

UNIVERSIDADE FEDERAL DE VIÇOSA  
*CAMPUS* DE RIO PARANAÍBA  
SISTEMAS DE INFORMAÇÃO

VICTOR JÚNIO ALCÂNTARA CARDOSO

**CLASSIFICAÇÃO DE ANEMIA FALCIFORME EM  
IMAGENS DE MICROSCOPIA UTILIZANDO REDES  
NEURAIS CONVOLUCIONAIS COMO EXTRATORES DE  
CARACTERÍSTICAS E CLASSIFICADORES  
CONVENCIONAIS**

RIO PARANAÍBA

2021

VICTOR JÚNIO ALCÂNTARA CARDOSO

CLASSIFICAÇÃO DE ANEMIA FALCIFORME EM IMAGENS DE  
MICROSCOPIA UTILIZANDO REDES NEURAI  
CONVOLUCIONAIS COMO EXTRATORES DE CARACTERÍSTICAS  
E CLASSIFICADORES CONVENCIONAIS

Monografia apresentada à Universidade Federal de Viçosa como parte das exigências para a obtenção do título de bacharel em Sistemas de Informação.

Orientador: Prof. Dr. João Fernando Mari

Coorientador: Prof<sup>ª</sup>. Me. Larissa Ferreira Rodrigues

RIO PARANAÍBA

2021

# Resumo

Anemia falciforme é uma doença que altera o formato dos eritrócitos, células responsáveis pelo transporte de oxigênio no corpo humano e pode ser identificada por meio da análise microscópica das células. A computação aplicada na medicina permite mais eficiência e menor custo no diagnóstico e no tratamento de doenças. Dessa forma, este trabalho tem como objetivo aplicar técnicas de visão computacional para classificar doença falciforme em imagens de microscopia utilizando os algoritmos *SVM* e Naive Bayes, a partir de características extraídas por três Redes Neurais Convolucionais (CNNs): ResNet, DenseNet e MobileNet. Foram realizados experimentos para comparar a performance de cada *CNN* como extrator de características e como classificador. Além disso, foi analisado o impacto das imagens segmentadas na classificação. Todos os experimentos foram validados por meio da validação cruzada *5-fold* nos conjuntos de treinamento e teste. Os resultados apontam que utilizar as imagens de eritrócitos segmentados e as características extraídas das *CNNs* com *SVM* linear tende a ser a melhor estratégia quando o custo computacional é importante, alcançando 96,80% de acurácia.

**Palavras-chaves:** anemia falciforme, classificação, aprendizado profundo, extração de características.

# Abstract

Sickle cell anemia is a disease that modifies the shape of erythrocytes, cells responsible for the transport of oxygen in the human body, and can be identified through microscopic analysis of the cells. Computing applied in medicine allows more efficiency and lower cost in the diagnosis and treatment of diseases. Thus, this work proposes to apply computer vision techniques to classify sickle cell disease in microscopy images using SVM and Naive Bayes algorithms, based on features extracted from three Convolutional Neural Networks (CNNs): ResNet, DenseNet and MobileNet. Experiments were performed to compare the performance of each CNN as a feature extractor and classifier. Also, the impact of the segmented images on the classification was analyzed. All experiments were validated using five-fold cross-validation over the training and test sets. Results point that using segmented images and features from the CNN with linear SVM tends to be the best performing strategy when the computational cost is important, achieving 96.80% of accuracy.

**Key-words:** sickle cell disease, classification, deep learning, feature extraction.

# Lista de ilustrações

|   |    |
|---|----|
| Figura 1 – Comparação entre células normais e falciformes. . . . .                                  | 9  |
| Figura 2 – Representação de uma imagem digital. . . . .   | 10 |
| Figura 3 – Representação de uma operação de convolução. . . . .                                     | 11 |
| Figura 4 – Representação da nova dimensão adicionada pelo <i>SVM</i> . . . . .                      | 12 |
| Figura 5 – Funcionamento da camada convolucional. . . . .   | 15 |
| Figura 6 – Comparação entre uma rede simples e uma rede residual. . . . .                           | 16 |
| Figura 7 – Representação de um bloco denso de cinco camadas. . . . .                                | 16 |
| Figura 8 – Representação da operação de convolução separável em profundidade. .                     | 17 |
| Figura 9 – Etapas do trabalho. . . . .  | 21 |
| Figura 10 – Conjunto de imagens . . . . .   | 21 |
| Figura 11 – Particionamento de imagens de treino e teste com <i>K-Fold</i> . . . . .                | 23 |
| Figura 12 – Exemplo de mapas de características extraídos pela DenseNet-169. . .                    | 26 |
| Figura 13 – Comparação entre as estratégias avaliadas ao utilizar imagens originais.                | 27 |
| Figura 14 – Comparação entre as estratégias avaliadas ao utilizar imagens segmen-<br>tadas. . . . . | 27 |
| Figura 15 – Tempo de execução dos melhores resultados obtidos. . . . .                              | 28 |

# Lista de tabelas

|          |   |    |
|----------|---|----|
| Tabela 1 | – Matriz de Confusão . . . . .  | 18 |
| Tabela 2 | – Média dos <i>folds</i> para cada classificador ao utilizar imagens originais. .                     | 24 |
| Tabela 3 | – Média dos <i>folds</i> para cada classificador ao utilizar imagens segmentadas.                     | 25 |
| Tabela 4 | – Média dos <i>folds</i> para cada arquitetura <i>CNN</i> e considerando imagens originais. . . . .   | 25 |
| Tabela 5 | – Média dos <i>folds</i> para cada arquitetura <i>CNN</i> e considerando imagens segmentadas. . . . . | 26 |
| Tabela 6 | – Comparação com os trabalhos relacionados. . . . .   | 29 |

# Sumário

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introdução</b>   | <b>7</b>  |
| 1.1      | Objetivo Geral  | 7         |
| 1.2      | Objetivos Específicos   | 8         |
| <b>2</b> | <b>Referencial Teórico</b>  | <b>9</b>  |
| 2.1      | Anemia Falciforme   | 9         |
| 2.2      | Processamento e Análise de Imagens  | 10        |
| 2.2.1    | Imagens Digitais  | 10        |
| 2.2.2    | Convolução  | 11        |
| 2.3      | Classificadores   | 12        |
| 2.3.1    | Máquinas de Vetor Suporte - ( <i>Support Vector Machines</i> - <i>SVM</i> ) | 12        |
| 2.3.2    | Naive Bayes   | 13        |
| 2.4      | Rede Neurais Convolucionais   | 13        |
| 2.4.1    | Extração de Características   | 14        |
| 2.4.2    | Arquiteturas  | 15        |
| 2.4.2.1  | ResNet  | 15        |
| 2.4.2.2  | DenseNet-169  | 16        |
| 2.4.2.3  | MobileNet   | 17        |
| 2.5      | Avaliação de Desempenho   | 17        |
| <b>3</b> | <b>Trabalhos Relacionados</b>   | <b>19</b> |
| <b>4</b> | <b>Material e Métodos</b>   | <b>21</b> |
| 4.1      | Conjunto de Imagens   | 21        |
| 4.2      | Ferramentas Computacionais  | 22        |
| 4.3      | Experimentos  | 22        |
| 4.4      | Validação   | 23        |
| <b>5</b> | <b>Resultados e Discussões</b>  | <b>24</b> |
| 5.1      | <i>CNNs</i> como extrator de características                                | 24        |
| 5.2      | <i>CNNs</i> treinadas com transferência de aprendizado                      | 25        |
| 5.3      | Comparação entre as melhores estratégias                                    | 28        |
| 5.4      | Comparação com trabalhos relacionados                                       | 29        |
| <b>6</b> | <b>Conclusões</b>   | <b>30</b> |
|          | <b>Referências</b>  | <b>31</b> |

# 1 Introdução

A anemia falciforme é uma doença hereditária causada pela deformidade das células dos eritrócitos (glóbulos vermelhos) responsáveis pelo transporte de oxigênio no corpo. O indivíduo afetado, apresenta essas células em um formato alongado, parecido com uma foice, que dá origem ao nome da doença. O portador sofre com diversos sintomas como crises de dor pela obstrução de vasos sanguíneos, fadiga, maior propensão à infarto do miocárdio, acidente vascular cerebral, doença renal crônica, embolia pulmonar, infecções que podem causar outros riscos à saúde (HOFFMAN et al., 2013) (KATO et al., 2018).

