



**Universidade Federal de Pernambuco**  
**Centro de Tecnologia e Geociências**  
**Departamento de Engenharia Biomédica**  
**Programa de Pós-Graduação em Engenharia Biomédica**

**PATRICIA ALVES GENUINO**

**SISTEMA INTELIGENTE PARA APOIO AO DIAGNÓSTICO DE DOENÇAS DE  
PELE USANDO IMAGENS E REDES NEURAIS ARTIFICIAIS PROFUNDAS**

**RECIFE 2024**

PATRICIA ALVES GENUINO

**SISTEMA INTELIGENTE PARA APOIO AO DIAGNÓSTICO DE DOENÇAS DE  
PELE USANDO IMAGENS E REDES NEURAIAS ARTIFICIAIS PROFUNDAS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Biomédica do Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Mestre em Engenharia Biomédica.

**Área de concentração:** Computação Biomédica.

**Orientador:** Prof<sup>o</sup>. Dr. Wellington Pinheiro dos Santos

**Coorientadora:** Prof<sup>a</sup>. Dra. . Juliana Carneiro Gomes

RECIFE 2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Genuino, Patrícia Alves.

Sistema inteligente para apoio ao diagnóstico de doenças de pele usando imagens e redes neurais artificiais profundas / Patrícia Alves Genuino. - Recife, 2025.

110f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências - CTG, Programa de Pós Graduação em Engenharia Biomédica, 2025.

Orientação: Wellington Pinheiro dos Santos.

Coorientação: Juliana Carneiro Gomes.

Inclui referências.

1. Sistema de apoio ao diagnóstico; 2. Hanseníase; 3. Câncer de pele; 4. Psoríase; 5. Artrite psoriásica; 6. Redes neurais artificiais profundas. I. Santos, Wellington Pinheiro dos. II. Gomes, Juliana Carneiro. III. Título.

UFPE-Biblioteca Central

PATRICIA ALVES GENUINO

**SISTEMA INTELIGENTE PARA APOIO AO DIAGNÓSTICO DE DOENÇAS DE  
PELE USANDO IMAGENS E REDES NEURAIAS ARTIFICIAIS PROFUNDAS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Biomédica do Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Mestre em Engenharia Biomédica.

Aprovado em: 26 de março de 2024

**BANCA EXAMINADORA**

---

Prof.<sup>o</sup>. Dr.. Wellington Pinheiro dos Santos (Orientador)  
Universidade Federal de Pernambuco

---

Prof.<sup>a</sup>. Dra. Juliana Carneiro Gomes (Coorientadora)  
Universidade Federal de Pernambuco

---

Prof.<sup>a</sup>. Dra. Maira Araujo de Santana (Examinadora Interna)  
Universidade Federal de Pernambuco

---

Prof.<sup>a</sup>. Dra. Giselle Machado Magalhães Moreno (Examinadora externa)  
Universidade Federal de Pernambuco

RECIFE 2024

## DEDICATÓRIAS

Dedico este momento sublime a Deus, cuja luz tem guiado meu caminho em toda minha jornada até aqui. Agradeço aos meus pais, meu marido, minha filha, meu filho e toda minha família, cujo carinho e apoio incondicionais foram fundamentais para que eu alcançasse esta etapa tão significativa em minha vida.

Gostaria de fazer uma dedicatória especial ao meu saudoso avô, José Antônio Alves, minha avó Maria da Paz Alves e meu pai Pedro Genuíno Pereira Filho. Embora não estejam mais entre nós fisicamente, suas presenças permanecem vivas em nossos corações e espíritos. Eles sempre sonharam em ver sua neta e filha não apenas alcançar a graduação, mas também conquistar outros títulos acadêmicos. Hoje, com imensa gratidão, estou aqui, realizando esse sonho que tanto almejaram.

Ao concluir com sucesso o meu mestrado, renovo minha dedicação a toda minha família. A jornada foi longa e desafiadora, mas graças ao apoio deles, alcancei este marco importante em minha vida acadêmica.

Que este momento marque não apenas uma conquista, mas também um novo começo. Agora, olho para o futuro com entusiasmo, sabendo que estou preparada para enfrentar novos desafios e contribuir de maneira significativa para a minha área de estudo.

Agradeço a todos que estiveram ao meu lado nesta jornada e celebro este momento com gratidão e alegria.

## AGRADECIMENTOS

Agradeço primeiramente a Deus pelo dom da vida e por ter me proporcionado chegar até aqui. Com Sua ajuda, venci todos os obstáculos e hoje celebro a vitória na minha segunda batalha profissional. Agradeço imensamente à minha família por todo o apoio, dedicação e paciência ao longo desses anos. Sem o amor e suporte de vocês, meu caminho teria sido muito mais árduo.

Quero fazer um agradecimento especial à minha mãe, Lucicleide, ao meu marido, Eduardo, à minha segunda mãe, Sandra, à minha tia Lúcia e aos meus filhos, Rickson e Jennifer. Vocês sempre compreenderam minhas ausências e caminharam ao meu lado durante esses dois anos, sendo fontes de força e amor incondicional.

Agradeço também aos professores que estiveram sempre dispostos a ajudar e contribuir para o meu aprendizado, em especial ao meu orientador, Prof. Dr. Wellington Pinheiro dos Santos, e à minha coorientadora, Prof<sup>ª</sup>. Dra. Juliana Carneiro Gomes. Seus ensinamentos e orientações foram fundamentais para meu crescimento acadêmico e profissional.

Não posso deixar de mencionar minha amiga, Dra. Mônica Marinho, que esteve ao meu lado apoiando-me e incentivando-me com sua torcida constante.

Além disso, gostaria de expressar minha gratidão a todo o departamento de engenharia biomédica e a todos que contribuíram para que este projeto se tornasse realidade. Cada um de vocês desempenhou um papel crucial nessa jornada e sou imensamente grata pela colaboração e apoio de todos.

Por fim, agradeço à minha instituição por me proporcionar esta oportunidade e por fornecer todas as ferramentas necessárias que me permitiram concluir este ciclo de forma satisfatória. Estou profundamente grata pela experiência e aprendizado que adquiri ao longo deste período.

## RESUMO

O diagnóstico precoce de doenças dermatológicas tem sido impulsionado pelo avanço da inteligência artificial, especialmente pelo aprendizado profundo combinado com algoritmos de classificação. Este estudo propõe uma estrutura de apoio ao diagnóstico, utilizando redes neurais artificiais profundas e algoritmos de classificação, visando a detecção precoce de condições dermatológicas. Foram exploradas diversas condições médicas, incluindo hanseníase, psoríase, artrite psoriásica e câncer de pele, cada uma com suas próprias características distintas. A hanseníase, por exemplo, é uma doença endêmica em regiões subdesenvolvidas, como o Brasil, onde o diagnóstico precoce é crucial para interromper a transmissão e prevenir complicações. Da mesma forma, a psoríase e a artrite psoriásica requerem detecção precoce para melhorar os resultados clínicos e a qualidade de vida dos pacientes. Além disso, o câncer de pele apresenta diferentes subtipos, como melanoma, carcinoma basocelular e carcinoma epidermóide, cada um exigindo uma abordagem específica para diagnóstico e tratamento eficazes. Foram utilizadas redes neurais artificiais profundas (VGG16, Inception v3 e ResNet-50) para a extração de atributos das imagens das condições dermatológicas. Além disso, a técnica de balanceamento SMOTE foi empregada durante a fase de treinamento para garantir uma distribuição adequada das classes e evitar viés nos resultados. Os conjuntos de dados foram divididos em treino e teste. Cinco categorias de diferenciais das doenças foram exploradas, avaliadas separadamente e em conjunto. Os resultados destacaram a rede Inception v3 como a mais eficaz para este propósito, combinada com o classificador SVM de grau 3, seguido pelo Random Forest com 500 árvores. Métricas de avaliação, incluindo acurácia, índice kappa, sensibilidade, especificidade e área sob a curva ROC, confirmaram a eficácia dessas combinações. Este estudo oferece uma contribuição significativa para o avanço no diagnóstico precoce de doenças dermatológicas, fornecendo boas perspectivas valiosas sobre a eficácia de diferentes abordagens de inteligência artificial e suas combinações. Esses resultados têm o potencial de impactar positivamente a prática clínica, oferecendo uma base sólida para o desenvolvimento de ferramentas de auxílio ao diagnóstico mais precisas, eficientes e acessíveis, melhorando assim os resultados clínicos e a qualidade de vida dos pacientes dermatológicos.

**Palavras-chave:** Sistema de apoio ao diagnóstico; Hanseníase; Câncer de pele; Psoríase; Artrite psoriásica; Redes neurais artificiais profundas.

## ABSTRACT

The early diagnosis of dermatological diseases has been propelled by the advancement of artificial intelligence, especially through deep learning combined with classification algorithms. This study proposes a diagnostic support framework utilizing deep artificial neural networks and classification algorithms, aiming at early detection of dermatological conditions. We explored various medical conditions, including leprosy, psoriasis, psoriatic arthritis, and skin cancer, each with its distinct characteristics. Leprosy, for example, is endemic in underdeveloped regions such as Brazil, where early diagnosis is crucial to interrupt transmission and prevent complications. Similarly, psoriasis and psoriatic arthritis require early detection to improve clinical outcomes and patients' quality of life. Additionally, skin cancer presents different subtypes such as melanoma, basal cell carcinoma, and squamous cell carcinoma, each demanding a specific approach for effective diagnosis and treatment. Deep neural networks (VGG16, Inception v3, and ResNet50) were used for extracting features from images of dermatological conditions. Furthermore, the SMOTE balancing technique was employed during the training phase to ensure proper class distribution and avoid bias in results. The datasets were split into training and testing sets. Five categories of disease differentials were explored, evaluated separately and in combination. The results highlighted the Inception v3 network as the most effective for this purpose, combined with a degree 3 SVM classifier, followed by Random Forest with 500 trees. Evaluation metrics, including accuracy, kappa index, sensitivity, specificity, and area under the ROC curve, confirmed the effectiveness of these combinations. This study offers a significant contribution to the advancement of early diagnosis of dermatological diseases, providing valuable insights into the effectiveness of different artificial intelligence approaches and their combinations. These results have the potential to positively impact clinical practice, offering a solid foundation for the development of more accurate, efficient, and accessible diagnostic aid tools, thereby improving clinical outcomes and the quality of life of dermatological patients.

**Keywords:** Diagnostic support system; Leprosy; Skin cancer; Psoriasis; Psoriatic arthritis; Deep artificial neural networks.

## LISTA DE FIGURAS

<b>Figura 1:</b> Fluxograma das etapas da metodologia	<b>21</b>
<b>Figura 2:</b> Amostras do banco de imagens que representam os diferentes tipos de hanseníase	<b>46</b>
<b>Figura 3:</b> Amostras do banco de imagens que representam os diferentes tipos de psoríase	<b>47</b>
<b>Figura 4:</b> Amostras do banco de imagens que representa a artrite psoriásica	<b>49</b>
<b>Figura 5:</b> Amostras do banco de imagens que representam os diferentes tipos de câncer de pele	<b>50</b>
<b>Figura 6:</b> Grupo de aprendizado profundo	<b>52</b>
<b>Figura 7:</b> Arquitetura CNN para classificação de imagens	<b>53</b>
<b>Figura 8:</b> A arquitetura do VGG	<b>54</b>
<b>Figura 9:</b> ImagiNet, aprendizagem profunda em comparação com o ser humano	<b>56</b>
<b>Figura 10:</b> Arquitetura Inception	<b>57</b>
<b>Figura 11:</b> Um bloco de construção do aprendizado residual	<b>58</b>
<b>Figura 12:</b> Fluxograma do experimento	<b>65</b>
<b>Figura 13:</b> Boxplot mostrando a distribuição da acurácia considerando os diferenciais 1,2,3,4, e 5	<b>79</b>
<b>Figura 14:</b> Boxplot mostrando a distribuição do índice kappa considerando os diferenciais 1,2,3,4, e 5	<b>80</b>
<b>Figura 15:</b> Boxplot mostrando a distribuição da sensibilidade considerando os diferenciais 1,2,3,4, e 5	<b>81</b>
<b>Figura 16:</b> Boxplot mostrando a distribuição da especificidade considerando os diferenciais 1,2,3,4, e 5	<b>82</b>
<b>Figura 17:</b> Boxplot mostrando a distribuição da área sob curva ROC considerando os diferenciais 1,2,3,4, e 5	<b>83</b>
<b>Figura 18:</b> Boxplot mostrando a distribuição da melhor acurácia de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5	<b>88</b>
<b>Figura 19:</b> Boxplot mostrando a distribuição do índice kappa de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5	<b>88</b>
<b>Figura 20:</b> Boxplot mostrando a distribuição da sensibilidade de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5	<b>89</b>
<b>Figura 21:</b> Boxplot mostrando a distribuição da especificidade de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5	<b>89</b>

<b>Figura 22:</b> Boxplot mostrando a distribuição da área sob a curva ROC de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5	<b>90</b>
<b>Figura 23:</b> Matriz de confusão referente ao primeiro diferencial no modo treino	<b>91</b>
<b>Figura 24:</b> Matriz de confusão referente ao primeiro diferencial no modo teste	<b>92</b>
<b>Figura 25:</b> Matriz de confusão referente ao segundo diferencial no modo treino	<b>93</b>
<b>Figura 26:</b> Matriz de confusão referente ao segundo diferencial no modo teste	<b>93</b>
<b>Figura 27:</b> Matriz de confusão referente ao terceiro diferencial no modo treino	<b>94</b>
<b>Figura 28:</b> Matriz de confusão referente ao terceiro diferencial no modo teste	<b>95</b>
<b>Figura 29:</b> Matriz de confusão referente ao quarto diferencial no modo treino	<b>96</b>
<b>Figura 30:</b> Matriz de confusão referente ao quarto diferencial no modo teste	<b>97</b>
<b>Figura 31:</b> Matriz de confusão referente ao quinto diferencial no modo treino	<b>98</b>
<b>Figura 32:</b> Legenda da figura 31	<b>99</b>
<b>Figura 33:</b> Matriz de confusão referente ao quinto diferencial no modo teste	<b>100</b>
<b>Figura 34:</b> Legenda da figura 33	<b>101</b>

## LISTA DE TABELAS

<b>Tabela 1:</b> Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças Dermatológicas, com Ênfase em Hanseníase	<b>24</b>
<b>Tabela 2:</b> Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças Dermatológicas, com ênfase cancer de pele	<b>27</b>
<b>Tabela 3:</b> Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças de Pele, com foco em Psoríase	<b>35</b>
<b>Tabela 4:</b> Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças de Pele, com foco em Artrite Psoriásica	<b>41</b>
<b>Tabela 5:</b> Cenários propostos para análise de imagens de diferentes doenças de pele	<b>61</b>
<b>Tabela 6:</b> Diversas doenças de pele com suas respectivas classes, divididas em diferenciais	<b>62</b>
<b>Tabela 7:</b> Representação da matriz de confusão	<b>66</b>
<b>Tabela 8:</b> Tabela de métricas utilizadas para avaliar o desempenho dos classificadores	<b>67</b>
<b>Tabela 9:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o primeiro diferencial	<b>69</b>
<b>Tabela 10:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG16 de acordo com o primeiro diferencial	<b>69</b>
<b>Tabela 11:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Resnet-50 de acordo com o primeiro diferencial	<b>70</b>
<b>Tabela 12:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o segundo diferencial	<b>71</b>
<b>Tabela 13:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG16 de acordo com o segundo diferencial	<b>71</b>
<b>Tabela 14:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o segundo diferencial	<b>72</b>
<b>Tabela 15:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o terceiro diferencial	<b>72</b>
<b>Tabela 16:</b> Métricas com Média e Desvio Padrão usando Rede Neural VGG 16, de acordo com o terceiro diferencial	<b>73</b>

<b>Tabela 17:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o terceiro diferencial	<b>73</b>
<b>Tabela 18:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o quarto diferencial	<b>75</b>
<b>Tabela 19:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG 16 de acordo com o quarto diferencial	<b>75</b>
<b>Tabela 20:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o quarto diferencial	<b>76</b>
<b>Tabela 21:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o quinto diferencial	<b>76</b>
<b>Tabela 22:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG 16 de acordo com o quinto diferencial	<b>77</b>
<b>Tabela 23:</b> Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o quinto diferencial	<b>78</b>
<b>Tabela 24:</b> Primeiro diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3	<b>84</b>
<b>Tabela 25:</b> Segundo diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3	<b>85</b>
<b>Tabela 26:</b> Terceiro diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3	<b>85</b>
<b>Tabela 27:</b> Quarto diferencial, rede neural Inception v3 com o classificador Random Forest de 500 árvores	<b>86</b>
<b>Tabela 28:</b> Quinto diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3	<b>87</b>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
<b>1.1</b>	<b>MOTIVAÇÃO, JUSTIFICATIVA E CARACTERIZAÇÃO DO PROBLEMA</b>	<b>15</b>
<b>1.2</b>	<b>OBJETIVOS</b>	<b>20</b>
1.2.1	Objetivo geral	20
1.2.2	Como objetivos específicos deste trabalho têm-se:	20
<b>1.3</b>	<b>METODOLOGIA</b>	<b>21</b>
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>23</b>
<b>2.1</b>	<b>TRABALHOS RELACIONADOS HANSENÍASE</b>	<b>23</b>
<b>2.2</b>	<b>TRABALHOS RELACIONADOS CÂNCER DE PELE</b>	<b>25</b>
<b>2.3</b>	<b>TRABALHOS RELACIONADOS PSORÍASE</b>	<b>31</b>
<b>2.4</b>	<b>TRABALHOS RELACIONADOS ARTRITE PSORIÁSICA</b>	<b>39</b>
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>43</b>
<b>3.1</b>	<b>HANSENÍASE</b>	<b>43</b>
<b>3.2</b>	<b>PSORÍASE (PSO)</b>	<b>46</b>
<b>3.3</b>	<b>ARTRITE PSORIÁSICA (PsA)</b>	<b>48</b>
<b>3.4</b>	<b>CÂNCER DE PELE</b>	<b>49</b>
<b>3.5</b>	<b>APRENDIZADO PROFUNDO</b>	<b>50</b>
<b>3.6</b>	<b>REDES NEURAIS CONVOLUCIONAIS (CNNs)</b>	<b>52</b>
<b>3.7</b>	<b>GRUPO DE GEOMETRIA VISUAL (VGG)</b>	<b>54</b>
<b>3.8</b>	<b>INCEPTION-V3</b>	<b>56</b>
<b>3.9</b>	<b>RESNET (REDE RESIDUAL)</b>	<b>57</b>
<b>3.10</b>	<b>BALANCEAMENTO DE DADOS</b>	<b>58</b>
<b>3.11</b>	<b>CLASSIFICAÇÃO</b>	<b>59</b>
3.11.1	Random Forest	59
3.11.2	SVM	59
<b>4</b>	<b>MATERIAIS E MÉTODOS</b>	<b>60</b>
<b>4.1</b>	<b>BASE DE DADOS</b>	<b>60</b>
<b>4.2</b>	<b>CENÁRIOS PROPOSTOS</b>	<b>60</b>
<b>4.3</b>	<b>EXTRAÇÃO DE ATRIBUTOS</b>	<b>63</b>
<b>4.4</b>	<b>TREINAMENTO E TESTE</b>	<b>64</b>
<b>4.5</b>	<b>ARRANJO EXPERIMENTAL</b>	<b>64</b>
<b>4.6</b>	<b>MÉTRICAS DE AVALIAÇÃO</b>	<b>65</b>

<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>68</b>
<b>5.1</b>	<b>AVALIAÇÃO DOS MODELOS DE REDE PROFUNDA</b>	<b>68</b>
5.1.1	Primeiro diferencial	68
5.1.2	Segundo diferencial	70
5.1.3	Terceiro diferencial	72
5.1.4	Quarto diferencial	74
5.1.5	Quinto diferencial	76
<b>5.2</b>	<b>REPRESENTAÇÃO GRÁFICA DOS RESULTADOS EM BOXPLOT NA FASE TREINO</b>	<b>78</b>
<b>5.3</b>	<b>FASE FINAL DOS EXPERIMENTOS (TESTE)</b>	<b>84</b>
5.3.2	Segundo diferencial	84
5.3.3	Terceiro diferencial	85
5.3.4	Quarto diferencial	86
5.3.5	Quinto diferencial	86
<b>5.4</b>	<b>REPRESENTAÇÃO GRÁFICA DOS RESULTADOS EM BOXPLOT DA FASE FINAL DE TESTE</b>	<b>87</b>
<b>5.5</b>	<b>MATRIZ DE CONFUSÃO</b>	<b>90</b>
5.5.1	Primeiro diferencial	91
5.5.2	Segundo diferencial	92
5.5.3	Terceiro diferencial	94
5.5.4	Quarto diferencial	95
5.5.5	Quinto diferencial	98
<b>6</b>	<b>CONCLUSÃO</b>	<b>103</b>
<b>7</b>	<b>REFERÊNCIAS</b>	<b>105</b>

# 1 INTRODUÇÃO

## 1.1 MOTIVAÇÃO, JUSTIFICATIVA E CARACTERIZAÇÃO DO PROBLEMA

A inteligência artificial (IA) está transformando a vida humana em diversas áreas. Na medicina, robôs realizam cirurgias com precisão. Sistemas de comunicação automatizados e a automação de serviços financeiros e bancários também são exemplos do impacto da IA. Com o uso de algoritmos avançados e grandes volumes de dados, esses sistemas podem propor ações eficazes e inovadoras (LOBO,2018). Os algoritmos baseados em redes neurais convolucionais em nuvem aprimoram consideravelmente a eficiência na classificação de imagens, detecção de objetos, segmentação e outras tarefas médicas. Em áreas da medicina que dependem de dados visuais, como a dermatologia, o reconhecimento e a análise de imagens por IA possibilitam diagnósticos rápidos e precisos, complementados pela avaliação profissional. A capacidade de aprendizado da IA permite identificar diferenças sutis nas características das lesões, como tamanho, textura e tonalidade, superando a precisão humana (LI et al., 2022).

A relevância deste trabalho está na contribuição para o diagnóstico precoce de doenças dermatológicas que causam impactos significativos na saúde dos pacientes. O diagnóstico precoce é fundamental para garantir um tratamento eficaz e reduzir possíveis complicações. Nesse sentido, a utilização de técnicas avançadas de inteligência artificial, como o aprendizado profundo, combinado com algoritmos de classificação, apresenta-se como uma abordagem promissora para o desenvolvimento de uma estrutura de apoio ao diagnóstico não invasivo.

Atualmente, há uma crescente escassez global de dermatologistas devido à alta incidência de condições de pele. Esta deficiência é especialmente grave em países em desenvolvimento e áreas remotas, onde a necessidade de mais instalações médicas, consultas especializadas e cuidados clínicos é urgente (Li et al., 2022). Além de atender às demandas dessas regiões desfavorecidas economicamente, a inteligência artificial (IA) tem a capacidade de oferecer diagnósticos rápidos, levando a abordagens terapêuticas mais variadas e acessíveis. A análise morfológica de lesões, uma base fundamental para diagnósticos dermatológicos, está sendo revolucionada pela tecnologia de reconhecimento de imagem e pelo amplo uso de smartphones em todo o mundo. Isso representa uma oportunidade transformacional para o diagnóstico e tratamento de doenças de pele na era atual (Li et al., 2022).

De acordo com LOBO (2018) a inteligência artificial (IA) tem o potencial de transformar a medicina e a educação médica, através da análise de grandes volumes de dados e algoritmos avançados, a IA pode melhorar a precisão dos diagnósticos, personalizar os trata-

mentos e acelerar a pesquisa médica. O autor destaca a importância de incorporar a IA no campo da medicina para melhorar a qualidade do cuidado ao paciente e aprimorar a formação dos profissionais de saúde.

De acordo com FOGEL e KVEDAR (2018) e HE et al. (2019), a utilização da inteligência artificial (IA) na área da saúde tem como objetivo aprimorar os cuidados de saúde, seja por meio da detecção precoce de doenças, diagnósticos mais precisos ou descoberta de novos tratamentos. O uso de métodos de aprendizado de máquina e, mais recentemente, de aprendizado profundo tem permitido o desenvolvimento de poderosas ferramentas de apoio ao diagnóstico por imagem e sinais. Essas ferramentas são capazes de extrair características das imagens e criar sistemas de classificação, contribuindo para a agilidade e confiabilidade dos diagnósticos.

Além disso, YU, BEAM e KOHANE (2018) destacam que a IA também pode desempenhar um papel importante na detecção precoce de doenças negligenciadas. Com a capacidade de analisar grandes volumes de dados e identificar padrões, a IA pode auxiliar na identificação dessas doenças, muitas vezes de baixa prevalência, permitindo uma intervenção mais rápida e eficaz.

