescimento desordenado de células biológicas. O Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA) estima que para 2020/2022 haverá ocorrência de cerca de 625 mil novos casos de câncer no Brasil para cada um dos anos (INCA, 2020a).

Para o câncer de pele, segundo o INCA, é esperado para cada ano do triênio 2020/2022, 83.770 em homens e de 93.160 em mulheres com um risco estimado de 80,12 novos casos para cada 100 mil homens e 86,65 novos casos a cada 100 mil mulheres (INCA, 2020b). O câncer de pele divide-se em não melanoma e melanoma. O tipo não melanoma é o de maior ocorrência entre homens e mulheres (BRAY et al., 2021), ao contrário do tipo melanoma, que tem menor incidência, mas uma inerente característica de crescimento acelerado, e maior letalidade (INCA, 2020b).

Atualmente, uma das técnicas utilizadas por dermatologistas para diagnosticar lesões da superfície da pele é a dermatoscopia. Tal técnica permite ao dermatologista realizar o exame de maneira não invasiva e segura. O procedimento é simples, o médico aplica um gel incolor para evitar a reflexão da pele, fato que permite uma melhor visualização, posteriormente, posiciona o dermatoscópio sobre a área de interesse. O equipamento permite uma ampliação de até 10 vezes no local de interesse (DERMATOSCOPY, 2020). A dermatoscopia vem sendo amplamente utilizada, pois permite ao médico uma visualização mais detalhada da lesão em relação ao olho nu.

Sendo assim, após a visualização da lesão com melhores detalhes, o

especialista pode utilizar alguns dos métodos de identificação para realizar o diagnóstico da lesão como, por exemplo, a regra ABCD, sendo (A) assimetria, (B) borda, (C) cor e (D) diâmetro (NACHBAR et al., 1994). Mesmo com equipamentos que auxiliam os exames, o diagnóstico é totalmente baseado na inspeção visual do especialista, e deve-se levar em consideração a experiência clínica do profissional, pois a interpretação é subjetiva mesmo para dermatologistas experientes (BINDER et al., 1995).

Nos últimos anos, têm ocorrido um crescente interesse na automatização da classificação de lesões de pele (Ganster et al., 2001; Situ et al., 2008; Barata et al., 2014; Ayan; Ünver, 2018; PEREZ et al., 2018; Mohamed; El-Behaidy, 2019; BISLA et al., 2019; Rashid et al., 2019; BISSOTO et al., 2019; MANGAROLIYA; SHAH, 2020; Nasr-Esfahani et al., 2016; Ayan; Ünver, 2018; Khan et al., 2019; CODELLA et al., 2015; GESSERT et al., 2018; de Oliveira Lima et al., 2019; Moldovan, 2019; Pai; Giridharan, 2019; Emara et al., 2019; REZVANTALAB et al., 2018; CHATURVEDI et al., 2020; ESTEVA et al., 2017; KAWAHARA; HAMARNEH, 2016; Bassi; Gomekar, 2019; SHI et al., 2019). Por meio de pesquisas realizadas na literatura, foi possível identificar algumas abordagens empregadas para este contexto. Identificou-se que a maioria dos trabalhos segue etapas como: aquisição das imagens, pré-processamento, segmentação, extração de características e classificação. No entanto, alguns exploram apenas algumas dessas etapas.

As abordagens da literatura evoluem desde as mais tradicionais, com o uso de morfologia explorando padrões de cor, textura, tamanho da lesão, entre outros; até as mais recentes envolvendo aprendizado profundo. Algumas abordagens apresentam menor complexidade, já outras demandam maior esforço computacional para alcançar bons resultados. Identificou-se que os esforços atualmente estão concentrados na utilização de abordagens baseadas em aprendizado profundo, explorando transferência de aprendizado a partir de arquiteturas tradicionais (tais como: Inception, ResNet, VGG, DenseNet, entre outras).

Apesar de resultados relevantes obtidos por abordagens de aprendizado profundo, ainda existe uma grande lacuna em relação a aplicação das mesmas. Isso ocorre justamente devido à subjetividade do diagnóstico de acordo com o especialista,

ou seja, de acordo com o contexto da lesão. Além disso, para que métodos de aprendizado profundo consigam uma convergência adequada é necessária uma quantidade considerável de amostras anotadas, bem como um considerável custo computacional para o treinamento do modelo.

Abordagens de aprendizado ativo possibilitam a seleção de um pequeno conjunto de amostras mais informativas a serem apresentadas ao especialista para anotação e utilizadas no aprendizado do classificador. Tais abordagens têm sido amplamente utilizadas e bem sucedidas em diversos contextos, permitindo obter acurácias elevadas mais rapidamente e com mínima intervenção do especialista no processo de aprendizado. Apesar de resultados significativos, algumas delas são ineficientes ou até mesmo inviáveis de serem aplicadas, dependendo do contexto e da quantidade de amostras.

Vale ressaltar que, de acordo com o levantamento bibliográfico realizado até o momento da escrita do presente texto, foi encontrado apenas um trabalho, o qual explora a abordagem de aprendizado ativo para o contexto do projeto. Além disso, outras estratégias para seleção de amostras mais informativas podem ser investigadas, bem como uma avaliação experimental mais extensiva. Dessa forma, o presente trabalho pretende avaliar o uso de algumas técnicas como aumento de dados e segmentação, investigando abordagens de aprendizado ativo mais eficazes e eficientes, de forma a obter classificadores mais robustos.

## 1.1 OBJETIVOS

### 1.1.1 OBJETIVO GERAL

O presente trabalho tem como objetivo a proposta, o desenvolvimento e a validação de métodos de aprendizado ativo, incluindo novas estratégias de seleção de amostras mais informativas, de forma a obter classificadores mais eficazes e eficientes para a análise e a classificação de lesões de pele.

### 1.1.2 OBJETIVOS ESPECÍFICOS

- Obtenção de conjuntos de dados de imagens de lesão de pele;
- Análise, extração e obtenção de características que melhor descrevem os conjuntos de imagens, por meio de descritores tradicionais e baseados em aprendizado profundo;
- Avaliação experimental extensiva para validação de métodos de aprendizado ativo mais adequados ao contexto de lesões de pele, por meio de diferentes critérios para a seleção de amostras mais informativas.

### 1.2 ORGANIZAÇÃO DO TEXTO

Este trabalho está estruturado nos seguintes capítulos: no Capítulo 2 são apresentados os conceitos fundamentais que envolvem abordagens de aprendizado profundo e de aprendizado ativo, concluindo o capítulo com alguns trabalhos existentes na literatura, elucidando a evolução dos mesmos. No Capítulo 3 é descrita a metodologia empregada neste trabalho de pesquisa, detalhando as fases e os métodos empregados. O Capítulo 4 discute os resultados obtidos e as contribuições relevantes no contexto de classificação de imagens de lesão de pele. O Capítulo 5 aponta as considerações finais e futuros trabalhos.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados conceitos e trabalhos relacionados para uma melhor compreensão deste trabalho. Os conceitos estão relacionados às redes neurais convolucionais (Seção 2.1) e abordagens de aprendizado ativo (Seção 2.2). Os trabalhos relacionados envolvem abordagens para classificação de lesões de pele encontrados na literatura (Seção 2.3).

### 2.1 REDE NEURAL CONVOLUCIONAL

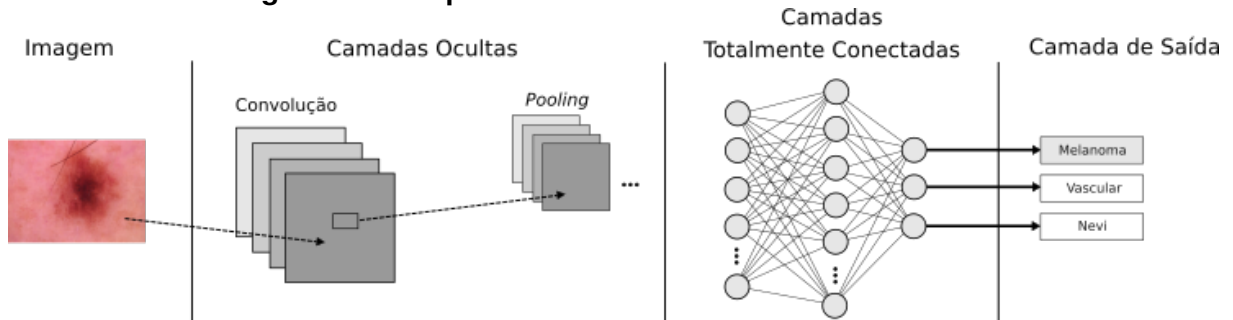
Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs) têm sido amplamente empregadas no contexto de visão computacional. As mesmas têm inspirações biológicas com intuito de emular as ativações cerebrais que ocorrem em humanos e animais (Lecun et al., 1998). As CNNs são constituídas analogamente em camadas que efetuam operações matemáticas lineares definidas como convoluções (GOODFELLOW et al., 2016a).

Especificamente, as CNNs são compostas por duas fases de processamento que consistem em extração de características e classificação. A Figura 1 ilustra um exemplo de uma CNN. Na fase de extração de características, camadas de convoluções e *pooling* são empregadas, seguida pela fase de classificação composta por uma rede de neurônios artificiais (totalmente conectados) para a classificação das características extraídas. Na sequência, uma descrição de cada fase e camada é apresentada.

- **Camadas Convolucionais:** são estruturas de processamento de dados que operam utilizando filtros de convolução, também denominados de *kernels* ou máscaras. A Figura 2 ilustra uma representação do processo de convolução.



**Figura 1: Exemplo de uma Rede Neural Convolucional.**



Fonte: Adaptado de (RAWAT; WANG, 2017)

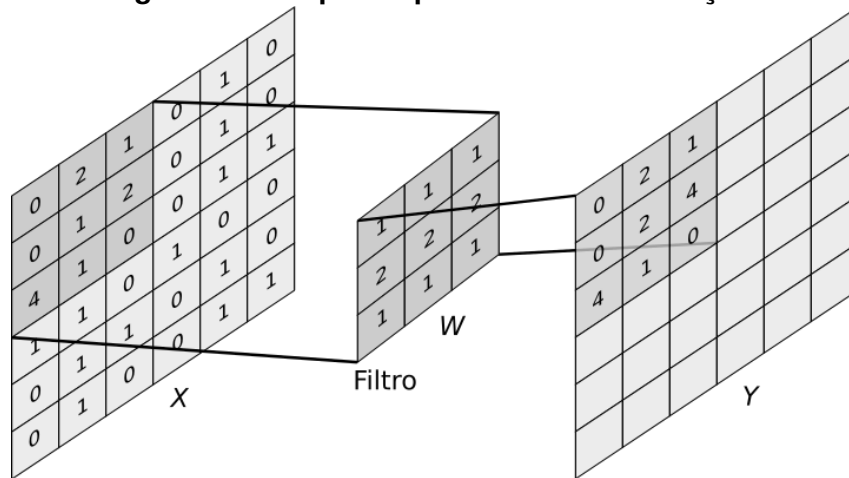
O filtro de convolução  $W$  é uma matriz de valores que representam pesos, sendo que tal matriz é transladada (“deslizada”) sobre a estrutura de entrada (e.g. imagem)  $X$ . A cada deslizamento é realizado cálculo para obter uma estrutura secundária chamada de mapa de características (*feature map*)  $Y$ . A Equação 1 apresenta formalmente a obtenção do mapa de características (RAWAT; WANG, 2017).

$$Y = f(W * X) \quad (1)$$

em que, o mapa de característica  $Y$  é obtido por meio do produto interno entre o filtro  $W$  e a matriz de dados (e.g, imagem) de entrada  $X$ ;  $f(x)$  trata-se de uma função de ativação não-linear (e.g. *ReLU*) para inserir o fator de não-linearidade ao processo e prover melhor tratamento para problemas não-convexos.

- **Camada de Pooling:** é responsável pela redução da dimensionalidade espacial das representações fornecidas pelo mapa de características, originadas a partir das camadas de convoluções, auxiliando na obtenção de invariância de pequenas translações de objetos contidos na imagem (GOODFELLOW et al., 2016b). Analogamente à camada de convolução também utiliza filtros. As operações dos filtros podem ser efetuadas utilizando diferentes abordagens (e.g. *max pooling*, *average pooling*, etc). Tais abordagens divergem somente no cálculo matemático efetuado. Por exemplo, filtros *max pooling* percorrem o mapa de características e extraem o maior valor contido na área determinada pelo filtro. Já o *average pooling*, consiste em obter a média aritmética dos

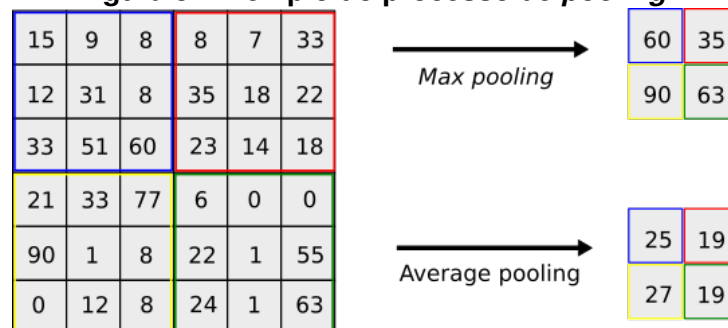
**Figura 2: Exemplo do processo de convolução.**



Fonte: Adaptado de (RAWAT; WANG, 2017)

valores encontrados na área do filtro (GOODFELLOW et al., 2016b). A Figura 3 ilustra a aplicação de um filtro de *pooling* de dimensões 3x3 utilizando as duas abordagens.

**Figura 3: Exemplo do processo de *pooling*.**



Fonte: Adaptado de (RAWAT; WANG, 2017)

- **Camada Totalmente Conectada:** após as camadas convolucionais e de *pooling*, características com diferentes granularidades de abstração (das mais simples às mais complexas) são identificadas. As mesmas são repassadas às camadas totalmente conectadas, responsáveis pela interpretação de tais características (GOODFELLOW et al., 2016b). A camada de entrada para essa etapa recebe os parâmetros advindos da camada de *pooling*, com a quantidade de neurônios de entrada igual à quantidade de parâmetros filtrados. Já a quantidade de neurônios da camada de saída depende da tarefa para a qual a CNN foi designada (i.e. da quantidade de classes do problema em questão)

(RAWAT; WANG, 2017).

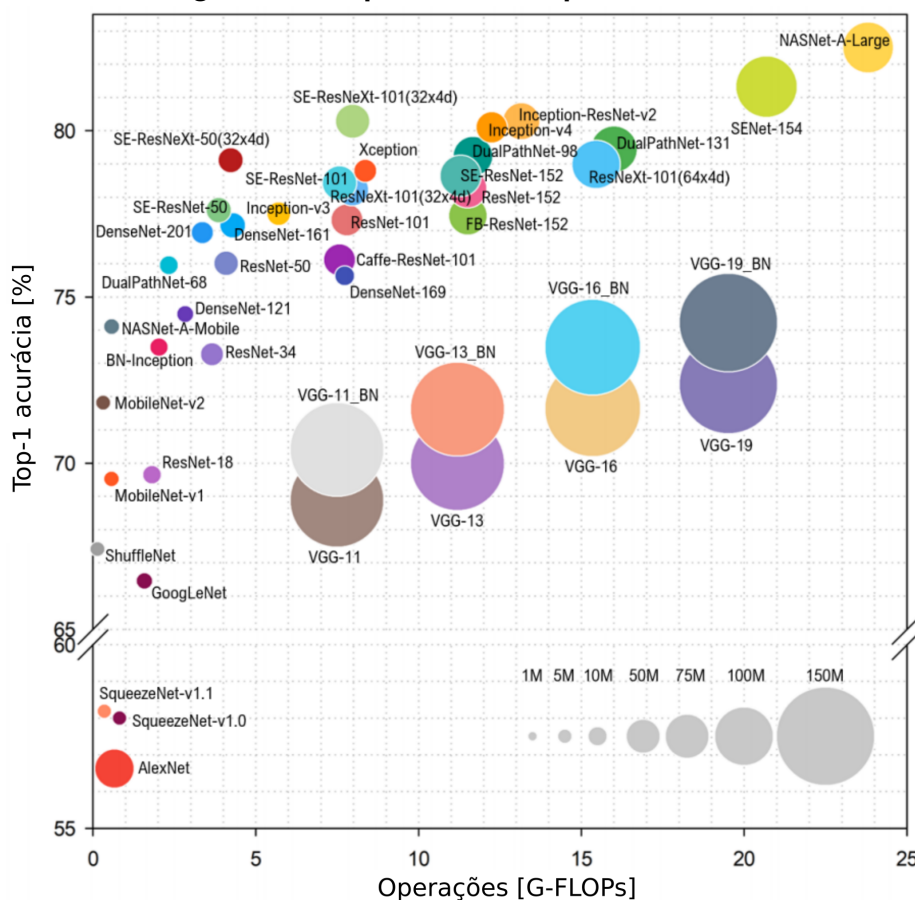
### 2.1.1 ARQUITETURAS CNN

Nos últimos anos a área de visão computacional tem acompanhado o crescente surgimento de arquiteturas CNNs (CANZIANI et al., 2016). Alguns fatores têm contribuído para isto, como avanços em *hardwares* para processamento computacional, conjunto de dados anotados, otimização de algoritmos e melhores arquiteturas CNNs (SZEGEDY et al., 2014). Projetos como *ImageNet* (RUSSAKOVSKY et al., 2015) têm fomentado a pesquisa no âmbito da imagem e visão, resultando em pesquisas e desenvolvimentos de arquiteturas CNNs cada vez melhores. A Figura 4 apresenta um gráfico de algumas arquiteturas existentes, sendo as mesmas comparadas de acordo com suas acurácias e quantidade de operações efetuadas para classificação de uma dada imagem.