A identificação rápida da doença falciforme é importante, pois o diagnóstico precoce aliado às medidas necessárias permite reduzir a mortalidade de 25% para 3% nos primeiros cinco anos de vida (KATO et al., 2018). Um dos procedimentos disponíveis para o diagnóstico precoce é o exame de eletroforese da hemoglobina. No Brasil, esse diagnóstico é realizado ainda na maternidade por meio do teste do pezinho (MELO-REIS et al., 2006). Entretanto, há diversas regiões emergentes onde esse tipo de diagnóstico não é realizado por falta de recursos e profissionais especializados. Por exemplo, em 2016, 100% dos hospitais do estado de Minas Gerais realizaram o teste do pezinho enquanto no estado do Amapá foram realizados apenas 55% (KATO et al., 2018).

As técnicas de processamento de imagens e aprendizado de máquina tem auxiliado na detecção de doenças e muitas vezes trazem soluções mais baratas e rápidas (WU; MERCHANT; CASTLEMAN, 2008). Portanto, a classificação automatizada da doença falciforme por meio das imagens de microscopia das células pode ser vantajosa para que esses testes sejam realizados com facilidade e em qualquer lugar.

Muitos algoritmos de classificação são utilizados nas tarefas de reconhecimento de imagens e têm aplicações em diversas áreas além de imagens médicas. Entretanto, esses algoritmos dependem de uma extração manual de características. Já as Redes Neurais Convolucionais (*Convolutional Neural Networks - CNN*) têm mostrado resultados significativos para a classificação de imagens, com uma vantagem sobre os classificadores convencionais: a própria *CNN* extrai as características, dispensando assim qualquer tipo de pré-processamento na imagem (GOODFELLOW; BENGIO; COURVILLE, 2016). Portanto, este trabalho propõe a mesclagem dessas tecnologias utilizando características aprendidas pela *CNN* em classificadores convencionais.

## 1.1 Objetivo Geral

O objetivo deste trabalho é construir um algoritmo capaz de classificar imagens microscópicas de eritrócitos em: saudáveis, com doença falciforme e outras deformações.



Além disso, investigar o desempenho dos classificadores tradicionais com as características extraídas por Redes Neurais Convolucionais.

## 1.2 Objetivos Específicos

- Explorar modelos *CNNs* como extratores de características não supervisionados para treinar classificadores convencionais.
- Selecionar as características mais relevantes para a tarefa de classificação.
- Comparar a performance dos classificadores convencionais ao utilizar as características extraídas pelas *CNNs*.
- Comparar a performance da classificação feita pela própria rede com a classificação realizada pelos classificadores convencionais.

## 2 Referencial Teórico

### 2.1 Anemia Falciforme

Na composição do sangue humano, existem as células chamadas de eritrócitos, também conhecido como hemácias. No interior da estrutura dessas células há a hemoglobina, uma proteína que é responsável pelo transporte de oxigênio pelo sistema circulatório do corpo humano (COSTA NICOLA CONRAN, 2013).

Os eritrócitos têm formato bicôncavo e possuem grande flexibilidade para realizar a passagem pelos vasos sanguíneos, entretanto podem ser acometidos por uma causa genética à uma deformidade em seu formato. Essa deformidade deixa as células alongadas e no formato de foice, diminuindo sua flexibilidade e, conseqüentemente, dificulta a passagem pelos vasos sanguíneos, prejudicando que o oxigênio chegue nos tecidos do corpo. Além disso, essas células deformadas têm um tempo de vida menor que as células normais, assim o portador da doença falciforme apresenta o número de eritrócitos sempre abaixo do normal (KATO et al., 2018). A Figura 1 ilustra eritrócitos normais e falciformes.

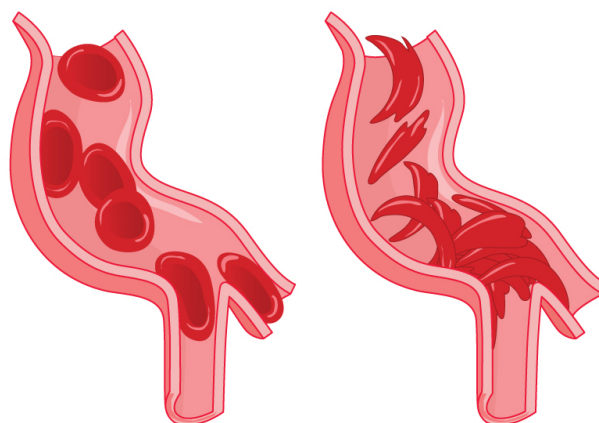


Figura 1 – Comparação entre células normais e falciformes.

Fonte: Children's Hospital Los Angeles (2017)

Por esses motivos, os principais sintomas da doença são: fadiga, crises de dores pela obstrução dos vasos sanguíneos, maior propensão à infartos, embolia pulmonar, infecções, entre outros (HOFFMAN et al., 2013).

Por meio do exame de eletroforese da hemoglobina é possível identificar modificações funcionais e estruturais relacionadas a síntese da hemoglobina e então proporcionar

o diagnóstico da doença falciforme. No Brasil esse exame é incluído na triagem neonatal conhecida como teste do pezinho, que é realizado na primeira semana de vida da criança. Porém, também é possível realizar uma análise morfológica das células por meio do hemograma, permitindo a visualização da deformidade (COSTA NICOLA CONRAN, 2013).

A anemia falciforme é uma doença incurável e é necessário que o paciente realize um acompanhamento médico durante toda a vida para controlar as crises de dor e normalizar a oxigenação do corpo (SERJEANT, 2013).

## 2.2 Processamento e Análise de Imagens

Nesta sessão são apresentados conceitos importantes de imagens digitais e a operação de convolução, que é utilizada para extração das características por meio das *CNNs*.

### 2.2.1 Imagens Digitais

Uma imagem pode ser definida como uma função bidimensional  $f(x, y)$  de intensidade da luz, tendo  $(x, y)$  como coordenadas no espaço e o valor de  $f$  em uma dada coordenada corresponde à intensidade de luz naquela posição (GONZALEZ; WOODS, 2006).

Portanto, uma imagem digital consiste na discretização desses valores, ou seja, um número finito de coordenadas espaciais e de valores de intensidades para que essa imagem possa ser representada em um computador. Considera-se então uma matriz cujas quantidades de linhas e colunas limitam a imagem no espaço, cada índice de linha e coluna representa um *pixel*, como mostrado na Figura 2, o menor elemento de uma imagem que contém o valor de intensidade daquela região, sendo limitado pelo número de bits que irá armazenar esse valor (GONZALEZ; WOODS, 2006).

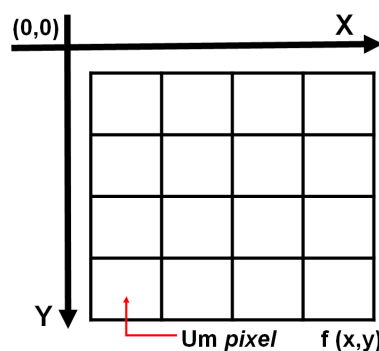


Figura 2 – Representação de uma imagem digital.

Fonte: Rodrigues (2018)

### 2.2.2 Convolução

A Convolução é uma operação matemática realizada *pixel a pixel* em uma imagem por meio de um filtro. O processo de convolução consiste em percorrer esse filtro ou máscara na extensão da imagem realizando a soma dos produtos em cada posição, esses resultados são os valores de uma nova imagem  $g(x, y)$ . Dada uma imagem  $f(x, y)$  e um filtro  $w(x, y)$  de tamanho  $m \times n$  podemos definir a convolução de acordo com a Equação 2.1.

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x - s, y - t), \quad (2.1)$$

em que  $a = \frac{(m-1)}{2}$  e  $b = \frac{(n-1)}{2}$ .

A Figura 3 ilustra uma operação de convolução. O tamanho e os valores dos filtros dependem dos objetivos do processamento da imagem, filtros chamados passa-alta realçam regiões de alta frequência da imagem como bordas e os filtros passa-baixa atenuam essas regiões, suavizando a imagem (GONZALEZ; WOODS, 2006). As Redes Neurais Convolucionais utilizam essa técnica para a extração das características como será discutido posteriormente na Seção 2.4.1.

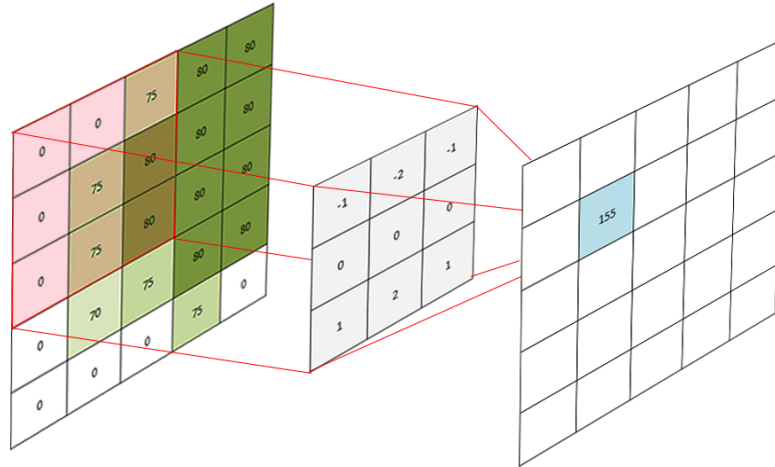


Figura 3 – Representação de uma operação de convolução.