De acordo com a Agência Fiocruz de Notícias (2013) as doenças negligenciadas são aquelas endêmicas, geralmente transmissíveis, que ocorrem em regiões subdesenvolvidas e não recebem investimentos adequados do Estado ou do mercado, seja devido à baixa prevalência dessas doenças ou simplesmente por afetarem as camadas menos favorecidas da população .

A hanseníase é uma doença crônica causada pela bactéria *Mycobacterium leprae*, capaz de afetar qualquer pessoa. Se não tratada na forma inicial, a doença geralmente evolui e se torna transmissível, resultando em incapacidades físicas (BRASIL, 2022). A hanseníase foi identificada em 1873 pelo cientista Armauer Hansen e é considerada uma das doenças mais antigas, com registros de casos há mais de 4000 anos, na China e na Índia. No entanto, ainda é uma doença complexa de erradicar, pois está fortemente associada à pobreza, falta de serviços de saúde e educação, além do estigma enfrentado pelos pacientes (SANTACROCE et al., 2021).

De acordo com o Ministério da Saúde (BRASIL, 2017) e (BRASIL, 2022), a transmissão da bactéria *Mycobacterium leprae* ocorre principalmente por meio de secreções das vias aéreas superiores e gotículas de saliva, especialmente durante o contato próximo com pacientes em estágios avançados da doença. Essa transmissão ocorre pelo ar e não por objetos utilizados pelo paciente. A hanseníase é caracterizada pela alteração da sensibilidade térmica,

dolorosa, tátil e da força muscular, além do surgimento de manchas na pele e nódulos, podendo resultar em incapacidades permanentes. É importante ressaltar que o tratamento adequado da hanseníase não apenas promove a cura da doença, mas também interrompe sua transmissão e prevenir incapacidades físicas (BRASIL, 2017;BRASIL, 2022).

O diagnóstico médico da hanseníase é realizado principalmente por meio da avaliação clínica do paciente, dependendo do reconhecimento dos sinais e sintomas, sendo possível apenas após a manifestação da doença. As lesões cutâneas geralmente são a primeira manifestação clínica observada. Caso seja necessário um exame laboratorial, o mais utilizado para auxiliar no diagnóstico é a baciloscopia, que envolve a coleta da serosidade cutânea. No entanto, esse exame pode apresentar resultados negativos nas fases iniciais da doença. O diagnóstico precoce é de suma importância para evitar a transmissão, complicações e deficiências (CHEN et al., 2022).

De acordo com a Organização Mundial da Saúde (1997), para fins de tratamento, a hanseníase é classificada em duas categorias principais: paucibacilar (PB), que apresenta poucos ou nenhum bacilo nos exames, e multibacilar (MB), que possui uma grande quantidade de bacilos.

No entanto, para uma melhor compreensão das formas clínicas da doença, neste trabalho será adotada a classificação de Madrid Sixth International Congress of Leprosy, 1953. Essa classificação divide a hanseníase em indeterminada (PB), tuberculóide (PB), dimorfa (MB) e virchowiana (MB), cada uma com suas características específicas (BRASIL, 2022).

A psoríase é uma doença crônica da pele caracterizada por manchas elevadas, vermelhas e escamosas em várias regiões do corpo humano (RAJ; LONDHE; SONAWANE, 2020). A segmentação das lesões de psoríase é uma etapa importante para o diagnóstico da doença, porém, atualmente, essa segmentação é feita manualmente por inspeção visual, o que é um processo demorado e subjetivo.

Pacientes com psoríase (PSO) frequentemente apresentam comorbidades, como condições cardiometabólicas e problemas de saúde mental. Além disso, cerca de 30% desses pacientes são diagnosticados com artrite psoriásica (PsA) (JIN et al., 2023).

Existem vários tipos de psoríase descritos na literatura. A psoríase em placa é a forma mais comum, caracterizada por manchas espessas e vermelhas na superfície da pele. A psoríase gutata aparece em forma de manchas vermelhas e coceiras em formato de gota de lágrima. Já a psoríase inversa afeta as dobras cutâneas, enquanto a psoríase pustulosa apresenta pústulas cheias de pus nas manchas escamosas vermelhas. A psoríase eritrodérmica é uma forma grave da doença, que causa inflamação e descamação intensa da pele (AIJAZ et al., 2022).

A psoríase pustular generalizada é uma condição cutânea autoinflamatória rara caracterizada por uma erupção generalizada de pústulas neutrofílicas estéreis (PUIG et al., 2023). A psoríase vulgar é responsável por aproximadamente 90% de todos os casos de psoríase (KIM et al., 2023).

A artrite psoriásica (PsA) é uma doença inflamatória crônica que afeta as articulações periféricas, articulações axiais e enteses (LIU et al., 2022). Essa doença musculoesquelética heterogênea foi descrita por Moll e Wright em 1973 como uma doença distinta com cinco subgrupos clinicamente sobrepostos, incluindo envolvimento articular axial e periférico (AHMAD et al., 2021). A PSA ocorre principalmente em associação com a psoríase, uma doença de pele, sendo que cerca de um terço dos indivíduos com psoríase desenvolvem PSA (LIU et al., 2022). Embora a psoríase geralmente precede o desenvolvimento da artrite, 20% dos pacientes apresentam artrite psoriásica antes da psoríase, e 15% têm manifestações cutâneas e articulares simultaneamente. A doença afeta igualmente homens e mulheres, sendo menos comum em asiáticos e negros e quase ausentes em populações nativas da América do Sul e Austrália (AHMAD et al., 2021).

A detecção precoce da PSA em pacientes com psoríase é crucial para o desfecho clínico e a qualidade de vida em longo prazo, porém, pode ser desafiadora devido à apresentação heterogênea da doença, com manifestações subclínicas em estágios iniciais (LIU et al., 2022).

O câncer de pele, no entanto, representa um grande fardo econômico para os serviços de saúde e pode ter impactos graves na saúde mental, especialmente por afetar áreas altamente visíveis do corpo. Devido à falta de conscientização sobre o rastreamento e características específicas das lesões, muitos casos são diagnosticados em estágios avançados, resultando em prognósticos desfavoráveis. Assim, é urgente desenvolver sistemas de inteligência artificial para auxiliar os clínicos nesta área (Li et al., 2022).

Os tecidos do corpo humano consistem em células normais que possuem a capacidade de se multiplicar de forma controlada. No entanto, as células cancerosas apresentam um crescimento descontrolado e anormal (INSTITUTO NACIONAL DO CÂNCER, 2011).

A pele é composta por três camadas principais: epiderme, derme e hipoderme. A maioria dos cânceres de pele tem início na epiderme, que é a camada mais superficial. Esses cânceres são classificados de acordo com as células que deram origem a eles, sendo os mais comuns o melanoma e os cânceres não melanoma, como o carcinoma basocelular e o carcinoma epidermoide (AMERICAN CANCER SOCIETY, 2019).

MACHADO et al. (2021) destacam a exposição à radiação ultravioleta solar como o principal fator de risco para o desenvolvimento do câncer de pele, incluindo o carcinoma ba-

so celular e o melanoma. Além disso, fatores genéticos, presença de nevos displásicos e tabagismo também são mencionados como riscos para o câncer de pele.

VILELA et al. (2021) ressaltam que, apesar de ser a segunda causa de mortalidade no Brasil, a maioria dos cânceres de pele tem um bom prognóstico quando diagnosticados precocemente. No país, estima-se uma alta incidência anual de casos de câncer de pele, sendo que aproximadamente 1 em cada 3 cânceres diagnosticados são de pele, com uma estimativa de 185,380 casos em 2020 e 4,120 óbitos registrados em 2018.

Atualmente, a aplicação da inteligência artificial (IA) na área médica enfrenta diversos desafios. Primeiramente, há uma escassez de conjuntos de dados de treinamento adequados e anotações manuais de qualidade, resultando em algoritmos de IA com precisão e praticidade insuficientes para atender às demandas clínicas diárias. Em segundo lugar, há uma desconexão entre os algoritmos de IA e os requisitos médicos reais, devido à falta de conjuntos de treinamento de alta qualidade e à incapacidade da IA de ser atualizada com vasta experiência ao longo do tempo. Além disso, a variedade de doenças dermatológicas e a falta de critérios uniformes de identificação dificultam o ensino da IA para diagnosticar múltiplas dermatoses, especialmente em casos raros onde há escassez de espécimes para treinamento. Esses desafios representam obstáculos significativos para o uso eficaz da IA na dermatologia (Li et al., 2022).

Este estudo tem como objetivo oferecer uma contribuição significativa para a área médica, fornecendo uma ferramenta de suporte ao diagnóstico precoce de doenças de pele, incluindo hanseníase, psoríase, artrite psoriásica e câncer de pele. A criação de uma estrutura de apoio ao diagnóstico não invasivo busca agilizar a detecção dessas condições, possibilitando um tratamento mais eficaz e reduzindo os impactos negativos na vida dos pacientes. A combinação de aprendizado profundo com técnicas de classificação visa desenvolver uma estrutura capaz de lidar com a complexidade e diversidades encontradas na prática clínica. Ao utilizar um modelo de aprendizado de máquina baseado em redes neurais profundas, pretende-se automatizar a análise de imagens fotográficas, facilitando a detecção e classificação das doenças de interesse.

Adicionalmente, na dermatologia, há o desafio das doenças raras, que se caracterizam pela baixa incidência e pela falta de espécimes disponíveis para o treinamento de modelos de aprendizado de máquina. Essa escassez de dados representa um obstáculo significativo para o avanço da aplicação da inteligência artificial nesse campo (Li et al., 2022).

Essa escassez representa um obstáculo significativo para a aplicação eficaz da inteligência artificial na dermatologia. Neste estudo, utilizamos um conjunto de dados composto

por 1840 imagens provenientes de dois atlas dermatológicos públicos, e cada doença abordada tem uma variação no número de imagens.

## 1.2 OBJETIVOS

### 1.2.1 Objetivo geral

O objetivo deste projeto é desenvolver um modelo de aprendizado de máquina baseado em uma arquitetura de rede neural profunda e um classificador específico para auxiliar no diagnóstico não invasivo de doenças de pele, utilizando a análise automatizada de imagens fotográficas.

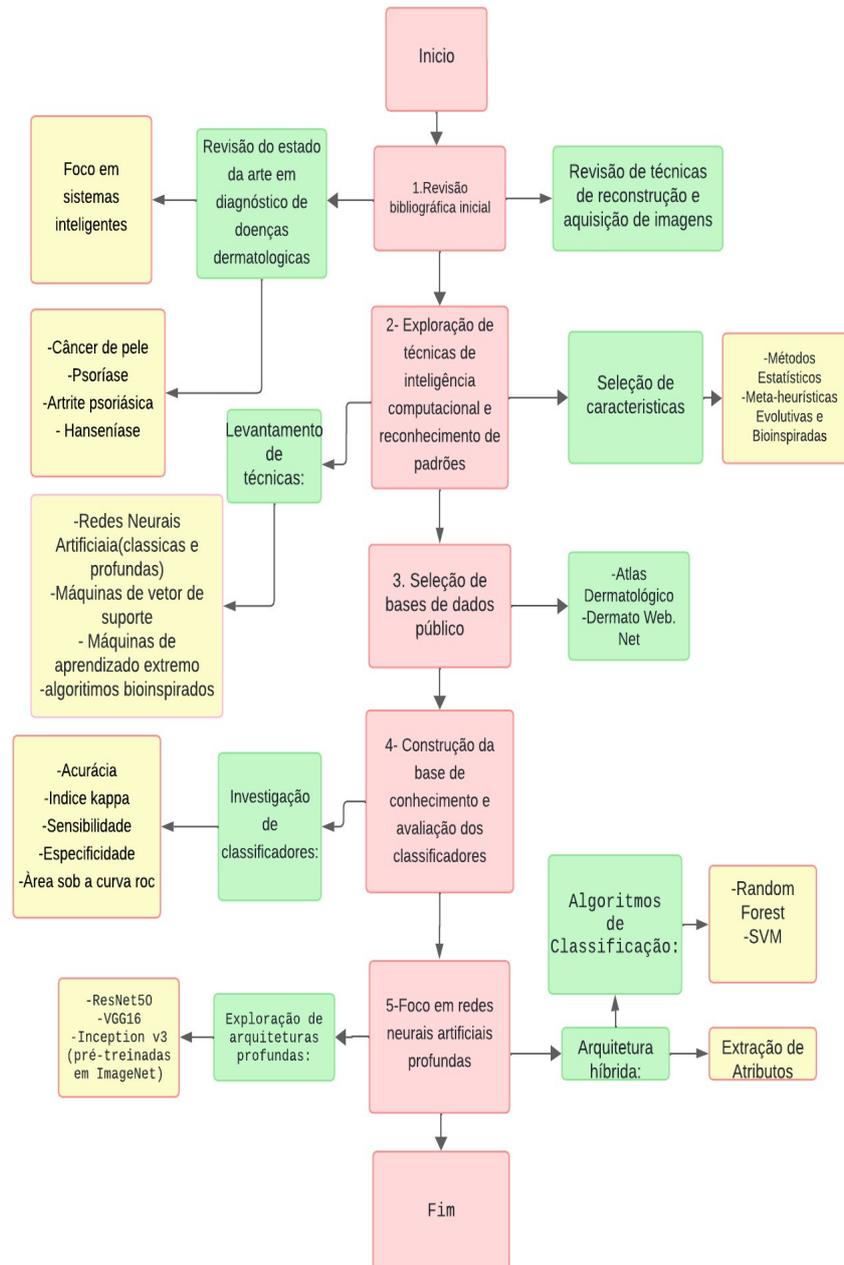
### 1.2.2 Como objetivos específicos deste trabalho têm-se:

- 1 Buscar e compilar bancos de imagens relacionadas a doenças dermatológicas específicas.
- 2 Avaliar diferentes modelos de arquiteturas, como redes neurais convolucionais pré-treinadas, para extrair atributos relevantes das imagens fotográficas.
- 3 Investigar e comparar diferentes classificadores, como árvores de decisão e máquinas de vetores de suporte (SVM), na identificação de doenças dermatológicas.
- 4 Classificar as imagens e obter resultados usando os classificadores selecionados.
- 5 Comparar os resultados obtidos com as técnicas do estado da arte, considerando métricas de desempenho como sensibilidade, especificidade e acurácia.
- 6 Divulgar os resultados por meio de publicações em periódicos e conferências.

### 1.3 METODOLOGIA

O projeto foi estruturado em seis etapas principais, conforme descrito e ilustrado no fluxograma na Figura 1.

Figura 1: Fluxograma das etapas da metodologia



Fonte: Próprio autor (2024).

De acordo com o fluxograma da Figura 1 na primeira fase, denominada Revisão Bibliográfica Inicial, foi realizada uma revisão abrangente sobre técnicas de reconstrução e a-

quisição de imagens, bem como suas aplicações nas áreas de Engenharia e Saúde. Além disso, foram revisadas as técnicas e o estado da arte para o diagnóstico de câncer de pele, psoríase, artrite psoriásica e hanseníase, com um foco especial em abordagens baseadas em sistemas inteligentes.

Na segunda etapa, Exploração de Técnicas de Inteligência Computacional e Reconhecimento de Padrões foram conduzidos um levantamento bibliográfico detalhado das técnicas de Inteligência Computacional e Reconhecimento de Padrões, com ênfase em classificadores baseados no paradigma conexionista. O estudo incluiu redes neurais artificiais clássicas e profundas, máquinas de vetor de suporte, máquinas de aprendizado extremo e algoritmos bioinspirados. Também foram exploradas técnicas de seleção de características, abrangendo métodos estatísticos e meta-heurísticas evolutivas e bioinspiradas.

Na terceira etapa, Seleção de Bases de Dados Públicos, foram selecionadas bases de dados públicas de imagens de lesões de pele, incluindo o Atlas Dermatológico e o Dermato Web. Net. Essas bases contêm um grande número de imagens que abrangem diversas condições dermatológicas e são fundamentais para o desenvolvimento e teste da tecnologia proposta.

Na quarta etapa, Construção da Base de Conhecimento e Avaliação de Classificadores, uma base de conhecimento foi construída e diferentes classificadores foram investigados com base em métricas de avaliação como acurácia, índice kappa, sensibilidade, especificidade e área sob a curva ROC. Esta fase é crucial para garantir que os classificadores selecionados sejam precisos e confiáveis.

Na quinta etapa, Foco em Redes Neurais Artificiais Profundas, o projeto concentrou esforços na exploração de arquiteturas de redes neurais artificiais profundas, especialmente redes convolutivas, devido ao desempenho comprovado dessas arquiteturas em tarefas de classificação de imagens complexas, incluindo aplicações médicas. Foram testadas as redes ResNet50, VGG16 e Inception v3, pré-treinadas na base de dados ImageNet.

Na sexta etapa, Abordagem Híbrida para Diagnóstico Precoce, usou-se inteligência artificial para diagnosticar doenças dermatológicas. A arquitetura híbrida inclui uma rede neural profunda para extrair atributos e um classificador simples na camada final para a classificação. Atributos extraídos foram utilizados em algoritmos como Random Forest e SVM, com balanceamento de dados para evitar viés nos resultados.

## 2 TRABALHOS RELACIONADOS

Foram examinadas as publicações mais recentes que abordam o suporte ao diagnóstico de doenças de pele, com foco nas aplicações de diagnóstico por imagem.

Houve uma priorização dos estudos que se dedicaram especificamente à hanseníase, câncer de pele, psoríase e artrite psoriásica, embora também tenham sido consideradas pesquisas relacionadas a outras condições dermatológicas, desde que incluíssem as doenças mencionadas.

### 2.1 TRABALHOS RELACIONADOS HANSENÍASE

O estudo conduzido por BAWEJA e PARHAR (2016) investiga o uso de uma arquitetura baseada em rede neural convolucional para o reconhecimento de lesões de hanseníase. Para esse propósito, um conjunto de imagens foi coletado, contendo tanto lesões de hanseníase quanto lesões não relacionadas à doença. As imagens foram obtidas a partir da base pública DermnetNZ e de resultados de pesquisa no Google, com um processo de tratamento manual. No total, foram utilizadas 150 imagens, distribuídas igualmente entre treinamento (60%), validação (20%) e teste (20%).

A arquitetura proposta consiste em duas camadas convolucionais, responsáveis por identificar características específicas nas imagens. Essas camadas estão conectadas a dois neurônios que fornecem a probabilidade de a imagem ser positiva ou negativa para hanseníase. O otimizador ADAM foi utilizado durante o treinamento, resultando em uma acurácia média de 91,6%.

O estudo conduzido por JOSHI et al. (2019) propôs um sistema para detecção de lesões na pele saudáveis ou relacionadas a condições como melanoma, eczema e hanseníase. O processo envolveu pré-processamento da imagem (aumento de contraste e conversão para escala de cinza), segmentação (*thresholding* global) e extração de atributos (avaliação da qualidade da imagem e matriz de concorrência de nível de cinza – GLCM). A classificação foi realizada utilizando uma árvore de decisão. O sistema incluiu uma interface gráfica do usuário (GUI) para visualização das imagens e dos resultados obtidos em cada etapa. A acurácia do sistema foi avaliada em 87% utilizando um conjunto de teste com 45 imagens.

No entanto, o estudo não forneceu informações sobre a quantidade total de imagens utilizadas e sua origem. Além disso, apenas a acurácia foi considerada como métrica de qualidade, não sendo utilizadas outras métricas para avaliar o desempenho do sistema.

Na Tabela 1 é apresentada uma síntese dos estudos analisados sobre Hanseníase e outras doenças de pele mencionadas nos estudos, os métodos diagnósticos empregados, as fontes de imagens utilizadas, os resultados alcançados e as observações relevantes feitas pelos autores.

Tabela 1: Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças Dermatológicas, com Ênfase em Hanseníase.

Autor (es)	Ano	Trabalho	Doenças incluídas	Técnicas Utilizadas	Base de imagens	Nº de imagens	Resultados e Observações
BAWEJA e PARHAR	2016	Uso de rede neural convolucional para reconhecimento de lesões de hanseníase.	Hanseníase	Rede Neural Convolucional	Base pública Derm-netz e resultados de pesquisa no Google.	150	Arquitetura de duas camadas convolucionais com otimizador ADAM resultando em acurácia média de 91,6%.
JOSHI et al.	2019	Sistema para detecção de lesões na pele saudáveis ou relacionadas a diversas condições.	Melanoma, Eczema, Hanseníase	Pré-processamento da imagem (aumento de contraste, conversão para escala de cinza), Segmentação (thresholding global), Extração de atributos (avaliação da qualidade da imagem, matriz GLCM), Classificação por árvore de decisão, Interface gráfica do usuário (GUI).	Não específica	Não especificado	O sistema foi avaliado com uma acurácia de 87% usando um conjunto de teste composto por 45 imagens. O estudo não mencionou a quantidade total nem a origem das imagens utilizadas. Além da acurácia, nenhuma outra métrica foi utilizada para avaliar o desempenho do sistema.

Fonte: Próprio auto (2024).

## 2.2 TRABALHOS RELACIONADOS CÂNCER DE PELE

De acordo com ESTEVA et al. (2017), o uso de redes neurais convolucionais na classificação de câncer de pele demonstrou resultados promissores. O estudo utilizou uma base de dados contendo 129.450 imagens clínicas, incluindo 3.374 imagens de dermatoscopia de 18 repositórios online com curadorias clínicas diferentes. Os pesquisadores desenvolveram uma rede neural convolucional profunda baseada na arquitetura GoogleNet Inception v3, que obteve desempenho comparável ao de dermatologistas em três tarefas principais de diagnóstico: classificação de melanoma, classificação de melanoma usando dermatoscopia e classificação de carcinoma.

Os resultados revelaram a eficácia das redes neurais convolucionais na área de dermatologia, alcançando uma precisão geral de  $72,1 \pm 0,9\%$  para condições gerais da pele e  $55,4 \pm 1\%$  para cânceres específicos. No entanto, é importante salientar que o diagnóstico de um dermatologista leva em consideração fatores além da análise visual e dermatoscópica de lesões isoladas. Assim, a abordagem ideal seria combinar a capacidade de classificação de imagens de lesões de pele por meio de redes neurais com a experiência clínica de um dermatologista.

Nesse sentido, é necessário realizar avaliações adicionais para verificar o desempenho desses sistemas em ambientes clínicos reais, levando em consideração as condições e práticas típicas nesses cenários (ESTEVA et al., 2017).

Segundo SHALU e KAMBOJ (2018), o sistema proposto para detecção de câncer de pele melanoma baseia-se na base de imagens MED-NODE, composta por 100 imagens de nevo (refere-se a uma pinta ou marca na pele) e 70 imagens de melanoma. O processo inclui etapas de pré-processamento, segmentação e extração de atributos. Diferentes classificadores foram avaliados, incluindo Naïve Bayes, árvore de decisão e k-vizinhos mais próximos (KNN).

Os resultados mostraram que a árvore de decisão obteve a melhor precisão, alcançando 82,35%. No entanto, observou-se que, de maneira geral, o sistema apresentou maior especificidade do que sensibilidade, indicando uma melhor capacidade em identificar casos benignos em comparação com casos de melanoma.

Portanto, é importante encontrar um entre o número de etapas de pré-processamento e a eficiência computacional, considerando as sensibilidades e especificidades desejadas para a detecção de melanoma (SHALU; KAMBOJ, 2018).

O estudo conduzido por ASTORINO et al. (2020) apresenta uma proposta de aplicação do Aprendizado de Múltiplas Instâncias (MIL – Multiple Instance Learning) para a detecção de melanoma. Em contraste com a abordagem de rotulação individual de instâncias, o objetivo

do MIL é discriminar entre conjuntos de itens positivos e negativos, conhecidos como “bolsas”. Uma bolsa é considerada positiva se ao menos uma de suas instâncias é positiva, e negativa se todas as instâncias forem negativas. O estudo utiliza uma base de dados composta por 40 imagens de melanoma e 80 imagens de nevos comuns obtidas da base de imagens PH2. As imagens foram adquiridas de pacientes classificados como fototipo II (pele clara) ou III (pele clara a moderada) de acordo com a classificação de Fitzpatrick, um sistema amplamente utilizado para categorizar diferentes tipos de pele humana.

Os resultados do estudo revelaram uma acurácia de 92,50%, sensibilidade de 97,50% e especificidade de 87,50% para a melhor configuração do algoritmo. A abordagem de MIL se mostra apropriada para a detecção de câncer de pele, uma vez que permite a extração de informações globais das imagens por meio da análise de sub-regiões, em vez de depender da análise da imagem como um todo. No entanto, os autores ressaltam que a inclusão de etapas adicionais de pré-processamento das imagens poderia aprimorar o desempenho do algoritmo. Além disso, é importante considerar que a base de imagens utilizada é limitada tanto em termos de quantidade quanto de diversidade de fototipos de pele.