As CNNs, treinadas para reconhecer uma grande quantidade de classes de amostras, oriundas do desafio do *ImageNet*, apresentam diversas arquiteturas. Tais CNNs possuem como principais diferenciais a profundidade (i.e. quantidades de camadas), quantidades de características extraídas e como estas características são extraídas. Por exemplo, a arquitetura AlexNet (KRIZHEVSKY et al., 2012), uma das primeiras arquiteturas resultantes do desafio do *ImageNet*, não é tão profunda quando comparada com arquiteturas mais recentes. A mesma possui um total de 8 camadas, sendo 5 convolucionais e 3 totalmente conectadas. Além disso, a rede utiliza camadas de *pooling* e a função *softmax*, responsável por retornar um vetor de mil posições na saída. A rede neural apresenta 60 milhões de parâmetros e 650.000 neurônios.

Com os avanços nas pesquisas, arquiteturas começaram a se tornar mais profundas, afim de melhorar a eficácia (e.g. acurácia, precisão, etc). Por exemplo, as arquiteturas VGG-16 e VGG19 (SIMONYAN; ZISSERMAN, 2014) possuem 16 e 19 camadas, respectivamente. Ambas arquiteturas são compostas por seis blocos de comandos, de forma que os primeiros blocos são formados por uma combinação de camadas de convolução e o último por camadas totalmente conectadas. A arquitetura VGG16 possui 13 camadas convolucionais e três totalmente conectadas e apresenta 138 milhões de parâmetros. A arquitetura VGG19 é composta por 16

Figura 4: Comparativo de arquiteturas CNNs



Fonte: Adaptado de (BIANCO et al., 2018).

camadas convolucionais e três camadas totalmente conectadas e possui 144 milhões de parâmetros. As arquiteturas VGG também utilizam filtros convolucionais de  $3 \times 3$  e camadas de *pooling*.

Com uma nova abordagem, Szegedy et al. (2014) propõem a arquitetura *Inception-v1*, com redução no número de parâmetros na rede e uma abordagem inovadora que consiste em blocos chamados *Inception*. Estes blocos ao invés de empilhar camadas convolucionais como nas arquiteturas anteriores, utilizam convoluções com diferentes filtros em paralelo, obtendo as saídas concatenadas e repassando adiante pela rede. A arquitetura *Inception-v1* apresenta 5 milhões de parâmetros e é composta por 22 camadas, sendo 9 blocos *Inception*. Na versão *Inception-v2* modificações foram realizadas nos blocos *Inception*, melhorando o tempo computacional ao substituir a convolução  $5 \times 5$  por duas convoluções  $3 \times 3$ .

Na versão *Inception-v3*, introduziu-se a utilização de *Batch Normalization* nas saídas das camadas totalmente conectadas. A arquitetura é composta por 48 camadas e possui 23 milhões de parâmetros (SZEGEDY et al., 2015). Em He et al. (2015) foram introduzidas duas abordagens *Inception-v4* e *Inception-ResNet*. No entanto para *Inception-v4* é mantido o conceito de blocos *Inception* com melhorias na arquitetura, fazendo-a mais uniforme e com mais blocos *Inception*. De outro modo, a abordagem empregada em *Inception-ResNet* introduz o conceito *Residual Network*, aumentando a velocidade durante o treinamento. As arquiteturas *Inception-v4* e *Inception-ResNet* são compostas por 43 e 55 milhões de parâmetros, respectivamente.

No sentido de arquiteturas mais profundas, a abordagem de *Residual Networks - ResNet*, apresentado por pesquisadores da *Microsoft Research*, trouxe o princípio aceito que arquiteturas mais profundas são capazes de aprender melhores representações (HE et al., 2015). O conceito *Residual Networks* consiste em que a entrada de um bloco de convolução é somada também à saída deste. Com este conceito foi possível elaborar arquiteturas com até 152 camadas (e.g. a ResNet-152 e a qual apresenta 60 milhões de parâmetros).

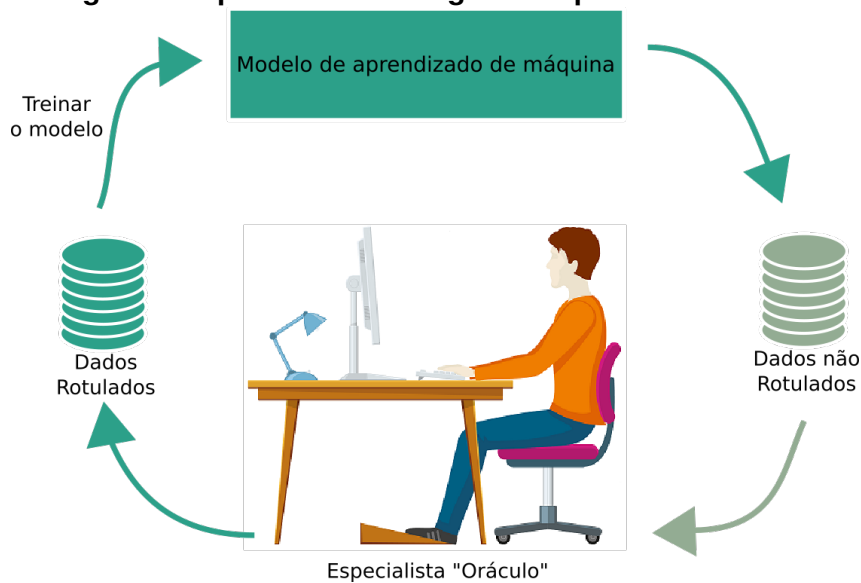
## 2.2 ABORDAGENS DE APRENDIZADO ATIVO

O processo de anotação manual de imagens biomédicas é custoso, pois demanda tempo e atenção por parte do especialista para evitar anotações inconsistentes. Anotações biomédicas devem ser precisas para que se possa desenvolver modelos automatizados com eficácia adequada para serem utilizados na rotina diária do especialista.

A abordagem de aprendizado ativo pode ser utilizada para reduzir o esforço do especialista no processo de anotação e acelerar o processo de aprendizado do classificador. Tal abordagem possibilita, em um processo iterativo, selecionar um conjunto pequeno de amostras mais informativas para serem anotadas por especialistas e utilizadas no treinamento do classificador (SETTLES, 2009). A Figura 5 ilustra o pipeline da abordagem de aprendizado ativo.

Na literatura diferentes estratégias de aprendizado ativo têm sido

**Figura 5: Pipeline da abordagem de aprendizado ativo.**



**Fonte: Autoria Própria**

desenvolvidas (Alajlan et al., 2014; LEWIS; CATLETT, 1994; SCHEFFER et al., 2001; SETTLES, 2009; SAITO et al., 2012; SAITO et al., 2013; SAITO et al., 2014). Um dos principais desafios dessas estratégias refere-se aos critérios adotados para seleção de amostras mais informativas (Huang et al., 2014). Os principais critérios são baseados em diversidade, incerteza e representatividade.

A estratégia de aprendizado ativo *Unsupervised Clustering* - UC (Alajlan et al., 2014) explora o critério de diversidade. Tal estratégia realiza o pré-processamento e a organização dos dados por meio do agrupamento (*clustering*) das amostras. Com as amostras organizadas em grupos, a estratégia UC seleciona um conjunto de amostras aleatórias de cada um dos grupos para anotação dos especialistas e, posteriormente, para o treinamento do classificador. Os autores adotam o critério de diversidade, priorizando a seleção de amostras de diferentes grupos, com o intuito de obter amostras de diferentes classes. No entanto, a seleção de amostras de um dado grupo ocorre de maneira aleatória, considerando todas as amostras daquele grupo. Critérios de incerteza e representatividade, os quais poderiam melhorar o processo de seleção das amostras, não foram considerados.

O critério de incerteza é adotado por técnicas clássicas como: *Least Confidence* (LC), *Margin Sampling* (MS), *Entropy* (EN) (SETTLES, 2012). A técnica

*Least Confidence* (LC) (SETTLES, 2012) é uma das mais simples dentre as clássicas.

A técnica *Least Confidence* (LC) (SETTLES, 2012) é uma das mais simples dentre as clássicas. A técnica LC consiste em priorizar as amostras com a maior incerteza em relação à classe mais provável, conforme definido por  $LC(x) = 1 - P(y'|x)$ , em que  $x$  é uma amostra e  $y'$  é o rótulo mais provável atribuído pelo classificador. Considerando um problema de classificação com três classes: melanoma, nevos e vascular. Após a classificação de uma dada amostra  $A_1$ , obtém-se as seguintes probabilidades: melanoma = 0.3, nevos = 0.4, vascular = 0.8. Para uma outra amostra  $A_2$  obtém-se: melanoma = 0.1, nevos = 0.3, vascular = 0.08. A partir dessas probabilidades, a classe mais provável para a amostra  $A_1$  é a vascular, com um valor de incerteza 0.2 ( $LC(A_1) = 1 - 0.8$ ). Para a amostra  $A_2$  a classe mais provável é a nevos com valor de incerteza de 0.7 ( $LC(A_2) = 1 - 0.3$ ). A amostra  $A_2$  apresenta o maior valor de incerteza, sendo assim será a primeira a ser selecionada. Uma desvantagem da técnica LC consiste em considerar a informação da melhor predição fornecida pelo classificador.

Já a técnica *Margin Sampling* (MS) (SCHEFFER et al., 2001), para seleção das amostras mais incertas, analisa as duas maiores probabilidades de rótulo ( $y'$  e  $y''$ ) de uma dada amostra ( $x$ ) fornecidas por uma determinada instância do classificador. A técnica MS prioriza amostras que apresentam uma margem menor entre as duas classes mais prováveis, conforme definido por  $MS(x) = P(y''|x) - P(y'|x)$ . A partir do mesmo exemplo mencionado anteriormente, em que para a amostra  $A_1$  obtém-se as seguintes probabilidades: melanoma = 0.3, nevos = 0.4, vascular = 0.8 e para a amostra  $A_2$  obtém-se: melanoma = 0.1, nevos = 0.3, vascular = 0.08. Sendo assim, para a amostra  $A_1$ , a classe mais provável é a vascular, seguida da segunda classe mais provável nevos, com o valor de incerteza, segundo a técnica MS, igual a 0.4 ( $MS(A_1) = 0.8 - 0.4$ ). Para a amostra  $A_2$ , tem-se as classes nevos e melanoma como as mais prováveis, com o valor de incerteza igual a 0.2 ( $MS(A_2) = 0.3 - 0.1$ ). Nesse caso, a amostra  $A_2$ , que apresenta o maior valor de incerteza, seria selecionada pela técnica MS. Apesar de considerar não apenas a informação da melhor predição do classificador, para problemas que apresentam uma grande quantidade de classes, muita informação é desconsiderada.

A técnica *Entropy* (EN) ou entropia (SETTLES, 2009) também explora o critério de incerteza. Para tanto, para seleção, prioriza a amostra com maior entropia em relação às probabilidades das amostras, conforme definido por  $EN(x) = -\sum_y P(y|x) \log P(y|x)$ . Diferentemente das técnicas anteriores, a técnica EN não favorece apenas as amostras para as quais apenas um dos rótulos seja altamente improvável. Considerando o mesmo exemplo mencionado anteriormente, em que para a amostra  $A_1$  têm-se as probabilidades: melanoma = 0.3, nevus = 0.4, vascular = 0.8 e para a amostra  $A_2$  as probabilidades: melanoma = 0.2, nevus = 0.3, vascular = 0.08. Após o cálculo de entropia para as amostras  $A_1$  e  $A_2$  obtém-se os valores 0.14 e 0.11, respectivamente. Sendo assim, a amostra  $A_1$  seria selecionada dado que apresenta o maior valor de entropia.

Outras técnicas de aprendizado ativo têm investigado o uso de diferentes critérios (como diversidade, incerteza e representatividade) em uma única estratégia de seleção de amostras. Por exemplo, técnicas como *Random Boundary Edges* (RBE) (SAITO et al., 2012), *Decreasing Boundary Edges* (DBE) (SAITO et al., 2013) e *Minimun-Spanning Tree Boundary Edges* (MST-BE) (SAITO et al., 2014).

As técnicas RBE, DBE e MST-BE consistem em realizar a redução e a organização do conjunto de amostras a priori, como um pré-processamento não supervisionado. As principais diferenças entre as técnicas consistem nas estratégias para organização das amostras. A redução adotada por tais técnicas baseia-se em agrupamento das amostras, por meio de qualquer algoritmo de *clustering*, sendo obtidas as amostras raízes e amostras de fronteira entre grupos. As amostras de fronteiras possibilitam selecionar as amostras mais difíceis (critério de incerteza) para o treinamento do classificador. Em seguida, o conjunto de amostras reduzido pode ser organizado, de acordo com o critério proposto por cada técnica de aprendizado ativo.

Inicialmente, a partir do agrupamento, a ideia é selecionar as raízes dos grupos, de forma a selecionar amostras representativas de todas as classes (critério de representatividade) mais rapidamente. Após a primeira iteração, as raízes dos grupos são exibidas ao especialista, o qual anota os rótulos. Tais amostras anotadas compõem o conjunto de treinamento da primeira instância do classificador.

Durante o processo de aprendizado, a partir da lista organizada previamente,



uma aresta de fronteira por vez é rotulada pelo classificador e as amostras (correspondentes às arestas) são selecionadas, caso elas recebam rótulos distintos (critério de diversidade). O conjunto de amostras selecionado a cada iteração é apresentado ao especialista que pode confirmar ou corrigir o pré-rótulo fornecido pela instância atual do classificador. Em seguida, tal conjunto de amostras é incorporado ao conjunto de treinamento, realiza-se o treinamento do classificador e uma nova instância é obtida. Esse processo continua até que o especialista esteja satisfeito com o desempenho do classificador.

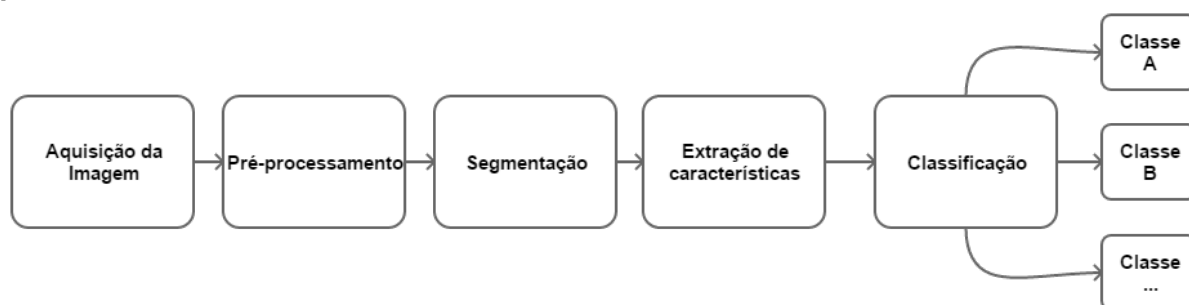
Tais técnicas não requerem a organização das amostras a cada iteração, dado que é proposta a redução e a organização a priori, como um pré-processamento realizado uma única vez. Além disso, a cada iteração, apenas um pequeno conjunto é classificado e selecionado, não ocorre a classificação de todas as amostras do conjunto. Sendo assim, tais técnicas de aprendizado ativo mostram-se adequadas para aplicações que exigem um tempo de resposta rápido e consideram grandes conjuntos de dados.

### 2.3 TRABALHOS RELACIONADOS

Vários trabalhos na literatura têm empreendido esforços em técnicas para automatizar a classificação de lesões de pele entre benignas, malignas (melanoma) e outros tipos de patologias. A Figura 6 ilustra as principais etapas adotadas nestes trabalhos. Inicialmente, é realizada a aquisição da imagem da lesão por um especialista, o qual utiliza ferramentas como o dermatoscópio para obter uma imagem focada. Após a aquisição da imagem, possíveis ruídos são eliminados na etapa de pré-processamento. Em seguida, a etapa de segmentação, apresentada em alguns trabalhos, consiste em segmentar a lesão, separando a mesma do fundo da imagem. Na etapa seguinte, características das lesões segmentadas são extraídas, de forma a discriminar cada categoria de lesões. Após a extração das características, é realizado o treinamento do classificador e a classificação das amostras.

Após o levantamento bibliográfico de alguns trabalhos que abordam técnicas para a classificação de lesões de pele, foi possível identificar cinco abordagens

**Figura 6: Principais etapas adotadas em trabalhos para a classificação de lesões de pele.**



**Fonte: Autoria própria**

gerais exploradas pelos mesmos. Os trabalhos têm explorado abordagens baseadas em morfologia e baseadas em aprendizado profundo. Dentre as baseadas em aprendizado profundo algumas envolvem propostas de aumento de dados, CNNs próprias ou CNNs tradicionais. A Tabela 1 apresenta os trabalhos identificados e organizados quanto às abordagens empregadas. Alguns dos trabalhos focam somente em uma determinada etapa, dentre as apresentadas na Figura 6, outros abrangem mais etapas e técnicas.