## 2.3 Classificadores

A classificação é um dos campos da base de conhecimento em aprendizado de máquina que pode ser aplicada em imagens digitais, que consiste na atribuição de um significado àquela imagem (GONZALEZ; WOODS, 2006). Existem diversos algoritmos que buscam realizar essa tarefa por meio das informações adquiridas da imagem, nessa seção abordaremos dois deles: *Support Vector Machine* e Naive Bayes, os quais serão descritos à seguir.

### 2.3.1 Máquinas de Vetor Suporte - (*Support Vector Machines - SVM*)

O algoritmo de classificação denominado Máquinas de Vetor e Suporte (do inglês *Support Vector Machines - SVM*) (CORTES; VAPNIK, 1995) é capaz de realizar a separação de dados lineares e não lineares. Dados linearmente separáveis são aqueles que dispostas as entradas em um plano, é possível traçar uma reta capaz de separar os elementos das classes. O *SVM*, por sua vez, para resolver problemas não linearmente separáveis projeta os dados de entrada em dimensões suficientemente altas para calcular um hiperplano que consiga separar os dados como representado na Figura 4.

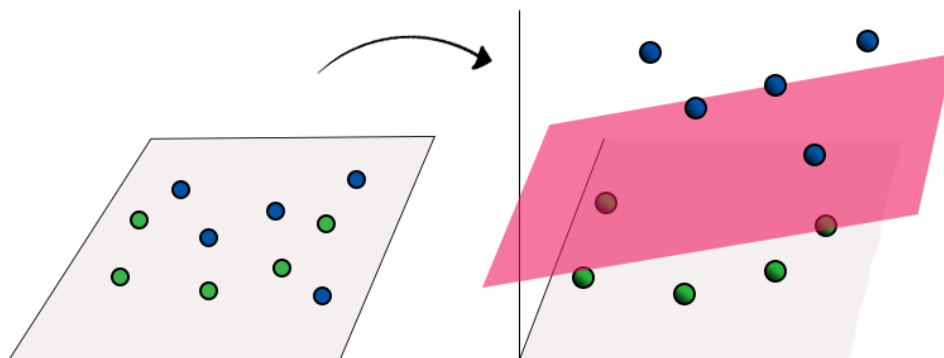


Figura 4 – Representação da nova dimensão adicionada pelo *SVM*.

Entretanto, pode haver infinitos hiperplanos possíveis capazes de separar os objetos, o *SVM* busca maximizar a margem de distância entre o hiperplano e os vetores de suporte, que são os objetos mais próximos ao hiperplano, mantendo uma distância máxima igual entre os vetores de cada classe. O *SVM* pode ter um maior custo computacional e conseqüentemente ser um algoritmo mais lento, porém tem boa capacidade de generalização e evita o *overfitting* dos dados de treinamento (HAN; KAMBER; PEI, 2011).

### 2.3.2 Naive Bayes

O Naive Bayes é um classificador probabilístico baseado no Teorema de Bayes, o qual classifica os objetos de acordo com o maior valor obtido por uma probabilidade condicional (DUDA; HART; STORK, 2000).

Seja  $X$  um objeto, com uma tupla de características, o qual será classificado, a probabilidade de que ele pertença à uma classe  $C$  é dada por  $P(C|X)$ . Por meio do Teorema de Bayes é gerada a Equação 2.2.

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (2.2)$$

Em que:

- $P(X|C)$  é a probabilidade do objeto ter determinadas características dado que ele pertence à classe  $C$ .
- $P(C)$  é a probabilidade a priori de qualquer objeto pertencer a classe  $C$ .
- $P(X)$  é a probabilidade a priori de um elemento no nosso conjunto de dados ter tais características.

Como  $P(X)$  é igual para todas as classes, será maximizado o valor de  $P(X|C)P(C)$ . Supondo que tenhamos  $m$  características no objeto  $X$ ,  $P(X|C)$  pode ser descrito como a Equação 2.3

$$P(X|C) = \prod_{k=1}^m P(X_k|C) \quad (2.3)$$

Caso  $X_k$  seja um atributo de valores contínuos, normalmente utiliza-se a Distribuição Gaussiana do conjunto de dados e a média e desvio padrão para estimar as probabilidades desses valores.

Considerando  $n$  classes o algoritmo classificará o elemento em uma classe  $C_i$  se e somente se:

$$P(X|C_i)P(C_i) > P(X|C_j)P(C_j) \text{ para } i \neq j, 1 \leq j \leq n \quad (2.4)$$

ou seja, o algoritmo classificará na classe onde o valor da probabilidade é máximo (HAN; KAMBER; PEI, 2011).

## 2.4 Rede Neurais Convolucionais

As Redes Neurais Convolucionais (do inglês *Convolutional Neural Networks* - *CNNs*) são modelos de Redes Neurais especializadas em processar dados com uma to-

pologia em grade, por exemplo, imagens que podem ser vistas como uma grade 2D de *pixels* (GOODFELLOW; BENGIO; COURVILLE, 2016). Esses modelos têm sido amplamente utilizados nas tarefas de reconhecimento e classificação de imagens e têm alcançado ótimos resultados, superando o estado da arte na maioria das vezes (SULTANA; SUFIAN; DUTTA, 2018).

O nome Redes Neurais Convolucionais deriva da principal operação realizada pelo algoritmo: a operação de convolução, discutida na Seção 2.2.2, que é aplicada nas imagens de entrada a fim de obter os mapas de características necessários para realizar a classificação (GOODFELLOW; BENGIO; COURVILLE, 2016).

Normalmente existem três blocos ou camadas principais presentes em uma *CNN*: camadas convolucionais, camadas de *pooling* e as camadas completamente conectadas. Dependendo da arquitetura pode não haver alguma dessas camadas ou ter outros módulos que substituam as funcionalidades de alguma dessas camadas básicas para funções específicas de cada modelo. O objetivo de cada camada é:

- Camada Convolucional: Realiza as operações de convolução e obtém os mapas de características.
- *Pooling*: Responsável por reduzir as características.
- Camada Completamente Conectada: As características extraídas são utilizadas como entrada para uma Rede de Neurônios para que seja realizada a classificação.

### 2.4.1 Extração de Características

Como dito anteriormente, a Camada Convolucional é responsável pela extração das características. Ela é composta por uma série de filtros que são aplicados em toda a imagem. Um filtro é uma matriz de pesos que indica o que a rede vai aprender. Sendo assim, após o processamento de uma imagem na camada convolucional de  $n$  filtros, há  $n$  mapas de características como saída, cada um obtido por meio de um filtro como ilustrado na Figura 5 (PONTI et al., 2017).

Os mapas de características normalmente passam por uma camada de *pooling* onde são reduzidos e, após uma série de convoluções, o tensor com os mapas de características é utilizado como entrada para uma camada completamente conectada semelhante ao *Multi-Layer Perceptron (MLP)* para a realização da classificação (PONTI et al., 2017). Nessa etapa, a estratégia desse trabalho consiste em não utilizar a camada completamente conectada das *CNNs* e sim aplicar os mapas de características como entrada para outros algoritmos de classificação.

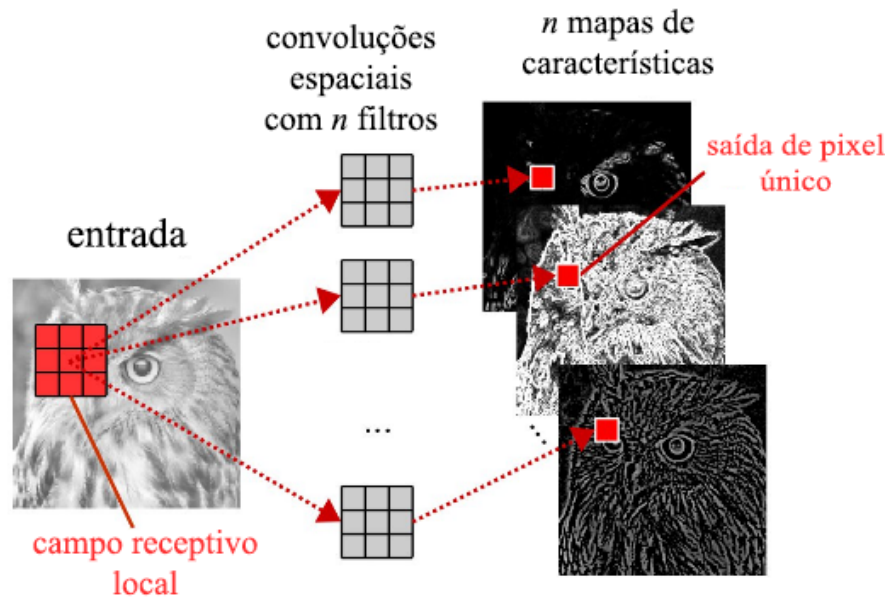


Figura 5 – Funcionamento da camada convolucional.