O estudo conduzido por PEREZ, AVILA e VALLE (2019) busca examinar os elementos que influenciam a seleção de uma arquitetura de rede neural profunda para a análise de lesões cutâneas, especialmente melanoma. O trabalho investiga nove arquiteturas diferentes, incluindo DenseNet, Dual Path Nets, Inception-v4, Inception-ResNet-v2, MobileNetV2, PNASNet, ResNet, SENet e Xception. Utilizando cinco divisões aleatórias dos dados da base de dados do ISIC Challenge 2017, contendo 2.750 imagens, foram criadas cinco combinações aleatórias para os conjuntos de treinamento (1.750), validação (500) e teste (500), com uma proporção aproximada de 60-20-20%.

O objetivo principal do estudo é analisar a capacidade de classificação entre melanoma e não melanoma (como queratose seborreica e nevo melanocítico). A estratégia proposta consiste em construir conjuntos contendo diversos modelos, com o intuito de melhorar e estabilizar as precisões obtidas, uma vez que as performances dos modelos de aprendizado de máquina costumam flutuar devido a fatores não controláveis, como a seleção do conjunto de treinamento e as condições aleatórias de inicialização das redes. Os resultados obtidos reproduzem as descobertas do desafio mencionado, indicando que as correlações anteriormente relatadas desaparecem completamente ao considerar apenas as redes com melhor desempenho. Sugere-se que fatores não controlados têm uma influência mais significativa do que a escolha da arquitetura de rede. Além disso, o estudo revela que nenhuma característica da rede (como a quantidade de parâmetros) apresenta correlação substancial com as métricas selecionadas, e é

difícil distinguir estatisticamente a maioria das redes umas das outras. Os modelos avaliados alcançaram valores de área sob a curva ROC entre 84% e 91%, aproximando-se da pontuação obtida pelo primeiro colocado no desafio(87,4%) (PEREZ;AVILA; VALLE, 2019).

Na tabela 2 apresenta-se uma síntese dos estudos analisados, abrangendo as doenças como câncer de pele , os métodos diagnósticos empregados, as fontes de imagens utilizadas, os resultados alcançados e as observações relevantes feitas pelos autores.

Tabela 2:Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças Dermatológicas, com ênfase em câncer de pele.

Autor(es)	Ano	Trabalhos	Doenças incluídas	Técnicas utilizadas	Base de imagens	Nº de imagens	Resultados e observações
ESTEVA et al.	2017	Uso de redes neurais convolucionais na classificação de câncer de pele.	Melanoma, Carcinoma e Condições Gerais da Pele.	Rede Neural Convolucional, Arquitetura GoogLeNet Inception v3.	129.450 imagens clínicas, incluindo 3.374 imagens de dermatoscopia de 18 repositórios online com curadorias clínicas diferentes.	129.450 imagens clínicas, incluindo 3.374 imagens de dermatoscopia .	Eficácia das redes neurais convolucionais na classificação de condições de pele, alcançando precisão geral de $72,1 \pm 0,9\%$ e $55,4 \pm 1\%$ para condições gerais e cânceres específicos, respectivamente.

SHA-LU; KAM-BOJ	2018	Sistema proposto para detecção de câncer de pele melanoma utilizando a base de imagens MED-NODE	Melanoma	Pré-processamento, Segmentação, Extração de atributos	Base de 170 imagens MED-NODE (100 imagens de nevo e 70 imagens de melanoma)	O sistema utilizou diferentes classificadores (Naïve Bayes, árvore de decisão, k-vizinhos mais próximos) e obteve melhor precisão com a árvore de decisão (82,35%). No entanto, apresentou maior especificidade do que sensibilidade, sugerindo uma melhor capacidade de identificar casos benignos em comparação com casos de melanoma.
-----------------	------	-------------------------------------------------------------------------------------------------	----------	-------------------------------------------------------	-----------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

ASTO- RINO et al.	2020	Aplicação do Aprendizado de Múltiplas Instâncias (MIL) para a detecção de melanoma	Melanoma	Aprendizado de Múltiplas Instâncias (MIL), aná- lise de sub- regiões	Base de imagens PH2 (40 imagens de melanoma e 80 imagens de nevos comuns)	120	O estudo alcançou uma alta precisão de 92,50%, sensibilidade de 97,50% e especificidade de 87,50% com a melhor configuração do algoritmo. A abordagem de Aprendizado por Instância Múltipla (MIL) permitiu a análise detalhada de sub-regiões das imagens, melhorando a detecção do câncer de pele. Os autores sugerem que a inclusão de etapas adicionais de pré-processamento das imagens poderia melhorar ainda mais o desempenho do algoritmo, dada a limitação da base de imagens em termos de quantidade e diversidade de fototipos de pele.
-------------------------	------	------------------------------------------------------------------------------------------------------	----------	-------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------	-----	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

PEREZ, 2019 AVILA, VALLE	Influência da seleção de arquiteturas de redes neurais profundas na análise de lesões cutâneas.	Melanoma, Queratose seborreica, Nevo melânico.	Redes neurais profundas (DenseNet, Dual Path Nets, Inception-v4, Inception-ResNet-v2, MobileNetV2, PNASNet, ResNet, SENet, Xception).	Base de dados do ISIC Challenge 2017 (2.750 imagens).	O estudo examinou nove arquiteturas diferentes de redes neurais profundas para a análise de lesões cutâneas, especialmente melanoma. Os resultados reproduzem as descobertas do desafio ISIC 2017, mostrando que as correlações anteriormente relatadas desaparecem completamente ao considerar apenas as redes com melhor desempenho. Os modelos avaliados alcançaram valores de área sob a curva ROC entre 84% e 91%, aproximando-se da pontuação obtida pelo primeiro colocado no desafio (87,4%).
--------------------------------	-------------------------------------------------------------------------------------------------	------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Fonte: Próprio auto (2024).

### 2.3 TRABALHOS RELACIONADOS PSORÍASE

De acordo com AIJAZ et al. (2022), a psoríase é uma doença inflamatória crônica da pele mediada pela resposta imune, afetando um grande número de indivíduos em todo o mundo. Estatísticas recentes indicam que aproximadamente 125 milhões de pessoas são afetadas pela psoríase. O uso de técnicas de aprendizagem profunda tem apresentado sucesso na predição de doenças de pele e também pode ser aplicado para a classificação dos diferentes tipos de psoríase. Nesse sentido, o estudo propõe uma aplicação baseada em aprendizagem profunda para a classificação eficaz de cinco tipos de psoríase: placa, gutata, inversa, pustulosa e eritrodérmica, além da predição de pele normal.

No estudo, foram utilizadas 172 imagens de pele normal do conjunto de dados BFL NTU e 301 imagens de psoríase do conjunto de dados Dermnet. As imagens de entrada foram submetidas a pré-processamento, incluindo aumento de dados, aprimoramento e segmentação, seguidos pela extração de características de cor, textura e forma. Foram aplicados dois algoritmos de aprendizagem profunda: redes neurais convolucionais (CNN) e memória de longo prazo (LSTM). Os modelos de classificação foram treinados com 80% das imagens. As previsões reportadas para a CNN e LSTM foram de 84,2% e 72,3%, respectivamente.

Um teste t de amostra emparelhada revelou diferenças significativas nas previsões geradas pelos dois algoritmos de aprendizagem profunda. Essas precisões relatadas no estudo evidenciam o potencial dessa aplicação de aprendizagem profunda para ser utilizada em outras áreas da dermatologia, a fim de melhorar a previsão de diagnósticos (AIJAZ et al., 2022).

O trabalho realizado por DE GOMA e DEVARAJ (2020) exemplifica uma aplicação que se concentra no reconhecimento de lesões de hanseníase e outras doenças de pele comuns nas Filipinas, como acne, dermatite atópica, queratose pilar, psoríase e verrugas. O estudo emprega classificadores baseados em redes neurais artificiais e máquina de vetores de suporte, após pré-processamento da imagem, que inclui etapas como transformação para escala de cinza e aplicação de filtro de mediana. Em seguida, a imagem é segmentada e atributos de cor, borda e textura são extraídos.

No estudo, foram avaliados dois classificadores para detecção de hanseníase, alcançando precisão e sensibilidade acima de 90%, com destaque para as redes neurais artificiais, que obtiveram 96,55% e 93,33%, respectivamente. É relevante mencionar que o algoritmo proposto tem a capacidade de normalizar os valores do pré-processamento de acordo com a cor da pele do paciente, o que contribuiu para os bons resultados obtidos. No entanto, algumas das

doenças consideradas no estudo possuem características em comum, o que pode dificultar a distinção entre elas.

No estudo realizado por DASH et al. (2019), foi proposto um método automatizado para segmentação de lesões de psoríase utilizando uma arquitetura de rede neural convolucional U-Net modificada, chamada de PsLSNet. O objetivo era alcançar uma segmentação precisa das lesões de psoríase a partir de imagens cutâneas. O conjunto de dados utilizado no estudo consistiu em 5.241 imagens de lesões de psoríase de 1.026 pacientes com psoríase. Essas imagens foram capturadas e anotadas por um dermatologista na Clínica e Centro de Pesquisa de Psoríase, Psoriatreat, localizada em Pune, Maharashtra, Índia.

A arquitetura da rede neural U-Net foi modificada para se adequar ao problema de segmentação de lesões de psoríase. Foram realizadas etapas de pré-processamento nas imagens, incluindo aumento de contraste e conversão para escala de cinza, a fim de melhorar a qualidade das imagens e facilitar a segmentação. O método proposto alcançou resultados promissores.

A precisão da segmentação das lesões de psoríase foi de 94,8%, indicando uma alta taxa de acerto na delimitação das áreas afetadas pela doença. Além disso, a sensibilidade do método foi de 89,6%, o que significa que ele foi capaz de identificar corretamente a grande maioria das lesões presentes nas imagens. A especificidade foi de 97,6%, indicando uma baixa taxa de falsos positivos. Esses resultados sugerem que o PsLSNet é uma abordagem eficaz para a segmentação automatizada de lesões de psoríase em imagens cutâneas. A utilização de uma arquitetura de rede neural convolucional modificada, acompanhado de um conjunto de dados significativo e anotações precisas, contribuiu para a obtenção desses resultados

AGGARWAL (2019) conduziu uma avaliação utilizando imagens de pele para identificação e diagnóstico diferencial de lesões de psoríase, acompanhada de outras quatro doenças dermatológicas: acne, dermatite atópica, impetigo e rosácea. O objetivo era investigar o efeito do aumento dos dados no desempenho dos modelos de aprendizado de máquina.

O estudo utilizou redes neurais convolucionais profundas (DCNNs) e um conjunto de dados composto por 332 imagens de pele para acne, 92 para dermatite atópica, 138 para impetigo, 280 para psoríase e 96 para rosácea. Essas imagens foram obtidas a partir de fontes de imagens dermatológicas de código aberto, incluindo o DermNet, Atlas de Dermatologia, Atlas Dermatológico Helênico e Google Images.

Ao aplicar o aumento dos dados, que consiste em técnicas de geração de novas imagens com base nas existentes, observou-se um impacto positivo no desempenho dos modelos de aprendizado de máquina. Cada uma das cinco doenças apresentou um aumento na Área Sob a

Curva (AUC) após o aumento dos dados. Em média, houve um aumento de 0,132 na AUC, com um desvio padrão de 0,033.

No caso específico da psoríase, a AUC melhorou significativamente para 0,87 após o aumento dos dados. Isso indica que o modelo de aprendizado de máquina foi capaz de realizar um diagnóstico diferencial mais preciso das lesões de psoríase com a utilização do aumento dos dados.

Portanto, os resultados deste estudo sugerem que o aumento dos dados pode ser uma estratégia eficaz para melhorar o reconhecimento de imagens e o desempenho de modelos de aprendizado de máquina no diagnóstico de doenças dermatológicas, incluindo a psoríase.

SHRIVASTAVA et al. (2016) realizaram uma avaliação da confiabilidade de um método desenvolvido em um estudo anterior para classificar imagens de pele de pacientes com psoríase como saudáveis ou doentes. O objetivo era verificar a precisão desse método em diferentes tamanhos de dados.

O estudo utilizou uma combinação de análise de componentes principais (PCA) e uma máquina de vetores de suporte (SVM) para a classificação das imagens. O conjunto de dados consistia em 540 imagens de pele, sendo 270 saudáveis e 270 de pacientes com psoríase de origem étnica indiana. Essas imagens foram capturadas e anotadas por um dermatologista na Clínica e Centro de Pesquisa de Psoríase, Psoriatreat, localizada em Pune, Maharashtra, Índia. A confiabilidade do método foi avaliada em 11 tamanhos de dados distintos, abrangendo diferentes proporções de imagens saudáveis e doentes. Os resultados demonstraram um índice médio de confiabilidade de 98,71% ao classificar as imagens de pele como saudáveis ou doentes. Isso indica que o método desenvolvido apresentou alta precisão na identificação da presença de psoríase nas imagens analisadas.

Portanto, com base nesses resultados, concluiu-se que o método desenvolvido apresenta uma confiabilidade significativa na classificação de imagens de pele de pacientes com psoríase como saudáveis ou doentes, independentemente do tamanho do conjunto de dados utilizado.

GEORGE, ALDEEN e GARNAVI (2020) desenvolveram um sistema assistido por computador semi-supervisionado para a pontuação automática da severidade do eritema em imagens de pele de pacientes com psoríase. O objetivo era criar um método preciso e automatizado para avaliar a gravidade dessa característica específica da psoríase nas imagens.

O sistema utilizou técnicas como aprendizado de dicionário baseado em *patch*, floresta aleatória, SVM (máquina de vetores de suporte) e impulsionamento (*boosting*) para realizar a classificação das imagens de acordo com a severidade do eritema. O conjunto de dados con-

sistia em 676 imagens de pele de psoríase, provenientes de 44 pacientes com a condição. Essas imagens foram obtidas por fotógrafos clínicos ao longo de três anos no Royal Melbourne Hospital, localizado na Austrália.

Os resultados obtidos pelo sistema foram avaliados por meio do escore F1, uma métrica que combina precisão e recall. O classificador florestal aleatório obteve um escore F1 de 0,71, demonstrando um desempenho satisfatório na pontuação automática da severidade do eritema nas imagens de pele de psoríase.

Portanto, o sistema desenvolvido pelos pesquisadores mostrou ser eficaz na classificação e pontuação automática da severidade do eritema em imagens de pele de pacientes com psoríase. Essa abordagem semi-supervisionada, combinando diferentes técnicas de aprendizado de máquina, oferece uma alternativa precisa e automatizada para a avaliação clínica dessa característica específica da doença.

A confiabilidade do método foi avaliada em 11 tamanhos de dados distintos, abrangendo diferentes proporções de imagens saudáveis e doentes. Os resultados demonstraram um índice médio de confiabilidade de 98,71% ao classificar as imagens de pele como saudáveis ou doentes. Isso indica que o método desenvolvido apresentou alta precisão na identificação da presença de psoríase nas imagens analisadas.

Portanto, com base nesses resultados, concluiu-se que o método desenvolvido apresenta uma confiabilidade significativa na classificação de imagens de pele de pacientes com psoríase como saudáveis ou doentes, independentemente do tamanho do conjunto de dados utilizado.

A Tabela 3 oferece uma síntese dos estudos relacionados ao apoio ao diagnóstico de doenças de pele, com ênfase na psoríase.

Tabela 3: Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças de Pele, com foco em Psoríase.

Autor(es)	Ano	Trabalho	Doenças incluídas	Técnicas utilizadas	Base de imagens	Nº de imagens	Resultados e Observações
AIJAZ et al.	2022	Aplicação de aprendizagem profunda para classificação de tipos de psoríase e pele normal.	Psoríase (Placa, Gutata, Inversa, Pustulosa, Eritro- de dérmica) e Normal.	Pré-processamento de dados, Aprimoramento, Segmentação, Extração de características (cor, textura, forma), Redes Neurais Convolucionais (CNN), Memória de Longo Prazo (LSTM).	Conjunto de dados BFL NTU (172 imagens de pele normal) e Conjunto de dados Dermnet (301 imagens de psoríase).	473	Precisões reportadas para CNN e LSTM foram de 84,2% e 72,3%, respectivamente. Teste t de amostra emparelhada revelou diferenças significativas nas precisões dos algoritmos.
DASH et al.	2019	Método automatizado para segmentação de lesões de psoríase utilizando a arquitetura de rede neural convolucional U-Net modificada, chamada de PsLS-Net.	Psoríase	Rede Neural Convolucional U-Net modificada, Pré-processamento (aumento de contraste, conversão para escala de cinza).	5.241 imagens de lesões de psoríase de 1.026 pacientes com psoríase.	5.241	O método proposto, PsLSNet, alcançou uma precisão de segmentação de 94,8%, sensibilidade de 89,6% e especificidade de 97,6%. Esses resultados sugerem que o PsLSNet é uma abordagem eficaz para a segmentação automatizada de lesões de psoríase em imagens cutâneas.

DE GO- MA; DE- VARAJ	2020	Reconhe- cimento de lesões de hanseníase e outras doenças de pele co- muns nas Filipinas	Hanseníase Acne, Dermatite Atópica, Queratose de Pilar, Pso Ver çãose, rugas	Pré-processamento da imagem (trans- formação para escala de cinza, aplicação de filtro de mediana), Segmentação, Extra- ção de atributos (cor, borda, textura), Clas- sificadores baseados em redes neurais artificiais e máquina de vetores de suporte	Não especi- ficado	Não espe- ci- ficado	Os classificadores baseados em redes neurais artificiais e máquina de vetores de suporte obteve preci- são e sensibilidade acima de 90% para detecção de hansenía- se.  Destaque para as redes neurais artifici- ais com 96,55% de precisão e 93,33% de sensibilidade.
----------------------------	------	------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------	-------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

AGGARWAL	2019	Avaliação do efeito do aumento dos dados de desempenho de modelos de aprendizado de máquina para identificação e diagnóstico diferencial de lesões de psoríase.	Psoríase, Acne, Dermite Atópica, Impetigo, Rosácea.	Aumento de dados, Redes Neurais Convolucionais Profundas (DCNNs).	de 332 imagens de pele para acne, 92 para dermatite atópica, 138 para impetigo, 280 para psoríase e 96 para rosácea, obtidas de fontes de imagens dermatológicas de código aberto.	Não cifi-	O aumento de dados teve um impacto positivo no desempenho dos modelos de aprendizado de máquina. Cada uma das cinco doenças apresentou um aumento na Área Sob a Curva (AUC) após o aumento dos dados. Em média, houve um aumento de 0,132 na AUC, com um desvio padrão de 0,033. No caso específico da psoríase, a AUC melhorou significativamente para 0,87 após o aumento dos dados, indicando que o modelo de aprendizado de máquina foi capaz de realizar um diagnóstico diferencial mais preciso das lesões de psoríase com a utilização do aumento dos dados.
----------	------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------	-------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

SHRIVASTAVA et al.	2016	Avaliação da confiabilidade de um método para classificar imagens de pele de pacientes com psoríase como saudáveis ou doentes	Psoríase, Pele Saudável	Análise de Componentes Principais (PCA), Máquina de Vetores de Suporte (SVM)	540 imagens de pele, sendo 270 saudáveis e 270 de pacientes com psoríase de origem étnica indiana, capturadas e anotadas por um dermatologista na Clínica e Centro de Pesquisa de Psoríase.	540	540 imagens de pele, sendo 270 saudáveis e 270 de pacientes com psoríase de origem étnica indiana, capturadas e anotadas por um dermatologista na Clínica e Centro de Pesquisa de Psoríase.
GEORGE, AL-DEEN e GARNAVI	2020	Desenvolvimento de um sistema assistido por computador semi-supervisionado para a pontuação automática da severidade do eritema em imagens de pele de pacientes com psoríase	Psoríase	Aprendizado de dicionário baseado em patch, Floresta Aleatória, SVM, Impulsioneamento (Boosting)	676 imagens de pele de psoríase, provenientes de 44 pacientes, obtidas ao longo de três anos no Royal Melbourne Hospital, Austrália	676	O classificador florestal aleatório obteve um escore F1 de 0,71, demonstrando um desempenho satisfatório na pontuação automática da severidade do eritema nas imagens de pele de psoríase.

Fonte: Próprio autor (2024).

## 2.4 TRABALHOS RELACIONADOS ARTRITE PSORIÁSICA

LOVE, CAI e KARLSON (2011), conduziram um estudo que abordou a validação do diagnóstico de artrite psoriásica em registros médicos eletrônicos por meio do uso de técnicas de processamento de linguagem natural. Os pesquisadores desenvolveram algoritmos de processamento de linguagem natural para analisar os registros médicos e identificar casos de artrite psoriásica com base em termos e padrões linguísticos específicos relacionados à condição. Os resultados demonstraram uma concordância significativa entre os diagnósticos obtidos pela análise de linguagem natural e os diagnósticos clínicos realizados por reumatologistas especializados. Concluiu-se que a abordagem de processamento de linguagem natural é uma estratégia eficaz e precisa para validar diagnósticos de artrite psoriásica em registros médicos eletrônicos, contribuindo para aprimorar a qualidade das informações registradas e facilitar a pesquisa e o tratamento da doença (LOVE et al., 2011).

Em um estudo recente conduzido por JALALI-NAJAFABADI et al. (2021), foram desenvolvidos modelos de seleção de características e predição de risco para artrite psoriásica (PsA) utilizando um conjunto de dados genéticos. O estudo empregou dados transversais de 1.462 casos de AP e 1.132 casos de psoríase cutânea (PsC), fazendo uso de alelos HLA de 2 dígitos imputados pelo algoritmo SNP2HLA. Também foi desenvolvido um método de estratificação para mitigar o impacto de possíveis fatores de confusão e ilustrar a influência desses fatores na seleção de recursos.

O conjunto de dados mitigado foi utilizado para treinar sete algoritmos supervisionados, empregando uma abordagem de validação cruzada aninhada estratificada, onde 80% dos dados foram utilizados para o treinamento e 20% foram reservados como conjunto de resistência para validação interna. Em seguida, os modelos de previsão de risco foram validados em um conjunto de dados proveniente do UK Biobank, contendo informações de 1.187 participantes, com um conjunto de recursos sobrepostos ao conjunto de treinamento.

A avaliação do desempenho dos métodos utilizou diversas métricas, incluindo a área sob a curva (AUC), acurácia, precisão, recall, score F1 e análise da curva de decisão (benefício líquido). O melhor modelo foi selecionado considerando três critérios: “menor número de subconjunto de recursos”, “AUC média máxima na validação cruzada aninhada” e boa generalização para o conjunto de dados do UK Biobank.

No conjunto de dados original, com mais de 100 iterações diferentes e sete métodos de seleção de recursos (FS), a variante genética HLA\_C\_06 foi identificada como a mais informativa. No entanto, após a mitigação dos dados, as características genéticas mais importantes,

com base na classificação, foram identificadas como HLA\_B\_27 pelos sete diferentes métodos de seleção de características, o que é consistente com análises anteriores que utilizaram métodos baseados em regressão. No entanto, a precisão preditiva desses recursos isolados após a mitigação foi moderada, com AUC de 0,54 na validação cruzada interna, AUC de 0,53 no conjunto de resistência interno e AUC de 0,55 no conjunto de dados externos.

Foi observado que a adição sequencial de recursos adicionais de HLA, com base na classificação, melhorou o desempenho do modelo de classificação Random Forest. Especificamente, a inclusão de 20 recursos de 2 dígitos selecionados pelo método Interaction Capping (ICAP) demonstrou um aumento no desempenho, com AUC de 0,61 na validação cruzada interna, AUC de 0,57 no conjunto de espera interno e AUC de 0,58 no conjunto de dados externos. Os resultados desse estudo destacam a aplicabilidade do método de estratificação para mitigar recursos de confusão e selecionar recursos relevantes em conjuntos de dados de alta dimensão, considerando potenciais fatos.

No estudo conduzido por NAVARINI et al. (2020), foi investigado o uso de algoritmos de aprendizado de máquina para prever o risco cardiovascular em pacientes com artrite psoriásica (PsA). O estudo utilizou o renomado Estudo Framingham do Coração como base de dados para treinar os modelos de aprendizado de máquina. Foi selecionado um número igual de pacientes com e sem eventos de doença cardiovascular para evitar viés no treinamento. Posteriormente, os modelos foram validados em um conjunto de dados específico de PsA contendo 155 pacientes.

Os resultados obtidos revelaram que os modelos de aprendizado de máquina superaram as calculadoras tradicionais de risco cardiovascular, como o Escore de Risco de Framingham, em termos de sensibilidade e área sob a curva. As áreas sob a curva variaram de 0,7584 a 0,8468 para os diferentes algoritmos de aprendizado de máquinas utilizadas. Além disso, a análise das características destacou a importância de fatores específicos da PsA, como a proteína C reativa e o Índice de Gravidade da Área Psoriática, na previsão do risco cardiovascular. Os modelos de aprendizado de máquina demonstraram uma sensibilidade significativamente maior em pacientes com PsA.