### 2.3.1 ABORDAGENS BASEADAS EM MORFOLOGIA

Abordagens baseadas em morfologia (Ganster et al., 2001; Situ et al., 2008; Barata et al., 2014; Pham et al., 2019) exploram características de cor, textura, forma, entre outras características inerentes às lesões (e.g. tamanho) e utilizam classificadores supervisionados tradicionais para classificação. A Tabela 2 apresenta de forma mais detalhada algumas informações referentes a tais propostas.

Alguns trabalhos têm proposta a classificação de lesões de pele por meio de características globais e locais presentes em lesões segmentadas (Ganster et al., 2001). Nesse caso, as características globais são obtidas com base em tamanho (área, perímetro, *polar measures* e *bounding rectangle*), forma (*formfactor*, *roundness*, *compactness*, *extent*) e cor (espaço HSI mínima, máxima, média e variância da intensidade e matiz). As características locais são baseadas na divisão da lesão em quatro quadrantes e categorizadas em três grupos: i-) forma e tamanho, ii-) quantização de cor e iii-) simetria. A classificação das amostras é realizada utilizando

**Tabela 1: Trabalhos que exploram automatização na classificação de lesões de pele, considerando diferentes abordagens, bases de dados, quantidades de classes, métricas de avaliação, tais como: SE (sensibilidade), ACC (acurácia), SP (especificidade), AUC (area under the ROC curve)**

Autores	Abordagem	Base de dados	Qtd. Classes	Métrica
Ganster et al. (2001)	Morfologia	Vienna General Hospital	2	SE 77.0 SP 84.0
			3	SE 87.0 SP 92.0
Situ et al. (2008)	Morfologia	Proprietária	2	ACC 82.0 SE 80.33 SP 78.5
Barata et al. (2014)	Morfologia	Hospital Pedro Hispano	2	SE 100.0 SP 75.0
Pham et al. (2019)	Morfologia	ISIC2016	2	AUC 73.37
		HAM10000	7	AUC 81.46 ACC 74.75
Codella et al. (2015)	CNN Tradicional	ISIC2017	3	ACC 93.1 SE 94.9 SP 92.8
Kawahara e Hamarneh (2016)	CNN Tradicional	<i>Edinburgh Dermofit Library</i>	10	ACC 79.5
Esteva et al. (2017)	CNN Tradicional	<i>ISIC 2016</i>	3	ACC 72.1
		+ <i>Edinburgh Dermofit Library</i>		
		+ <i>Stanford Hospital</i>	9	ACC 55.4
Rezvantab et al. (2018)	CNN Tradicional	HAM10000 + PH2	8	AUC 98.79
Gessert et al. (2018)	CNN Tradicional	HAM10000	7	ACC 85.1
Khan et al. (2019)	CNN Tradicional	ISBI2016	2	ACC 90.2 SE 90.5 SP 99.2
		ISBI2017	3	ACC 95.6 SE 95.5 SP 95.0
		HAM10000	7	ACC 89.8 SE 89.71 SP 89.71
de Oliveira Lima et al. (2019)	CNN Tradicional	HAM10000	7	ACC 93.0 SE 92.0 SP 93.0
Moldovan (2019)	CNN Tradicional	HAM10000	7	ACC 85.0
Pai e Giridharan (2019)	CNN Tradicional	HAM10000	7	ACC 78.0
Emara et al. (2019)	CNN Tradicional	HAM10000	7	ACC 86.17
Chaturvedi et al. (2020)	CNN Tradicional	HAM10000	7	ACC 83.1 SE 83.0 SP 89.0
Nasr-Esfahani et al. (2016)	CNN Própria	<i>Dept. of Dermatology of the University Medical Groningen</i>	2	ACC 81.0 SE 81.0 SP 80.0
Demyanov et al. (2016)	CNN Própria	ISIC2016	2	ACC 88.0
Bassi e Gomekar (2019)	CNN Própria	HAM10000	7	ACC 65.8
	CNN Tradicional			ACC 82.8
Mangaroliya e Shah (2020)	CNN Própria	HAM10000	7	ACC 89.0 SE 76.0 SP 82.0
Ayan e Ünver (2018)	Aumento de Dados	ISBI2016	2	ACC 81.0
Perez et al. (2018)	Aumento de Dados	ISBI2017 + PH2	2	AUC 88.2
Mohamed e El-Behaidy (2019)	Aumento de Dados	HAM10000	7	ACC 97.0
Bisla et al. (2019)	Aumento de Dados	ISBI2017 + PH2	2	AUC 88.0
Rashid et al. (2019)	Aumento de Dados	HAM10000	7	ACC 86.0
Bissoto et al. (2019)	Aumento de Dados	HAM10000	7	ACC 84.7
Shi et al. (2019)	Aprendizado Ativo + CNN Tradicional	ISIC 2017	2	ACC 86.0 SE 47.9 SP 90.8

**Fonte: Autoria própria**

o classificador *k-Nearest Neighbors* (*k*-NN).

Em Situ et al. (2008) explora-se a classificação de melanoma a partir da distribuição de alguns padrões locais nas lesões. As imagens são segmentadas e cada amostra foi representada com base na técnica *Bag-of-Features* - BoF (O'HARA; DRAPER, 2011). As características são extraídas baseadas em filtros *Wavelet* e *Gabor*. Para classificação foram considerados os classificadores *Support*

**Tabela 2: Trabalhos que exploram abordagens baseadas em morfologia para classificação de lesões de pele.**

Autores	Segmentação	Extração	Classificação	Métrica
Ganster et al. (2001)	Sim	Características globais e locais	$k$ -NN	SE 87.0 SP 92.0 (3 classes) SE 77.0 SP 84.0 (2 classes)
Situ et al. (2008)	Sim	Características locais	SVM	ACC 82.0 SE 80.33 SP 78.5 SE 100.0 SP 75.0 (locais)
Barata et al. (2014)	Sim	Características globais e locais	$k$ -NN	SE 96.0 SP 80.0 (globais) AUC 73.37
Pham et al. (2019)	Não	Características globais	RF	AUC 81.46 ACC 74.75

**Fonte: Autoria própria**

*Vector Machine* (SVM) e *Naive Bayes* (NB). O melhor resultado obtido foi utilizando características *Wavelet* e o classificador SVM com o *kernel Generalized Intersection Histogram* (GHI).

Uma avaliação de quais tipos de características, globais ou locais, são mais discriminativas para a classificação de melanoma são apresentadas em Barata et al. (2014). A partir de imagens segmentadas são obtidas características baseadas em textura (*histograms of gradients* - HOG) e baseadas em cor (escalas *Red Green Blue* (RGB), *Hue Saturation Value* (HSV), *Hue Saturation Intensity* (HSI), *International Commission on Illumination* (CIE) (Tkalcic; Tasic, 2003) e *Opponent Color Space* (OCS) (Bratkova et al., 2009), considerando histogramas. Características globais foram obtidas considerando toda a área da lesão e as locais foram utilizando a técnica BoF. Para classificação foram considerados os classificadores *AdaBoost*, SVM e  $k$ -NN. Ambas características globais e locais alcançam resultados significativos, sendo as locais levemente melhores. No entanto, o tempo de extração das características locais é maior em comparação ao das globais. Em ambos os tipos de características, as baseadas em cor apresentam desempenho significativamente melhor em relação às baseadas em textura.

Já em Pham et al. (2019), os autores realizaram um comparativo de características considerando extratores de cor (escala HSV, HOG), textura (*Local Binary Pattern* - LBP) e similaridade (*Scale-Invariant Feature Transform* - SIFT), considerando também as combinações entre as características. Para a classificação foram utilizados diferentes classificadores como SVM, *Logistic Regression* - LR,

*Random Forest* - RF, *AdaBoost*, *Balanced Bagging* e *Balanced RF*. O melhor resultado foi obtido utilizando normalização linear como pré-processamento, características na escala de cor HSV e o classificador *Balanced RF*.

### 2.3.2 ABORDAGENS BASEADAS EM APRENDIZADO PROFUNDO

Segundo Pathan et al. (2018), nos períodos que antecedem 2016, a literatura é voltada para o processamento clássico de imagem, como segmentação por limiar, extração de características baseada em forma, textura, cor, dentre outros. Com o avanço das pesquisas, outras abordagens ganharam destaque.

Abordagens baseadas em aprendizado profundo têm se mostrado importantes no processo de classificação de imagens de lesão de pele, sendo atualmente o estado da arte (Celebi et al., 2019).

Um dos cenários mais difundidos consiste na utilização de arquiteturas tradicionais existentes na literatura (e.g. ResNet, VGG, AlexNet, Inception, DenseNet, dentre outras) treinadas em bases de dados contendo milhões de imagens como o *Image Large Scale Visual Recognition Challenge* - ImageNet (RUSSAKOVSKY et al., 2014). Estas arquiteturas pré-treinadas têm obtido bons resultados na classificação de determinados conjuntos, sendo então passíveis de utilização para efetuar a transferência de aprendizado (*transfer learning*) para outros contextos de imagens. A Tabela 3 apresenta de forma mais detalhada algumas informações referentes aos trabalhos que abrangem esse cenário.

Tais trabalhos da literatura que consideram CNNs tradicionais com transferência de aprendizado (CODELLA et al., 2015; KAWAHARA; HAMARNEH, 2016; ESTEVA et al., 2017; REZVANTALAB et al., 2018; GESSERT et al., 2018; Khan et al., 2019; de Oliveira Lima et al., 2019; Moldovan, 2019; Pai; Giridharan, 2019; Emara et al., 2019; CHATURVEDI et al., 2020) têm utilizado duas estratégias para o treinamento e a classificação. A primeira explora classificação *end2end*, i.e. a arquitetura com seus pesos de treinamento anteriores para classificar um conjunto de dados. A segunda explora a extração de características profundas, utilizando as mesmas para o treinamento de classificadores tradicionais.

**Tabela 3: Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando CNNs tradicionais com transferência de aprendizado, para classificação de lesões de pele.**

Autores	Segmentação	Aumento	Arquitetura	Classificação	Métrica
Codella et al. (2015)	Não	Não	Caffe	SVM	ACC 93.1 SE 94.9 SP 92.8
Kawahara e Hamarneh (2016)	Não	Sim	AlexNet	end2end	ACC 79.5
Esteva et al. (2017)	Não	Sim	InceptionV3	end2end	ACC 72.1
Rezvantab et al. (2018)	Não	Sim	DenseNet201	end2end	AUC 98.79
			Inceptionv3		AUC 98.70
			InceptionResNetv2		AUC 98.20
			ResNet-152		AUC 98.61
Gessert et al. (2018)	Não	Sim	Densenet + ResNeXt + PolyNet + SENet	SVM	ACC 85.1
Khan et al. (2019)	Não	Não	ResNet-50 + ResNet-101	SVM	ACC 90.2 SE 90.5 SP 99.2
					ACC 95.6 SE 95.5 SP 95.0
					ACC 89.8 SE 89.71 SP 89.71
de Oliveira Lima et al. (2019)	Não	Não	InceptionV3	MLP	ACC 93.0 SE 92.0 SP 93.0
Moldovan (2019)	Não	Sim	DenseNet121	end2end	ACC 85.0
Pai e Giridharan (2019)	Não	Não	VGGNet	end2end	ACC 78.0
Emara et al. (2019)	Não	Sim	InceptionV4	end2end	ACC 86.17
Chaturvedi et al. (2020)	Não	Sim	MobileNet	end2end	ACC 83.1 SE 83.0 SP 89.0

Outros trabalhos (MANGAROLIYA; SHAH, 2020; Nasr-Esfahani et al., 2016; Demyanov et al., 2016; Bassi; Gomekar, 2019) têm apresentado a proposta de arquiteturas profundas próprias, realizando o treinamento do zero. A Tabela 4 apresenta algumas informações referentes a estes trabalhos.

**Tabela 4: Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando CNNs próprias, para classificação de lesões de pele.**

Autores	Segmentação	Aumento	Método	Classificação	Métrica
Nasr-Esfahani et al. (2016)	Sim	Sim	Treinamento do zero	End2end	ACC 81.0
Demyanov et al. (2016)	Não	Sim	Treinamento do zero	End2end	ACC 88.0
Bassi e Gomekar (2019)	Não	Sim	Treinamento do zero	End2end	ACC 65.8
			VGG16		ACC 82.8
			VGG19		ACC 79.4
			InceptionV3		ACC 74.8
			MobileNet		ACC 76.0
VGG19 adaptada	ACC 70.0				
Mangaroliya e Shah (2020)	Não	Sim	Treinamento do zero	End2end	ACC 89.0

**Fonte: Autoria própria.**

Em Nasr-Esfahani et al. (2016), os autores propõem inicialmente ações de pré-processamento, ajustando a iluminação dos canais de cores e segmentação utilizando o *k-means*. A arquitetura CNN elaborada não apresenta grande profundidade comparada com outras arquiteturas existentes, possuindo 2 camadas convolucionais e uma camada *max pooling* após cada camada convolucional. A camada totalmente conectada possui 100 neurônios e a final 2 neurônios responsáveis pela ativação das classes Melanoma e Nevus (termo médico para verruga). O modelo foi treinado por 20.000 épocas.

Já em Demyanov et al. (2016) foi proposta uma arquitetura CNN composta por 3 camadas convolucionais e uma camada *max pooling* após cada camada convolucional. A camada totalmente conectada consiste em 128 neurônios e a final 2 neurônios (responsáveis pela ativação das classes Glóbulos Regulares e Rede Típica). O trabalho teve como proposta identificar padrões nas imagens denominados clinicamente como Glóbulos Regulares e Rede Típica. O modelo foi treinado por 50 épocas.

Os autores, em Bassi e Gomekar (2019), além de desenvolverem uma arquitetura própria, exploraram também outras abordagens. Inicialmente, por meio de uma modificação na arquitetura VGG19, informações dos pacientes (e.g. idade, sexo, confirmação do tipo e do local da lesão no corpo) são concatenadas. Além disso, por meio de transferência de aprendizado, utilizam as arquiteturas VGG16, VGG19, InceptionV3 e MobileNet, empregando também o aumento de dados.

Em Mangaroliya e Shah (2020), é proposto um modelo com 20 camadas, sendo composto por 4 camadas convolucionais, e após cada camada de convolução uma camada de *max pooling* e *batch normalization*. É considerada também a técnica *dropout*, após cada camada *batch normalization*. A camada totalmente conectada possui 256 neurônios, seguida de 1 camada *dropout* que antecede a camada final, a qual contém 7 neurônios responsáveis pela ativação da classe. O modelo foi treinado em poucas épocas (30).

Um dos fatores que contribuem para bons resultados, obtidos pelas abordagens de aprendizado profundo, é a quantidade de amostras utilizadas durante a fase de treinamento. Em alguns trabalhos da literatura (Ayan; Ünver, 2018; PEREZ

et al., 2018; Mohamed; El-Behaidy, 2019; BISLA et al., 2019; Rashid et al., 2019; BISSOTO et al., 2019), são utilizadas abordagens para aumento de dados, afim de contornar a falta de amostras naturais (Mikołajczyk; Grochowski, 2018). A Tabela 5 apresenta as técnicas consideradas para a geração de imagens sintéticas.

As principais abordagens adotadas para o aumento de dados consistem em realizar operações simples (Ayan; Ünver, 2018; PEREZ et al., 2018; Mohamed; El-Behaidy, 2019) ou aplicar *Generative Adversarial Networks* (GANs) (BISLA et al., 2019; Rashid et al., 2019; BISSOTO et al., 2019). Ambas abordagens têm possibilitado melhorias significativas no desempenho das arquiteturas. As GANs conseguem gerar imagens sintéticas realistas, apesar de necessitarem um alto poder computacional para atingirem convergência adequada.

**Tabela 5: Trabalhos que exploram abordagens baseadas em aprendizado profundo, considerando aumento de dados, para classificação de lesões de pele.**

Autores	Segmentação	Abordagem	Técnica	Métrica
Ayan e Ünver (2018)	Não	Operações simples	Rotação em ângulos <i>zoom</i> giros verticais/horizontais cortes randômicos deslocamentos	ACC 81.0
Perez et al. (2018)	Não	Operações simples	Modificações nas cores rotação, giros, cortes randômicos remoção randômica de partes distorções união de imagens combinações de operações	AUC 88.2
Mohamed e El-Behaidy (2019)	Não	Operações simples	giro vertical/horizontal rotação angular <i>zoom</i> deslocamento largura e altura	ACC 93.0 ACC 97.0
Bisla et al. (2019)	Sim	GANs	DCGAN	AUC 88.0
Rashid et al. (2019)	Não	GANs	GAN	ACC 86.1
Bissoto et al. (2019)	Não	GANs	pix2pixHD GAN	AUC 84.7

**Fonte: Autoria própria.**

Por meio do levantamento dos esforços encontrados na literatura, foi possível identificar que existe apenas um trabalho (SHI et al., 2019) envolvendo a abordagem de aprendizado ativo no contexto de lesão de pele. Em Shi et al. (2019), foi explorada a abordagem de aprendizado ativo para reduzir o custo de anotação em análise de lesões de pele. Para tanto, foram selecionadas amostras mais informativas baseadas em critérios de diversidade e incerteza para o treinamento do classificador. Nesse trabalho foi considerada a arquitetura ResNet101. A abordagem proposta pelos



autores, com apenas 50% do conjunto de treinamento, atingiu o desempenho de abordagens do estado-da-arte, obtendo resultados equivalentes ou superiores em relação às abordagens que requerem o conjunto de treinamento completo (anotado).