Fonte: Adaptado de Ponti et al. (2017)

## 2.4.2 Arquiteturas

### 2.4.2.1 ResNet

A primeira arquitetura considerada neste trabalho é a ResNet, ganhadora do *Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)* (RUSSAKOVSKY et al., 2015), uma competição de classificação de imagens realizada pelo *ImageNet* que conta com uma base de imagens com cerca de 14 milhões de imagens. A competição se resume em classificar imagens em 1000 classes diferentes e é realizada anualmente (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

A ResNet apresenta uma abordagem de blocos residuais, ou seja, quando a imagem passa por alguma camada a rede constrói um mapa residual, esses resíduos da saída de camadas anteriores são passados para camadas à frente a fim de manter algumas características por meio das conexões de atalho. Formalmente, pode-se dizer que, dado uma entrada  $x$ , ao passar por uma camada de pesos há uma saída do mapa subjacente  $H(x)$ , então define-se  $F(x)$  como sendo  $H(x) - x$ . Dessa forma é obtido o mapeamento original por meio de  $F(x) + x$  (HE et al., 2016). A Figura 6 mostra a arquitetura completa da ResNet.



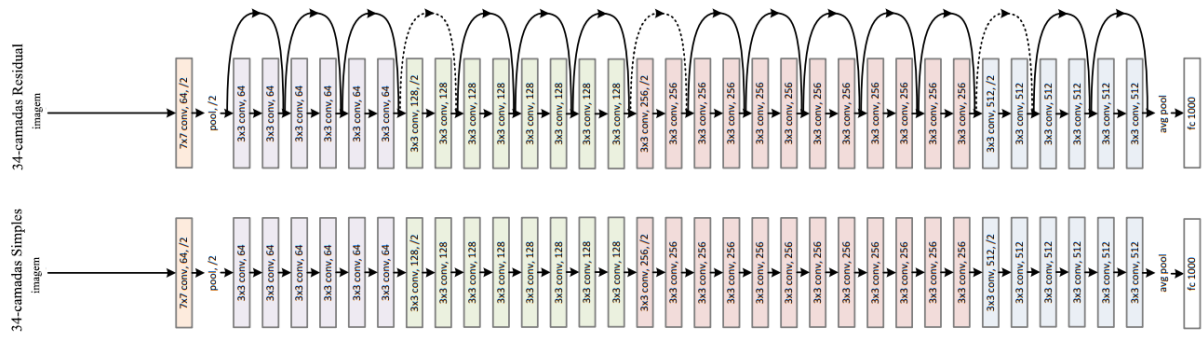


Figura 6 – Comparação entre uma rede simples e uma rede residual.

Fonte: Adaptado de He et al. (2016)

#### 2.4.2.2 DenseNet-169

A segunda arquitetura utilizada foi a DenseNet, proposta por Huang et al. (2017). A principal diferença da sua implementação é o fato de possuir camadas densas, ou seja, a saída de cada camada é interligada como entrada de todas as camadas subsequentes e não somente a camada imediatamente à frente. Segundo o autor, essa abordagem traz vantagens como o fortalecimento da propagação das características e sua reutilização. Uma vez que um mapa de características é extraído em uma camada todas as próximas o utilizarão como entrada e causa uma redução do número de parâmetros. A representação dessa arquitetura pode ser visualizada na Figura 7.

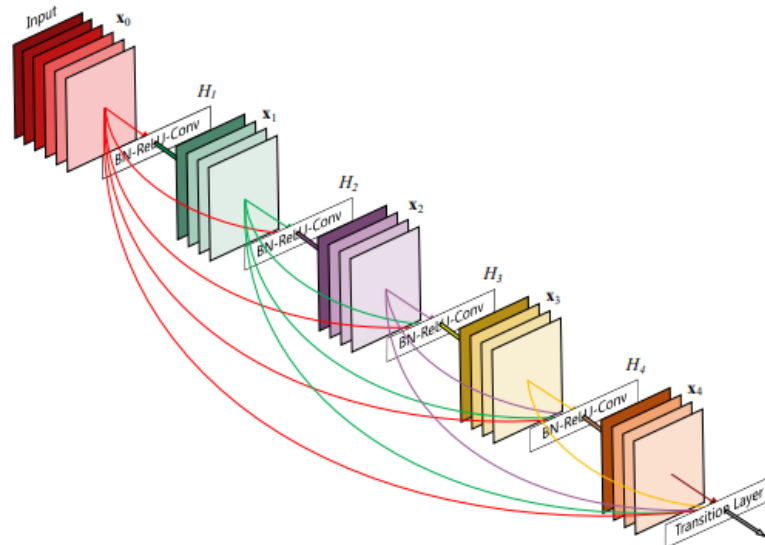


Figura 7 – Representação de um bloco denso de cinco camadas.

Fonte: Huang et al. (2017)

### 2.4.2.3 MobileNet

Esta rede desenvolvida por Howard et al. (2017) tem como objetivo criar uma arquitetura mais leve. Como o próprio nome sugere, a MobileNet é passível de ser utilizada em dispositivos móveis e sistemas embarcados. Para isso, sua arquitetura é baseada em técnicas de convoluções separáveis por profundidade, que reduz significativamente o número de parâmetros e consequentemente o custo computacional. Uma representação da operação de convolução separável em profundidade pode ser vista na Figura 8.

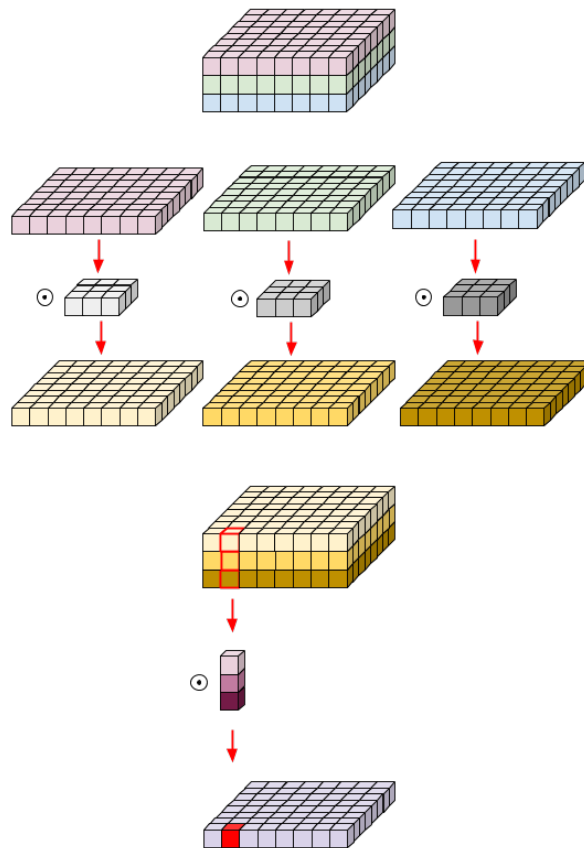


Figura 8 – Representação da operação de convolução separável em profundidade.

Fonte: Pandey (2018)

## 2.5 Avaliação de Desempenho

A performance da classificação foi avaliada por meio de uma matriz de confusão, responsável por mostrar as frequências de classificação para cada classe (COSTA; CESAR-JR., 2009). A Tabela 1 mostra a representação geral de uma matriz de confusão, em que são obtidos os valores:

- Verdadeiro Positivo: o objeto é classificado como pertencente à uma classe  $C$  e

realmente pertence.

- Verdadeiro Negativo: o objeto é classificado como não pertencente à uma classe  $C$  e realmente não pertence.
- Falso Positivo: o objeto é classificado em uma classe  $C$ , porém não pertence àquela classe.
- Falso Negativo: o objeto é classificado como não pertencente a uma classe  $C$ , porém pertence àquela classe.