No entanto, é importante ressaltar que o tamanho do conjunto de dados de PsA utilizado neste estudo foi limitado, o que pode afetar a generalização dos resultados. Portanto, pesquisas futuras com conjuntos de dados maiores e mais específicos para a PsA são necessárias para validar ainda mais a eficácia desses modelos de aprendizado de máquina nessa população. Apesar dessa limitação, os resultados obtidos destacam o potencial desses algoritmos para melhorar a previsão do risco cardiovascular em pacientes com PsA, oferecendo novas

perspectivas no campo da medicina personalizada. Essa abordagem pode auxiliar na identificação de pacientes de alto risco e na implementação de estratégias preventivas mais eficazes para essa população.

A Tabela 4 apresenta uma síntese dos trabalhos relacionados ao apoio ao diagnóstico de doenças de pele, com ênfase na artrite psoriásica. Detalham-se as doenças abordadas, os métodos empregados, as fontes de imagens utilizadas, os resultados alcançados e as observações pertinentes feitas pelo autor.

Tabela 4: Panorama de Estudos Relacionados ao Apoio ao Diagnóstico de Doenças de Pele, com foco em Artrite Psoriásica.

Autor (es)	Ano	Trabalho	Doenças incluídas	Técnicas utilizadas	Base de imagens	Nº de imagens	Resultados e Observações
LOVE, CAI e KARLSON	2011	Validação do diagnóstico de artrite psoriásica em registros médicos eletrônicos por meio do uso de técnicas de processamento de linguagem natural	Artrite Psoriásica	Processamento de Linguagem Natural	Registros Médicos Eletrônicos	Não especificado	Houve uma concordância significativa entre os diagnósticos obtidos pela análise de linguagem natural e os diagnósticos clínicos realizados por reumatologistas especializados. A abordagem foi considerada eficaz e precisa.

JALALI-NAJA-FABADI et al.	2021	Desenvolvimento de modelos de seleção de características e predição de risco para artrite psoriásica (PsA) utilizando dados genéticos	Artrite psoriásica, Psoríase Cutânea	Seleção de características, Predição de risco	Dados genéticos	1.462 AP e 1.132 PsC	Utilização de alelos HLA de 2 dígitos imputados pelo algoritmo SNP2HLA. Desenvolvimento de métodos de estratificação para mitigar impacto de fatores de confusão.
NAVARINI et al.	2020	Investigação do uso de algoritmos de aprendizado de máquina para prever o risco cardiovascular em pacientes com artrite psoriásica (PsA)	Artrite Psoriásica	Algoritmos de aprendizado de máquina	Estudo Framingham do Coração	Não especificado	Os modelos superaram as calculadoras tradicionais de risco cardiovascular, como o Escore de Risco de Framingham, em termos de sensibilidade e área sob a curva.

Fonte: Próprio autor (2024).

### 3 FUNDAMENTAÇÃO TEÓRICA

#### 3.1 HANSENÍASE

A hanseníase foi identificada em 1873 pelo cientista Armauer Hansen e é considerada uma das doenças mais antigas, com registros de casos há mais de 4000 anos, na China e na Índia. No entanto, ainda é uma doença desafiadora de erradicar devido à forte relação com a pobreza, falta de acesso a serviços de saúde e educação, além do estigma associado aos pacientes (SANTACROCE et al., 2021).

De acordo com DE SOUZA et al. (2021), a hanseníase é uma doença infecciosa que afeta a pele e os nervos periféricos, resultando em estigmas e incapacidades que limitam a participação comunitária e o engajamento social. Segundo o relatório da OMS de 2020, houve 208.619 novos casos de hanseníase registrados globalmente em 2018, com uma taxa de prevalência correspondente a 0,2 em cada 10.000 indivíduos. O Brasil apresentou 27.893 novos casos em 2019, sendo o segundo país com o maior número de registros da doença no mundo. Além disso, o Brasil registrou o maior número de casos com deformidades, seguido pela Índia.

A meta estabelecida pela Organização Mundial da Saúde (OMS) é eliminar a hanseníase e interromper sua transmissão em nível nacional ou regional até 2030. Para alcançar esse objetivo, é de suma importância realizar a detecção ativa de casos e o rastreamento de contatos. Portanto, é imperativo melhorar as ferramentas de diagnóstico para identificar novos casos da doença de maneira precoce e precisa (DE SOUZA et al., 2021).

A hanseníase é uma doença crônica provocada pela infecção da bactéria *Mycobacterium leprae*, que pode afetar indivíduos de qualquer faixa etária. Quando não tratada em estágios iniciais, a doença tende a progredir e se tornar transmissível. Essa progressão é caracterizada por um processo gradual e lento, podendo levar ao desenvolvimento de incapacidades físicas (BRASIL, 2017, p. 8; BRASIL, 2022).

Embora o período de incubação da doença não seja precisamente conhecido, estima-se que, em média, dura cerca de cinco anos. Existem relatos de casos nos quais os sintomas surgiram após apenas um ano do contato suspeito, enquanto em outros a incubação pode se prolongar por até 20 anos ou mais. Essa compreensão aprofundada da dinâmica de transmissão e incubação da hanseníase é crucial para o diagnóstico precoce e o controle eficaz da doença (BRASIL, 2022).

A hanseníase é transmitida pela bactéria *Mycobacterium leprae* através de secreções das vias aéreas superiores e gotículas de saliva, principalmente durante o contato próximo com

pacientes que estão em fases avançadas da doença e ainda não iniciaram o tratamento. A transmissão ocorre pelo ar e não por meio de objetos utilizados pelo paciente. Os principais sintomas da hanseníase incluem alterações na sensibilidade térmica, dor, tato e força muscular, bem como o surgimento de manchas na pele e nódulos. Se não tratada adequadamente, a doença pode levar a incapacidades permanentes. No entanto, o tratamento adequado da hanseníase possibilita a cura, interrompe a transmissão da doença e previne a ocorrência de incapacidades físicas (BRASIL, 2017; BRASIL, 2022).

O diagnóstico médico da hanseníase é geralmente baseado na avaliação clínica do paciente, que depende do reconhecimento dos sinais e sintomas característicos da doença. As lesões cutâneas costumam ser a primeira manifestação clínica observada. Em alguns casos, pode ser necessário realizar exames laboratoriais para auxiliar no diagnóstico, sendo a baciloscopia o mais comumente utilizado. Esse exame envolve a coleta de serosidade cutânea para análise microscópica. No entanto, é importante destacar que a baciloscopia pode apresentar resultados negativos nas fases iniciais da doença. Portanto, o diagnóstico precoce é de extrema importância para evitar a transmissão da doença, complicações e deficiências a longo prazo (CHEN et al., 2022).

De acordo com o estudo de SOUZA et al. (2021), o diagnóstico convencional da hanseníase é baseado na avaliação clínica, incluindo sinais como lesões cutâneas, espessamento dos nervos e detecção do patógeno em amostras de pele ou tecido histológico. Diferentes classificações, como a Classificação de Madrid e a classificação operacional da OMS, são utilizadas para auxiliar no entendimento e tratamento da doença. A Classificação de Madrid divide os pacientes em quatro categorias, enquanto a classificação operacional da OMS divide-os em hanseníase paucibacilar e hanseníase multibacilar, com base no número de lesões cutâneas.

Para fins de tratamento, a hanseníase é classificada em duas categorias principais pela Organização Mundial da Saúde: paucibacilar (PB), que apresenta poucos ou nenhum bacilo nos exames, e multibacilar (MB), que possui uma grande quantidade de bacilos (BRASIL, 2022).

De acordo com o Ministério da Saúde (BRASIL, 2022), o Brasil teve um papel central na criação de uma classificação lógica para a hanseníase, com base nos estudos pioneiros de Rabbello (1936), que revelaram pela primeira vez a natureza contínua da doença, variando conforme a resposta imune do indivíduo à infecção. Isso levou à elaboração da Classificação de Madri (1953), que enfatiza os aspectos clínicos e complementares do diagnóstico, considerando a morfologia das lesões cutâneas e as manifestações neurológicas como critérios fundamentais. Essa classificação identifica dois extremos estáveis e opostos da doença (formas tuberculóide e virchowiana), bem como formas intermediárias e instáveis (hanseníase dimorfa)

e uma forma inicial com sintomas clínicos sutis (forma indeterminada). Internacionalmente, a Classificação de Ridley e Jopling (1966) é adotada, mantendo os princípios da Classificação de Madri, mas baseando-se especialmente nos aspectos histopatológicos das lesões para categorizar as formas clínicas da doença.

A hanseníase apresenta diferentes formas clínicas, cada uma com características distintas. A hanseníase indeterminada Figura 2 (b) é uma fase inicial da doença que pode passar despercebida. Nessa fase, a lesão de pele geralmente é uma única mancha mais clara do que a pele ao redor, sem alterações de relevo, bordas mal delimitadas e secas (BRASIL, 2017; BRASIL, 2022).

A hanseníase tuberculóide Figura 2 (d) é frequentemente caracterizada por uma placa elevada em relação à pele ao redor, que pode ser totalmente anestésica ou apresentar bordas elevadas, bem delimitadas e um centro claro em forma de anel ou círculo. Quando afeta mais de um nervo, a apresentação da doença é geralmente assimétrica (BRASIL, 2017; BRASIL, 2022).

A hanseníase dimorfa esta forma apresenta características tanto da hanseníase tuberculóide Figura 2 (d) quanto da Lepromatosa Figura 2 (c). Os pacientes podem apresentar lesões cutâneas e comprometimento neurológico moderado. É a forma mais comum da doença e se caracteriza por múltiplas manchas avermelhadas ou esbranquiçadas na pele, com bordas elevadas e mal delimitadas na periferia. Também pode apresentar múltiplas lesões bem delimitadas, semelhantes à forma tuberculóide, porém com uma borda externa pouco definida (BRASIL, 2017; BRASIL, 2022).

A hanseníase virchowiana também conhecida como Lepromatosa Figura 2 (c) é a forma mais contagiosa da doença. Os sinais e sintomas podem ser discretos, pois o paciente geralmente apresenta lesões difusas, simétricas e de limites imprecisos. Desta forma, é comum observar a pele avermelhada, seca, infiltrada e com os poros dilatados (BRASIL, 2017; BRASIL, 2022).

O diagnóstico diferencial é essencial para distinguir a hanseníase de outras doenças que podem apresentar sinais e sintomas semelhantes durante o exame clínico. Dado o amplo espectro de manifestações dermatológicas e neurológicas da hanseníase, é fundamental realizar uma avaliação criteriosa para evitar erros diagnósticos e contribuir para uma identificação precisa da doença (KUNDAKCI & ERDEM, 2019).

No processo de diagnóstico diferencial, são considerados diversos aspectos específicos para cada tipo de hanseníase de acordo com a Figura 2(a), (b), (c) e (d). Para o diagnóstico diferencial da hanseníase indeterminada Figura 2 (a), algumas condições que devem ser conside-

radas são: Pitiríase Alba, Dermatite Seborreica, Nevo Acromico, Nevo Anêmico e Vitiligo. No caso da hanseníase tuberculóide, é importante distinguir de condições como Dermatofitose do Corpo, Granuloma Anular, Sarcoidose, Tuberculose Cutânea Verrucosa e Psoríase. Para a hanseníase dimorfa, algumas condições que podem ser confundidas incluem Esclerodermia Linear, Dermatofitose do Corpo, Parapsoríase em Placas e Psoríase. Por fim, para a hanseníase virchowiana, é importante considerar o diagnóstico diferencial com condições como Neurofibromatose, Xantoma Tuberoso e Condrodermatite Nodular Crônica da Hélice.

Figura 2 : Amostras do banco de imagens que representam os diferentes tipos de hanseníase.



(a) Intermediária

(b) Indeterminada

(c) Lepromatosa

(d) Tuberculóide

Fonte: (SILVA, 1995).

### 3.2 PSORÍASE (PSO)

A psoríase é uma doença crônica da pele que tem manchas elevadas, vermelhas e escamosas sobre as diferentes regiões do corpo humano. A segmentação das lesões de psoríase é uma etapa pré-requisito para o diagnóstico da psoríase. Na prática atual, a segmentação da lesão da psoríase é feita manualmente por inspeção visual, o que é um processo demorado e subjetivo. (RAJ; LONDHE; SONAWANE, 2020).

De acordo com JIN et al. (2023), Pacientes com psoríase (PSO) frequentemente apresentam diversas comorbidades, tais como condições cardiometabólicas e problemas de saúde mental. Além disso, é comum que cerca de 30% desses pacientes sejam diagnosticados com artrite psoriática (PsA). De acordo com AIJAZ et al.(2022), segue abaixo alguns tipos de psoríase:

1. Psoríase em placa. A psoríase em placas tende a aparecer na superfície da pele na forma de manchas espessas e vermelhas como mostrado na Figura 3 (d).
2. Psoríase gutata. A psoríase gutata é uma forma de infecção da pele que aparece na superfície da pele em manchas vermelhas e coceiras em forma de gota de lágrima ,conforme ilustrado na Figura 3 (C).

3. Psoríase inversa ou invertida. A psoríase inversa, também conhecida como psoríase oculta, é uma forma de psoríase que tende a infectar as dobras cutâneas, áreas em que uma região da pele se esfrega contra outra região da pele, como mostrado na Figura 3 (a).

4. Psoríase pustulosa. Saliências brancas cheias de pus dentro ou ao redor de manchas escamosas vermelhas são representativas da psoríase.

5. Psoríase eritrodérmica. Sendo um dos tipos graves de psoríase, a psoríase eritrodérmica envolve inflamação com erupções cutâneas descamativas que queimam, de acordo com a Figura 3 (b).

A psoríase pustular generalizada Figura 3 (f) é uma doença cutânea autoinflamatória rara caracterizada por erupção generalizada de pústulas neutrofilicas estéreis (PUIG et al;2023).

Segundo Kim et al. (2023), a maioria das lesões de psoríase está relacionada à pele, sendo a psoríase vulgar responsável por aproximadamente 90% de todos os casos.

Conforme ilustrado na Figura 3 (e), a psoríase palmo-plantar afeta as palmas das mãos e as solas dos pés. Para a psoríase ungueal, pode-se fazer referência à Figura 3 (g), que mostra os efeitos da doença nas unhas.

Figura 3: Amostras do banco de imagens que representam os diferentes tipos de psoríases



(a) Invertida



(b) Eritrodérmica



(c) Gutata



(d) Psoríase em placas



(e) Palmo e planta



(f) Pustular



(g) Ungueal

Fonte: (Dermato Web.net; SILVA, 1995).

### 3.3 ARTRITE PSORIÁSICA (PsA)

A artrite psoriásica (PsA) Figura 4, é uma doença reumática inflamatória que pode afetar as articulações periféricas, articulações axiais e enteses (LIU et al., 2022). A artrite psoriática (PsA) é uma doença musculoesquelética inflamatória heterogênea crônica. As manifestações inespecíficas e muitas vezes sutis tornam o diagnóstico precoce e o tratamento subsequente desafiadores. Foi descrita por Moll e Wright em 1973 como uma doença distinta com 5 sub-grupos clinicamente sobrepostos que incluíam envolvimento articular axial e periférico. As manifestações da doença variam de alguns locais de entesite a envolvimento grave de numerosas articulações periféricas, levando a deformidades graves denominadas artrite mutilante. (AHMAD et al., 2021). Podemos observar imagens desta doença na Figura 4 (a),(b) e (c).

A PsA ocorre em grande parte em associação com a doença de pele psoríase (PSO), com cerca de um terço dos indivíduos com PsO desenvolvendo PSA (LIU et al., 2022). A psoríase geralmente precede o desenvolvimento de artrite; no entanto, 20% dos pacientes apresentam PsA antes da psoríase e 15% terão manifestações cutâneas e articulares simultaneamente. Tanto homens quanto mulheres são igualmente afetados, e é menos comum em asiáticos e negros e quase ausente em populações nativas da América do Sul e Austrália (AHMAD et al., 2021).

A detecção precoce do PsA em pacientes com PSO é um importante determinante do desfecho clínico e da qualidade de vida a longo prazo do paciente, mas pode ser desafiadora devido à apresentação heterogênea do PsA, com apenas manifestações subclínicas em estágios iniciais da doença (LIU et al., 2022).

Figura 4: Amostras do banco de imagens que representa a artrite psoriásica.



(a) Artrite psoriásica

(b) Artrite psoriásica

(c) Artrite psoriásica

Fonte: (Dermato Web.net; SILVA ,1995).

### 3.4 CÂNCER DE PELE

Os tecidos do corpo humano consistem em células normais que possuem a capacidade de se multiplicar por meio de um processo natural e contínuo. O ciclo celular da maioria das células normais envolve crescimento, multiplicação e morte de maneira ordenada. No entanto, o crescimento das células cancerosas difere do crescimento das células normais, pois elas se multiplicam de forma descontrolada em vez de morrerem, resultando na formação de novas células anormais.

O câncer é caracterizado pela perda do controle da divisão celular e pela capacidade de migrar para outros tecidos, processo conhecido como metástase. Os diferentes tipos de câncer correspondem aos vários tipos de células presentes no corpo. Portanto, o câncer de pele ocorre devido ao crescimento anormal e descontrolado das células da pele (INSTITUTO NACIONAL DO CÂNCER, 2011).

A pele é composta por três camadas principais: epiderme, derme e hipoderme. A maioria dos cânceres de pele tem início na epiderme, que é a camada mais superficial. Esses cânceres são classificados de acordo com as células que deram origem a eles, sendo os mais comuns o melanoma Figura 5 (c) e os cânceres não melanoma, como o carcinoma basocelular Figura 5 (a) e o carcinoma epidermoide Figura 5 (b) (AMERICAN CANCER SOCIETY, 2019).

De acordo com Machado et al. (2021), a exposição à radiação ultravioleta solar é considerada o principal fator de risco para o desenvolvimento do câncer de pele, e estudos têm mostrado a relação entre a exposição solar intensa e intermitente e o CBC e melanoma. Além disso, fatores como características genéticas, presença de nevos displásicos e tabagismo também são mencionados como riscos para o câncer de pele (MACHADO et al., 2021).

De acordo com VILELA et al. (2021) apesar de as neoplasias serem a segunda causa de mortalidade no Brasil, a maioria dos cânceres de pele possui um bom prognóstico quando

diagnosticados e tratados precocemente, no Brasil, a situação epidemiológica reflete a realidade global, apresentando uma estimativa anual de 185.380 casos em 2020, com 4.120 óbitos registrados em 2018. O câncer de pele é a neoplasia com maior incidência anual no país, com uma incidência estimada de 2 a 3 milhões de casos do tipo não melanoma e 132.000 casos do tipo melanoma por ano . Estima-se que 1 em cada 3 cânceres diagnosticados seja de pele.

Figura 5: Amostras do banco de imagens que representam os diferentes tipos de câncer de pele.



( a) Carcinoma basal de células



(b) Carcinoma de células escamosa



(C) Melanoma

Fonte: (SILVA, 1995).

### 3.5 APRENDIZADO PROFUNDO

O sistema proposto neste estudo tem como objetivo lidar com imagens que apresentam lesões de pele, o que requer o tratamento adequado dessas imagens. Para que algoritmos de aprendizado de máquina possam ser aplicados, é necessário que as imagens sejam convertidas

em uma representação numérica, o que pode ser feito por meio da criação de um mapa de características.

Dessa forma, as informações originais da imagem são preservadas, permitindo que esse mapa de características evidencie as partes mais relevantes das imagens.

De acordo com COPPIN (2010), o aprendizado está intrinsecamente ligado à inteligência, pois um sistema que é capaz de aprender a realizar uma determinada tarefa pode ser considerado inteligente. O processo de aprendizagem envolve a aquisição de novas formas de conhecimento, o desenvolvimento motor, habilidades cognitivas, organização do conhecimento e a descoberta de novos fatos e teorias por meio da observação e experimentação.

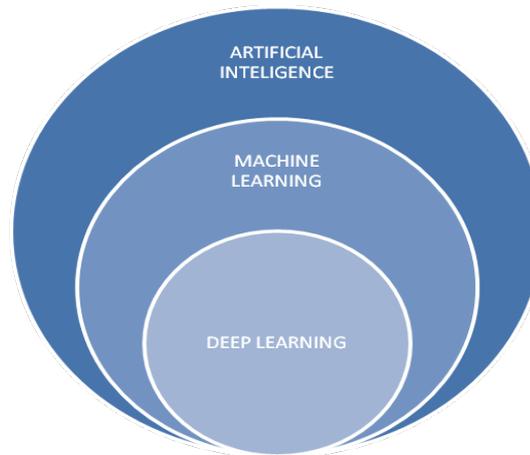
Desde os primórdios da era dos computadores, pesquisadores têm se empenhado em incorporar algumas dessas capacidades em sistemas computacionais. No entanto, esse tem sido um desafio significativo para a área de Inteligência Artificial (IA).

O estudo e a modelagem dos processos de aprendizagem em computadores, assim como suas diversas manifestações, são o objetivo principal do campo de estudo do aprendizado de máquina (SANTOS, 2005).

Segundo SANTOS (2005, citado por MITCHELL, 1997), o aprendizado de máquina consiste em adquirir conhecimento por meio da experiência, ou seja, à medida que uma tarefa é executada, o sistema aprende a melhor forma de resolvê-la. Além disso, o aprendizado de máquina envolve a estruturação do conhecimento existente para alcançar uma compreensão mais profunda do processo de aprendizagem.

O uso do aprendizado de máquina (ML) se expandiu amplamente na pesquisa e tem sido aplicado em diversas áreas, como mineração de texto, detecção de spam, recomendação de vídeos, classificação de imagens e recuperação de conceitos multimídia. Nesses contextos, o aprendizado profundo (DL), uma subcategoria do ML (Figura 6), tem sido amplamente empregado (ALZUBAIDI et al., 2021).

Figura 6: Grupo de aprendizado profundo



Fonte: Adaptado de ALZUBAIDI et al., 2021.

O DL se inspira nos padrões de processamento de informações que podem ser observados no cérebro humano. Uma das principais características do DL é sua capacidade de aprender automaticamente a partir dos dados, sem necessidade de regras definidas por humanos. Em vez disso, utiliza grandes conjuntos de dados para mapear as entradas em rótulos específicos. Essa abordagem é possível através do uso de várias camadas de algoritmos conhecidas como redes neurais artificiais (RNAs), cada uma delas oferecendo uma interpretação diferente dos dados fornecidos a elas (ALZUBAIDI et al., 2021). Os tipos mais famosos de redes de aprendizagem profunda são: redes neurais recursivas (RvNNs), RNNs e CNNs (ALZUBAIDI et al., 2021).

### 3.6 REDES NEURAI CONVOLUCIONAIS (CNNs)

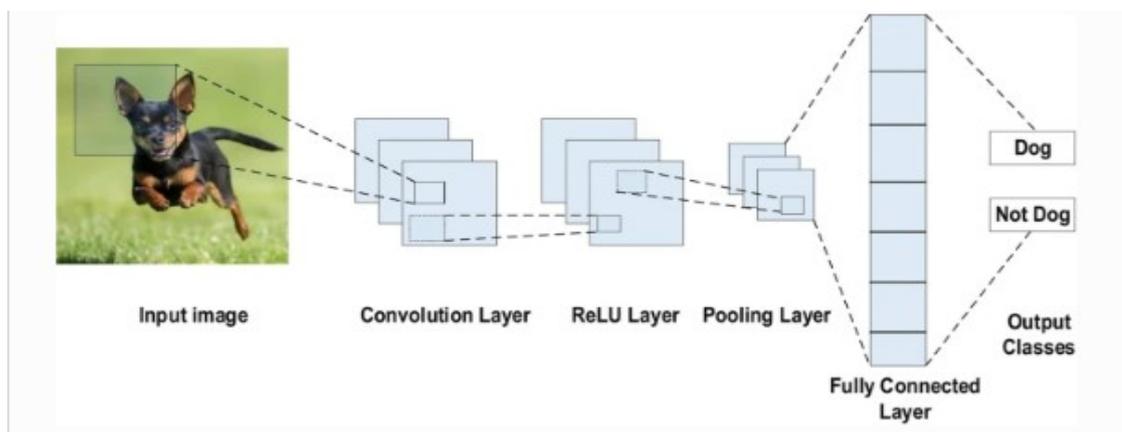
Na área do aprendizado profundo (DL), as redes neurais convolucionais (CNNs) emergem como o algoritmo mais proeminente e amplamente utilizado. Uma das principais vantagens das CNNs sobre suas antecessoras é sua capacidade de identificar automaticamente características relevantes sem a necessidade de intervenção humana. Essas redes têm encontrado aplicações extensivas em diversos campos, como visão computacional, processamento de fala, reconhecimento facial, entre outros. A estrutura das CNNs foi inspirada nos neurônios presentes nos cérebros de humanos e animais, assemelhando-se a uma rede neural convencional (ALZUBAIDI et al., 2021).