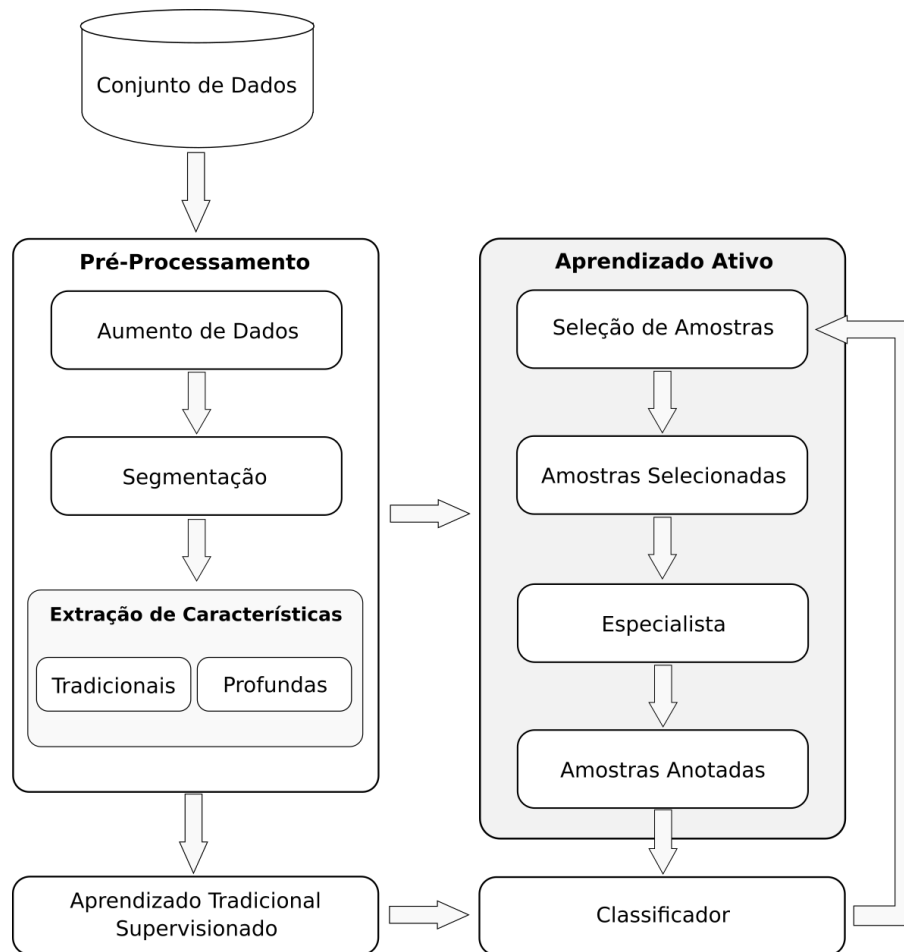
### 3 METODOLOGIA

O presente trabalho explora a abordagem de aprendizado ativo para a classificação de lesões de pele, de forma a selecionar amostras mais informativas para o processo de aprendizado. Além disso, realiza um estudo para obtenção das melhores características que descrevam cada lesão de pele, bem como os classificadores mais adequados para tal contexto. A Figura 7 ilustra as etapas realizadas no presente trabalho.

Como conjunto de dados inicial foi considerado o *Human Against Machine with 10000 training images* (HAM10000) (TSCHANDL et al., 2018). O mesmo possui 10.015 imagens dermatoscópicas coletadas pelo departamento de Dermatologia da Universidade Médica de Vienna, Austria, e da prática clínica do professor Cliff Rosendahl em Queensland, Austrália (TSCHANDL et al., 2018).

Este conjunto de dados é o mesmo utilizado no evento ISIC 2018 (CODELLA et al., 2018a), bem como está dividido em (7) lesões de pele encontradas na prática da dermatologia. Trata-se de um conjunto de dados desbalanceado em que apenas uma das classes representa mais de 60% de todo o conjunto. A Tabela 6 apresenta o tipo (maligno ou benigno) para cada categoria de lesão, bem como a distribuição das amostras em cada lesão, considerando o conjunto de dados original e o conjunto de dados com aumento de dados. A Figura 8 ilustra exemplos de imagens do conjunto de dados HAM10000 a partir de cada uma das classes.

Como observado na literatura, o desbalanceamento pode impactar no processo de aprendizado em problemas de classificação. Nesse contexto, para equilibrar o desbalanceamento do conjunto de dados foram aplicadas estratégias de **aumento de dados** para as imagens dos tipos *akiec*, *bcc*, *bkl*, *df*, *mel* e *vasc*, com exceção para a classe *nv*, visto que é a de maior predominância de amostras naturais.



**Figura 7: Pipeline da abordagem proposta para classificação de lesões de pele, utilizando técnicas de aprendizado profundo e ativo**

**Fonte: Autoria própria**

Para tanto, inicialmente, a partir das amostras originais, foram geradas novas amostras por meio da aplicação de simples operações. Dentre as operações passíveis de serem aplicadas para modificação e geração das imagens estão as seguintes: redimensionamento na largura e altura; rotações na imagem até  $180^\circ$ ; mudança na orientação horizontal e vertical; aumento na imagem (zoom); dentre outras. A Figura 9 apresenta exemplos de imagens do conjunto  $D_1$  obtidas a partir do aumento dos dados para cada uma das classes (com exceção da classe  $nv$ , a qual apresenta uma maior quantidade de amostras).

Após a etapa de aumento de dados e a obtenção do conjunto de dados  $D_1$ , foram avaliadas (2) estratégias de **segmentação**, sendo obtidos dois conjuntos de dados  $D_2$  e  $D_3$ , respectivamente. A primeira estratégia de segmentação considera

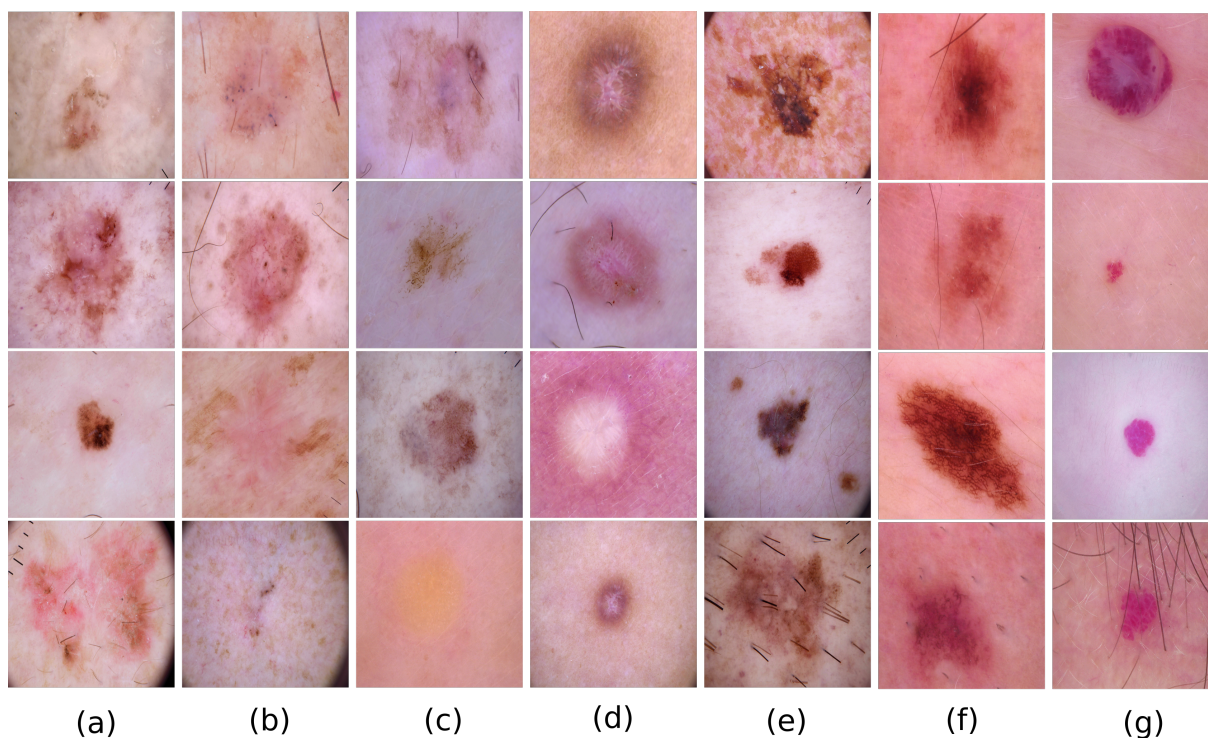
Abreviação - Patologia	Tipo	Original	Aumento
<i>akiec - Actinic keratoses and intraepithelial carcinoma/Bowen's disease</i>	Maligno	327	5.659
<i>bcc - Basal cell carcinoma</i>	Maligno	514	5.654
<i>bkl - Benign keratosis-like lesions. Solar lentigines/seborrheic</i>	Benigno	1099	6.045
<i>df - Dermatofibroma</i>	Benigno	115	4.650
<i>mel - Melanoma</i>	Maligno	1.113	5.865
<i>nv - Melanocytic nevi</i>	Benigno	6.705	6.705
<i>vasc - Vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc)</i>	Benigno	142	5.730

**Tabela 6: Descrição do conjunto de dados HAM10000, incluindo o tipo (maligno ou benigno) e a distribuição das amostras em cada categoria de lesão, considerando o conjunto de dados original (com 10.015 amostras) e o conjunto após a aplicação da estratégia de aumento de dados (com 40.308 amostras)**

o modelo CNN U-Net proposto por (RONNEBERGER et al., 2015). Para treinamento deste modelo como mecanismo de segmentação semântica, foi utilizado o conjunto de dados do *ISIC Challenge 2017* (CODELLA et al., 2018b) que possui 1.991 amostras de imagens de lesões de pele anotadas com *ground truth*. A arquitetura requer poucas épocas de treinamento para obter resultados consideráveis, sendo treinada por 30 épocas. A Figura 10 apresenta exemplos a partir de cada uma das classes de imagens segmentadas do conjunto  $D_2$  após a obtenção das máscaras binárias resultantes do modelo U-Net.

A segunda estratégia de segmentação utiliza máscaras binárias de segmentação disponíveis em (TSCHANDL, 2020) para segmentação do conjunto  $D_1$ . Estas máscaras, segundo (TSCHANDL et al., 2019), são resultados da segmentação semântica utilizando FCN e revisão manual do especialista. A Figura 11 apresenta exemplos a partir de cada uma das classes de imagens segmentadas do conjunto  $D_3$  após a utilização das máscaras disponibilizadas por (TSCHANDL et al., 2019).

Após a obtenção dos (3) conjuntos de dados considerados nos experimentos, é realizada a etapa de **extração de características**, de forma a obter aquelas que melhor descrevam cada lesão de pele. Para este trabalho foram exploradas propriedades inerentes às imagens, utilizando descritores de baixo nível baseados em cor, textura e forma (i.e., *handcrafted features*), bem como o uso de transferência de

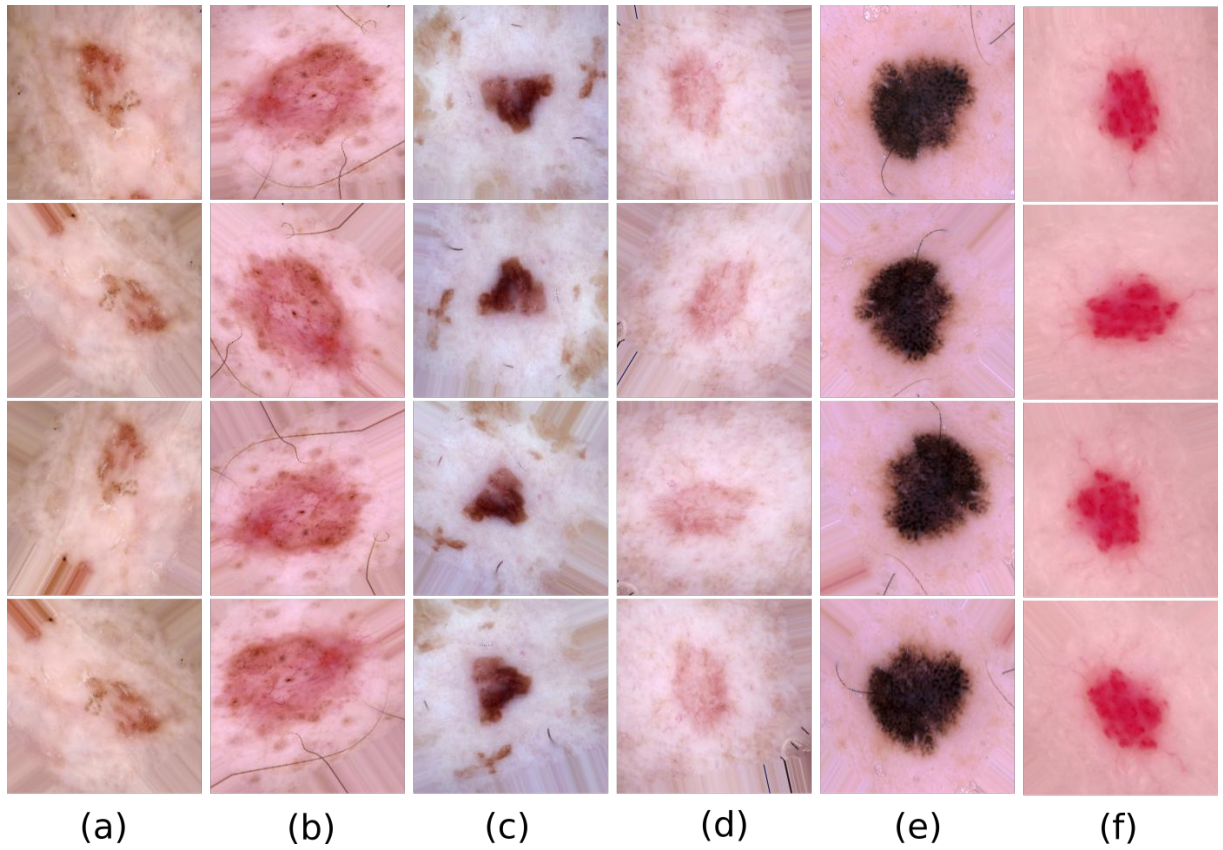


**Figura 8: Exemplos de imagens do conjunto de dados HAM10000 a partir de cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc.**

aprendizado para arquiteturas CNNs encontradas na literatura treinadas no *ImageNet* para obtenção das *deep features*. A Tabela 7 elenca os descritores que foram explorados.

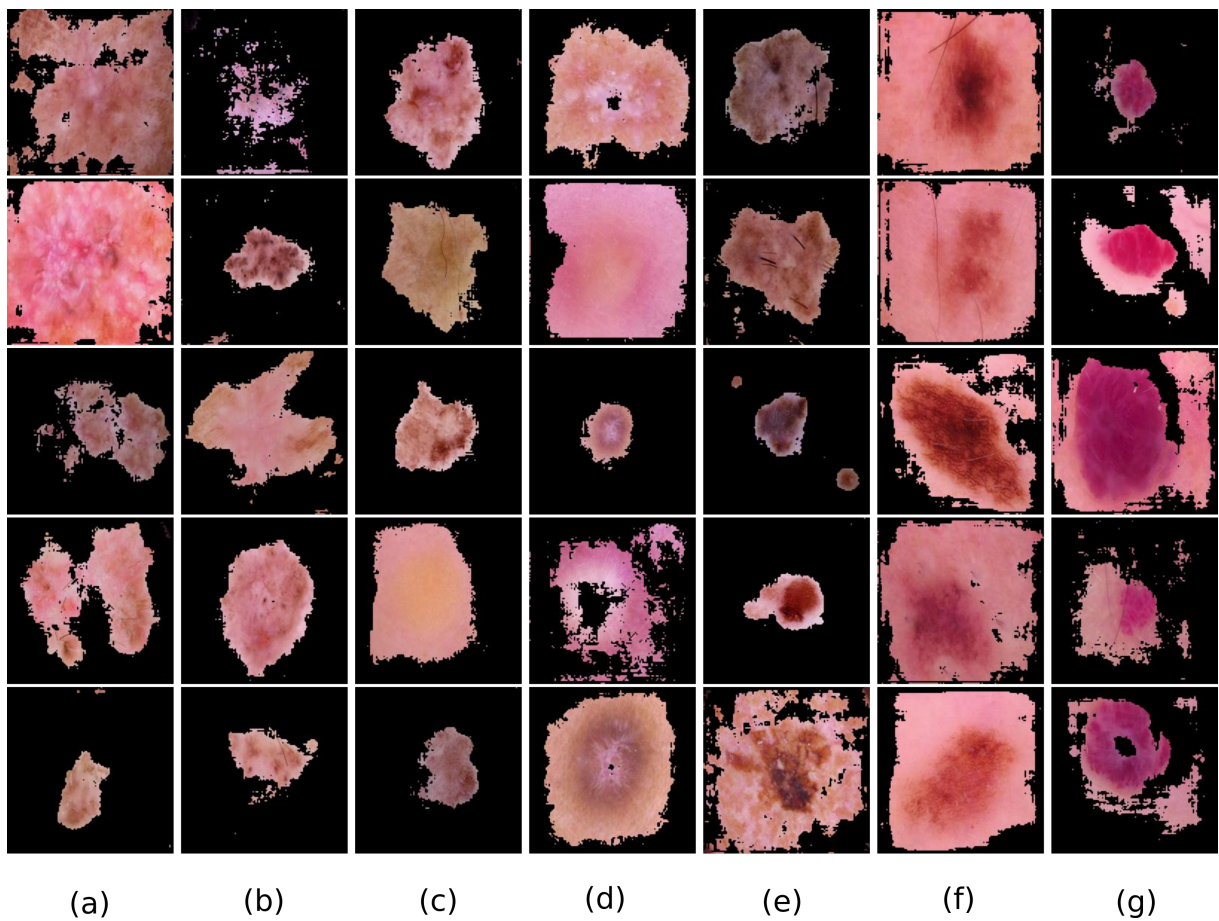
A partir dos vetores de características obtidos na etapa de extração, inicialmente, foram avaliados os melhores cenários e configurações por meio da **abordagem supervisionada tradicional**, de forma a serem considerados posteriormente. Para tanto, além de avaliar quais características (*handcrafted* ou *deep*) melhor descrevem as lesões, foram avaliados quais os **classificadores** de padrões apresentam melhores resultados, bem como as combinações (extrator-classificador) mostram-se mais apropriadas.