Tabela 1 – Matriz de Confusão

|      |     | PREVISTO                 |                         |
|------|-----|--------------------------|-------------------------|
|      |     | Sim                      | Não                     |
| REAL | Sim | Verdadeiro Positivo (VP) | Falso Negativo(FN)      |
|      | Não | Falso Positivo(FP)       | Verdadeiro Negativo(VN) |

Representação geral de uma matriz de confusão.

À partir dos valores obtidos na matriz de confusão são realizados os cálculos das principais métricas para avaliar o desempenho da classificação em:

- Acurácia: Razão dos valores verdadeiros sobre todos os valores.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.5)$$

- Precisão: Razão das classificações corretas positivas sobre todas as amostras classificadas como positivo.

$$Precisão = \frac{VP}{VP + FP} \quad (2.6)$$

- *Recall* : Razão das amostras positivas classificadas corretamente sobre o total de amostras positivas.

$$Recall = \frac{VP}{VP + FN} \quad (2.7)$$

- *F1-Score*: Média ponderada da precisão e sensibilidade.

$$F1-Score = \frac{2VP}{2VP + FP + FN} \quad (2.8)$$

### 3 Trabalhos Relacionados

Neste capítulo, são discutidos os trabalhos que apresentam tópicos comuns com este e suas peculiaridades. Os trabalhos em questão utilizam a mesma base de dados do presente trabalho.

Gual-Arnau, Herold-García e Simó (2015) apresentaram a primeira proposta de classificação do conjunto de dados *erythrocytesIDB* e propuseram um método de extração de características com descritores morfológicos baseados na geometria integral e usaram o *K-Nearest Neighbors (KNN)* para realizar a classificação de eritrócitos. A acurácia alcançada foi de 96,16% utilizando o descritor da fórmula de Crofton.

Rodrigues, Naldi e Mari (2016) propuseram uma abordagem de segmentação das imagens, ou seja, a separação da célula normal, falciforme e com outras deformações do fundo da imagem. A partir disso, realizaram a extração das características morfológicas das mesmas considerando perímetro, área, circularidade, entre outras. Após, realizaram a seleção das características buscando aumentar a performance nos três classificadores utilizados: *Support Vector Machine (SVM)*, Naive Bayes e *K-Nearest Neighbors*. Foi obtido um resultado de 94,60% com o *SVM* e utilizando oito características morfológicas.

Faria, Rodrigues e Mari (2018) apresentaram uma simples e eficiente combinação dos descritores *SIFT* e *SURF*, empilhando os descritores de pontos de controle em uma única matriz, usou o *SVM* e o *Multi-Layer Perceptron (MLP)* para classificação. A abordagem proposta obteve 93,67% de acurácia com o *MLP*.

Silva, Rodrigues e Mari (2020) aplicaram o algoritmo de otimização Bayesiana para encontrar quais são as melhores estratégias de aumento de dados para o conjunto de imagens *erythrocytesIDB* e obtiveram uma acurácia de 92,54%.

A técnica de transferência de aprendizado foi aplicada por Alzubaidi et al. (2020) em três conjuntos de imagens diferentes. Os autores propuseram três modelos *CNNs* e utilizaram dois conjuntos de imagens de eritrócitos para treinar as *CNNs* propostas. Obtiveram resultados superiores 98% para as três *CNNs* propostas e considerando treinamento baseado em transferência de aprendizado, extração de características e classificação com *SVM*.

Paz-Soto et al. (2020) utilizaram funções integrais baseadas em geometria para extrair características do contorno das células do *dataset erythrocytesIDB* e utilizaram uma rede neural artificial como classificador, alcançando uma acurácia de 98,40%.

O presente trabalho se diferencia dos anteriormente citados ao propor uma classificação sem realizar uma extração manual das características, deixando essa função para as

Redes Neurais Convolucionais a fim de automatizar o processo e avaliar o desempenho com as imagens segmentadas que envolvem pré processamento. Também foi possível verificar o impacto das características obtidas pelas *CNNs* em comparação aos trabalhos anteriores que em sua maioria utilizaram características baseadas em descritores morfológicos.

## 4 Material e Métodos

O presente capítulo descreve o material e método necessários para realização do trabalho. As etapas são sumarizadas na Figura 9.

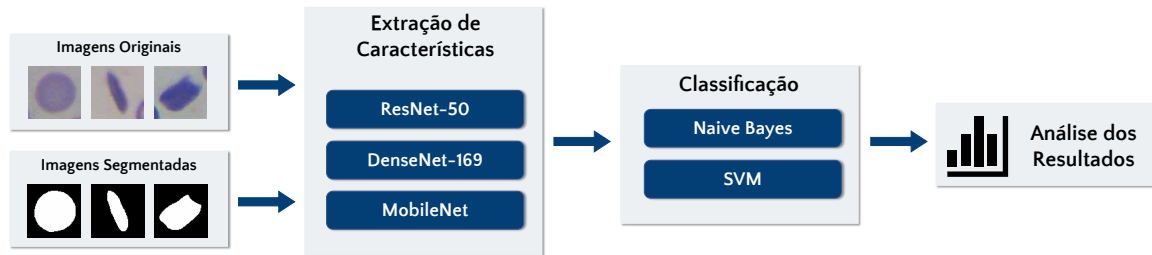


Figura 9 – Etapas do trabalho.

### 4.1 Conjunto de Imagens

Para a realização dos experimentos foi utilizada a base de dados *erythrocytesIDB*<sup>1</sup> disponibilizada publicamente pela *University of the Balearic Islands*. É composta por 626 imagens microscópicas de eritrócitos individuais divididos em 202 imagens de células normais, 211 imagens que apresentam deformidade causada pela doença falciforme e 213 imagens de células com outros tipos de deformidades. A Figura 10 (a) mostra exemplos de imagens para cada classe do conjunto de dados. Também foram utilizadas as imagens segmentadas utilizando o método proposto por Rodrigues, Naldi e Mari (2016) (Figura 10 (b)).

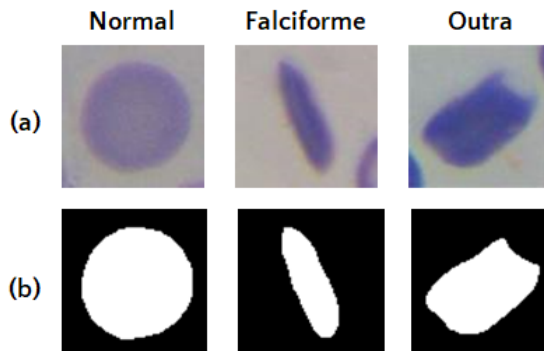


Figura 10 – (a) Exemplos de imagens disponíveis no conjunto de dados *erythrocytesIDB*.  
(b) Exemplos de imagens segmentadas.

<sup>1</sup> Disponível em: <<http://erythrocytesidb.uib.es>>

A técnica de aumento de dados foi aplicada nas imagens utilizadas no treinamento, ou seja, foram aplicadas modificações a fim de aumentar o número de imagens e melhorar o treinamento das *CNNs*. Como proposto por Silva, Rodrigues e Mari (2020) as estratégias escolhidas foram inversões horizontais e inversões verticais durante o treinamento.

## 4.2 Ferramentas Computacionais

Para o desenvolvimento do método proposto foi utilizada a biblioteca Keras (CHOLLET et al., 2015) (versão 2.3.1) e TensorFlow (ABADI et al., 2015) como *backend*. O Keras é um *framework* para aprendizado profundo escrito em *Python* e foi utilizado para a implementação das arquiteturas *CNNs* apresentadas na Seção 2.4.2.

Também foram utilizadas outras bibliotecas presentes no *Python*: o *scipy*, em que está presente o *numpy* para manipulação de *arrays* e o *matplotlib* para visualização dos dados. Além do *scikit-learn* para aprendizado de máquina.

Os experimentos foram realizados na plataforma em nuvem Google *Colab* que permite a utilização de máquinas para executar códigos *python* de forma gratuita. Segue abaixo as especificações da máquina utilizada:

- CPU Intel Xeon 2.20 GHz;
- 12 GB de memória RAM;
- GPU Tesla T4 16 GB;
- Sistema Operacional: Ubuntu 18.04.3 LTS.

## 4.3 Experimentos

Foram avaliadas três arquiteturas *CNNs*: ResNet, DenseNet e MobileNet. Elas foram responsáveis pela extração das características das imagens como explicado na Seção 2.4.1. Após, os mapas de características gerados por cada uma foi utilizado como entrada para os dois algoritmos de classificação: *SVM* e Naive Bayes (Seção 2.3). Esses algoritmos classificaram as imagens entre normal, falciforme e com outras deformidades. Este processo foi realizado com as imagens do conjunto de dados original e também com as imagens segmentadas. Por fim, foi realizada uma análise comparativa entre a classificação com algoritmos supervisionados e classificação com *CNNs* em termos de acurácia, precisão, *recall*, *F1-Score* e tempo de treinamento.