Uma das vantagens fundamentais das redes neurais convolucionais (CNNs), quando comparadas aos seus antecessores, é sua capacidade de detectar automaticamente características significativas sem intervenção humana direta. Esse atributo tem sido um dos principais motivos para a ampla adoção e popularidade das CNNs nas aplicações mencionadas (ALZUBAIDI et al., 2021).

Nos últimos 10 anos, várias arquiteturas de redes neurais convolucionais (CNNs) surgiram e desempenharam um papel crucial na melhoria significativa do desempenho em várias aplicações. As modificações realizadas nessas arquiteturas incluíram reformulação estrutural, técnicas de regularização, otimização de parâmetros e outras abordagens inovadoras. O avanço mais notável no desempenho das CNNs ocorreu com o aumento da profundidade das redes, permitindo a aprendizagem de representações mais complexas dos dados. A reorganização das unidades de processamento e o desenvolvimento de novos blocos arquiteturais também contribuíram significativamente para esse progresso (ALZUBAIDI et al., 2021).

Uma Rede Neural Convolucional (RNC) é composta por três principais tipos de camadas: a camada convolucional, a camada de agrupamento (*pooling*) e a camada totalmente conectada (SOUZA et al., 2020). Podemos verificar a arquitetura CNN na Figura 7.

Figura 7: Arquitetura CNN para classificação de imagens.



Fonte: (ALZUBAIDI et al., 2021).

As camadas convolucionais são responsáveis por aplicar filtros sobre a imagem de entrada para obter uma série de mapas de características, sendo um mapa para cada filtro utilizado. Esses mapas de características destacam diferentes padrões e características presentes nas imagens (SOUZA et al., 2020).

As camadas de agrupamento entram em ação para reduzir a amostra dos mapas de características obtidas. Essa redução é realizada por meio de operações de agrupamento, como o max pooling, que mantém as características mais proeminentes dos mapas enquanto reduz a sua dimensão espacial (SOUZA et al., 2020).

Camada totalmente conectada, onde os mapas de características resultantes das camadas convolucionais e de agrupamento são tratados como vetores e passam por uma ou mais camadas densamente conectadas, que efetuam a classificação final da imagem em uma ou mais categorias pré-definidas (SOUZA et al., 2020).

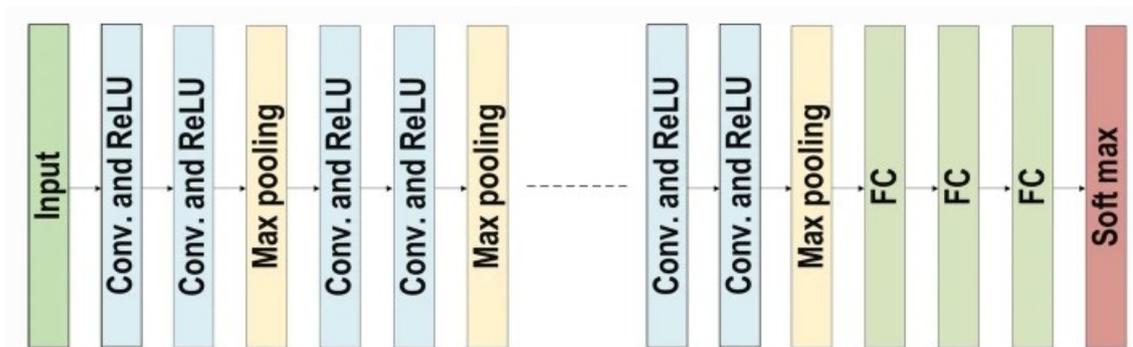
### 3.7 GRUPO DE GEOMETRIA VISUAL (VGG)

Após a comprovação da eficácia das CNNs no campo do reconhecimento de imagens, Simonyan e Zisserman propuseram um princípio de design simples e eficiente para essas redes, conhecido como Grupo de Geometria Visual (VGG) Figura 8 . Esse projeto inovador marcou um importante avanço na área (ALZUBAIDI et al., 2021).

A VGG16 é uma arquitetura de rede neural convolucional que conquistou uma boa colocação na competição *ImageNet* em 2014. Amplamente considerada uma das melhores arquiteturas de modelos de visão, a VGG16 alcançou uma precisão de 90,1% para o top-5 e 71,3% para o top-1 no conjunto de dados *ImageNet* (RAY; GUPTA; AL, 2020).

A arquitetura de rede VGG utiliza camadas convolucionais 3 x 3 empilhadas umas sobre as outras com profundidade crescente. Quando referimos VGG16, significa que a rede é composta por 16 camadas (SILVA, 2018).

Figura 8: A arquitetura do VGG



Fonte: (ALZUBAIDI et al., 2021).

*ImageNet* é um extenso e diversificado banco de dados de imagens que abrange uma ampla variedade de categorias. Ele é composto por 12 "subárvores" principais, cada uma representando grupos específicos de objetos, como mamíferos, aves, peixes, répteis, anfíbios, veículos, móveis, instrumentos musicais, formações geológicas, ferramentas, flores e frutas (DENG et al., 2009). Essas subárvores contêm um total de 21.841 conjuntos de sinônimos (*synsets*) e impressionantes 14.197.122 milhões de imagens. o *ImageNet* contém *synsets* de alta qualidade (precisão de aproximadamente 99%) e imagens de resolução total com um tamanho médio em torno de  $400 \times 350$  pixels (DENG et al., 2009).

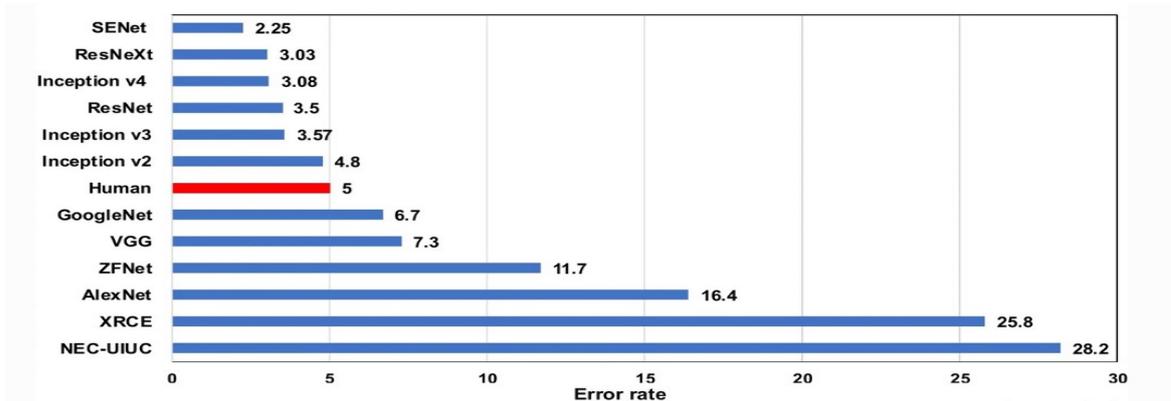
Devido à sua abrangência e tamanho, o *ImageNet* se torna uma ferramenta valiosa para inúmeras aplicações de reconhecimento visual, incluindo o reconhecimento de objetos, classificação de imagens e localização de objetos. A diversidade e quantidade de dados disponíveis no *ImageNet* possibilitam o treinamento de modelos de aprendizado de máquina mais precisos e robustos, que podem ser aplicados em diversos campos, desde tecnologias de reconhecimento de objetos em dispositivos móveis até sistemas de vigilância baseados em visão computacional (DENG et al., 2009).

Desde 2010, como parte do *Pascal Visual Object Challenge*, tem sido realizado anualmente o *ImageNet Large-Scale Visual Recognition Challenge* (ILSVRC) Figura 9 . Essa competição utiliza um subconjunto do ImageNet, composto por aproximadamente 1000 imagens em cada uma das 1000 categorias (Krizhevsky et al., 2017).

O ImageNet contém *synsets* de alta qualidade (precisão de aproximadamente 99%) e imagens de resolução total com um tamanho médio em torno de  $400 \times 350$  pixels (DENG et al., 2009).

Os resultados demonstraram que, em muitos casos, os modelos de aprendizado profundo superaram significativamente a precisão humana na tarefa de classificação de imagens de acordo com a Figura 9 (ALZUBAIDI et al., 2021).

Figura 9: ImagiNet, aprendizagem profunda em comparação com o ser humano.



Fonte: (ALZUBAIDI et al., 2021).

Redes neurais profundas têm sido amplamente utilizadas em uma variedade de tarefas de aprendizado de máquina, desde reconhecimento de imagem até processamento de linguagem natural. No entanto, esses modelos são propensos a problemas como *overfitting* (sobreajuste ou superajuste) e instabilidade no treinamento (COSTA, 2019). Para lidar com esses desafios, técnicas de regularização, como *dropout* e normalização em lote, foram desenvolvidas e incorporadas em arquiteturas de rede neural.

O *dropout* é uma técnica de regularização que funciona desligando aleatoriamente um número especificado de neurônios em uma camada durante o treinamento. Isso força a rede a aprender representações mais robustas e reduz a dependência de neurônios individuais ou combinações específicas de neurônios (BUZUTI; THOMAZ, 2018).

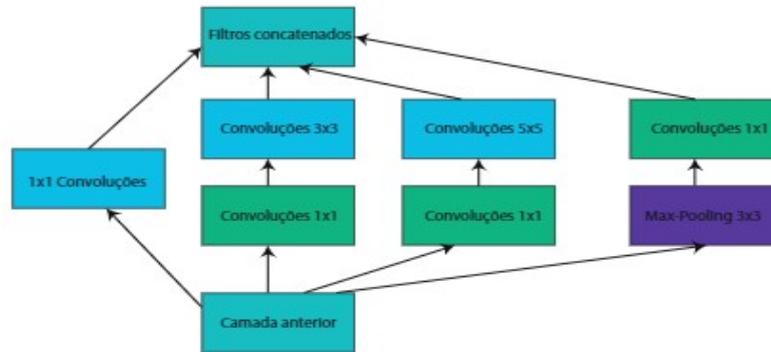
A normalização em lote é uma técnica que normaliza a entrada de cada camada para ter uma média próxima de zero e uma variância próxima de um. Isso ajuda a estabilizar o treinamento e reduz a sensibilidade dos modelos à inicialização dos pesos, permitindo o uso de taxas de aprendizado mais altas (IOFFE; SZEGEDY, 2015).

### 3.8 INCEPTION-V3

A arquitetura Inception apresentada na Figura 10 conquistou a vitória no ILSVRC em 2014. Inception-v3 é uma versão específica dessa arquitetura voltada para a tarefa de classificação de imagens, tendo sido treinada no conjunto de dados ImageNet. No ano de 2015, essa rede apresentou um bom desempenho com um custo computacional relativamente baixo, graças à redução dos parâmetros estimados pela rede. Essa otimização resultou em um melhor

desempenho computacional durante o treinamento em comparação com as redes VGG (SILVA, 2018).

Figura 10: Arquitetura Inception



Fonte: (SOUZA et al., 2020).

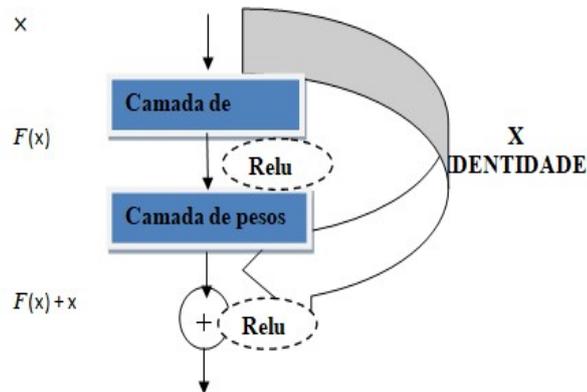
### 3.9 RESNET (REDE RESIDUAL)

A ResNet (Figura 11) é uma arquitetura inovadora de rede neural convolucional desenvolvida pela Microsoft, que se destacou ao vencer o prestigiado concurso ILSRVC em 2015, utilizando o conjunto de dados *ImageNet*. De forma surpreendente, a ResNet demonstrou um desempenho superior ao dos seres humanos na tarefa de classificação de imagens (SOUZA et al., 2020).

Um desafio crítico enfrentado ao trabalhar com redes mais profundas é a degradação do desempenho à medida que a profundidade da rede aumenta. A precisão do modelo satura e, posteriormente, decai rapidamente, apesar do potencial de representação adicional que camadas mais profundas deveriam oferecer. No entanto, a degradação observada não é atribuída ao *overfitting* e, curiosamente, adicionar mais camadas à rede leva a um aumento no erro de treinamento (SOUZA et al., 2020). Para superar esse problema, a ResNet foi projetada utilizando blocos residuais. Estes blocos apresentam conexões de salto (*skip connections*), permitindo que a rede aprenda as diferenças residuais entre as camadas.

Essa abordagem inovadora facilitou o treinamento de redes extremamente profundas, pois as conexões de salto proporcionam um caminho direto para que as informações originais sejam transmitidas através da rede, preservando a representação essencial (SOUZA et al., 2020).

Figura 11: Um bloco de construção do aprendizado residual



Fonte: (SOUZA et al., 2020), adaptado.

### 3.10 BALANCEAMENTO DE DADOS

Ao considerar que a quantidade de informações para cada classe (por exemplo, doenças) é discrepante, o uso desequilibrado da base de dados pode prejudicar a construção e as previsões dos modelos. Portanto, é necessário balancear os dados.

Existem duas abordagens possíveis para esse balanceamento: (1) subamostragem, que envolve usar a classe minoritária como referência e excluir dados das demais; ou (2) sobreamostragem, que consiste em utilizar a classe majoritária como referência e adicionar dados às demais. A primeira opção é mais simples, mas está associada à perda de informação, o que impacta negativamente no desempenho dos classificadores. Por outro lado, na sobreamostragem, apesar do custo computacional geralmente maior, nenhuma informação é descartada.

Desse modo, a sobreamostragem de Minoria Sintética (*SMOTE*) é adotada como uma técnica reconhecida na literatura para abordar o desequilíbrio, gerando dados adicionais para a classe minoritária. Essa abordagem se baseia na criação de exemplos ao longo das linhas que conectam um ponto a alguns de seus  $K$ -vizinhos mais próximos. Então, o gerador de dados sintéticos não apenas copia os dados existentes, mas cria dados novos que se assemelham em forma ou valores aos dados já presentes (ELREEDY E ATIYA ;2019).

### 3.11 CLASSIFICAÇÃO

Um classificador é um algoritmo que possui a capacidade de atribuir uma amostra a uma categoria específica, dentro de certos limites. Para habilitar essa capacidade, é crucial empregar um conjunto de dados para treinar o modelo, permitindo que ele aprenda padrões e características das diferentes categorias. Uma vez treinado, o classificador pode fazer previsões sobre a categoria à qual uma nova amostra pertence com base nas informações aprendidas durante o treinamento.

#### 3.11.1 Random Forest

Florestas Aleatórias representam um método de aprendizado em conjunto aplicado a tarefas como classificação e regressão. Esse método envolve a construção de múltiplas árvores de decisão durante o treinamento e a utilização do modo das classes para classificação ou da média das previsões individuais das árvores para regressão. Ao criar essa coleção de árvores, cada uma treinada em subconjuntos diferentes dos dados, as Florestas Aleatórias mitigam o risco de superajuste (VELAZQUEZ et al., 2023).

Por meio dessa abordagem, as Florestas Aleatórias demonstram resiliência ao desafio de superajuste, proporcionando uma maior generalização para novos dados não vistos. Esse processo de construção e média das árvores individuais reduz a probabilidade de que o modelo se ajuste excessivamente aos detalhes específicos do conjunto de treinamento (VELAZQUEZ et AL. , 2023).

Neste estudo, a abordagem de *Random Forest* será empregada com variações no número de árvores, incluindo 100, 200, 300, 400 e 500 árvores. Isso permitirá uma avaliação abrangente do desempenho do modelo em diferentes configurações, explorando o impacto da quantidade de árvores no processo de classificação ou regressão. A variação no número de árvores contribuirá para uma análise mais profunda da capacidade de generalização e estabilidade do *Random Forest*, proporcionando *insights* valiosos para a seleção do número ideal de árvores com base nos resultados obtidos.

#### 3.11.2 SVM

O SVM, ou "*Support Vector Machine*" (Máquina de Vetores de Suporte), é um algoritmo de aprendizado de máquina amplamente aplicado em tarefas de classificação e regressão. Ele

se destaca na criação de fronteiras de decisão que efetivamente separam diferentes classes de dados em espaços de múltiplas dimensões (LORENA; CARVALHO, 2007).

A essência do SVM reside na busca por um hiperplano ótimo que melhor demarque as classes de dados, ao mesmo tempo em que maximize a margem entre elas. Pontos de dados cruciais, chamados vetores de suporte, que ficam próximos da margem, influenciam na determinação da posição deste hiperplano. Ademais, o SVM tem a capacidade de empregar transformações de *kernel* para projetar os dados em espaços de características mais complexas. Essas transformações permitem a separação de classes que, em seu espaço original, não seriam linearmente distintas.

No contexto deste estudo vão ser explorados o SVM em conjunto com diferentes tipos de *kernels*, incluindo os polinomiais de graus 1, 2 e 3, assim como o kernel RBF (Função de Base Radial). Essas operações matemáticas têm o propósito de ampliar a capacidade do SVM para lidar tanto com relações simples como complexas entre as classes. O kernel polinomial de grau 1 procura por relações mais diretas, enquanto os graus 2 e 3 introduzem uma crescente complexidade. O kernel RBF, altamente versátil, identifica padrões intrincados nos dados.

## 4 MATERIAIS E MÉTODOS

### 4.1 BASE DE DADOS

Atlas Dermatológico, criado por Samuel Freire da Silva, M.D., em homenagem ao Mestre e Professor Delso Bringel Calheiros, é uma base pública que possui uma extensa coleção de 12.269 imagens de doenças dermatológicas em geral. Essa base foi atualizada pela última vez em 24 de janeiro de 2024 (SILVA, 1995).

O “DERMATOWEB.NET” também oferece um atlas com mais de 5.300 fotografias de alta qualidade, representando a maioria das dermatoses comuns, menos comuns e tumores de pele (CASANOVA et al., 2009). Essas imagens são uma valiosa ferramenta para profissionais da área da dermatologia, permitindo o estudo e a identificação dessas condições dermatológicas.

### 4.2 CENÁRIOS PROPOSTOS

Os cenários propostos na Tabela 5 foram desenvolvidos para aprofundar a compreensão das características visuais das doenças de pele, aproveitando os recursos dos algoritmos e da

inteligência artificial para analisar e interpretar as imagens dermatológicas. Por meio da aplicação de técnicas avançadas de aprendizado de máquina e processamento de imagens, buscamos uma abordagem objetiva e precisa na identificação de padrões e características específicas associadas a cada condição dermatológica. Essa abordagem permite uma análise mais detalhada e refinada das imagens, possibilitando uma identificação mais precisa das doenças de pele em suas diversas variações e tipos.

Na Tabela 5 apresenta os cenários propostos para análise de imagens de diferentes doenças de pele, juntamente com suas variações ou tipos e o número correspondente de imagens disponíveis para cada categoria. As doenças incluídas são: câncer de pele (carcinoma basocelular, carcinoma de células escamosas e melanoma), hanseníase (lepra borderline, lepra lepromatosa e lepra tuberculóide), psoríase (em placas, eritrodérmica, gutata, invertida, palmar e plantar, pustular e ungueal) e artrite psoriásica.

Tabela 5: Cenários propostos para análise de imagens de diferentes doenças de pele.

Doenças	Variações/Tipos	Nº de imagens
	Carcinoma basocelular	391
Câncer de pele	Carcinoma de células escamosas	134
	Melanoma	120
Hanseníase	Leprosy- Bordeline	145
	Leprosy- Lepromatous	276
	Leprosy- Tuberculoid	232
Psoríase	Psoríase - em placas	236
	Psoríase- eritrodérmica	43
	Psoríase- Gutata	74
	Psoríase-invertida	24
	Psoríase-palmo e plantar	26
	Psoríase- pustular	86

	Psoríase Ungueal	41
Artrite psoriásica	Artrite psoriásica	12
	Total	1840

Fonte: (SILVA, 1995; Dermato web.net).

Na tabela 6, estão apresentadas diversas bases de dados de imagens de doenças de pele, cada uma com suas respectivas classes. As bases estão divididas em diferenciais, permitindo a análise de cada classe separadamente e, em seguida, todas juntas.

Tabela 6: Diversas doenças de pele com suas respectivas classes, divididas em diferenciais.

Base	Quantidade de classes	Classes
#Primeiro diferencial – diversas doenças	4	Câncer de pele
		Hanseníase
		Psoríase
		Artrite psoriásica
# Segundo diferencial – Câncer de pele	3	Carcinoma basocelular
		Carcinoma de células escamosas
		Melanoma
# Terceiro diferencial - Hanseníase	3	Leprosy- Bordeline
		Leprosy- Lepromatous
		Leprosy- Tuberculoid
# Quarto diferencial – Psoríase	7	Psoríase - em placas
		Psoríase- eritrodérmica
		Psoríase- Gutata
		Psoríase-invertida
		Psoríase-palmo e plantar
		Psoríase- pustular
		Psoríase
Ungueal		

# Quinto diferencial - todas as variações incluindo artrite psoriásica	14	*Todas as variações
*combinando todas as variações		

Fonte: Próprio autor (2024).

### 4.3 EXTRAÇÃO DE ATRIBUTOS

A extração de atributos das imagens é uma etapa fundamental no processo de análise de dados em muitas aplicações de visão computacional, incluindo o diagnóstico de doenças de pele por meio de imagens dermatológicas. Neste estudo, utilizamos o software Weka 3.8 em conjunto com a biblioteca WekaDeeplearning4j para realizar essa extração de atributos.

WekaDeeplearning4j é uma ferramenta que proporciona um ambiente conveniente para o treinamento e teste de modelos de deep learning, os quais são implementados no Deeplearning4j a partir do Weka (LANG et al., 2019).

O procedimento segue uma série de passos bem definidos que podemos observar na Tabela 12. Primeiro, é necessário criar um arquivo .arff contendo a enumeração das imagens e suas respectivas categorias, servindo como entrada para o processo de extração de atributos. Em seguida, no software Weka, seleciona-se o filtro D14jMlpFilter, responsável por realizar a extração de atributos por meio da rede neural profunda.

O próximo passo envolve a escolha da arquitetura específica da rede neural profunda a ser utilizada, feita através do parâmetro "Pretrained zoo model", que permite selecionar entre diferentes arquiteturas pré-treinadas. Depois, especifica-se que as imagens são do tipo "imagem" e indica-se o diretório onde estão armazenadas.

Com todas as configurações definidas, aplica-se o filtro selecionado, que extrai os atributos das imagens de acordo com a arquitetura da rede neural escolhida. A extração dos atributos das imagens, um procedimento fundamental para as análises subsequentes, foi realizada após a conclusão das etapas anteriores.

Este estudo empregou três arquiteturas distintas de redes neurais profundas: Inception V3, VGG16 e ResNet. Cada uma dessas arquiteturas demonstra capacidades únicas na extração de características das imagens dermatológicas. A Inception V3 revelou-se capaz de extrair 2049 atributos distintos das imagens, enquanto a VGG16 produziu um conjunto mais amplo, totalizando 4097 características únicas. Por fim, a ResNet, assim como a Inception V3, extraiu

2049 características. Esses atributos são essenciais para alimentar os classificadores de aprendizado de máquina, possibilitando a identificação e classificação das diferentes condições dermatológicas com base nas características visuais das imagens.