Sendo assim, em seguida, foram exploradas as **abordagens de aprendizado ativo** existentes na literatura (*Decreasing Boundary Edges (DBE)*, *Entropy (EN)*, *Least Confidence (LC)*, *Margin Sampling (MS)*, *Minimum-Spanning Tree Boundary Edges (MST-BE)*, *Root Distance-Based Sampling (RDS)*) (SAITO et al., 2013; SETTLES, 2012, 2012, 2012; SAITO et al., 2014; SAITO et al., 2015) (descritas na Seção 4.3) para avaliar a utilização das mesmas no contexto de lesões de pele. Cada estratégia

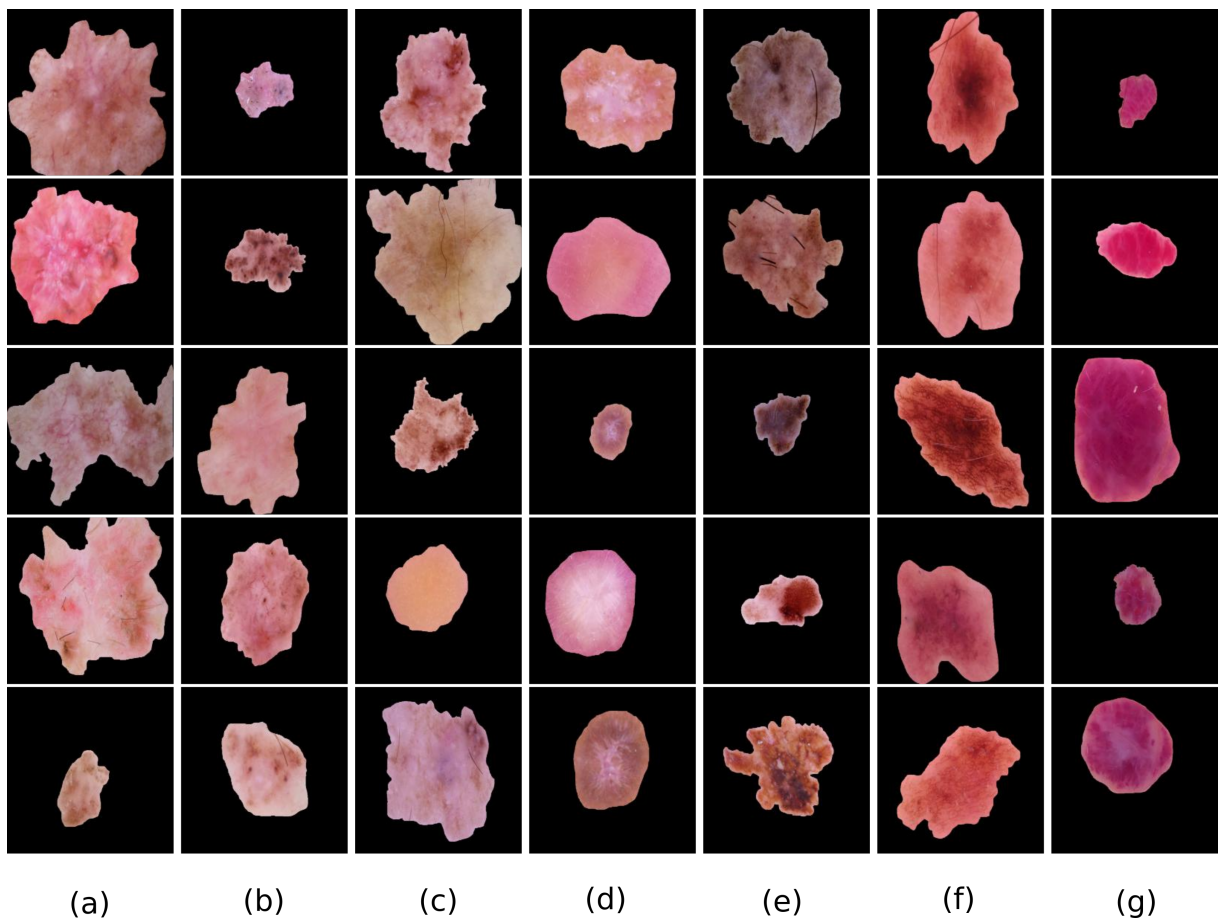


**Figura 9: Exemplos de imagens do conjunto de dados  $D_1$  com aplicação de aumento de dados em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) vasc (com exceção da classe *nv*). A primeira imagem de cada coluna, corresponde à imagem original**





**Figura 10:** Exemplos de imagens do conjunto de dados  $D_2$  com aplicação de segmentação utilizando a arquitetura U-Net em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc.



**Figura 11:** Exemplos de imagens do conjunto de dados  $D_3$  com aplicação de segmentação com revisão manual do especialista em cada uma das classes: (a) akiec, (b) bcc, (c) bkl, (d) df, (e) mel, (f) nv, (g) vasc.



<b>Extrator</b>	<b>Referência</b>	<b>Categoria</b>	<b>Características</b>
ACC	(HUANG et al., 1997)	Cor	768
BIC	(STEHILING et al., 2002)	Cor	128
CEDD	(CHATZICHRISTOFIS; BOUTALIS, 2008a)	Cor	144
FCTH	(CHATZICHRISTOFIS; BOUTALIS, 2008b)	Cor e Textura	192
GCH	(STRICKER; ORENGO, 1995)	Cor	66
HaralickColor	(HARALICK et al., 1973)	Color	15
JCD	(CHATZICHRISTOFIS; BOUTALIS, 2008a)	Cor e Textura	168
LBP	(GUO et al., 2010)	Textura	256
LCH	(SMITH; CHANG, 1996)	Cor	264
Moments	(GRAF, 2015)	Textura	4
MPOC	(GRAF, 2015)	Textura	18
DenseNet121	(HUANG et al., 2016)	Genérica	1.024
InceptionV3	(SZEGEDY et al., 2015)	Genérica	2.048
MobileNet	(HOWARD et al., 2017)	Genérica	1.024
ResNet50	(HE et al., 2015)	Genérica	2.048
VGG16	(SIMONYAN; ZISSERMAN, 2014)	Genérica	512
VGG19	(SIMONYAN; ZISSERMAN, 2014)	Genérica	512

**Tabela 7: Descritores de imagens, tipos e quantidades de características extraídas.**

de aprendizado ativo adota critérios para seleção das amostras mais informativas, de forma a acelerar o processo de aprendizado do **classificador** e minimizar a interação do **especialista** no processo de anotação das amostras a cada iteração do aprendizado.

## 4 EXPERIMENTOS E RESULTADOS

Para a validação das estratégias de seleção da abordagem de aprendizado ativo foi realizada uma avaliação experimental, considerando diferentes cenários, além dos conjuntos de dados descritos no Capítulo 3. Para tanto, foram realizados experimentos preliminares, incluindo diferentes extratores de características e a abordagem de aprendizado supervisionado tradicional. Sendo assim, as melhores configurações e combinações (extrator-classificador) são aplicadas nos experimentos envolvendo a abordagem proposta.

### 4.1 DESCRIÇÃO DOS EXPERIMENTOS

Foi adotado o mesmo protocolo experimental, de maneira a realizar comparações justas entre as abordagens de aprendizado. Os resultados foram obtidos a partir de uma média de 5 execuções, a partir dos conjuntos de amostras gerados aleatoriamente, de forma a compor os conjuntos de aprendizado e de teste. Foi realizada a divisão estratificada, mantendo a mesma proporção (80% e 20%) de amostras de cada classe em cada um dos conjuntos (de aprendizado e de teste, respectivamente).

A partir da obtenção das características (por meio dos extratores apresentados na Tabela 7), o primeiro experimento considera a abordagem de aprendizado supervisionado tradicional, a qual exige o conjunto inteiro anotado. Para tal experimento, foram utilizados os classificadores  $k$ -NN (COVER; HART, 1967), MLP (TAUD; MAS, 2018), RF (BREIMAN, 2001) e SVM (HEARST et al., 1998), bem como com sua respectiva parametrização com valores padrão. Por meio desse experimento, é possível obter um *baseline* para os demais, indicando qual classificador

supervisionado apresenta um melhor desempenho, bem como quais características descrevem melhor os conjuntos.

O segundo experimento explora a abordagem de aprendizado ativo para seleção das amostras mais informativas, com base em diferentes critérios (incerteza, diversidade, representatividade) adotados para a seleção. As estratégias avaliadas foram *Decreasing Boundary Edges* (DBE) (SAITO et al., 2013), *Entropy* (EN) (SETTLES, 2012), *Least Confidence* (LC) (SETTLES, 2012), *Margin Sampling* (MS) (SETTLES, 2012), *Minimum-Spanning Tree Boundary Edges* (MST-BE) (SAITO et al., 2014), *Root Distance-Based Sampling* (RDS) (SAITO et al., 2015). Além disso, foi também avaliada a estratégia *Rand*, em que amostras aleatórias são selecionadas a cada iteração. Para cada uma das estratégias são selecionadas por iteração uma quantidade de amostras equivalente a duas vezes a quantidade de classes existentes nos conjuntos de dados (de forma a serem justas as comparações entre as estratégias, bem como de forma a obter amostras representativas de cada classe desde a primeira iteração, conforme sugerido em (SAITO et al., 2015)). Por exemplo, para os conjuntos de dados ( $D_1$ ,  $D_2$  e  $D_3$ ), os quais apresentam 7 classes, são selecionadas 14 amostras a cada iteração. Tal quantidade de amostras selecionadas a cada iteração é a mesma quantidade de grupos gerados pelo agrupamento considerado por algumas estratégias de aprendizado ativo. Diferentes algoritmos de agrupamento podem ser aplicados. Nos experimentos foi utilizado o algoritmo k-means (HARTIGAN; WONG, 1979). Além dos valores de acurácias médias, são consideradas outras métricas para avaliação (e.g. precisão, revocação, F1-Score, quantidades de amostras corrigidas pelo especialista, tempo para seleção das amostras para cada iteração do aprendizado).

Para facilitar as análises, as comparações entre as estratégias de aprendizado ativo foram realizadas considerando triplas separadas por hífen, indicando as técnicas utilizadas em cada etapa de extração, seleção e classificação. Por exemplo, a tripla BIC-MS-RF, indica que foram considerados o extrator BIC, a estratégia de aprendizado ativo MS e o classificador RF.

Para tais experimentos foi utilizada a configuração de hardware composta por processador Intel Core i5-9400F, 16 gigabyte de memória RAM e placa de vídeo

NVIDIA Geforce GTX 1050 Ti. A linguagem de programação Python 3.0 foi adotada para a codificação e execução dos experimentos, tal como frameworks utilizados popularmente como Scikit-learn (PEDREGOSA et al., 2011), TensorFlow (ABADI et al., 2015) e dependências.

## 4.2 ANÁLISE E DISCUSSÃO DOS RESULTADOS

**Tabela 8: Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias  $\pm$  desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados  $D_1$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco**

Extrator	RF	$k$ -NN	MLP	SVM
ACC	<b>91.01<math>\pm</math>0.48*</b>	79.29 $\pm$ 0.52	45.40 $\pm$ 0.18	40.21 $\pm$ 0.76
BIC	<b>90.20<math>\pm</math>0.88*</b>	76.73 $\pm$ 0.70	46.23 $\pm$ 0.92	38.62 $\pm$ 0.90
CEDD	<b>75.62<math>\pm</math>1.00</b>	<b>75.09<math>\pm</math>0.58</b>	50.38 $\pm$ 0.76	45.11 $\pm$ 0.60
FCTH	<b>75.09<math>\pm</math>0.94</b>	71.21 $\pm$ 0.52	41.46 $\pm$ 1.12	37.84 $\pm$ 0.52
GCH	<b>87.44<math>\pm</math>0.12</b>	78.52 $\pm$ 0.64	46.31 $\pm$ 1.02	37.25 $\pm$ 0.48
HaralickColor	<b>75.53<math>\pm</math>0.34</b>	40.89 $\pm$ 0.94	24.40 $\pm$ 0.56	20.32 $\pm$ 0.42
JCD	<b>79.59<math>\pm</math>1.00</b>	<b>78.8<math>\pm</math>0.82</b>	50.84 $\pm$ 0.48	44.72 $\pm$ 0.44
LBP	<b>48.25<math>\pm</math>0.60</b>	44.73 $\pm$ 1.16	44.41 $\pm$ 0.84	38.47 $\pm$ 0.74
LCH	<b>69.60<math>\pm</math>0.74</b>	63.66 $\pm$ 1.02	46.84 $\pm$ 0.90	38.17 $\pm$ 0.80
Moments	<b>41.68<math>\pm</math>1.04</b>	34.92 $\pm$ 0.74	24.73 $\pm$ 0.62	21.28 $\pm$ 0.60
MPOC	<b>79.06<math>\pm</math>0.54</b>	58.42 $\pm$ 0.58	28.23 $\pm$ 0.82	26.48 $\pm$ 1.04
DenseNet121	75.77 $\pm$ 0.20	<b>85.65<math>\pm</math>0.80</b>	72.82 $\pm$ 0.80	61.72 $\pm$ 1.22
InceptionV3	64.78 $\pm$ 1.12	<b>79.83<math>\pm</math>0.60</b>	70.08 $\pm$ 1.02	58.51 $\pm$ 0.90
MobileNet	71.68 $\pm$ 0.88	<b>89.03<math>\pm</math>0.92</b>	74.78 $\pm$ 1.34	<u>63.22<math>\pm</math>1.38</u>
ResNet50	76.40 $\pm$ 1.16	<b>88.80<math>\pm</math>0.68</b>	77.51 $\pm$ 0.74	<u>64.76<math>\pm</math>1.50</u>
VGG16	72.16 $\pm$ 0.64	<b>85.34<math>\pm</math>0.38</b>	71.50 $\pm$ 0.54	60.79 $\pm$ 0.52
VGG19	71.46 $\pm$ 1.18	<b>84.02<math>\pm</math>0.50</b>	71.33 $\pm$ 0.36	61.44 $\pm$ 0.52
Xception	64.93 $\pm$ 0.96	<b>76.81<math>\pm</math>0.34</b>	70.90 $\pm$ 0.44	59.35 $\pm$ 1.62

Com relação ao primeiro experimento realizado, envolvendo a abordagem de aprendizado tradicional, observa-se que os extratores tradicionais, mais especificamente o extrator BIC apresentou um melhor desempenho para todos os conjuntos  $D_1$ ,  $D_2$  e  $D_3$  (Tabelas 8-10), respectivamente). Além disso, tais extratores extraem uma quantidade de características inferior em relação às arquiteturas profundas (e.g. para o conjunto  $D_1$ , o extrator BIC atinge cerca de 90,20% de acurácia com apenas 128 características, enquanto que a arquitetura MobileNet

atinge cerca de 89,03% de acurácia com 1.024 características). É possível também observar que, de forma geral, o classificador RF apresentou os melhores resultados de acurácias em comparação aos demais ( $k$ -NN, MLP e SVM) para todos os conjuntos de dados, considerando os extratores tradicionais. Percebe-se que o classificador  $k$ -NN apresentou maiores acurácias, em relação ao RF, para todos os conjuntos de dados, considerando as características obtidas por meio das arquiteturas profundas.