## 4.4 Validação

Para a avaliação geral dos resultados da classificação, foi utilizada a técnica de validação cruzada (KOHAVI, 1995) escolhida pelo fato de fornecer resultados mais confiáveis do que as demais técnicas de validação. Foi aplicado o método *K-fold*, representado na Figura 11, sendo os dados de treinamento divididos em  $k$  subconjuntos de tamanhos iguais e a cada interação, um subconjunto é utilizado para a realização do teste e todos os  $k - 1$  conjuntos são utilizados no treinamento. Neste trabalho foi considerado  $k = 5$ .

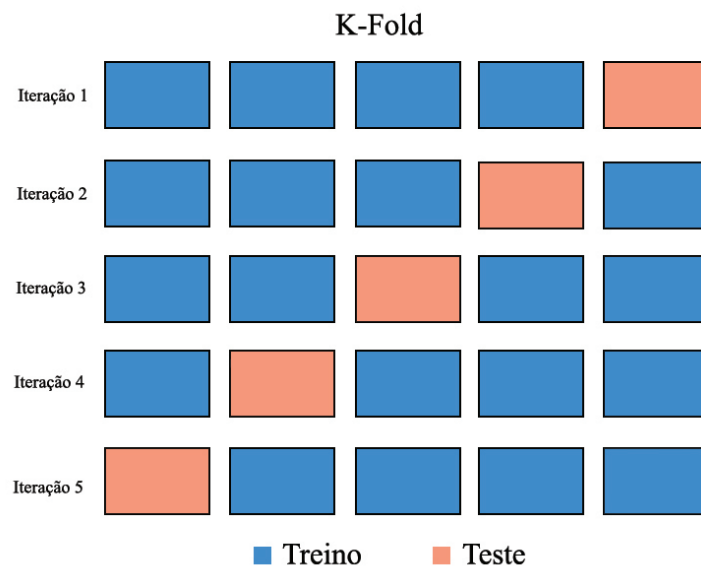


Figura 11 – Particionamento de imagens de treino e teste com *K-Fold*.

Com o resultado de cada *fold* foi calculada a média aritmética dos índices derivados da matriz de confusão para então obter o resultado final da classificação.



## 5 Resultados e Discussões

Neste capítulo, os resultados experimentais são apresentados e discutidos. Para todos os experimentos realizados, foram consideradas *CNNs* pré-treinadas com *ImageNet* e aumento de dados com inversões horizontais e inversões verticais.

### 5.1 *CNNs* como extrator de características

Foram realizados experimentos para avaliar a performance das *CNNs* como extratores de características. Inicialmente, foram consideradas as imagens originais, sem nenhum tipo de pré-processamento e as características extraídas por meio das *CNNs* foram utilizadas como vetor de características para os classificadores convencionais: Naive Bayes e *SVM* linear, com constante definida empiricamente em 2,9.

A Tabela 2 mostra a performance de cada classificador, sendo o melhor resultado obtido ao extrair as características com a arquitetura MobileNet e classificação com *SVM*, alcançando 91,21% de acurácia. Ao considerar a extração de características com a ResNet-50, percebe-se que ambos os classificadores avaliados não alcançaram bons resultados.

Tabela 2 – Média dos *folds* para cada classificador ao utilizar imagens originais.

|                     | Acurácia (%) |              | Precisão (%) |              | Recall (%) |              | F1-Score (%) |              |
|---------------------|--------------|--------------|--------------|--------------|------------|--------------|--------------|--------------|
|                     | Bayes        | SVM          | Bayes        | SVM          | Bayes      | SVM          | Bayes        | SVM          |
| <b>DenseNet-169</b> | 82,60        | 90,59        | 83,47        | 91,23        | 82,67      | 90,66        | 82,82        | 90,61        |
| <b>ResNet-50</b>    | 59,10        | 68,05        | 61,30        | 70,26        | 59,73      | 67,88        | 52,91        | 64,69        |
| <b>MobileNet</b>    | 85,30        | <b>91,21</b> | 87,48        | <b>91,42</b> | 85,30      | <b>91,28</b> | 85,39        | <b>91,07</b> |

Posteriormente, foram realizados experimentos a fim de verificar como as imagens segmentadas com o método proposto por Rodrigues, Naldi e Mari (2016) pode impactar na extração de características. Os resultados experimentais, mostrados na Tabela 3, demonstram que as imagens segmentadas aprimoraram a performance de ambos os classificadores avaliados. É evidente que ao classificar as imagens segmentadas, para todas as *CNNs* utilizadas como extrator de características, o classificador Naive Bayes foi capaz de obter uma acurácia superior a 90% e o classificador *SVM* uma acurácia superior a 95%.

Tabela 3 – Média dos *folds* para cada classificador ao utilizar imagens segmentadas.

|                     | Acurácia (%) |              | Precisão (%) |              | Recall (%) |              | F1-Score (%) |              |
|---------------------|--------------|--------------|--------------|--------------|------------|--------------|--------------|--------------|
|                     | Bayes        | SVM          | Bayes        | SVM          | Bayes      | SVM          | Bayes        | SVM          |
| <b>DenseNet-169</b> | 95,37        | 95,84        | 95,74        | 96,00        | 95,38      | 95,87        | 95,44        | 95,88        |
| <b>ResNet-50</b>    | 90,26        | 96,32        | 90,98        | 96,55        | 90,26      | 96,36        | 90,41        | 96,33        |
| <b>MobileNet</b>    | 94,26        | <b>96,80</b> | 94,69        | <b>96,88</b> | 94,26      | <b>96,84</b> | 94,31        | <b>96,81</b> |

Ao considerar a classificação com *SVM* e extração com a ResNet-50, os resultados aumentaram de 68.05% ao utilizar imagens originais (Tabela 2) para 96,32% ao utilizar imagens segmentadas (Tabela 3). Esse resultado sugere que a ResNet-50 foi capaz de extrair importantes padrões de forma. Finalmente, o melhor resultado foi obtido ao considerar extração de características com MobileNet e imagens segmentadas, alcançando 96,80% de acurácia.

## 5.2 CNNs treinadas com transferência de aprendizado

Para realizar uma comparação com os resultados apresentados na Seção 5.1, foi realizado um experimento que utiliza cada *CNN* como classificador, ou seja, as características são extraídas pela rede durante o treinamento e após ocorre a classificação. Novamente, foi avaliada a performance de cada *CNN* quando treinada com imagens originais e imagens segmentadas, as quais foram redimensionadas para  $224 \times 224$  *pixels* (tamanho de entrada permitido por cada *CNN*). O treinamento foi realizado considerando 50 épocas, taxa de aprendizado de  $1 \times 10^{-4}$ , otimização com gradiente estocástico (*SGD*) e coeficiente *momentum* de 0,9.

As Tabelas 4 e 5 mostram os resultados obtidos ao classificar as imagens originais e segmentadas, respectivamente. É possível notar que o uso de imagens segmentadas trouxe melhorias para as arquiteturas ResNet-50 e MobileNet. A arquitetura DenseNet-169 entretanto, obteve o mesmo resultado, em termos de acurácia, para ambos os casos (imagens originais e segmentadas). Esse resultado sugere que a arquitetura DenseNet-169 é capaz de lidar bem com padrões de textura ou forma e não requer pré-processamento, conforme pode ser visto na Figura 12 que ilustra os resultados das convoluções aplicadas considerando uma imagem original como entrada.