#### 4.4 TREINAMENTO E TESTE

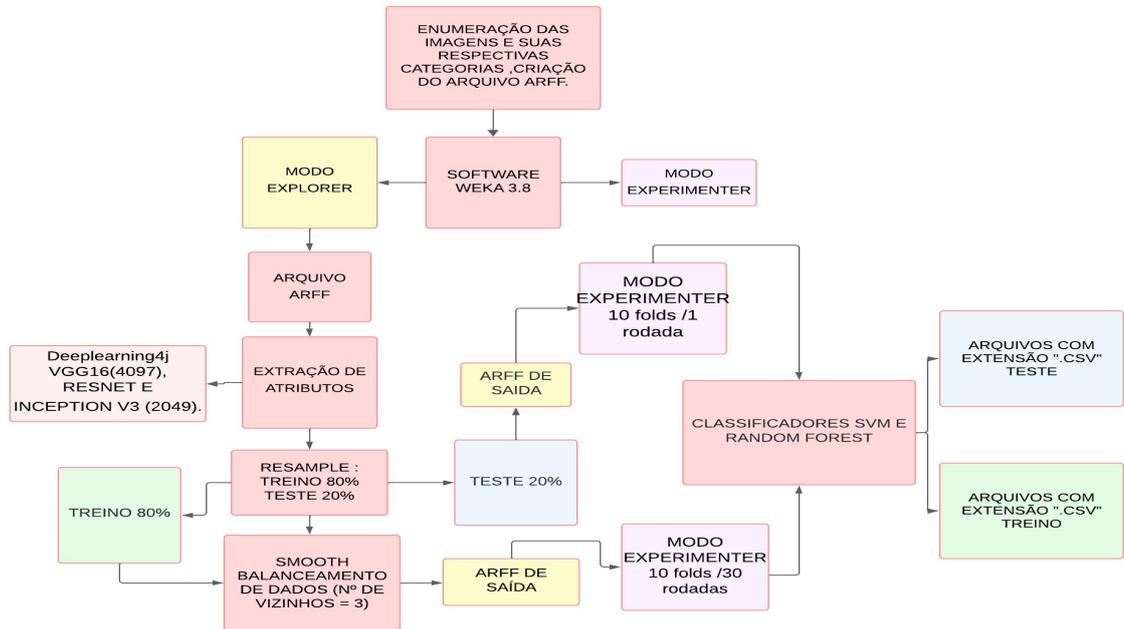
Executar a separação das instâncias em dois grupos distintos, sendo um conjunto de teste abrangendo 20% das instâncias totais, e um conjunto de treino incorporando os 80% restantes, empregando o método Resample como mostrado na Figura 12 . Esse método, reconhecido por sua habilidade em criar distribuições equilibradas das classes presentes nos dados, assegura uma alocação imparcial das instâncias em ambos os conjuntos, garantindo uma avaliação justa e representativa do desempenho dos modelos de redes neurais profundas na classificação de imagens.

Os resultados apresentados neste estudo foram obtidos a partir da realização de múltiplas repetições do experimento no modo "experimenter" do software WEKA. Para cada configuração experimental, realizamos um número específico de repetições, rodando 30 vezes com validação cruzada de 10 folds, garantindo a robustez e confiabilidade dos resultados. As médias e os desvios padrão foram calculados a partir dos resultados das repetições, proporcionando uma representação estatística precisa do desempenho dos modelos. Além disso, utilizamos boxplots para visualizar a distribuição dos resultados das repetições e identificar possíveis variabilidades entre elas. Essa abordagem nos permitiu realizar uma análise abrangente e robusta dos resultados obtidos.

#### 4.5 ARRANJO EXPERIMENTAL

Todos os procedimentos experimentais foram conduzidos utilizando o software Weka 3.8 ,e os resultados decorrentes desses experimentos foram posteriormente exportados para arquivos com extensão ".csv", a fim de possibilitar uma análise detalhada de acordo com o fluxograma mostrado na Figura 12 .

Figura 12: Fluxograma do experimento.



Fonte:Próprio autor (2024).

#### 4.6 MÉTRICAS DE AVALIAÇÃO

As seguintes métricas foram escolhidas para avaliar o desempenho dos modelos após a etapa de classificação. As expressões matemáticas de cada uma delas podem ser vistas na Tabela 8.

Acurácia, mede a correção geral da classificação e é calculada como a proporção de amostras classificadas corretamente em relação ao número total de amostras (SANTOS et al., 2016; FERREIRA et al., 2023).

Recall (Sensibilidade ou Taxa de Verdadeiros Positivos) mede a capacidade do modelo em identificar corretamente as instâncias positivas e é calculado como a proporção de verdadeiros positivos em relação à soma de verdadeiros positivos e falsos negativos (SANTOS et al., 2016; FERREIRA et al., 2023).

Especificidade (Taxa de Verdadeiros Negativos) mede a capacidade do modelo em identificar corretamente as instâncias negativas e é calculada como a proporção de verdadeiros negativos em relação à soma de verdadeiros negativos e falsos positivos (SANTOS et al., 2016; FERREIRA et al., 2023).

Área Sob a Curva ROC (AUC-ROC) quantifica a capacidade do modelo em discriminar entre instâncias positivas e negativas em diferentes limiares de classificação. É calculada como a área sob a curva da característica de operação do receptor (SANTOS et al., 2016; FERREIRA et al., 2023).

Índice Kappa é uma medida de concordância entre as classificações do modelo e as classificações reais, corrigida pela probabilidade de concordância ao acaso. É útil quando há desequilíbrio nas classes de predição (SANTOS et al., 2016; FERREIRA et al., 2023).

Matriz de Confusão fornece uma representação tabular das previsões do modelo e as etiquetas reais, mostrando as contagens de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos, Tabela 7.

Tabela 7: Representação da matriz de confusão.

	Classe Real: Positivo	Classe Real: Negativo
Classe Predita: Positivo (P)	Verdadeiro Positivo (VP) Falso Negativo Negativo (N) (FN)	Falso Positivo (FP) Verdadeiro Negativo (VN)

Fonte: (SOUZA et al., 2020), adaptado.

De acordo com a Tabela 7, a matriz de confusão é composta por quatro elementos.

**Verdadeiro Positivo (VP):** São as instâncias que foram corretamente classificadas como pertencentes à classe positiva pelo modelo. Em nosso contexto, os VPs representam as imagens que foram corretamente identificadas como pertencentes a uma determinada categoria (SOUZA et al., 2020).

**Falso Positivo (FP):** São as instâncias que foram incorretamente classificadas como pertencentes à classe positiva pelo modelo, quando na verdade não são. Os FPs representam as imagens que foram erroneamente atribuídas a uma categoria específica pelo modelo (SOUZA et al., 2020).

**Falso Negativo (FN):** São as instâncias que pertencem à classe positiva, mas foram erroneamente classificadas como pertencentes à classe negativa pelo modelo. Os FNs indicam as imagens que deveriam ter sido atribuídas a uma categoria específica, mas foram negligenciadas pelo modelo (SOUZA et al., 2020).

Verdadeiro Negativo (VN): São as instâncias que foram corretamente classificadas como pertencentes à classe negativa pelo modelo. Os VNs representam as imagens que foram corretamente identificadas como não pertencentes a uma categoria específica (SOUZA et al., 2020).

Tabela 8: Tabela de métricas utilizadas para avaliar o desempenho dos classificadores.

MÉTRICAS	EXPRESSÕES MATEMÁTICAS
ACURÁCIA	$\rho_v = (nTP + nTN) / (nTP + nTN + nFP + nFN)$
ÍNDICE KAPPA	$\kappa = (\rho_v - \rho_z) / (1 - \rho_z)$
SENSIBILIDADE	$\rho_{TP} = nTP / (nTP + nFN)$
ESPECIFICIDADE	$\rho_{TN} = nTN / (nTN + nFP)$
ÁREA DA CURVA ROC	$AUC = \int_0^1 \rho_{TP} d\rho_{TN}$

Fonte: (SANTOS et al., 2016; FERREIRA et al., 2023) ,adaptado.

O termo  $\rho_v$  representa a Acurácia, definida como o número de verdadeiros positivos (nTP) mais verdadeiros negativos (nTN) dividido pelo total de predições (nTP + nTN + nFP + nFN).

O símbolo  $\kappa$  é o Índice Kappa, utilizado para medir a concordância entre as predições do modelo e as observações reais, ajustado para a concordância esperada ao acaso.

O  $\rho_{TP}$  denota a Sensibilidade, ou Taxa de Verdadeiros Positivos, calculada como nTP dividido pelo total de verdadeiros positivos mais falsos negativos (nTP + nFN).

O  $\rho_{TN}$  representa a Especificidade, ou Taxa de Verdadeiros Negativos, calculada como nTN dividido pelo total de verdadeiros negativos mais falsos positivos (nTN + nFP).

AUC refere-se à Área sob a Curva ROC, que quantifica a capacidade discriminativa do modelo através da integral da sensibilidade em função da especificidade ao longo de todos os possíveis limiares de decisão.

## 5 RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentados e discutidos todos os resultados obtidos ao longo dos experimentos. Para apresentação dos dados, serão utilizadas tabelas exibindo médias e desvios padrões, juntamente com gráficos boxplots para cada métrica definida na tabela 7. A análise será conduzida inicialmente nas redes profundas responsáveis pela extração de atributos das imagens, seguida pela avaliação das diferentes configurações do classificador. Ao final, será selecionada a combinação mais eficaz de rede profunda e classificador, levando em consideração a problemática tratada.

### 5.1 AVALIAÇÃO DOS MODELOS DE REDE PROFUNDA

Ao analisar as bases de dados e a distribuição das imagens em classes de acordo com a tabela 6, torna-se evidente a importância de explorar e compreender as nuances de cada categoria de doença. Isso significa que cada condição dermatológica possui características visuais distintas, e compreender essas diferenças é fundamental para desenvolver modelos de classificação precisos e robustos.

#### 5.1.1 Primeiro diferencial

No primeiro diferencial das Tabelas 9, 10 e 11, focamos em uma variedade de condições dermatológicas, incluindo diferentes tipos de câncer de pele, hanseníase, psoríase e artrite psoriásica. Utilizamos a Tabela 6 como referência para a classificação de cada condição.

As Tabelas apresentam as métricas de média e desvio padrão para acurácia, especificidade, sensibilidade, área sob a curva ROC e índice Kappa. As análises foram realizadas considerando:

- Random Forest com variações na quantidade de árvores (100, 200, 300, 400 e 500).
- SVM com diferentes configurações de kernel (polinomial de grau 1, 2 e 3; e Kernel RBF).

Os atributos das imagens foram extraídos utilizando as redes neurais profundas Inception V3(Tabela 9) ,VGG16 (Tabela 10) e ResNet 50 (Tabela 11), e os dados foram balanceados pelo método SMOTE com  $k = 3$  vizinhos.Essas métricas ajudam a avaliar o desempenho dos classificadores em identificar corretamente as diversas condições dermatológicas, permitindo uma comparação detalhada e a escolha do modelo mais eficaz.

Tabela 9: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o primeiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	85,1244	2,4600	0,8013	0,0328	0,8594	0,0509	0,9241	0,0214	0,9614	0,0127
RANDOM FOREST 200	86,5233	2,1906	0,8180	0,0292	0,8657	0,0477	0,9283	0,0207	0,9645	0,0123
RANDOM FOREST 300	86,7306	2,3661	0,8231	0,0315	0,8682	0,0476	0,9292	0,0204	0,9654	0,0121
RANDOM FOREST 400	86,9113	2,3761	0,8254	0,0316	0,8683	0,0475	0,9291	0,0207	0,9660	0,0120
RANDOM FOREST 500	87,0493	2,4037	0,8273	0,0320	0,8691	0,0467	0,9295	0,0207	0,9662	0,0119
SVM (KERNEL DE GRAU 1)	88,0743	2,0979	0,8410	0,0279	0,8716	0,0487	0,9427	0,0174	0,9500	0,0146
SVM (KERNEL DE GRAU 2)	90,0446	1,9578	0,8672	0,0261	0,8909	0,0455	0,9557	0,0162	0,9585	0,0143
SVM (KERNEL DE GRAU 3)	90,4983	1,9601	0,8733	0,0261	0,8892	0,0459	0,9584	0,0155	0,9581	0,0141
SVM ( KERNEL RBF)	88,6000	2,2574	0,8480	0,0301	0,8757	0,0460	0,9508	0,0173	0,9510	0,0146

Fonte: Próprio autor (2024).

Com base nessas métricas apresentadas na Tabela 9, o SVM (Kernel de Grau 3) é o melhor classificador entre os listados. Ele obteve altas pontuações em acurácia, índice Kappa, sensibilidade, especificidade e área sob a curva roc.

Tabela 10: Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG16 de acordo com o primeiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	81,9937	2,5001	0,7599	0,0333	0,8117	0,0552	0,9079	0,0216	0,9431	0,0151
RANDOM FOREST 200	82,3922	2,4666	0,7652	0,0328	0,8125	0,0558	0,9100	0,0218	0,9449	0,0148
RANDOM FOREST 300	82,4743	2,4012	0,7663	0,0320	0,8134	0,0548	0,9105	0,0217	0,9453	0,0148
RANDOM FOREST 400	82,5611	2,3681	0,7674	0,0315	0,8158	0,0546	0,9105	0,0216	0,9456	0,0147
RANDOM FOREST 500	82,5402	2,4049	0,7672	0,0320	0,8152	0,0550	0,9107	0,0217	0,9458	0,0146
SVM (KERNEL DE GRAU 1)	85,5269	2,4227	0,8070	0,0323	0,8365	0,0549	0,9354	0,0204	0,9413	0,0152
SVM (KERNEL DE GRAU 2)	84,9768	2,3775	0,7996	0,0317	0,8285	0,0545	0,9349	0,0187	0,9413	0,0155
SVM (KERNEL DE GRAU 3)	84,3604	2,8395	0,7914	0,0378	0,8353	0,0643	0,9244	0,0348	0,9385	0,0170
SVM ( KERNEL RBF)	83,7688	2,3022	0,7835	0,0306	0,8334	0,0476	0,9174	0,0206	0,9244	0,0172

Fonte: Próprio autor (2024).

A análise das métricas na Tabela 10, revela que os classificadores SVM (Kernel de Grau 1), notadamente, destacam-se como os mais eficazes para abordar o problema em questão. Apresentando acurácias significativamente elevadas, índices Kappa consistentemente altos e desempenho notável em sensibilidade, especificidade e área sob a curva ROC, os SVM (Kernel de Grau 1) demonstram uma robustez distintiva em sua capacidade de classificação. Di-

ante desses resultados, a opção pelo SVM (Kernel de Grau 1) emerge como uma escolha sólida e respaldada pelos indicadores científicos avaliados.

Tabela 11: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Resnet-50 de acordo com o primeiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	89,5675	2,1188	0,8608	0,0282	0,8962	0,0437	0,9550	0,0176	0,9837	0,0072
RANDOM FOREST 200	89,8028	2,1548	0,8638	0,0289	0,8977	0,0435	0,9568	0,0173	0,9845	0,0068
RANDOM FOREST 300	89,8723	2,1633	0,8649	0,0288	0,8998	0,0430	0,9574	0,0174	0,9849	0,0067
RANDOM FOREST 400	89,9385	2,1487	0,8658	0,0286	0,9005	0,0440	0,9573	0,0174	0,9849	0,0067
RANDOM FOREST 500	89,9756	2,1025	0,8663	0,0280	0,9020	0,0439	0,9576	0,0174	0,9851	0,0066
SVM (KERNEL DE GRAU 1)	86,6861	2,2971	0,8224	0,0306	0,8516	0,0511	0,9419	0,0192	0,9324	0,0208
SVM (KERNEL DE GRAU 2)	88,3355	2,2821	0,8444	0,0304	0,8810	0,0480	0,9477	0,0186	0,9437	0,0187
SVM (KERNEL DE GRAU 3)	88,5161	2,2031	0,8468	0,0293	0,8820	0,0455	0,9487	0,0187	0,9439	0,0189
SVM ( KERNEL RBF)	76,9429	2,4475	0,6925	0,0326	0,7410	0,0589	0,9315	0,0197	0,8974	0,0233

Fonte: Próprio autor (2024).

Analisando as métricas apresentadas na Tabela 11, o Random Forest 500 destaca-se como o melhor classificador entre as opções avaliadas. Apresenta a mais alta acurácia, índice Kappa consistente, sensibilidade, especificidade e área sob a curva ROC notavelmente superior em comparação com os outros classificadores. Esses resultados indicam que o Random Forest 500 é a escolha mais robusta e eficaz para resolver o problema em questão.

### 5.1.2 Segundo diferencial

Na seção dedicada ao segundo diferencial Tabelas 12,13 e 14 , exploraremos o câncer de pele, abordando especificamente três tipos principais: Carcinoma basocelular, Carcinoma de células escamosas e Melanoma, de acordo com a Tabela 6 . As Tabelas apresentam as métricas de média e desvio padrão para acurácia, especificidade, sensibilidade, área sob a curva ROC e índice Kappa. As análises foram realizadas considerando:

- Random Forest com variações na quantidade de árvores (100, 200, 300, 400 e 500).
- SVM com diferentes configurações de kernel (polinomial de grau 1, 2 e 3; e Kernel RBF).

Os atributos das imagens foram extraídos utilizando as redes neurais profundas Inception V3(Tabela 12) ,VGG16 (Tabela 13) e ResNet 50 (Tabela 14), e os dados foram balanceados pelo método SMOTE com  $k = 3$  vizinhos.

Tabela 12: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o segundo diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	91,1476	2,9746	0,8672	0,0446	0,8157	0,0734	0,9654	0,0236	0,9709	0,0157
RANDOM FOREST 200	91,5752	7,1748	0,8826	0,0399	0,8314	0,0652	0,9721	0,0209	0,9776	0,0134
RANDOM FOREST 300	92,2502	5,4308	0,8882	0,0396	0,8384	0,0646	0,9740	0,0199	0,9797	0,0129
RANDOM FOREST 400	92,4964	5,4728	0,8919	0,0404	0,8416	0,0670	0,9758	0,0192	0,9807	0,0124
RANDOM FOREST 500	92,5712	5,4197	0,8930	0,0387	0,8443	0,0626	0,9758	0,0192	0,9815	0,0122
SVM (KERNEL DE GRAU 1)	93,6191	5,3075	0,9088	0,0331	0,8402	0,0632	0,9891	0,0129	0,9103	0,0357
SVM (KERNEL DE GRAU 2)	94,2223	5,8908	0,9223	0,0309	0,8535	0,0591	0,9957	0,0095	0,9245	0,0339
SVM (KERNEL DE GRAU 3)	94,5365	5,3586	0,9225	0,0329	0,8541	0,0640	0,9963	0,0089	0,9288	0,0324
SVM ( KERNEL RBF)	92,2886	2,6373	0,8843	0,0395	0,8504	0,0590	0,9621	0,0248	0,9166	0,0306

Fonte: Próprio autor (2024).

De acordo com os resultados da Tabela 12, o classificador SVM com kernel de grau 3 obteve a maior acurácia média, alcançando 94,54%. Além disso, apresentou um índice Kappa médio de 0,92, indicando um bom nível de concordância entre as previsões do modelo e os resultados reais.

Tabela 13: Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG16 de acordo com o segundo diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	88,4851	3,2610	0,8272	0,0488	0,7715	0,0769	0,9554	0,0263	0,9553	0,0202
RANDOM FOREST 200	89,2875	3,2331	0,8393	0,0484	0,7812	0,0768	0,9611	0,0237	0,9620	0,0182
RANDOM FOREST 300	89,5620	3,2427	0,8434	0,0486	0,7874	0,0765	0,9619	0,0234	0,9641	0,0178
RANDOM FOREST 400	89,8122	3,3002	0,8471	0,0494	0,7904	0,0784	0,9642	0,0228	0,9650	0,0176
RANDOM FOREST 500	89,8160	3,2669	0,8472	0,0489	0,7901	0,0772	0,9645	0,0226	0,9659	0,0169
SVM (KERNEL DE GRAU 1)	86,0127	3,6136	0,7901	0,0541	0,7679	0,0815	0,9158	0,0379	0,8464	0,0419
SVM (KERNEL DE GRAU 2)	87,0321	3,2510	0,8059	0,0494	0,7680	0,0795	0,9321	0,0328	0,8556	0,0393
SVM (KERNEL DE GRAU 3)	86,3688	4,1888	0,7955	0,0628	0,7661	0,0869	0,9255	0,0547	0,8550	0,0427
SVM ( KERNEL RBF)	81,6297	3,8702	0,7244	0,0580	0,6423	0,0841	0,9247	0,0346	0,8090	0,0441

Fonte: Próprio autor (2024).

O classificador com o melhor desempenho médio em termos de acurácia de acordo com a Tabela 13, foi o Random Forest com 500 árvores, atingindo 89,82%. Já em relação ao índice Kappa, que mede a concordância entre as previsões do modelo e os resultados reais, o Random Forest com 500 árvores também obteve o maior valor médio, com 0,84. Esses resultados indicam que o Random Forest com 500 árvores é o mais eficaz para este conjunto de dados

Tabela 14: Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o segundo diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	86,6583	3,7021	0,7998	0,0555	0,7634	0,0794	0,9345	0,0362	0,9311	0,0280
RANDOM FOREST 200	87,4977	3,5636	0,8124	0,0534	0,7734	0,0728	0,9412	0,0332	0,9382	0,0258
RANDOM FOREST 300	87,8336	3,4569	0,8175	0,0518	0,7780	0,0735	0,9431	0,0316	0,9406	0,0255
RANDOM FOREST 400	88,0367	3,3292	0,8205	0,0499	0,7811	0,0741	0,9445	0,0293	0,9418	0,0250
RANDOM FOREST 500	88,0332	3,3326	0,8205	0,0499	0,7790	0,0726	0,9453	0,0297	0,9425	0,0248
SVM (KERNEL DE GRAU 1)	80,3410	3,8438	0,7051	0,0576	0,6488	0,0857	0,9008	0,0379	0,7921	0,0476
SVM (KERNEL DE GRAU 2)	83,4650	3,789,	0,7519	0,0568	0,7168	0,0780	0,9029	0,0371	0,8162	0,0469
SVM (KERNEL DE GRAU 3)	84,9955	3,6429	0,7749	0,0546	0,7316	0,0795	0,9157	0,0350	0,8267	0,0469
SVM ( KERNEL RBF)	63,2976	5,8755	0,4521	0,0728	0,6382	0,0793	0,7549	0,0537	0,7183	0,0493

Fonte: Próprio autor (2024).

Analisando a Tabela 14 ,o classificador Random Forest com 400 árvores obteve a maior acurácia média de 88,04%, enquanto o SVM com kernel de grau 3 obteve a maior média de índice Kappa, com 0,77. Esses resultados sugerem que o Random Forest com 400 árvores é o mais eficaz em termos de acurácia média, enquanto o SVM com kernel de grau 3 apresenta a melhor concordância média com os resultados reais.

### 5.1.3 Terceiro diferencial

Na abordagem diferenciada da Hanseníase Tabelas 15,16 e 17, concentramo-nos em três categorias distintas: *leprosy borderline*, *leprosy lepromatous* e *leprosy tuberculoid*.

Cada uma dessas classificações representa uma manifestação única da Hanseníase, permitindo uma análise mais aprofundada e uma compreensão mais completa das variações dessa condição dermatológica. Utilizando a tabela 6 como referência para classificação em cada classe.

Tabela 15: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o terceiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	76,2421	5,2042	0,6436	0,0780	0,8769	0,0692	0,8737	0,0523	0,9492	0,0265
RANDOM FOREST 200	78,2862	4,7397	0,6743	0,0711	0,8907	0,0665	0,8825	0,0518	0,9565	0,0242
RANDOM FOREST 300	79,2999	4,5676	0,6895	0,0685	0,8998	0,0645	0,8853	0,0512	0,9595	0,0231
RANDOM FOREST 400	79,7669	4,4934	0,6965	0,0674	0,9006	0,0628	0,8886	0,0484	0,9606	0,0221
RANDOM FOREST 500	79,8725	4,5358	0,6980	0,0680	0,9028	0,0630	0,8899	0,0486	0,9617	0,0219
SVM (KERNEL DE GRAU 1)	80,2076	4,7491	0,7031	0,071	0,9215	0,0561	0,8765	0,0527	0,9247	0,0316
SVM (KERNEL DE GRAU 2)	83,0771	4,4050	0,7461	0,0660	0,9500	0,0482	0,8936	0,0480	0,9378	0,0284
SVM (KERNEL DE GRAU 3)	83,8432	4,4396	0,7576	0,0665	0,9592	0,0453	0,8986	0,0487	0,9407	0,0280
SVM ( KERNEL RBF)	78,6466	4,8050	0,6797	0,0720	0,8833	0,0660	0,8720	0,0508	0,8957	0,0371

Fonte: Próprio autor (2024).

De acordo com a análise da Tabela 15, o classificador SVM com kernel de grau 3 apresentou a maior acurácia média de 83,84%, seguido pelo SVM com kernel de grau 2, com 83,08%. Em relação ao índice Kappa, o SVM com kernel de grau 3 também obteve a maior média, com 0,75, seguido pelo SVM com kernel de grau 2, com 0,74. Esses resultados sugerem que o SVM com kernel de grau 3 é o mais eficaz em termos de acurácia média e concordância média com os resultados reais.