**Tabela 9: Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias  $\pm$  desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados  $D_2$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco**

Extrator	RF	$k$ -NN	MLP	SVM
ACC	<b>84.03<math>\pm</math>0.64</b>	74.84 $\pm$ 10.46	42.76 $\pm$ 1.24	36.09 $\pm$ 0.98
BIC	<b>87.43<math>\pm</math>0.42*</b>	75.64 $\pm$ 0.62	44.14 $\pm$ 0.62	33.68 $\pm$ 1.90
CEDD	<b>64.63<math>\pm</math>0.66</b>	<b>64.38<math>\pm</math>0.52</b>	49.98 $\pm$ 0.46	44.74 $\pm$ 0.44
FCTH	<b>67.34<math>\pm</math>0.66</b>	64.53 $\pm$ 0.56	45.99 $\pm$ 0.74	41.42 $\pm$ 0.96
GCH	<b>84.13<math>\pm</math>0.68</b>	76.95 $\pm$ 0.42	40.53 $\pm$ 0.98	29.32 $\pm$ 2.64
HaralickColor	<b>69.00<math>\pm</math>0.80</b>	45.62 $\pm$ 0.56	34.62 $\pm$ 0.94	31.51 $\pm$ 1.14
JCD	<b>71.34<math>\pm</math>0.98</b>	69.56 $\pm$ 0.70	50.66 $\pm$ 0.64	44.26 $\pm$ 1.10
LBP	<b>46.95<math>\pm</math>0.58</b>	43.67 $\pm$ 0.78	25.96 $\pm$ 0.64	19.47 $\pm$ 1.40
LCH	<b>67.04<math>\pm</math>0.80</b>	48.07 $\pm$ 0.82	41.38 $\pm$ 1.14	28.63 $\pm$ 1.12
Moments	<b>33.64<math>\pm</math>0.92</b>	28.84 $\pm$ 0.94	24.62 $\pm$ 0.30	22.15 $\pm$ 0.76
MPOC	<b>69.63<math>\pm</math>0.74</b>	45.57 $\pm$ 0.62	29.00 $\pm$ 0.86	26.63 $\pm$ 1.44
DenseNet121	66.72 $\pm$ 0.92	<b>75.83<math>\pm</math>0.76</b>	<u>67.10<math>\pm</math>0.66</u>	<u>56.87<math>\pm</math>1.06</u>
InceptionV3	54.04 $\pm$ 0.66	<b>63.59<math>\pm</math>0.66</b>	59.95 $\pm$ 1.42	48.20 $\pm$ 1.64
MobileNet	58.43 $\pm$ 0.78	<b>73.15<math>\pm</math>0.48</b>	64.67 $\pm$ 1.10	53.44 $\pm$ 0.94
ResNet50	64.68 $\pm$ 1.18	<b>74.25<math>\pm</math>0.66</b>	<u>67.06<math>\pm</math>0.78</u>	53.16 $\pm$ 1.38
VGG16	60.80 $\pm$ 0.78	<b>69.28<math>\pm</math>0.76</b>	62.98 $\pm$ 0.74	52.77 $\pm$ 0.78
VGG19	59.72 $\pm$ 1.00	<b>67.85<math>\pm</math>0.88</b>	62.51 $\pm$ 0.56	52.23 $\pm$ 1.10
Xception	55.02 $\pm$ 1.26	<b>62.75<math>\pm</math>1.00</b>	61.81 $\pm$ 1.22	48.77 $\pm$ 1.18

Realizando comparações de desempenho entre os conjuntos de dados  $D_1$ ,  $D_2$  e  $D_3$  (ver Tabelas 8-10), pode-se observar que o conjunto  $D_3$  apresentou um melhor desempenho entre todos os conjuntos (conforme Tabela 10). Vale destacar que tal conjunto consiste de imagens, as quais foram inicialmente segmentadas automaticamente e posteriormente receberam avaliação e correção manual do especialista. Analizando os resultados apresentados a partir dos conjuntos  $D_1$  e  $D_2$ , pode-se perceber que, nesse caso, a segmentação (das imagens do conjunto  $D_2$ ), não

Tabela 10: Resultados considerando a abordagem de aprendizado supervisionado tradicional. Média de acurácias  $\pm$  desvio padrão obtidos pelos classificadores ( $k$ -NN, MLP, RF e SVM), utilizando as características obtidas pelos extratores tradicionais e arquiteturas profundas, para o conjunto de dados  $D_3$ . Os melhores resultados (i.e. classificadores) para cada extrator são destacados em negrito. Os melhores resultados (i.e. extratores) para cada classificador são apresentados sublinhados. Os melhores resultados (i.e. maiores acurácias) obtidos são apresentados com asterisco

Extractor	RF	$k$ -NN	SVM	MLP
ACC	<b>86.98<math>\pm</math>0.24</b>	83.68 $\pm$ 0.50	45.27 $\pm$ 0.56	40.48 $\pm$ 1.98
BIC	<b>92.13<math>\pm</math>0.24*</b>	82.02 $\pm$ 0.44	46.29 $\pm$ 0.60	40.54 $\pm$ 2.16
CEDD	<b>71.51<math>\pm</math>0.78</b>	69.54 $\pm$ 1.02	55.49 $\pm$ 0.52	50.12 $\pm$ 0.70
FCTH	<b>77.67<math>\pm</math>0.98</b>	74.24 $\pm$ 1.26	52.38 $\pm$ 0.60	47.24 $\pm$ 0.82
GCH	<b>90.32<math>\pm</math>0.50</b>	84.40 $\pm$ 0.70	41.06 $\pm$ 0.30	34.50 $\pm$ 2.30
HaralickColor	<b>82.39<math>\pm</math>0.34</b>	80.56 $\pm$ 0.58	41.97 $\pm$ 0.66	37.87 $\pm$ 0.94
JCD	<b>79.54<math>\pm</math>0.92</b>	77.23 $\pm$ 0.76	57.28 $\pm$ 0.46	51.45 $\pm$ 0.44
LBP	<b>40.98<math>\pm</math>0.64</b>	36.22 $\pm$ 0.78	23.30 $\pm$ 2.18	20.66 $\pm$ 0.30
LCH	<b>69.14<math>\pm</math>0.82</b>	51.64 $\pm$ 0.60	41.16 $\pm$ 0.52	33.41 $\pm$ 4.52
Moments	<b>47.50<math>\pm</math>0.68</b>	40.28 $\pm$ 0.96	25.10 $\pm$ 0.50	24.45 $\pm$ 0.52
MPOC	<b>82.70<math>\pm</math>0.56</b>	72.61 $\pm$ 0.76	39.91 $\pm$ 0.72	36.04 $\pm$ 1.00
DenseNet121	79.37 $\pm$ 0.72	<b>90.14<math>\pm</math>0.34</b>	74.42 $\pm$ 1.04	63.08 $\pm$ 1.12
InceptionV3	65.27 $\pm$ 1.22	<b>81.72<math>\pm</math>0.22</b>	69.02 $\pm$ 0.90	53.85 $\pm$ 2.02
MobileNet	70.17 $\pm$ 0.98	<b>86.50<math>\pm</math>0.60</b>	71.56 $\pm$ 0.70	58.07 $\pm$ 1.50
ResNet50	78.64 $\pm$ 1.08	<b>88.54<math>\pm</math>0.68</b>	<u>77.16<math>\pm</math>0.64</u>	<u>64.77<math>\pm</math>0.44</u>
VGG16	73.81 $\pm$ 1.20	<b>86.16<math>\pm</math>0.52</b>	69.26 $\pm$ 1.48	58.76 $\pm$ 1.16
VGG19	72.43 $\pm$ 0.76	<b>84.27<math>\pm</math>0.42</b>	68.12 $\pm$ 0.98	56.53 $\pm$ 1.66
Xception	68.99 $\pm$ 0.44	<b>82.54<math>\pm</math>0.58</b>	70.91 $\pm$ 0.82	55.70 $\pm$ 0.88

possibilitou a obtenção de maiores acurácias. O conjunto  $D_2$ , formado por imagens segmentadas automaticamente pela arquitetura U-Net (Figura 10), não apresentou melhores resultados em relação ao conjunto  $D_1$ , o qual consiste de imagens sem segmentação – considerando o *background* das mesmas (Figura 9). Isso pode ter ocorrido devido a problemas no processo de segmentação totalmente automatizado em algumas amostras do conjunto. Apesar de algumas lesões apresentarem boa qualidade na segmentação, outras não apresentaram uniformidade em obter a região de interesse (conforme Figura 10).

A Tabela 11 apresenta outras métricas de avaliação além da acurácia, tais como precisão, sensibilidade e F1-Score obtidas pelas melhores configurações (i.e. combinação extrator-classificador) para cada conjunto de dados. É possível observar que o conjunto  $D_3$  (formado por imagens segmentadas automaticamente e com correções manuais realizadas pelo especialista) apresentou os melhores resultados

considerando todas as métricas. Sendo assim, a abordagem de aprendizado ativo proposta deve ser avaliada utilizando tal conjunto (Seção 4.3).

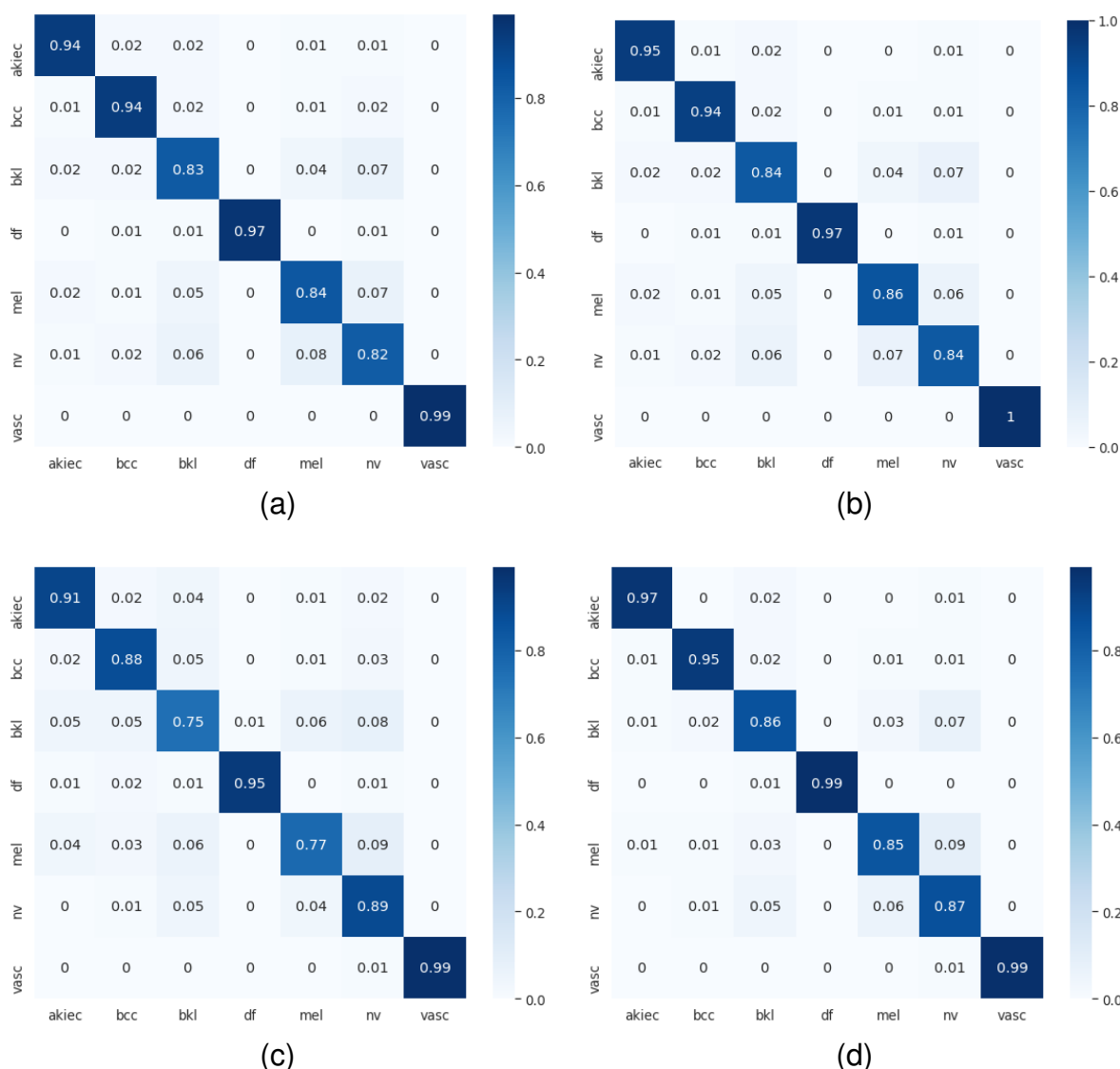
Além dos resultados gerais apresentados a partir das métricas de avaliação mencionadas, foram mensurados os desempenhos por classe, conforme matrizes de confusão (Figura 12). Analisando as matrizes de confusão obtidas pela abordagem de aprendizado supervisionada tradicional (Figura 12), é possível notar resultados significativos (i.e. uma maior concentração de valores na diagonal principal da matriz) para todos os conjuntos ( $D_1$ ,  $D_2$  e  $D_3$ ), considerando suas respectivas melhores combinações extrator-classificador. No entanto, ainda há certa confusão entre algumas classes, e.g. entre as classes *mel* e *nv*; *bkl* e *nv*; *mel* e *bkl*. Observa-se uma maior assertividade para a classe *vasc*. Pode-se observar também que o conjunto  $D_3$  apresentou menos confusões entre as classes em comparação aos demais conjuntos ( $D_1$  e  $D_2$ ). Na Figura 13(a)-(d) são apresentados exemplos de imagens das classes (*bkl*, *mel*, *nv* e *vasc*, respectivamente). Nota-se a elevada similaridade entre as classes *mel*, *nv* e *bkl* (Figura 13(a)-13(c), respectivamente), ao contrário da classe *vasc* (Figura 13(d)).

**Tabela 11: Resultados considerando a abordagem de aprendizado supervisionado tradicional. Métricas de avaliação (acurácia, precisão, sensibilidade, F1-Score) obtidas pelas melhores configurações (i.e. combinação extrator-classificador) para cada conjunto de dados ( $D_1$ ,  $D_2$ ,  $D_3$ ). Os melhores resultados obtidos são destacados em negrito**

Conjunto de dados	Extrator	Classificador	Acurácia	Precisão	Sensibilidade	FScore
$D_1$	ACC	RF	91.01±0.48	91.43	91.45	91.43
$D_1$	BIC	RF	90.02±0.88	90.65	90.67	90.65
$D_2$	BIC	RF	87.43±0.42	87.96	87.74	87.78
$D_3$	<b>BIC</b>	<b>RF</b>	<b>92.13±0.24</b>	<b>92.63</b>	<b>92.50</b>	<b>92.56</b>

### 4.3 APRENDIZADO ATIVO

A abordagem de aprendizado supervisionado tradicional requer uma grande quantidade de amostras rotuladas durante o processo de treinamento. Os resultados apresentados anteriormente na Seção 4.2 foram obtidos por meio de experimentos que consideraram todas as amostras (100%) do conjunto de dados para treinamento dos classificadores. Com outra perspectiva, a abordagem de aprendizado ativo tem como base selecionar amostras mais informativas, com o objetivo de obter um

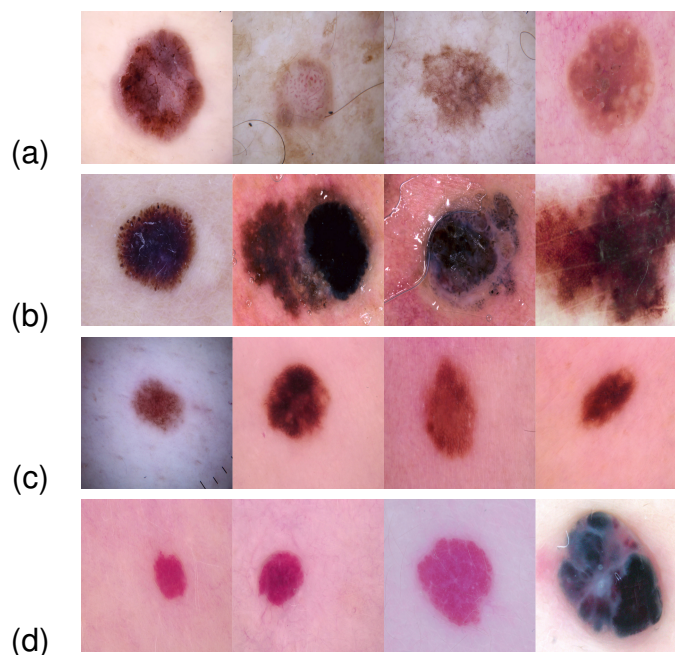


**Figura 12: Resultados considerando a abordagem de aprendizado supervisionado tradicional. Matrizes de confusão correspondentes às melhores configurações (i.e. combinação extrator-classificador) para cada conjunto de dados: (a)  $D_1$  com ACC-RF, (b)  $D_1$  com BIC-RF, (c)  $D_2$  com BIC-RF e (d)  $D_3$  com BIC-RF.**

classificador que atinja maiores acurácias com uma menor quantidade de amostras rotuladas, de forma a acelerar o processo de aprendizado e minimizar os esforços e as interações com o especialista no processo de anotação (correção) das amostras.