Tabela 4 – Média dos *folds* para cada arquitetura *CNN* e considerando imagens originais.

| <b>CNN</b>   | <b>Acurácia (%)</b> | <b>Precisão (%)</b> | <b>Recall (%)</b> | <b>F1-Score (%)</b> |
|--------------|---------------------|---------------------|-------------------|---------------------|
| DenseNet-169 | <b>96,65</b>        | <b>96,89</b>        | <b>96,71</b>      | <b>96,62</b>        |
| ResNet-50    | 92,65               | 93,13               | 92,70             | 92,56               |
| MobileNet    | 87,70               | 88,73               | 87,85             | 87,11               |

Tabela 5 – Média dos *folds* para cada arquitetura *CNN* e considerando imagens segmentadas.

| <i>CNN</i>   | Acurácia (%) | Precisão (%) | <i>Recall</i> (%) | <i>F1-Score</i> (%) |
|--------------|--------------|--------------|-------------------|---------------------|
| DenseNet-169 | <b>96,65</b> | <b>96,78</b> | <b>96,67</b>      | <b>96,65</b>        |
| ResNet-50    | 95,37        | 93,50        | 95,42             | 95,38               |
| MobileNet    | 95,54        | 95,62        | 95,59             | 95,54               |

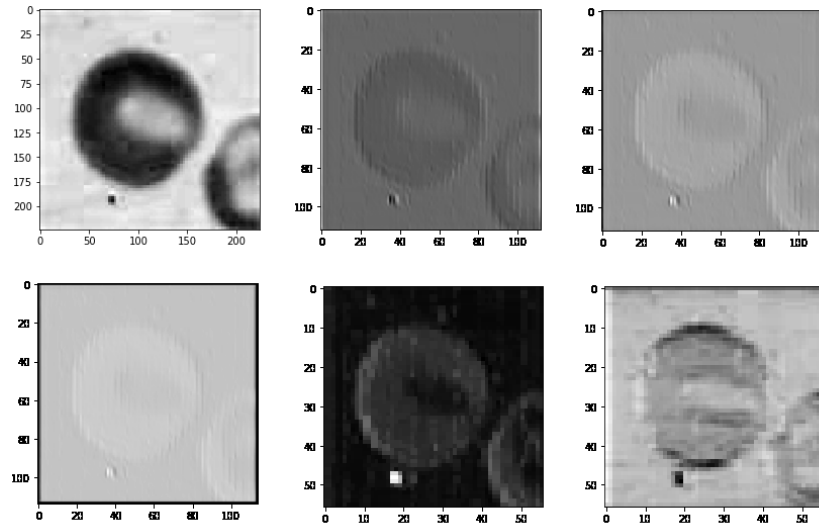


Figura 12 – Exemplo de mapas de características extraídos pela DenseNet-169.

Os gráficos apresentados nas Figuras 13 e 14 mostram a acurácia média de classificação ao considerar cada estratégia explorada no presente trabalho. É possível observar que quando as imagens originais são utilizadas, somente a MobileNet foi ligeiramente melhor ao ser testada como extrator de características e classificação com *SVM* linear. Para as demais *CNNs* avaliadas, a classificação pela própria rede foi melhor. Esse resultado sugere que as *CNNs* lidam melhor com padrões de forma ao classificar os eritrócitos. Ao analisar o resultado obtido com as imagens segmentadas observa-se que na maioria dos casos a *CNN* funcionou melhor como extrator de características e classificação com *SVM*.

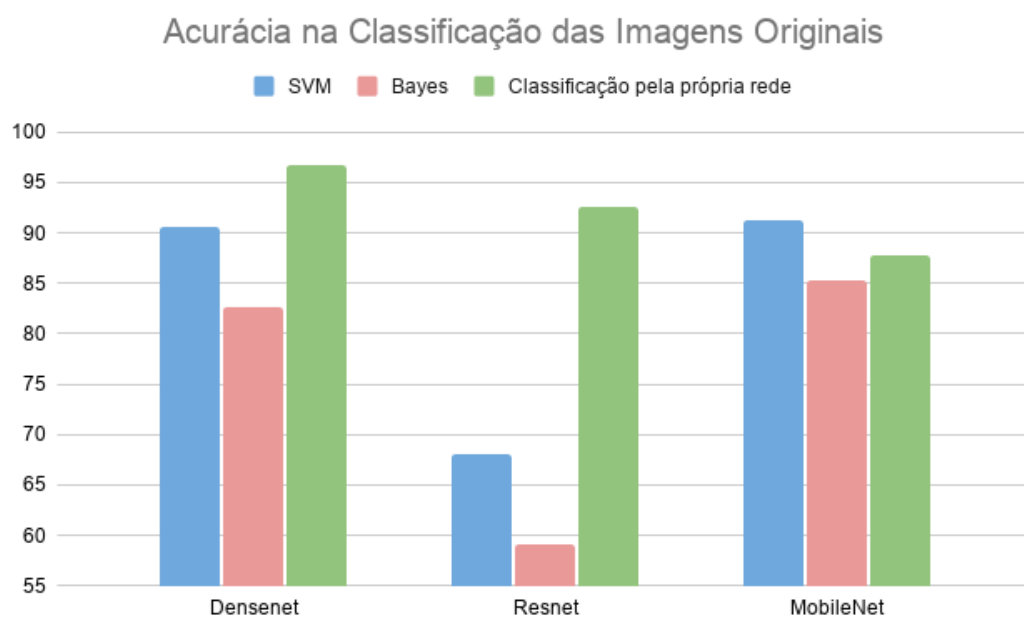


Figura 13 – Comparação entre as estratégias avaliadas ao utilizar imagens originais.

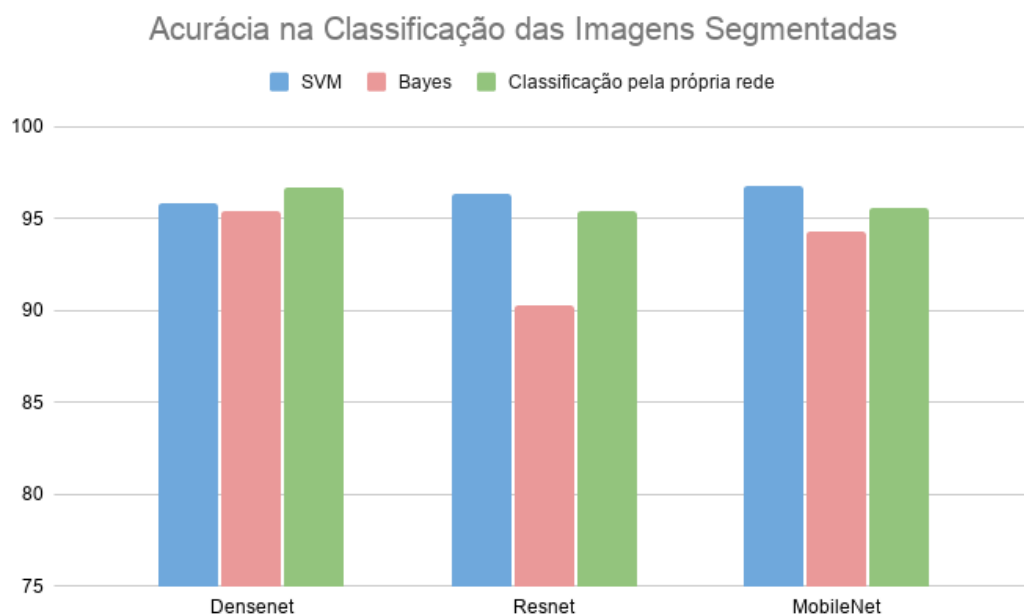


Figura 14 – Comparação entre as estratégias avaliadas ao utilizar imagens segmentadas.

### 5.3 Comparação entre as melhores estratégias

Ao realizar uma análise comparativa entre as melhores estratégias, as quais foram apresentadas anteriormente, observa-se que o melhor resultado foi obtido ao classificar as imagens de eritrócitos com *SVM* e com as características extraídas da MobileNet treinada com imagens segmentadas, alcançando 96,80% de acurácia. Entretanto, a DenseNet apresentou uma performance muito parecida chegando a 96,65% para ambos os conjuntos de imagens. Esse resultado tão semelhante permite levar em consideração o tempo de execução como métrica de comparação. Para treinar e classificar com a DenseNet, é necessário 58 minutos e 55 segundos. Em contrapartida, realizar a extração de características com a compacta rede MobileNet e classificação com *SVM* exige apenas 4 minutos e 43 segundos, conforme ilustra a Figura 15.

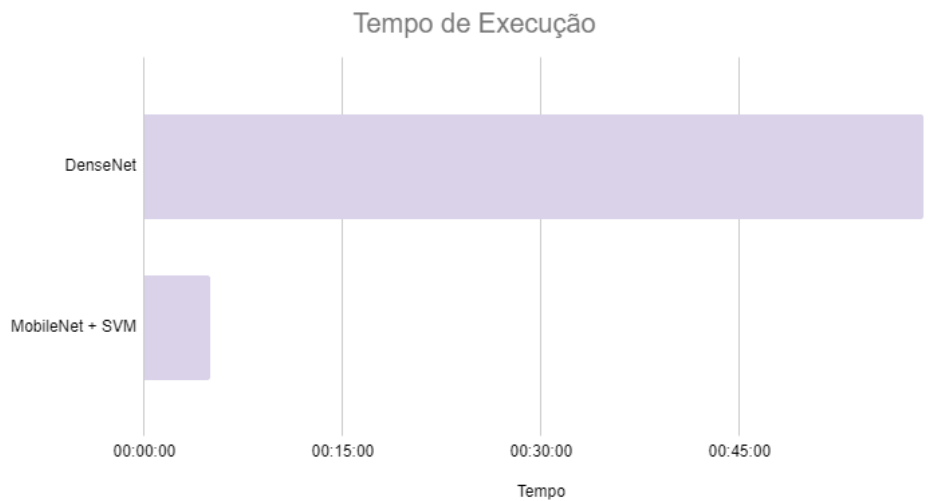


Figura 15 – Tempo de execução dos melhores resultados obtidos.