Tabela 16: Métricas com Média e Desvio Padrão usando Rede Neural VGG 16, de acordo com o terceiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	74,0933	5,1989	0,6114	0,0779	0,8390	0,0794	0,8834	0,0481	0,9382	0,0308
RANDOM FOREST 200	75,2206	5,2687	0,6283	0,0790	0,8451	0,0774	0,8925	0,0427	0,9442	0,0295
RANDOM FOREST 300	75,3884	5,1431	0,6308	0,0771	0,8445	0,0757	0,8938	0,0452	0,9461	0,0287
RANDOM FOREST 400	75,7057	5,0393	0,6355	0,0755	0,8440	0,0768	0,8960	0,0436	0,9471	0,0284
RANDOM FOREST 500	75,8067	5,1104	0,6371	0,0766	0,8446	0,0773	0,8963	0,0435	0,9476	0,0284
SVM (KERNEL DE GRAU 1)	70,6350	5,1525	0,5595	0,0772	0,8192	0,0824	0,8292	0,0579	0,8565	0,0472
SVM (KERNEL DE GRAU 2)	72,9637	5,0414	0,5944	0,0756	0,8422	0,0753	0,8397	0,0563	0,8683	0,0450
SVM (KERNEL DE GRAU 3)	72,4724	5,2685	0,5871	0,0790	0,8542	0,0817	0,8213	0,0681	0,8692	0,0454
SVM ( KERNEL RBF)	70,1961	5,3940	0,5529	0,0809	0,774	0,0836	0,8441	0,0525	0,8399	0,0474

Fonte: Próprio autor (2024).

Com base nos resultados obtidos na Tabela 16, o classificador Random Forest com 500 árvores se destacou, alcançando uma acurácia média de 75,81%, um índice Kappa médio de 0,63 e sensibilidade média de 0,84. Esses valores sugerem que o modelo é capaz de realizar previsões precisas e consistentes para a classificação das amostras.

Tabela 17: Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50. de acordo com o terceiro diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	65,3203	5,1038	0,4798	0,0765	0,7977	0,0851	0,8574	0,0526	0,9141	0,0413
RANDOM FOREST 200	66,8147	5,3625	0,5022	0,0804	0,8110	0,0860	0,8671	0,0499	0,9240	0,0392
RANDOM FOREST 300	67,5454	5,4423	0,5132	0,0816	0,8137	0,0863	0,8726	0,0499	0,9286	0,0376
RANDOM FOREST 400	68,0639	5,3372	0,5210	0,0800	0,8181	0,08590	0,8773	0,0501	0,9306	0,0381
RANDOM FOREST 500	68,3006	5,3793	0,5245	0,0806	0,8209	0,0864	0,8754	0,0509	0,9320	0,0378
SVM (KERNEL DE GRAU 1)	59,3732	5,4485	0,3905	0,0817	0,7795	0,0939	0,7624	0,0643	0,8152	0,0493
SVM (KERNEL DE GRAU 2)	57,0228	5,4959	0,3553	0,0824	0,7486	0,0932	0,7744	0,0639	0,8051	0,0522
SVM (KERNEL DE GRAU 3)	57,9946	5,4506	0,3699	0,0817	0,765	0,0900	0,7788	0,0628	0,8149	0,0516
SVM ( KERNEL RBF)	41,6446	5,1165	0,1256	0,0764	0,7936	0,1157	0,3715	0,1506	0,5946	0,0628

Fonte: Próprio autor (2024).

Com base nos resultados apresentados na Tabela 17, observamos que o classificador Random Forest com 500 árvores obteve o melhor desempenho em termos de acurácia média, atingindo 68,30%. Além disso, apresentou um índice Kappa médio de 0,52 e sensibilidade mé-

dia de 0,82. Esses valores indicam uma capacidade razoável do modelo em realizar previsões precisas, embora ainda haja espaço para melhorias. Por outro lado, o classificador SVM com kernel RBF teve a pior performance, com uma acurácia média de apenas 41,64% e uma sensibilidade média de 0,79. Isso sugere que esse modelo pode não ser adequado para lidar com o problema em questão.

O classificador SVM com kernel RBF teve a pior performance devido a sua baixa capacidade de generalização para o conjunto de dados utilizado. O desempenho do SVM depende significativamente da escolha adequada de hiperparâmetros, como parâmetro de regularização (C) e parâmetro do kernel (gamma). Se esses hiperparâmetros não forem ajustados corretamente, o SVM pode sofrer de overfitting (sobreajuste) ou underfitting (subajuste), levando a uma baixa capacidade de generalização e, conseqüentemente, a resultados menos satisfatórios em termos de acurácia e sensibilidade. Além disso, o kernel RBF é sensível à escala dos dados, o que pode afetar negativamente o desempenho do classificador se os dados não estiverem devidamente normalizados. Portanto, essas questões podem ter contribuído para o desempenho inferior do SVM com kernel RBF em comparação com o Random Forest.

#### 5.1.4 Quarto diferencial

No quarto diferencial Tabela 18,19 e 20, destacamos sete subcategorias que abrangem as diversas manifestações da psoríase: Psoríase em placas, psoríase eritrodérmica, psoríase gutata, psoríase invertida, psoríase palmo e plantar, psoríase pustular e psoríase ungueal.

Cada uma dessas variantes representa uma faceta única da psoríase, proporcionando uma visão abrangente das diferentes apresentações clínicas dessa condição dermatológica. Este é o quarto diferencial em nossa análise clínica, fornecendo insights valiosos para uma abordagem mais específica e personalizada. Utilizando a tabela 6 como referência para classificação em cada classe.

Tabela 18: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o quarto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	95,4146	7,1923	0,9535	0,0192	0,7789	0,0982	0,9953	0,0067	0,9853	0,0088
RANDOM FOREST 200	95,5449	8,7503	0,9585	0,0185	0,7918	0,0967	0,9970	0,0048	0,9912	0,0065
RANDOM FOREST 300	95,7308	8,7660	0,9606	0,0183	0,7998	0,0965	0,9977	0,0047	0,9930	0,0053
RANDOM FOREST 400	96,1377	7,2206	0,9619	0,0174	0,8040	0,0926	0,9979	0,0041	0,9939	0,0049
RANDOM FOREST 500	96,1658	7,2245	0,9622	0,0175	0,8040	0,0930	0,9983	0,0038	0,9945	0,0047
SVM (KERNEL DE GRAU 1)	95,5338	1,2296	0,9688	0,0145	0,8223	0,0841	0,9997	0,0014	0,9176	0,0428
SVM (KERNEL DE GRAU 2)	95,3638	1,3269	0,9704	0,0145	0,8262	0,0848	0,9999	0,0008	0,9166	0,0427
SVM (KERNEL DE GRAU 3)	95,2742	1,2265	0,9658	0,0152	0,7990	0,0890	0,9999	0,0008	0,9082	0,0436
SVM (KERNEL RBF)	95,8515	1,1270	0,9690	0,0157	0,8489	0,0860	0,9961	0,0056	0,9392	0,0364

Fonte: Próprio autor (2024).

Os resultados da Tabela 18 indicam que o classificador *Random Forest* apresentou uma acurácia média superior em comparação com os classificadores SVM em todas as variações testadas. Especificamente, o *Random Forest* com 500 árvores obteve a maior acurácia média de 96,53%, com um desvio padrão de 7,22%. Além disso, este mesmo classificador também teve o índice Kappa médio alto, indicando um excelente nível de concordância entre as previsões do modelo e as classes reais.

O SVM com *kernel* de grau 1 apresentou o terceiro melhor desempenho em termos de acurácia média. O SVM com *kernel* RBF, apesar de ter uma acurácia média semelhante ao *Random Forest*, mostrou uma sensibilidade maior e especificidade médias mais baixas.

Tabela 19: Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG 16 de acordo com o quarto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	94,2245	1,7579	0,9326	0,0205	0,7076	0,1029	0,9909	0,0085	0,9773	0,0142
RANDOM FOREST 200	94,4358	1,7609	0,9350	0,0205	0,7126	0,1045	0,9914	0,0088	0,9804	0,0125
RANDOM FOREST 300	94,5274	1,6970	0,9361	0,0197	0,7192	0,0936	0,9912	0,0088	0,9813	0,0122
RANDOM FOREST 400	94,4995	1,7526	0,9358	0,0204	0,7176	0,0979	0,9915	0,0087	0,9818	0,0118
RANDOM FOREST 500	94,5174	1,7400	0,9360	0,0203	0,7196	0,0999	0,9912	0,0090	0,9820	0,0119
SVM (KERNEL DE GRAU 1)	95,2769	7,1485	0,9518	0,0174	0,7853	0,0929	0,9910	0,0085	0,9074	0,0413
SVM (KERNEL DE GRAU 2)	95,4908	7,1824	0,9543	0,0183	0,7915	0,0957	0,9928	0,0078	0,9131	0,0404
SVM (KERNEL DE GRAU 3)	95,5153	1,6842	0,9476	0,0196	0,9021	0,0447	0,9909	0,0094	0,9021	0,0447
SVM (KERNEL RBF)	94,0919	1,8842	0,9310	0,0219	0,7298	0,1056	0,9862	0,0116	0,8977	0,0450

Fonte: Próprio autor (2024)

Na tabela 19, os resultados indicam que o classificador SVM com *kernel* de grau 3 obteve uma acurácia média de 95,52%, com um desvio padrão de 1,68%. Em termos de índice Kappa médio, sensibilidade média, especificidade média e área sob a curva ROC média, o SVM (*Kernel* de Grau 3) alcançou valores de 0,95, 0,90, 0,99 e 0,90, respectivamente.

Tabela 20: Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o quarto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	92,7346	1,9641	0,9152	0,0229	0,6265	0,1085	0,9885	0,0098	0,9538	0,0195
RANDOM FOREST 200	93,2413	1,8481	0,9211	0,0215	0,6387	0,1049	0,9910	0,0084	0,9627	0,0171
RANDOM FOREST 300	93,3203	1,9334	0,9220	0,0225	0,6404	0,1091	0,9914	0,0083	0,9649	0,0165
RANDOM FOREST 400	93,4832	1,9050	0,9239	0,0222	0,6452	0,1093	0,9920	0,0078	0,9661	0,0162
RANDOM FOREST 500	93,5010	1,9000	0,9241	0,0221	0,6458	0,1077	0,9920	0,0080	0,9668	0,0160
SVM (KERNEL DE GRAU 1)	91,4614	2,0253	0,9003	0,0236	0,5689	0,1118	0,9834	0,0122	0,8054	0,0544
SVM (KERNEL DE GRAU 2)	92,4065	1,9657	0,9114	0,0229	0,6771	0,1037	0,9732	0,0155	0,8400	0,0520
SVM (KERNEL DE GRAU 3)	92,8853	1,9304	0,9169	0,0225	0,6799	0,1026	0,9765	0,0146	0,8428	0,0529
SVM ( KERNEL RBF)	70,0558	3,2021	0,6506	0,0373	0,2323	0,0981	0,9690	0,0156	0,7434	0,0558

Fonte: Próprio autor (2024).

Os resultados da Tabela 20 mostram que o classificador *Random Forest* com 500 árvores obteve a maior acurácia média 93,50% em comparação com outros modelos. Além disso, apresentou um índice *Kappa* médio de 0,92 , indicando um bom grau de concordância além do esperado por acaso.

Em termos de sensibilidade e especificidade, o *Random Forest* 500 também obteve valores favoráveis, com médias de 0,92 e 0,64 , respectivamente. Isso sugere que o modelo tem uma boa capacidade de identificar verdadeiros positivos e verdadeiros negativos.

Por outro lado, o SVM com *kernel* RBF teve um desempenho significativamente inferior, com uma acurácia média de apenas 70,06% e sensibilidade de 0,65. Isso indica que esse modelo pode não ser adequado para esta tarefa específica de classificação.

### 5.1.5 Quinto diferencial

No quinto diferencial Tabela 21, 22 e 23 , abordamos a amplitude total das variações consideradas em nossa análise. Com um total de 14 variações, essa categoria engloba todas as nuances e distintas apresentações das condições estudadas. Utilizando a tabela 6 como referência para classificação em cada classe.

Tabela 21: Métricas com Média e Desvio Padrão usando Rede Neural Profunda Inception V3 de acordo com o quinto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	91,6613	1,1673	0,9110	0,0101	0,6431	0,0848	0,9859	0,0063	0,9570	0,0138
RANDOM FOREST 200	93,0563	1,1416	0,9258	0,0097	0,6810	0,0810	0,9882	0,0057	0,9648	0,0120
RANDOM FOREST 300	93,5943	1,0863	0,9320	0,0093	0,6980	0,0815	0,9891	0,0053	0,9670	0,0114
RANDOM FOREST 400	93,8590	1,0377	0,9338	0,0111	0,7070	0,0790	0,9893	0,0052	0,9682	0,0111
RANDOM FOREST 500	94,0243	1,0168	0,9365	0,0085	0,7153	0,0785	0,9896	0,0051	0,9689	0,0109
SVM (KERNEL DE GRAU 1)	94,8156	9,2939	0,9441	0,0100	0,7815	0,0761	0,9918	0,0043	0,9541	0,0192

SVM (KERNEL DE GRAU 2)	96,0033	8,0627	0,9575	0,0069	0,7997	0,0669	0,9951	0,0034	0,9588	0,0186
SVM (KERNEL DE GRAU 3)	96,1066	8,0414	0,9586	0,0069	0,7979	0,0678	0,9949	0,0034	0,9584	0,0193
SVM ( KERNEL RBF)	91,4317	1,1537	0,9084	0,0099	0,7435	0,0725	0,9814	0,0068	0,9524	0,0166

Fonte: Próprio autor (2024)

Analisando os resultados da tabela 21, indicam que o classificador SVM com kernel de grau 3 obteve a maior acurácia média 96,11% em comparação com outros modelos. Além disso, apresentou o maior índice *Kappa* médio 0,95, indicando um excelente grau de concordância além do esperado por acaso.

Em termos de sensibilidade e especificidade, o SVM com *kernel* de grau 3 também obteve os valores mais altos, com médias de 0,80 e 0,95, respectivamente. Isso sugere que o modelo tem uma excelente capacidade de identificar verdadeiros positivos e verdadeiros negativos. O SVM de grau 3 superou todos os outros modelos .

Tabela 22: Métricas com Média e Desvio Padrão usando Rede Neural Profunda VGG 16 de acordo com o quinto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	90,1622	1,2659	0,8940	0,0136	0,6002	0,0816	0,9805	0,0071	0,9433	0,0192
RANDOM FOREST 200	90,9810	1,1889	0,9028	0,0128	0,6173	0,0841	0,9820	0,0069	0,9480	0,0182
RANDOM FOREST 300	91,2900	1,1443	0,9060	0,0124	0,6257	0,0812	0,9830	0,0067	0,9496	0,0173
RANDOM FOREST 400	91,3788	1,1594	0,9086	0,0124	0,6266	0,0841	0,9832	0,0065	0,9506	0,0170
RANDOM FOREST 500	91,4798	1,1241	0,9082	0,0121	0,6299	0,0852	0,9831	0,0064	0,9511	0,0171
SVM (KERNEL DE GRAU 1)	89,5522	1,2445	0,8874	0,0134	0,6638	0,0716	0,9743	0,0079	0,9192	0,0217
SVM (KERNEL DE GRAU 2)	90,1736	1,1741	0,8941	0,0126	0,6628	0,0761	0,9772	0,0073	0,9201	0,0217
SVM (KERNEL DE GRAU 3)	89,6207	1,2831	0,8882	0,0138	0,6568	0,0815	0,9754	0,0095	0,9191	0,0218
SVM ( KERNEL RBF)	83,9645	1,5005	0,8273	0,0161	0,4727	0,0847	0,9722	0,0077	0,8946	0,0259

Fonte: Próprio autor (2024).

Com base nos resultados apresentados na Tabela 22, o classificador *Random Forest* com 500 árvores obteve a maior acurácia média 91,48% entre todos os modelos, além do maior índice *Kappa* médio 0,90 , indicando um bom grau de concordância além do esperado por acaso. Em termos de sensibilidade e especificidade, o *Random Forest* também teve os valores mais altos, com médias de 0,90 e 0,95, respectivamente.

O classificador SVM com *kernel* de grau 2 ficou em segundo lugar em termos de acurácia média 90,17% e índice *Kappa* médio 0,89 , além de apresentar sensibilidade e especificidade razoavelmente altas.

Por outro lado, o SVM com *kernel* RBF teve a menor acurácia média (83,96%) e o menor índice *Kappa* médio 0,82 , juntamente com sensibilidade e especificidade mais baixas em comparação com os outros modelos.

Tabela 23: Métricas com Média e Desvio Padrão usando Rede Neural Profunda ResNet 50 de acordo com o quinto diferencial.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 100	87,9897	1,2950	0,8706	0,0139	0,5365	0,0876	0,9822	0,0072	0,9266	0,0212
RANDOM FOREST 200	89,2768	1,2402	0,8845	0,0133	0,5640	0,0891	0,9850	0,0061	0,9359	0,0184
RANDOM FOREST 300	89,7451	1,2490	0,8895	0,0134	0,5767	0,0907	0,9861	0,0059	0,9388	0,0178
RANDOM FOREST 400	89,9434	1,5868	0,8925	0,0110	0,5780	0,0873	0,9862	0,0060	0,9404	0,0172
RANDOM FOREST 500	90,0903	1,2657	0,8932	0,0136	0,5785	0,0853	0,9866	0,0058	0,9411	0,0170
SVM(KERNEL DE GRAU 1)	41,5152	2,9200	0,3152	0,0345	0,8165	0,0628	0,7707	0,0329	0,6548	0,0205
SVM(KERNEL DE GRAU 2)	41,1317	3,4293	0,3269	0,0387	0,6381	0,0796	0,8526	0,0282	0,8558	0,0282
SVM(KERNEL DE GRAU 3)	41,4843	3,3692	0,3305	0,0381	0,6432	0,0791	0,8519	0,0278	0,8560	0,0278
SVM ( KERNEL RBF)	29,2312	1,9530	0,1231	0,0252	0,9698	0,0278	0,3084	0,0325	0,6548	0,0205

Fonte: Próprio autor (2024).

Com base nos resultados apresentados na tabela 23, o classificador *Random Forest* com 500 árvores apresentou a maior acurácia média 90,09% entre todos os modelos, além do maior índice *Kappa* médio 0,89, indicando um bom grau de concordância além do esperado por acaso. Em termos de sensibilidade e especificidade, o *Random Forest* também teve os valores mais altos, com médias de 0,57 e 0,98 , respectivamente.

Por outro lado, os classificadores SVM com diferentes kernels tiveram desempenho muito inferior, com acurácia média em torno de 41% e índices *Kappa* médios abaixo de 0.35 ,especificamente, o SVM com *kernel* RBF teve a pior acurácia média 29,23% e sensibilidade muito baixa 0,12.

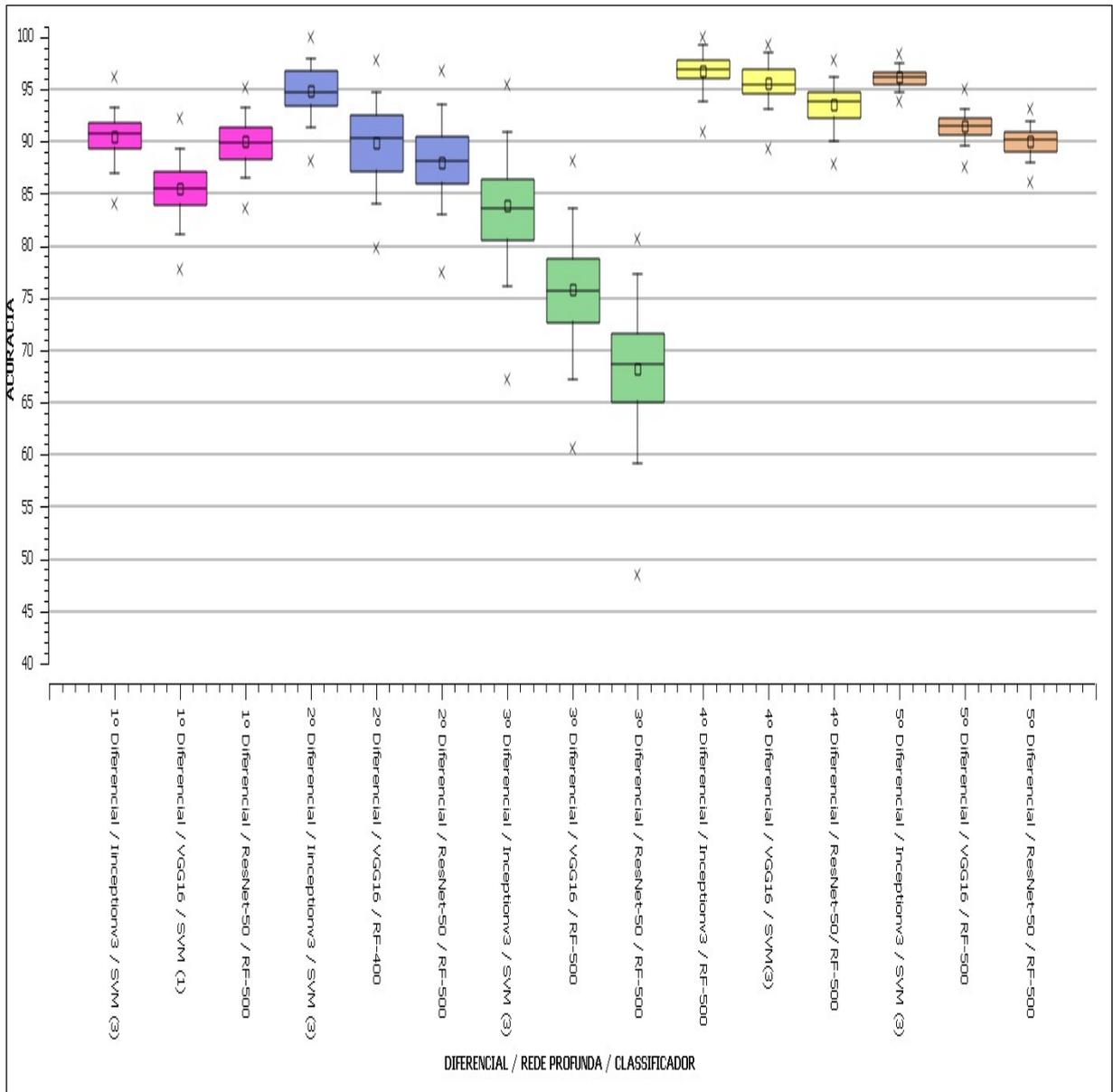
Esses resultados sugerem que o SVM com *kernel* de grau 3 apresenta um desempenho robusto e consistente na classificação das amostras, com boa capacidade de generalização para novos dados. A alta acurácia combinada com índices *Kappa* significativos e sensibilidade e especificidades equilibradas indicam que esse classificador é eficaz na discriminação entre classes positivas e negativas, bem como na minimização de falsos positivos e falsos negativos.

## 5.2 REPRESENTAÇÃO GRÁFICA DOS RESULTADOS EM BOXPLOT NA FASE TREINO

Os resultados da melhor rede neural e do melhor classificador são representados graficamente em boxplot nas Figuras 13, 14 ,15 ,16 e 17 de acordo com cada diferencial destacado na Tabela 6. Esses resultados foram analisados utilizando métricas como acurácia, índice *kappa*, especificidade, sensibilidade e área sob a curva ROC. A análise foi realizada com base nos conjuntos de dados representados nas tabelas 9 a 23.