Considerando os experimentos relacionados à abordagem de aprendizado ativo (Figura 14), para avaliação das estratégias de seleção (DBE, EN, LC, MS, MST-BE, Rand e RDS), fixou-se o extrator BIC e o classificador RF para todos os conjuntos de dados dado que os mesmos apresentaram um melhor desempenho no experimento anterior (com exceção do conjunto  $D_1$  em que o extrator ACC também apresentou



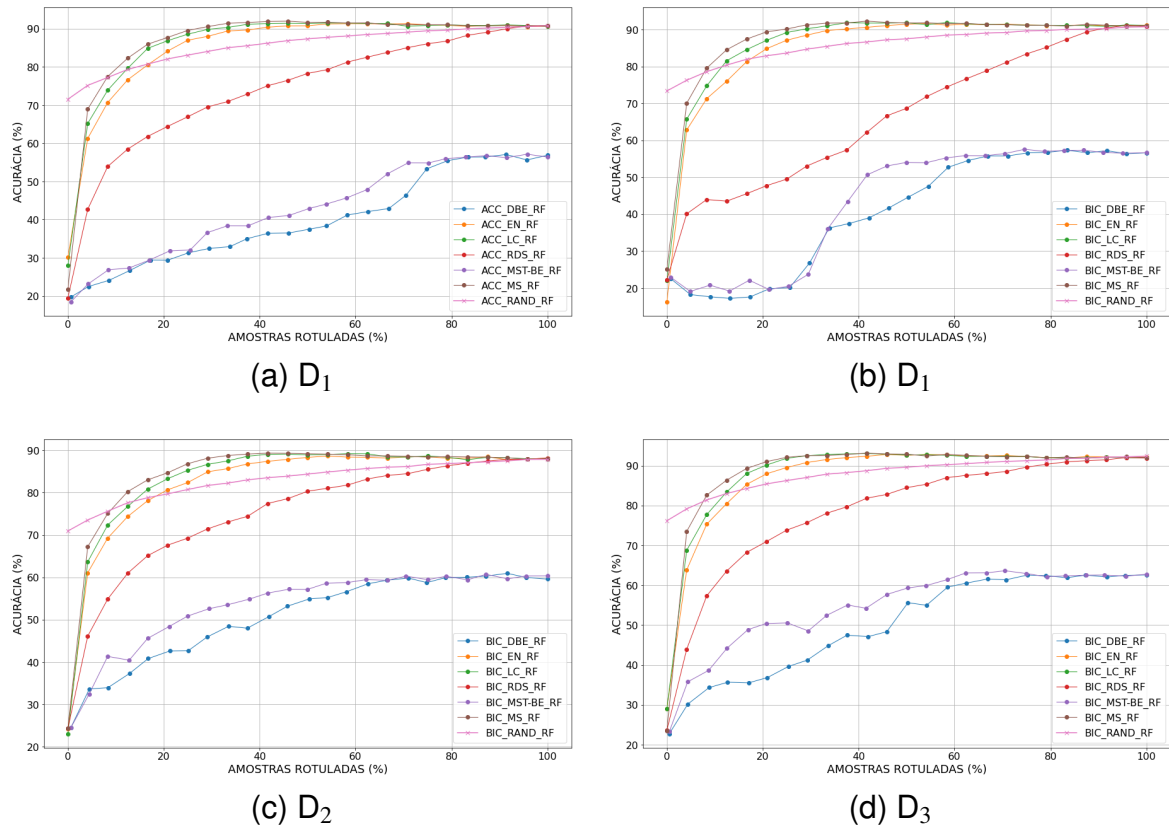


**Figura 13: Exemplos de imagens apresentando similaridade/dissimilaridade visual (maiores/menores confusões) entre as classes: (a) bkl, (b) mel, (c) nv e (d) vasc**

resultados equivalentes ao BIC, sendo também considerado no experimento com a abordagem de aprendizado ativo).

Dentre as estratégias de aprendizado ativo, MS apresentou as melhores acurácias de classificação. É possível também observar que as estratégias de seleção (DBE, EN, LC, MS, MST-BE e RDS) possibilitam atingir elevadas acurácias mais rapidamente (menos iterações de aprendizado) e com uma menor quantidade de amostras rotuladas, para todos os conjuntos (Figuras 14(a)-14(d)), em comparação com a abordagem de aprendizado supervisionado tradicional. Tal abordagem requer uma quantidade maior (100%) de amostras rotuladas para atingir acurácias equivalentes.

Por exemplo, avaliando o conjunto  $D_1$ , a estratégia MS conseguiu alcançar acurácias de  $92.30\% \pm 0.10$  e  $92.35\% \pm 0.10$ , utilizando o classificador RF e os extratores ACC e BIC, respectivamente (Figuras 14(a) e 14(b)) com apenas 40-45% dos dados. Considerando a abordagem de aprendizado supervisionada tradicional (Tabela 8) é possível verificar que o classificador RF, considerando os extratores ACC e BIC, alcançou acurácias de  $91.01\% \pm 0.48$  e  $90.20\% \pm 0.88$ , respectivamente, utilizando todo o conjunto de treinamento.

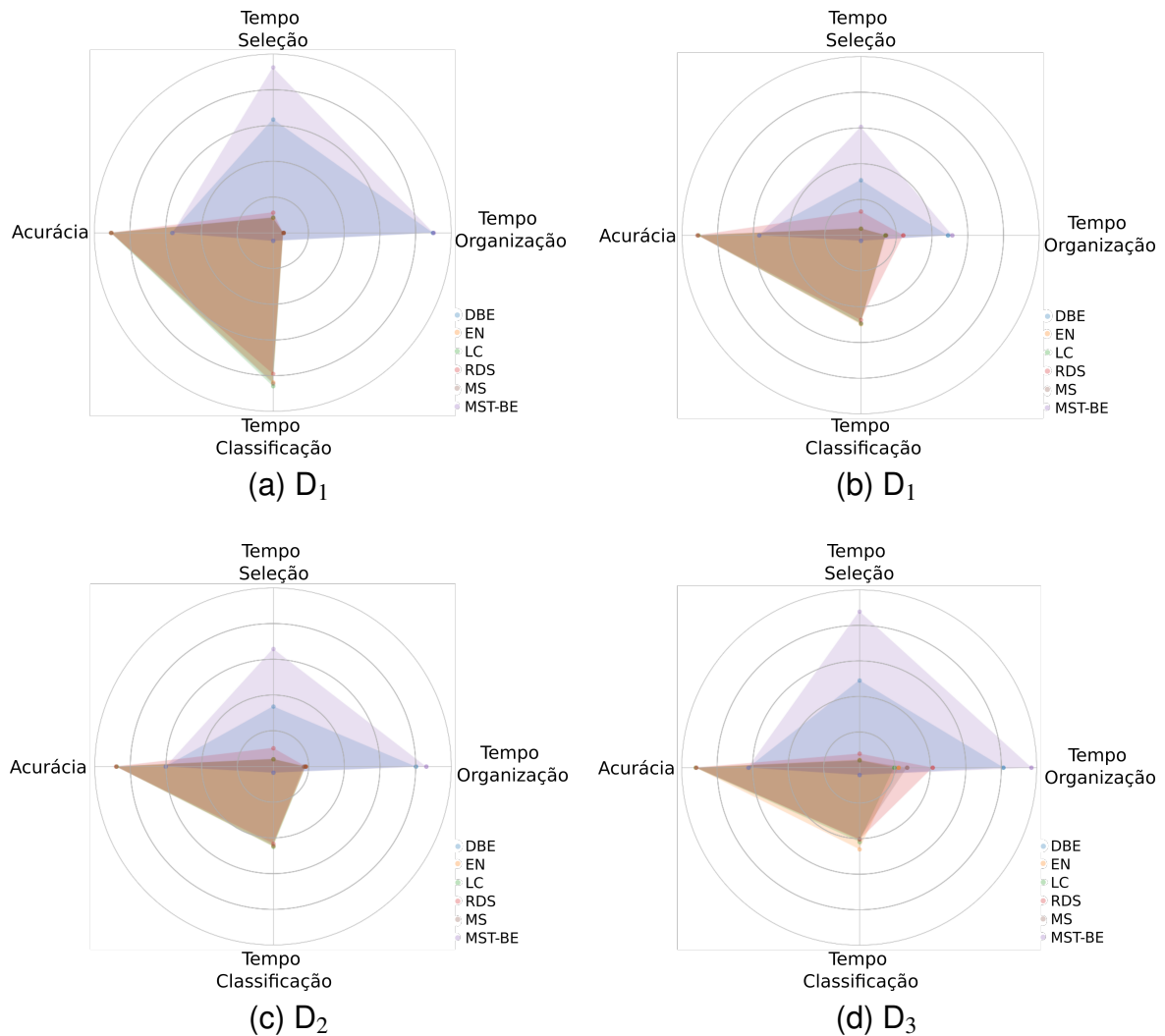


**Figura 14: Resultados considerando a abordagem de aprendizado ativo. Acurácias a cada porcentagem de amostras rotuladas obtidas pelas estratégias de aprendizado ativo DBE, EN, LC, MS, MST-BE, RAND e RDS com o classificador RF para os conjuntos de dados e respectivas configurações extrator-classificador. (a)  $D_1$ , ACC-RF, (b)  $D_1$ , BIC-RF, (c)  $D_2$ , BIC-RF e (d)  $D_3$ , BIC-RF**

Para o conjunto  $D_2$ , a abordagem de aprendizado ativo MS também obteve resultados superiores (acurácia de  $89.88\% \pm 0.90$  com apenas 43% de amostras) em relação à abordagem supervisionada tradicional (a qual atingiu  $87.43\% \pm 0.42$  de acurácia requerendo 100% do conjunto rotulado).

Utilizando o conjunto  $D_3$  (Figura 14(d)), a estratégia de seleção MS possibilitou uma redução de 65% na quantidade de amostras rotuladas necessárias para o treinamento do classificador, em comparação com a abordagem de aprendizado supervisionado tradicional. A estratégia MS alcançou uma acurácia de 93.24% com apenas 35% (11.494 amostras) do conjunto de treinamento. Enquanto que, na abordagem supervisionada tradicional, o classificador atingiu acurácias de 92.13% sendo necessário 100% (32.242 amostras) do conjunto de treinamento.

Além das acurácias, outras métricas foram consideradas para avaliar o desempenho das estratégias de seleção, a cada porcentagem de amostras, tais



**Figura 15: Resultados considerando a abordagem de aprendizado ativo. Acurácias e tempos de organização, seleção e classificação obtidas pelas estratégias de aprendizado ativo DBE, EN, LC, MS, MST-BE e RDS, para os conjuntos de dados e respectivas configurações extrator-classificador. (a)  $D_1$ , ACC-RF, (b)  $D_1$ , BIC-RF, (c)  $D_2$ , BIC-RF e (d)  $D_3$ , BIC-RF.**

como tempos para seleção, organização e classificação (Figura 15). Percebe-se que, em geral, as estratégias (EN, LC e MS) requerem menos tempos para seleção das amostras a cada iteração do aprendizado, por exemplo, em média 0.27, 0.25 e 0.24 segundos, respectivamente, considerando o conjunto  $D_3$  (Figura 15(d)). Diferentemente, as estratégias DBE, MST-BE e RDS necessitam de tempos maiores para seleção, para todos os conjuntos, uma vez que realizam o pré-processamento, envolvendo os processos de redução e de organização das amostras baseadas em grafos.

## 5 CONSIDERAÇÕES FINAIS

Com base nos trabalhos (Ayan; Ünver, 2018; PEREZ et al., 2018; Mohamed; El-Behaidy, 2019; BISLA et al., 2019; Rashid et al., 2019; BISSOTO et al., 2019; MANGAROLIYA; SHAH, 2020; Nasr-Esfahani et al., 2016; Khan et al., 2019; CODELLA et al., 2015; GESSERT et al., 2018; de Oliveira Lima et al., 2019; Moldovan, 2019; Pai; Giridharan, 2019; Emara et al., 2019; REZVANTALAB et al., 2018; CHATURVEDI et al., 2020; ESTEVA et al., 2017; KAWAHARA; HAMARNEH, 2016; Bassi; Gomekar, 2019) encontrados na literatura, envolvendo o contexto de lesão de pele, foi possível identificar que a partir do ano de 2016, há uma tendência das pesquisas utilizarem abordagens de aprendizado profundo.

Dentre tais abordagens, aquelas que apresentam propostas de arquiteturas próprias, treinadas do zero, atingiram resultados significativos. No entanto, os mesmos podem ser melhorados com a proposta de arquiteturas mais elaboradas e mais profundas, além de utilizar uma quantidade maior de amostras.

Abordagens que consideram arquiteturas já consolidadas, treinadas em massivas quantidades de dados como a *ImageNet* e por meio de transferência de aprendizado, ajuste fino ou como extratores de características apresentaram melhores resultados. Além disso, a geração de imagens sintéticas proposta pelas abordagens para o aumento de dados também alcançaram melhores resultados em relação ao uso de pequenas quantidades de imagens no processo de treinamento do classificador.

Existem alguns esforços com a proposta de métodos eficazes de segmentação de imagens de lesão de pele. No entanto, nos trabalhos da literatura em que o foco da pesquisa consiste no processo de classificação, tais abordagens de segmentação são pouco exploradas.

Apesar dos resultados significativos observados, as arquiteturas de aprendizado profundo são, atualmente o estado da arte para visão computacional. Entretanto, uma desvantagem na utilização das mesmas denota-se pela exigência de elevadas quantidades de amostras anotadas para obterem uma boa capacidade de generalização. Quanto aos esforços relacionados às abordagens de aprendizado ativo neste contexto, os mesmos são ínfimos.

Nesse sentido, o presente trabalho apresenta contribuições significativas para o contexto de lesão de pele. Inicialmente, foi explorada a abordagem de aprendizado supervisionada tradicional, avaliando diferentes cenários e configurações de extratores (tradicionais e profundos) e classificadores. Também foi avaliada a aplicação de segmentação nas imagens a partir da perspectiva de segmentação totalmente automatizada e segmentação com revisão de especialista. Foi possível identificar, de forma geral, que o extrator BIC obteve os melhores resultados em extrair padrões das imagens de lesão de pele, nos 3 conjuntos de dados avaliados. Na sequência, observou-se que a abordagem de aprendizado ativo se destaca em conseguir alcançar resultados equivalentes e/ou superiores com uma menor quantidade de amostras. Para trabalhos futuros, pretende-se desenvolver uma nova técnica de aprendizado ativo, explorando critérios de seleção como: diversidade, incerteza e representatividade.

## REFERÊNCIAS

ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; GOODFELLOW, I.; HARP, A.; IRVING, G.; ISARD, M.; JIA, Y.; JOZEFOWICZ, R.; KAISER, L.; KUDLUR, M.; LEVENBERG, J.; MANÉ, D.; MONGA, R.; MOORE, S.; MURRAY, D.; OLAH, C.; SCHUSTER, M.; SHLENS, J.; STEINER, B.; SUTSKEVER, I.; TALWAR, K.; TUCKER, P.; VANHOUCHE, V.; VASUDEVAN, V.; VIÉGAS, F.; VINYALS, O.; WARDEN, P.; WATTENBERG, M.; WICKE, M.; YU, Y.; ZHENG, X. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.

Alajlan, N.; Pasolli, E.; Melgani, F.; Franzoso, A. Large-scale image classification using active learning. **IEEE Geoscience and Remote Sensing Letters**, v. 11, n. 1, p. 259–263, 2014.

Ayan, E.; Ünver, H. M. Data augmentation importance for classification of skin lesions via deep learning. In: **2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)**. [S.l.: s.n.], 2018. p. 1–4.

Barata, C.; Ruela, M.; Francisco, M.; Mendonça, T.; Marques, J. S. Two systems for the detection of melanomas in dermoscopy images using texture and color features. **IEEE Systems Journal**, v. 8, n. 3, p. 965–979, set. 2014.

Bassi, S.; Gomekar, A. Deep learning diagnosis of pigmented skin lesions. In: **2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)**. [S.l.: s.n.], 2019. p. 1–6.

BIANCO, S.; CADENE, R.; CELONA, L.; NAPOLETANO, P. Benchmark analysis of representative deep neural network architectures. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 6, p. 64270–64277, 2018. ISSN 2169-3536. Disponível em: <<http://dx.doi.org/10.1109/ACCESS.2018.2877890>>.

BINDER, M.; SCHWARZ, M.; WINKLER, A.; STEINER, A.; KAIDER, A.; WOLFF, K.; PEHAMBERGER, H. Epiluminescence microscopy: a useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists. **Archives of dermatology**, American Medical Association, v. 131, n. 3, p. 286–291, 1995.

BISLA, D.; CHOROMANSKA, A.; STEIN, J. A.; POLSKY, D.; BERMAN, R. S. Skin lesion segmentation and classification with deep learning system. **CoRR**, abs/1902.06061, 2019. Disponível em: <<http://arxiv.org/abs/1902.06061>>.