Apesar do ótimo tempo de execução, o resultado obtido pela extração de características leva em consideração as imagens segmentadas e com isso é incluída uma etapa de pré-processamento nas imagens. Portanto, é necessário avaliar a situação para definir qual a melhor abordagem para o problema de classificação de doença falciforme em imagens de microscopia, levando em consideração dois fatores principais: processamento e tempo. Situações em que o custo computacional é primordial, é recomendável utilizar MobileNet com *SVM*. Já em situações que não há pré-processamento disponível e o custo computacional não é importante, utilizar a arquitetura DenseNet é a melhor opção.

## 5.4 Comparação com trabalhos relacionados

A Tabela 6 traz uma comparação entre o melhor resultado alcançado nesta pesquisa e os resultados relatados na literatura, sendo que todos os trabalhos consideraram o conjunto de imagens *erythrocytesIDB*. O melhor resultado apresentado nesta pesquisa, obtido ao extrair características com MobileNet e classificação com *SVM* linear, é superior ou bem próximo dos resultados apresentados na literatura.

Tabela 6 – Comparação com os trabalhos relacionados.

| Referência                              | Acurácia (%) |
|---|--------------|
| Rodrigues, Naldi e Mari (2016)          | 94,59        |
| Gual-Arnau, Herold-García e Simó (2015) | 96,10        |
| Faria, Rodrigues e Mari (2018)          | 93,67        |
| Silva, Rodrigues e Mari (2020)          | 92,54        |
| Paz-Soto et al. (2020)                  | 98,40        |
| Alzubaidi et al. (2020)                 | 99,98        |
| <b>Presente trabalho (2021)</b>         | <b>96,80</b> |

## 6 Conclusões

O presente trabalho avaliou a estratégia de utilização de *CNNs* para extrair características de imagens de eritrócitos e realizar sua classificação entre saudável, falciforme ou com outras deformidades. Também foi possível investigar a performance das *CNNs* e classificadores convencionais utilizando imagens segmentadas. O resultado de 96,80% de acurácia alcançado ao extrair características com MobileNet e classificação com *SVM* demonstra que essa abordagem pode auxiliar na identificação de doença falciforme em imagens de microscopia.

Embora utilizar imagens segmentadas requer uma etapa de pré-processamento, a abordagem proposta exige menos tempo para realizar a tarefa de classificação quando comparada com a classificação realizada pelas próprias *CNNs* e obtém resultados com acurácias muito próximas e com um tempo de execução aproximadamente 14 vezes mais rápido.

Para trabalhos futuros, pretende-se considerar abordagens de otimização dos parâmetros das *CNNs* e dos classificadores, testar os métodos propostos neste trabalho para classificar outros conjuntos de imagens médicas e avaliar outras arquiteturas *CNNs* e outros classificadores.

# Referências

- ABADI, M. et al. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.
- ALZUBAIDI, L. et al. Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. **Electronics**, Multidisciplinary Digital Publishing Institute, v. 9, n. 3, p. 427, 2020.
- CHILDREN'S HOSPITAL LOS ANGELES. **What is Sickle Cell Disease?** 2017. Disponível em: <<https://www.chla.org/blog/rn-remedies/what-sickle-cell-disease>>. Acesso em: 01 jul. 2020.
- CHOLLET, F. et al. **Keras**. 2015. <<https://keras.io>>.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Springer, v. 20, n. 3, p. 273–297, 1995.
- COSTA, L. d. F.; CESAR-JR., R. M. **Shape Classification and Analysis: Theory and Practice**. 2nd. ed. Boca Raton, FL, USA: CRC Press, Inc., 2009. ISBN 0849379296, 9780849379291.
- COSTA NICOLA CONRAN, K. Y. F. F. **Tratado de Hematologia, Capítulo 27**. São Paulo: Atheneu, 2013.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification (2nd Edition)**. New York, NY, USA: Wiley-Interscience, 2000. ISBN 0471056693.
- FARIA, L. C. de; RODRIGUES, L. F.; MARI, J. F. Cell classification using handcrafted features and bag of visual words. In: **XIV Workshop de Visão Computacional**. Ilhéus, BA, Brazil: WVC, 2018.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. USA: Prentice-Hall, Inc., 2006. ISBN 013168728X.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. USA: MIT Press, 2016. <<http://www.deeplearningbook.org>>.
- GUAL-ARNAU, X.; HEROLD-GARCÍA, S.; SIMÓ, A. Erythrocyte shape classification using integral-geometry-based methods. **Medical & biological engineering & computing**, Springer, v. 53, n. 7, p. 623–633, 2015.
- HAN, J.; KAMBER, M.; PEI, J. **Data Mining: Concepts and Techniques**. 3rd. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN 0123814790, 9780123814791.
- HE, K. et al. Deep residual learning for image recognition. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Las Vegas, NV, USA: IEEE, 2016. p. 770–778.



- HOFFMAN, R. et al. **Hematology: Basic Principles and Practice**. Saunders/Elsevier, 2013. (Churchill Livingstone). ISBN 9781437729283. Disponível em: <<https://books.google.com.br/books?id=a1estSuaQ6kC>>.
- HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017.
- HUANG, G. et al. Densely connected convolutional networks. In: **The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Honolulu, HI, USA: IEEE, 2017.
- KATO, G. J. et al. Sick cell disease. **Nature Reviews Disease Primers**, Nature Publishing Group, v. 4, n. 1, p. 1–22, 2018.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995. (IJCAI'95), p. 1137–1143. ISBN 1-55860-363-8.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Neural Information Processing Systems**, v. 25, 01 2012.
- MELO-REIS, P. R. et al. A importância do diagnóstico precoce na prevenção das anemias hereditárias. **Revista Brasileira de Hematologia e Hemoterapia**, SciELO Brasil, v. 28, n. 2, p. 149–152, 2006.
- PANDEY, A. **Depth-wise Convolution and Depth-wise Separable Convolution**. 2018. Disponível em: <<https://medium.com/@zurister/depth-wise-convolution-and-depth-wise-separable-convolution-37346565d4ec>>. Acesso em: 01 abr. 2021.
- PAZ-SOTO, Y. et al. Automatic classification of erythrocytes using artificial neural networks and integral geometry-based functions. In: **2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. Porto de Galinhas, Brazil: IEEE, 2020. p. 156–163.
- PONTI, M. A. et al. Everything you wanted to know about deep learning for computer vision but were afraid to ask. In: **2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T)**. Niterói, Brazil: IEEE, 2017. p. 17–41. ISSN 2474-0705.
- RODRIGUES, L. F. **Comparação entre redes neurais convolucionais e técnicas de pré-processamento para classificar células HEp-2 em imagens de imunofluorescência**. Dissertação (Mestrado) — Universidade Federal de Viçosa, 2018. Disponível em: <<https://locus.ufv.br//handle/123456789/25441>>.
- RODRIGUES, L. F.; NALDI, M. C.; MARI, J. F. Morphological analysis and classification of erythrocytes in microscopy images. In: **Anais do XII Workshop de Visão Computacional**. Campo Grande, MS, Brazil: WVC, 2016.
- RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. **International journal of computer vision**, Springer, v. 115, n. 3, p. 211–252, 2015.

SERJEANT, G. R. The natural history of sickle cell disease. **Cold Spring Harbor perspectives in medicine**, Cold Spring Harbor Laboratory Press, v. 3, n. 10, p. a011783, 2013.

SILVA, M. da; RODRIGUES, L.; MARI, J. F. Optimizing data augmentation policies for convolutional neural networks based on classification of sickle cells. In: **Anais do XVI Workshop de Visão Computacional**. Porto Alegre, RS, Brasil: SBC, 2020. p. 46–51. Disponível em: <<https://sol.sbc.org.br/index.php/wvc/article/view/13479>>.

SULTANA, F.; SUFIAN, A.; DUTTA, P. Advancements in image classification using convolutional neural network. In: **2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)**. Kolkata, India: IEEE, 2018. p. 122–129.

WU, Q.; MERCHANT, F.; CASTLEMAN, K. **Microscope Image Processing**. 1st. ed. USA: Academic Press, Inc., 2008. ISBN 012372578X.