BISSOTO, A.; PEREZ, F.; VALLE, E.; AVILA, S. Skin lesion synthesis with generative adversarial networks. **CoRR**, abs/1902.03253, 2019. Disponível em: <<http://arxiv.org/abs/1902.03253>>.

Bratkova, M.; Boulos, S.; Shirley, P. orgb: A practical opponent color space for computer graphics. **IEEE Computer Graphics and Applications**, v. 29, n. 1, p. 42–55, 2009.

BRAY, F.; ZNAOR, A.; CUEVA, P.; KORIR, A.; SWAMINATHAN, R.; ULLRICH, A.; WANG, S. A.; PARKIN, D. M. Planning and developing population-based cancer registration in low-or middle-income settings. 2021. Disponível em: <<https://www.iarc.fr/news-events/planning-and-developing-population-based-cancer-registration-in-low-and-middle-income-settings-as-an-e-book/>>.

BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001.

CANZIANI, A.; PASZKE, A.; CULURCIELLO, E. **An Analysis of Deep Neural Network Models for Practical Applications**. 2016.

Celebi, M. E.; Codella, N.; Halpern, A. Dermoscopy image analysis: Overview and future directions. **IEEE Journal of Biomedical and Health Informatics**, v. 23, n. 2, p. 474–478, 2019.

CHATURVEDI, S. S.; GUPTA, K.; PRASAD, P. S. Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using mobilenet. **Springer**, v. 1141, 2020. Disponível em: <[https://doi.org/10.1007/978-981-15-3383-9\\_15](https://doi.org/10.1007/978-981-15-3383-9_15)>.

CHATZICHRISTOFIS, S. A.; BOUTALIS, Y. S. Cedd: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In: SPRINGER. **International Conference on Computer Vision Systems**. [S.l.], 2008. p. 312–322.

CHATZICHRISTOFIS, S. A.; BOUTALIS, Y. S. Fcth: Fuzzy color and texture histogram—a low level feature for accurate image retrieval. In: IEEE. **Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on**. [S.l.], 2008. p. 191–196.

CODELLA, N.; CAI, J.; ABEDINI, M.; GARNAVI, R.; HALPERN, A.; SMITH, J. R. Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images. In: ZHOU, L.; WANG, L.; WANG, Q.; SHI, Y. (Ed.). **Machine Learning in Medical Imaging**. Cham: Springer International Publishing, 2015. p. 118–126. ISBN 978-3-319-24888-2.

CODELLA, N. C. F.; GUTMAN, D.; CELEBI, M. E.; HELBA, B.; MARCHETTI, M. A.; DUSZA, S. W.; KALLOO, A.; LIOPYRIS, K.; MISHRA, N.; KITTLER, H.; AL. et. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). **2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)**, IEEE, Apr 2018. Disponível em: <<http://dx.doi.org/10.1109/ISBI.2018.8363547>>.

CODELLA, N. C. F.; GUTMAN, D.; CELEBI, M. E.; HELBA, B.; MARCHETTI, M. A.; DUSZA, S. W.; KALLOO, A.; LIOPYRIS, K.; MISHRA, N.; KITTLER, H.; AL. et. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). **2018 IEEE 15th International**

**Symposium on Biomedical Imaging (ISBI 2018)**, IEEE, Apr 2018. Disponível em: <<http://dx.doi.org/10.1109/ISBI.2018.8363547>>.

COVER, T. M.; HART, P. E. Nearest neighbor pattern classification. **IEEE Trans. Inf. Theory**, v. 13, p. 21–27, 1967.

de Oliveira Lima, J. P.; Silva de Araújo Filho, L. C.; Santos da Siva, F.; Seródio Figueiredo, C. M. Pigmented dermatological lesions classification using convolutional neural networks ensemble mediated by multilayer perceptron network. **IEEE Latin America Transactions**, v. 17, n. 11, p. 1902–1908, 2019.

Demyanov, S.; Chakravorty, R.; Abedini, M.; Halpern, A.; Garnavi, R. Classification of dermoscopy patterns using deep convolutional neural networks. In: **2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)**. [S.l.: s.n.], 2016. p. 364–368.

DERMATOSCOPY. **Dermatoscopy**. 2020. Disponível em: <<http://dermoscopy.org>>.

Emara, T.; Afify, H. M.; Ismail, F. H.; Hassanien, A. E. A modified inception-v4 for imbalanced skin cancer classification dataset. In: **2019 14th International Conference on Computer Engineering and Systems (ICCES)**. [S.l.: s.n.], 2019. p. 28–33.

ESTEVA, A.; KUPREL, B.; NOVOA, R. A.; KO, J.; SEETTER, S. M.; BLAU, H. M.; THRUN, S. Dermatologist-level classification of skin cancer with deep neural networks. In: . [s.n.], 2017. v. 542. Disponível em: <<https://doi.org/10.1038/nature21056>>.

Ganster, H.; Pinz, P.; Rohrer, R.; Wildling, E.; Binder, M.; Kittler, H. Automated melanoma recognition. **IEEE Transactions on Medical Imaging**, v. 20, n. 3, p. 233–239, March 2001.

GESSERT, N.; SENTKER, T.; MADESTA, F.; SCHMITZ, R.; KNIEP, H.; BALTRUSCHAT, I.; WERNER, R.; SCHLAEFER, A. **Skin Lesion Diagnosis using Ensembles, Unscaled Multi-Crop Evaluation and Loss Weighting**. 2018.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning (Adaptive Computation and Machine Learning series)**. [S.l.]: The MIT Press, 2016. ISBN 0262035618.

GRAF, F. **JFeatureLib v1.6.3**. out. 2015. Disponível em: <<http://dx.doi.org/10.5281/zenodo.31793>>.

GUO, Z.; ZHANG, L.; ZHANG, D. Rotation invariant texture classification using LBP variance (LBPV) with global matching. **PR**, Elsevier, v. 43, n. 3, p. 706–719, 2010.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, SMC-3, n. 6, p. 610–621, Nov 1973. ISSN 0018-9472.



HARTIGAN, J. A.; WONG, M. A. Algorithm as 136: A k-means clustering algorithm. **Journal of the Royal Statistical Society. Series C (Applied Statistics)**, [Wiley, Royal Statistical Society], v. 28, n. 1, p. 100–108, 1979. ISSN 00359254, 14679876. Disponível em: <<http://www.jstor.org/stable/2346830>>.

HE, K.; ZHANG, X.; REN, S.; SUN, J. **Deep Residual Learning for Image Recognition**. 2015.

HEARST, M.; DUMAIS, S.; OSMAN, E.; PLATT, J.; SCHOLKOPF, B. Support vector machines. **Intelligent Systems and their Applications, IEEE**, v. 13, p. 18 – 28, 08 1998.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. **MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications**. 2017.

HUANG, G.; LIU, Z.; MAATEN, L. van der; WEINBERGER, K. Q. **Densely Connected Convolutional Networks**. 2016.

HUANG, J.; KUMAR, S. R.; MITRA, M.; ZHU, W.-J.; ZABIH, R. Image indexing using color correlograms. In: IEEE. **Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on**. [S.l.], 1997. p. 762–768.

Huang, S.; Jin, R.; Zhou, Z. Active learning by querying informative and representative examples. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 36, n. 10, p. 1936–1949, 2014.

INCA, I. N. D. C. **ABC do Câncer - Abordagens Básicas para o Controle do Câncer**. 2020. Disponível em: <<https://www.inca.gov.br/publicacoes/livros/abc-do-cancer-abordagens-basicas-para-o-controle-do-cancer>>.

INCA, I. N. D. C. **Estimativa 2020: incidência de câncer no Brasil**. 2020. Disponível em: <<https://www.inca.gov.br/publicacoes/livros/estimativa-2020-incidencia-de-cancer-no-brasil>>.

INSTITUTE, N. C. **What Is Cancer?** out. 2019. <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>. Acessado em 31/10/2019.

KAWAHARA, J.; HAMARNEH, G. Multi-resolution-tract cnn with hybrid pretrained and skin-lesion trained layers. **Springer**, v. 10019, 2016.

Khan, M. A.; Javed, M. Y.; Sharif, M.; Saba, T.; Rehman, A. Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. In: **2019 International Conference on Computer and Information Sciences (ICCIS)**. [S.l.: s.n.], 2019. p. 1–7.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGESS, C. J. C.; BOTTOU, L.; WEINBERGER, K. Q. (Ed.). **Advances in Neural Information Processing Systems 25**. Curran Associates, Inc., 2012. p. 1097–1105. Disponível em: <<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>.

Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, nov. 1998. ISSN 1558-2256.

LEWIS, D. D.; CATLETT, J. Heterogeneous uncertainty sampling for supervised learning. In: COHEN, W. W.; HIRSH, H. (Ed.). **Machine Learning Proceedings 1994**. San Francisco (CA): Morgan Kaufmann, 1994. p. 148 – 156. ISBN 978-1-55860-335-6. Disponível em: <<http://www.sciencedirect.com/science/article/pii/B978155860335650026X>>.

MANGAROLIYA, K.; SHAH, M. **Dermatologist vs Neural Network**. 2020.

Mikołajczyk, A.; Grochowski, M. Data augmentation for improving deep learning in image classification problem. In: **2018 International Interdisciplinary PhD Workshop (IIPhDW)**. [S.l.: s.n.], 2018. p. 117–122.

Mohamed, E. H.; El-Behaidy, W. H. Enhanced skin lesions classification using deep convolutional networks. In: **2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS)**. [S.l.: s.n.], 2019. p. 180–188.

Moldovan, D. Transfer learning based method for two-step skin cancer images classification. In: **2019 E-Health and Bioengineering Conference (EHB)**. [S.l.: s.n.], 2019. p. 1–4.

NACHBAR, F.; STOLZ, W.; MERKLE, T.; COGNETTA, A. B.; VOGT, T.; LANDTHALER, M.; BILEK, P.; BRAUN-FALCO, O.; PLEWIG, G. The abcd rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. **Journal of the American Academy of Dermatology**, v. 30, n. 4, p. 551 – 559, 1994. ISSN 0190-9622. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0190962294700613>>.

Nasr-Esfahani, E.; Samavi, S.; Karimi, N.; Soroushmehr, S. M. R.; Jafari, M. H.; Ward, K.; Najarian, K. Melanoma detection by analysis of clinical images using convolutional neural network. In: **2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. [S.l.: s.n.], 2016. p. 1373–1376.

O'HARA, S.; DRAPER, B. A. Introduction to the bag of features paradigm for image classification and retrieval. **CoRR**, abs/1101.3354, 2011. Disponível em: <<http://arxiv.org/abs/1101.3354>>.

Pai, K.; Giridharan, A. Convolutional neural networks for classifying skin lesions. In: **TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)**. [S.l.: s.n.], 2019. p. 1794–1796.

PATHAN, S.; PRABHU, K.; SIDDALINGASWAMY, P. Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review. **Biomedical Signal Processing and Control**, Elsevier BV, v. 39, p. 237–262, jan. 2018. ISSN 1746-8094.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.;

DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEREZ, F.; VASCONCELOS, C.; AVILA, S.; VALLE, E. Data augmentation for skin lesion analysis. **CoRR**, abs/1809.01442, 2018. Disponível em: <<http://arxiv.org/abs/1809.01442>>.

Pham, T. C.; Tran, G. S.; Nghiem, T. P.; Doucet, A.; Luong, C. M.; Hoang, V. A comparative study for classification of skin cancer. In: **2019 International Conference on System Science and Engineering (ICSSE)**. [S.l.: s.n.], 2019. p. 267–272.

Rashid, H.; Tanveer, M. A.; Aqeel Khan, H. Skin lesion classification using gan based data augmentation. In: **2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. [S.l.: s.n.], 2019. p. 916–919.

RAWAT, W.; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural computation**, v. 29, p. 2352–2449, set. 2017. ISSN 1530-888X.

REZVANTALAB, A.; SAFIGHOLI, H.; KARIMIJESHNI, S. Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. **CoRR**, abs/1810.10348, 2018. Disponível em: <<http://arxiv.org/abs/1810.10348>>.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. **CoRR**, abs/1505.04597, 2015. Disponível em: <<http://arxiv.org/abs/1505.04597>>.

RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M. S.; BERG, A. C.; LI, F. Imagenet large scale visual recognition challenge. **CoRR**, abs/1409.0575, 2014. Disponível em: <<http://arxiv.org/abs/1409.0575>>.

RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision (IJCV)**, v. 115, n. 3, p. 211–252, 2015.

SAITO, P.; REZENDE, P.; FALCAO, A.; SUZUKI, C. N.; GOMES, J. F. Reduction and organization selector selection and classification non-annotated dataset annotated dataset training classifier user learning cycle. In: . [S.l.: s.n.], 2012.

SAITO, P. T.; REZENDE], P. J. [de; FALCãO, A. X.; SUZUKI, C. T.; GOMES, J. F. An active learning paradigm based on a priori data reduction and organization. **Expert Systems with Applications**, v. 41, n. 14, p. 6086 – 6097, 2014. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417414002012>>.

SAITO, P. T.; SUZUKI, C. T.; GOMES, J. F.; REZENDE, P. J. de; aO, A. X. F. Robust active learning for the diagnosis of parasites. **Pattern Recogn.**, Elsevier Science

Inc., USA, v. 48, n. 11, p. 3572–3583, nov. 2015. ISSN 0031-3203. Disponível em: <<https://doi.org/10.1016/j.patcog.2015.05.020>>.

SAITO, P. T. M.; REZENDE, P. J. de; FALCAO, A. X.; SUZUKI, C. T. N.; GOMES, J. F. A data reduction and organization approach for efficient image annotation. In: **Proceedings of the 28th Annual ACM Symposium on Applied Computing**. New York, NY, USA: Association for Computing Machinery, 2013. (SAC '13), p. 53–57. ISBN 9781450316569. Disponível em: <<https://doi.org/10.1145/2480362.2480374>>.

SCHEFFER, T.; DECOMAIN, C.; WROBEL, S. Active hidden markov models for information extraction. In: **Proceedings of the 4th International Conference on Advances in Intelligent Data Analysis**. Berlin, Heidelberg: Springer-Verlag, 2001. (IDA '01), p. 309–318. ISBN 3540425810.

SETTLES, B. Active learning literature survey. n. 1648, 2009. Disponível em: <<http://burrsettles.com/pub/settles.activelearning.pdf>>.

SETTLES, B. **Active Learning**. [S.l.]: Morgan and Claypool Publishers, 2012. (Synthesis Lectures on Artificial Intelligence and Machine Learning).

SHI, X.; DOU, Q.; XUE, C.; QIN, J.; CHEN, H.; HENG, P.-A. **An Active Learning Approach for Reducing Annotation Cost in Skin Lesion Analysis**. 2019.

SIMONYAN, K.; ZISSERMAN, A. **Very Deep Convolutional Networks for Large-Scale Image Recognition**. 2014.

Situ, N.; Yuan, X.; Chen, J.; Zouridakis, G. Malignant melanoma detection by bag-of-features classification. In: **2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. [S.l.: s.n.], 2008. p. 3110–3113.

SMITH, J. R.; CHANG, S.-F. Local color and texture extraction and spatial query. In: **Intl. Conference on Image Processing**. [S.l.: s.n.], 1996. v. 3, p. 1011–1014.

STEHLING, R. O.; NASCIMENTO, M. A.; FALCÃO, A. X. A compact and efficient image retrieval approach based on border/interior pixel classification. In: **Intl. Conference on Information and Knowledge Management**. [S.l.: s.n.], 2002. p. 102–109.

STRICKER, M. A.; ORENGO, M. Similarity of color images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Storage and Retrieval for Image and Video Databases III**. [S.l.], 1995. v. 2420, p. 381–393.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. **Going Deeper with Convolutions**. 2014.

SZEGEDY, C.; VANHOUCHE, V.; IOFFE, S.; SHLENS, J.; WOJNA, Z. Rethinking the inception architecture for computer vision. **CoRR**, abs/1512.00567, 2015. Disponível em: <<http://arxiv.org/abs/1512.00567>>.

TAUD, H.; MAS, J. Multilayer perceptron (mlp). In: **Geomatic Approaches for Modeling Land Change Scenarios**. [S.l.]: Springer, 2018. p. 451–455.

Tkalcic, M.; Tasic, J. F. Colour spaces: perceptual, historical and applicational background. In: **The IEEE Region 8 EUROCON 2003. Computer as a Tool**. [S.l.: s.n.], 2003. v. 1, p. 304–308 vol.1.

TSCHANDL, P. **HAM10000 Lesion Segmentations**. 2020. Disponível em: <<https://www.kaggle.com/tschandl/ham10000-lesion-segmentations>>.

TSCHANDL, P.; ROSENDAHL, C.; KITTLER, H. The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions. **CoRR**, abs/1803.10417, 2018. Disponível em: <<http://arxiv.org/abs/1803.10417>>.

TSCHANDL, P.; SINZ, C.; KITTLER, H. Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. **Computers in Biology and Medicine**, v. 104, p. 111–116, 2019. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S001048251830372X>>.

BoF BIC CNN DBE EN GHI PHOG GAN HOG HSV HSI INCA CIE  $k$ -NN LC  
LR MS MLP MST-BE NB OCS PPGI RBE RF RGB SIFT SVM UC