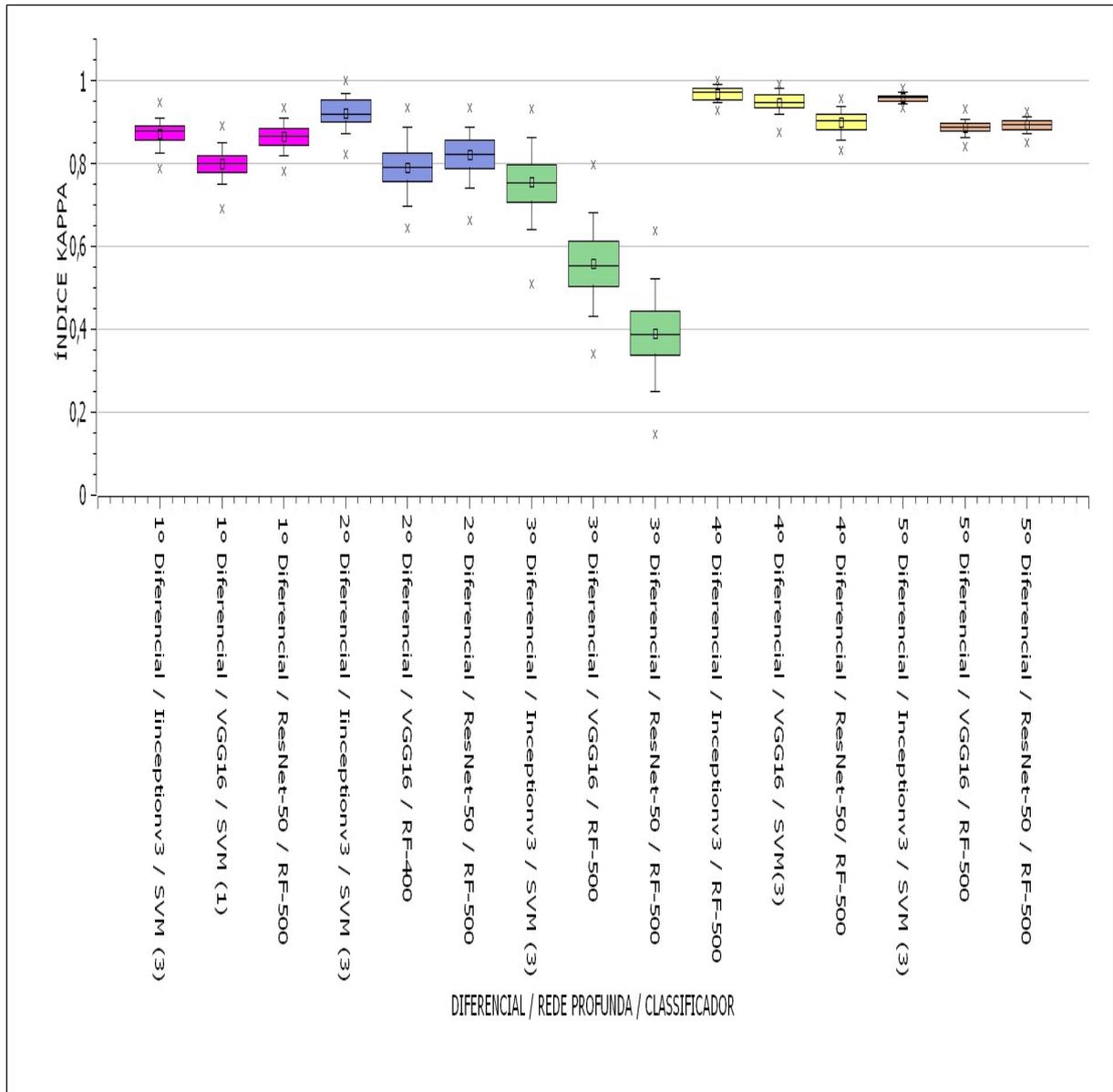
Considerando a extração de atributos das imagens pelas redes neurais profundas Inception v3, VGG 16, ResNet 50 e pelos classificadores *Random Forest* e SVM, todos os dados foram balanceados usando *SMOTE* com  $k = 3$  vizinhos. Acurácia representada em gráfico boxplot .

Figura 13: Boxplot mostrando a distribuição da acurácia considerando os diferenciais 1,2,3,4, e 5.



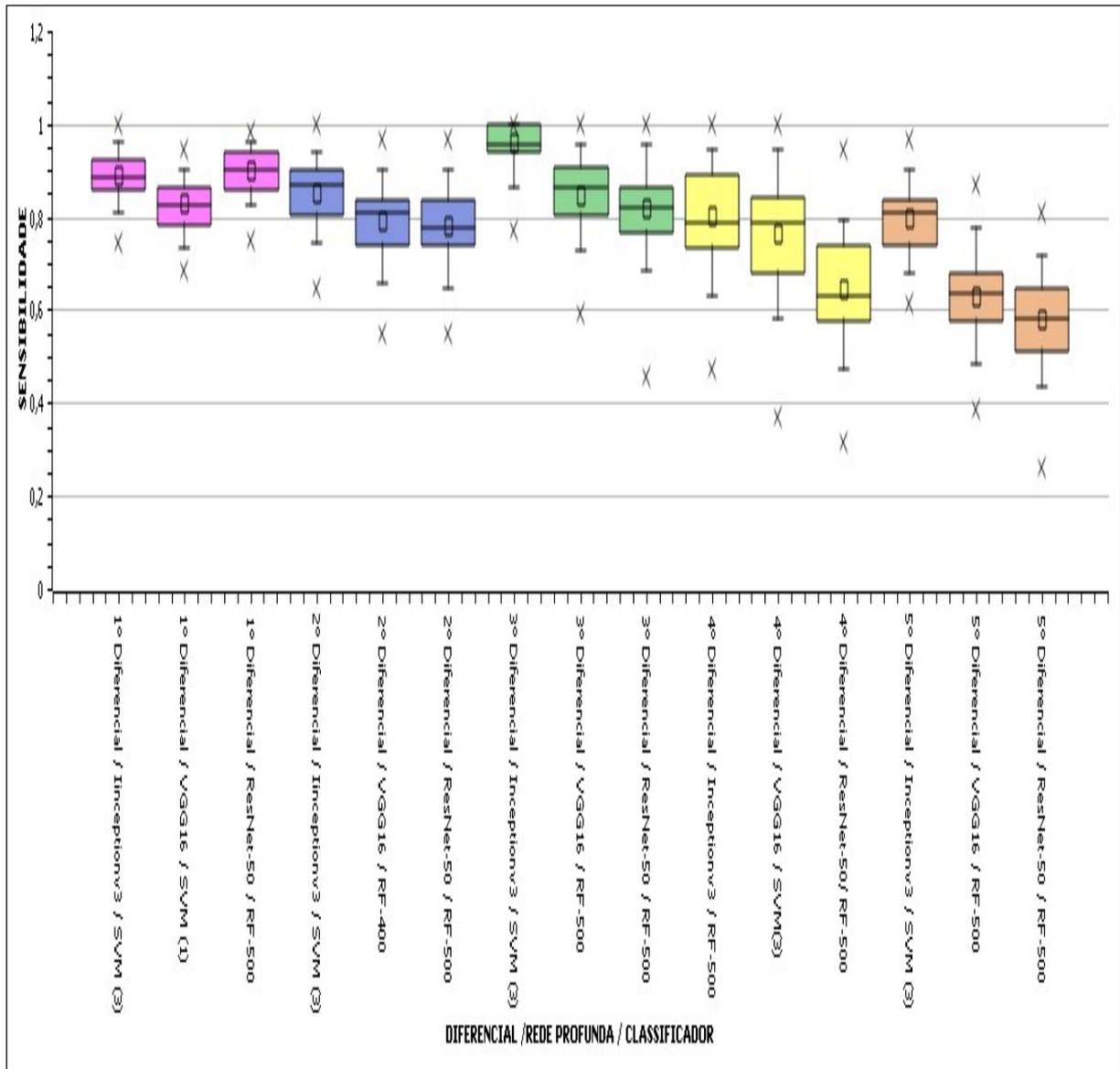
Fonte: Próprio autor (2024).

Figura 14: Boxplot mostrando a distribuição do índice *kappa* considerando os diferenciais 1,2,3,4, e 5.



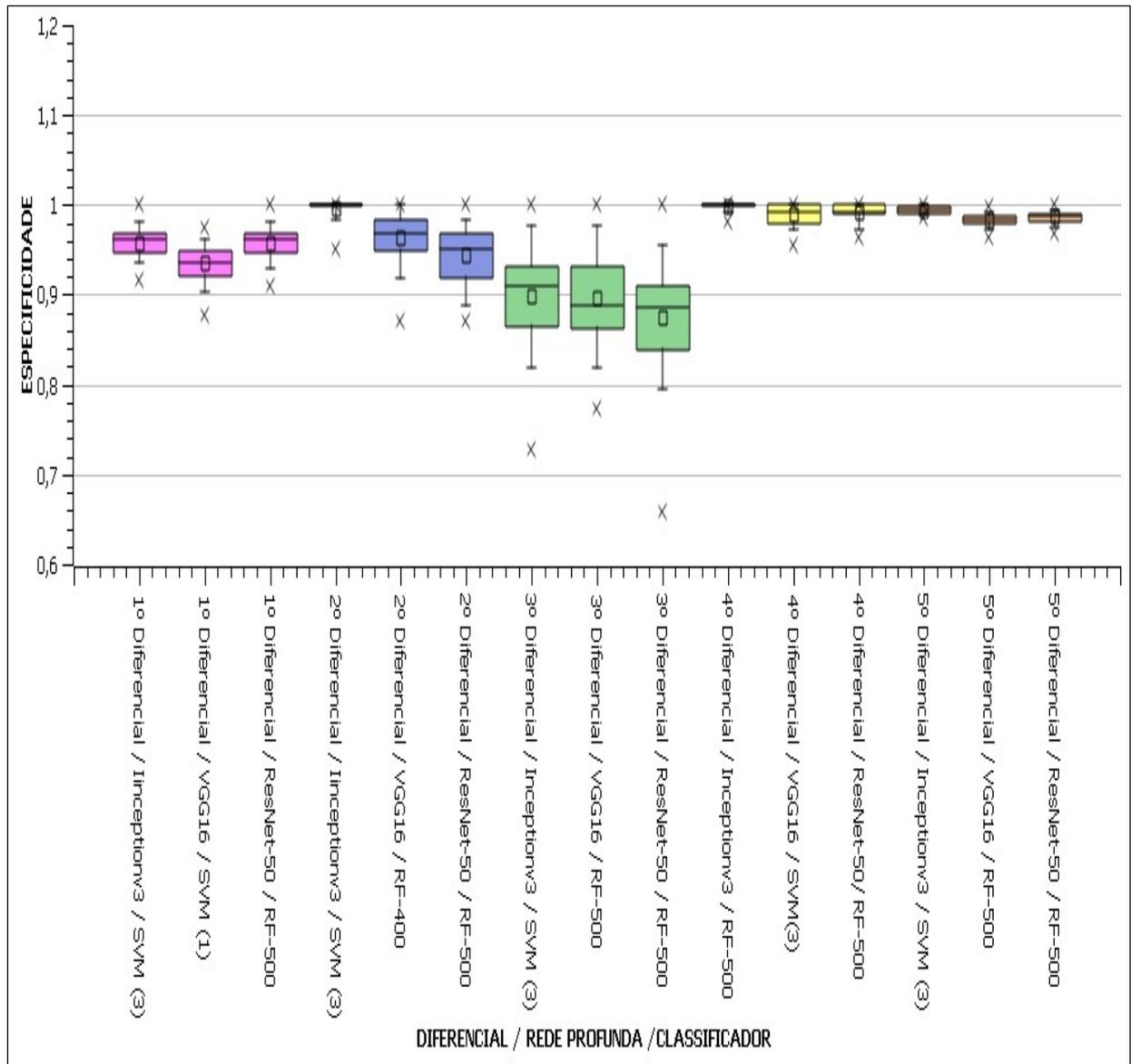
Fonte: Próprio autor (2024).

Figura 15: Boxplot mostrando a distribuição da sensibilidade considerando os diferenciais 1,2,3,4, e 5.



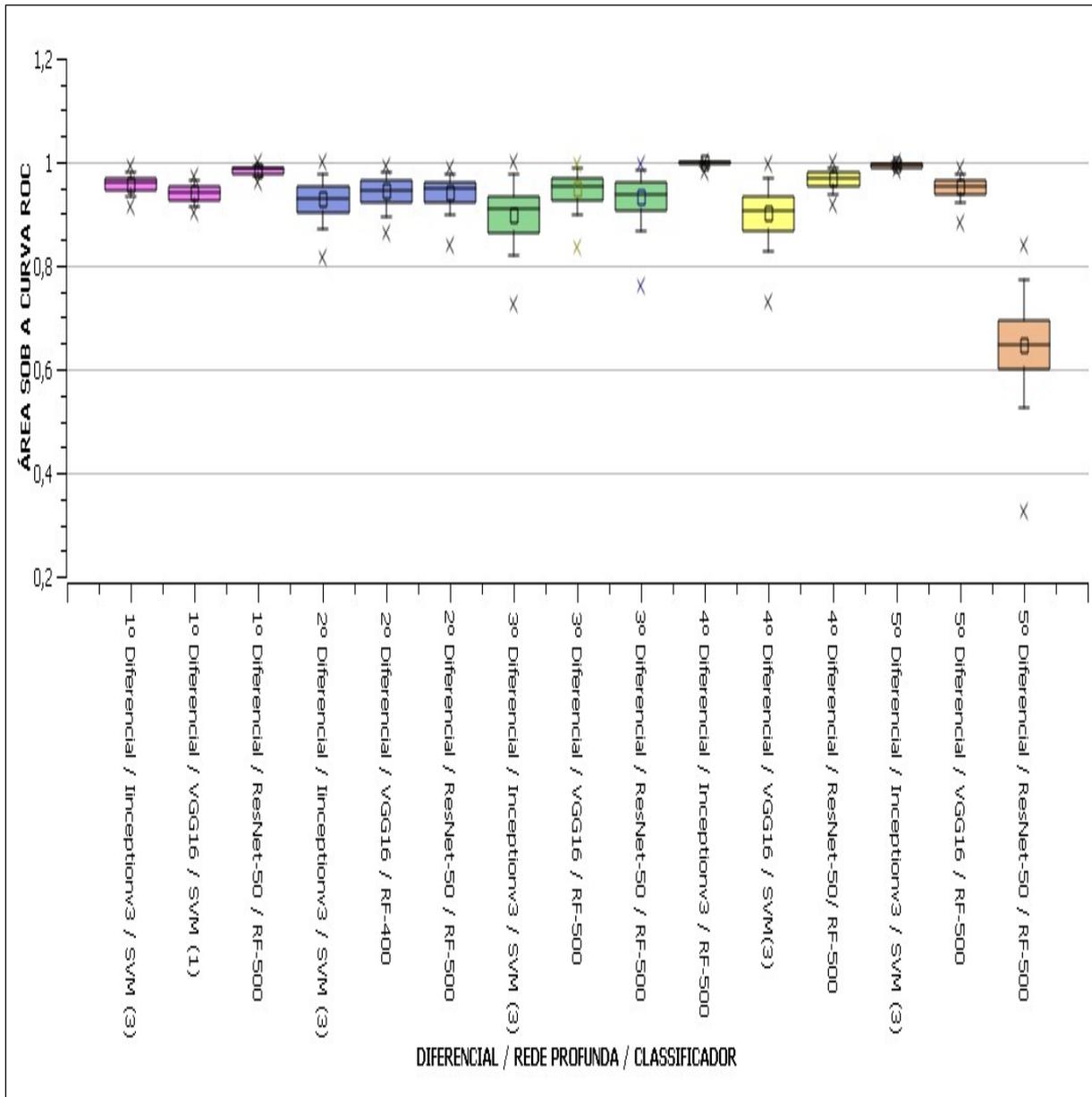
Fonte: Próprio autor (2024).

Figura 16: Boxplot mostrando a distribuição da especificidade considerando os diferenciais 1,2,3,4, e 5.



Fonte: Próprio autor (2024).

Figura 17: Boxplot mostrando a distribuição da área sob curva ROC considerando os diferenciais 1,2,3,4, e 5.



Fonte: Próprio autor (2024)

Notavelmente, os boxplots Figuras (13, 14, 15, 16 e 17) revelaram que a combinação da arquitetura Inception V3 com o classificador SVM de *kernel* de grau 3, seguido de *Random Forest* com 500 árvores, obteve os melhores resultados em comparação com outras combinações de redes profundas e classificadores. Essa análise gráfica reforça ainda mais a eficácia dessa abordagem como o melhor modelo para a detecção e classificação das doenças dermatológicas consideradas neste estudo.

### 5.3 FASE FINAL DOS EXPERIMENTOS (TESTE)

A fase final dos experimentos (teste) tem como objetivo submeter esse modelo selecionado a um teste prático, onde será avaliado o seu desempenho na classificação de imagens previamente não utilizadas. Essa etapa é crucial para verificar a capacidade do modelo em generalizar para novos dados e manter sua precisão diante de imagens desconhecidas.

#### 5.3.1 Primeiro diferencial

Conforme apresentado na Tabela 24, o estudo analisou o desempenho do classificador SVM com *kernel* de grau 3 em um conjunto de dados específico. Os resultados revelaram que o classificador obteve uma acurácia média de 79,71%, o que significa que ele foi capaz de prever corretamente cerca de 79,71% das amostras. Além disso, o índice kappa médio foi calculado em 0,6941, indicando uma concordância significativa entre as previsões do modelo e as observações reais. A sensibilidade média foi estimada em 0,8371, demonstrando a capacidade do modelo em identificar corretamente os verdadeiros positivos, enquanto a especificidade média foi de 0,9041, indicando sua habilidade em identificar corretamente os verdadeiros negativos. A área sob a curva ROC média foi determinada como 0,8896, destacando a capacidade do modelo em distinguir entre as classes positiva e negativa. Os desvios padrão associados a cada métrica fornecem insights sobre a consistência dos resultados em diferentes contextos. Esses detalhes fornecem uma compreensão abrangente do desempenho do classificador SVM com *kernel* de grau 3 e sua utilidade em diversas aplicações práticas.

Tabela 24: Primeiro diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
SVM (KERNEL DE GRAU 3)	79,7147	8,8220	0,6941	1,3330	0,8371	1,7170	0,9041	0,0622	0,8896	0,0208

Fonte: Próprio autor (2024).

#### 5.3.2 Segundo diferencial

De acordo com os resultados apresentados na Tabela 25, observa-se que o classificador SVM com *kernel* de grau 3, combinado com a arquitetura Inception V3, alcançou uma acurácia média de aproximadamente 62,63%, indicando um desempenho moderado. O índice *Kappa* médio de 0,1979 sugere uma concordância além do esperado pelo acaso, embora não seja

muito alta. A sensibilidade média foi alta, atingindo 90%, o que demonstra a capacidade do classificador em identificar corretamente a maioria dos casos positivos. Por outro lado, a especificidade média foi relativamente baixa, registrando 32,33%, indicando dificuldades do classificador em identificar corretamente os casos negativos. A área sob a curva ROC média foi de 61,16%, evidenciando um desempenho moderado na capacidade de discriminar entre as classes positivas e negativas.

Tabela 25: Segundo diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
SVM (KERNEL DE GRAU 3)	62,6282	7,3898	0,1979	0,1705	0,9000	0,0790	0,3233	0,1826	0,6116	0,0725

Fonte: Próprio autor (2024)

### 5.3.3 Terceiro diferencial

Para o terceiro diferencial, Tabela 26, o classificador SVM com kernel de grau 3, em conjunto com a arquitetura Inception V3, obteve uma acurácia média de 59,23%, com um desvio padrão de 9,63%. O índice Kappa médio foi de 0,3512, acompanhado de um desvio padrão de 0,1502. Em relação à sensibilidade, a média foi de 0,3833, com um desvio padrão de 0,2490, enquanto a especificidade média foi de 0,93, com um desvio padrão de 0,0823. A área sob a curva ROC média foi de 0,7586, apresentando um desvio padrão de 0,1177.

Esses resultados sugerem um desempenho moderado do classificador SVM com kernel de grau 3, em conjunto com a arquitetura Inception V3, na tarefa de classificação para o diferencial 3. No entanto, a presença de um alto desvio padrão em várias métricas indica uma variabilidade significativa nos resultados, o que pode refletir uma certa instabilidade ou dificuldade do modelo em generalizar para diferentes conjuntos de dados.

Tabela 26: Terceiro diferencial, rede neural Inception v3 com o classificador SVM de kernel de grau 3.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
SVM (KERNEL DE GRAU 3)	59,2307	9,6281	0,3512	0,1502	0,3833	0,2490	0,93	0,0823	0,7586	0,1177

Fonte: Próprio autor (2024)

### 5.3.4 Quarto diferencial

Os resultados obtidos para o diferencial 4, Tabela 27, demonstram que o classificador alcançou uma sensibilidade de 100%, indicando sua capacidade de identificar corretamente todos os casos positivos. No entanto, a especificidade de 0 revela uma dificuldade em classificar corretamente os casos negativos. A acurácia média de 45,63% reflete um desempenho geral abaixo do esperado.

A análise da área sob a curva ROC média, que atingiu 73,48%, sugere que o classificador foi capaz de discriminar moderadamente bem entre as classes. Esses achados destacam a importância de avaliar não apenas a acurácia global, mas também outras métricas de desempenho, como sensibilidade, especificidade e área sob a curva ROC, para uma compreensão abrangente do desempenho do modelo.

O uso da arquitetura Inception V3 em conjunto com o classificador *Random Forest* de 500 árvores revelou-se uma abordagem promissora, embora os resultados indiquem áreas passíveis de melhoria ou refinamento.

Tabela 27: Quarto diferencial, rede neural Inception v3 com o classificador *Random Forest* de 500 árvores.

CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
RANDOM FOREST 500	45,6363	4,3661	0	0	1	0	0	0	0,7348	0,1630

Fonte: Próprio autor (2024).

### 5.3.5 Quinto diferencial

O modelo mais promissor para o diferencial 5, conforme revelado pela Tabela 28, consiste no emprego da arquitetura de rede neural profunda Inception V3 em conjunto com o classificador SVM utilizando kernel de grau 3. Os resultados obtidos demonstram que esse modelo alcançou uma acurácia média de aproximadamente 41,10%, indicando um desempenho que, embora modesto, destaca-se pela relevância na classificação das amostras.

Apesar do índice Kappa médio de 0,2949 indicar uma concordância além do esperado pelo acaso, seu valor ainda é consideravelmente baixo, sugerindo margem para aprimoramentos na tarefa de classificação.

A sensibilidade média registrou um valor elevado de 94,64%, o que evidencia a capacidade do classificador em identificar corretamente a maioria dos casos positivos. Entretanto, a espe-

cificidade média apresentou-se relativamente inferior, alcançando 68,40%, o que sugere que o modelo enfrentou desafios na correta identificação dos casos negativos.

Por fim, a área sob a curva ROC média, de 86,61%, aponta para um desempenho moderado na capacidade de discriminação entre as classes positivas e negativas.

Tabela 28: Quinto diferencial , rede neural Inception v3 com o classificador SVM de kernel de grau 3.

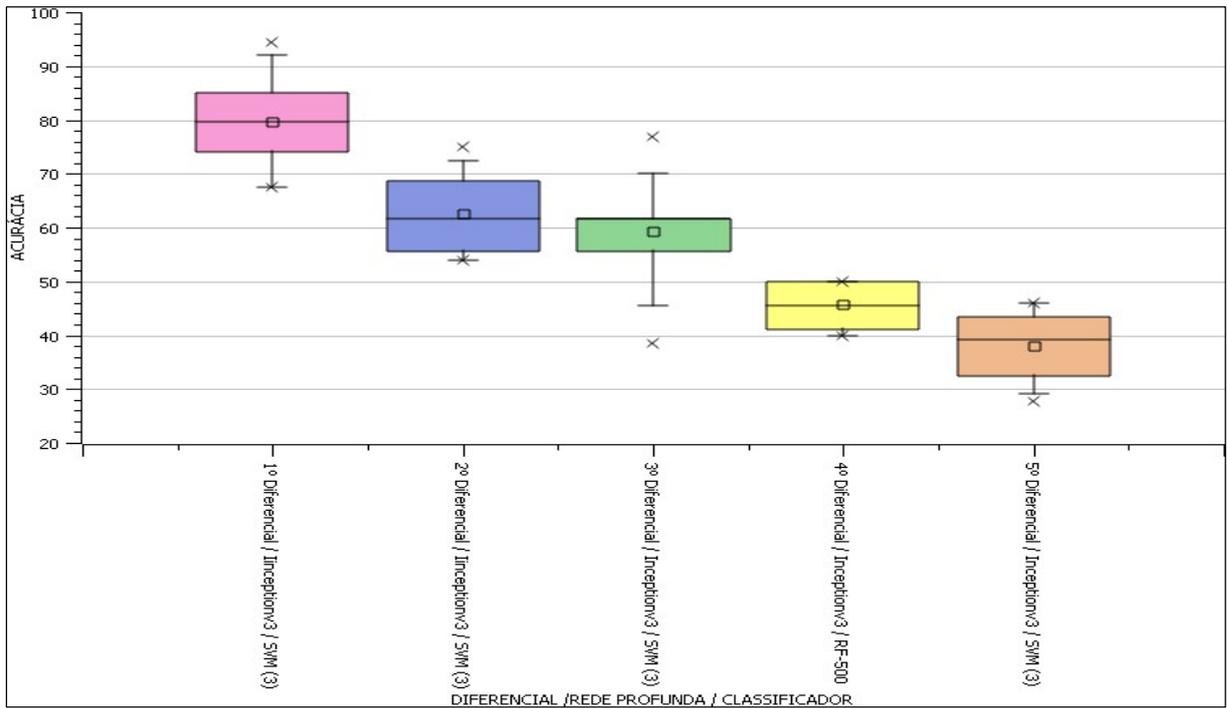
CLASSIFICADORES	ACURÁCIA (%)		ÍNDICE KAPPA		SENSIBILIDADE		ESPECIFICIDADE		ÁREA SOB A CURVA ROC	
	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão	Média	Desvio padrão
SVM (KERNEL DE GRAU 3)	41,0998	6,4062	0,2949	0,0847	0,9464	0,0910	0,6840	0,1455	0,8661	0,0655

Fonte: Próprio autor (2024)

#### 5.4 REPRESENTAÇÃO GRÁFICA DOS RESULTADOS EM BOXPLOT DA FASE FINAL DE TESTE

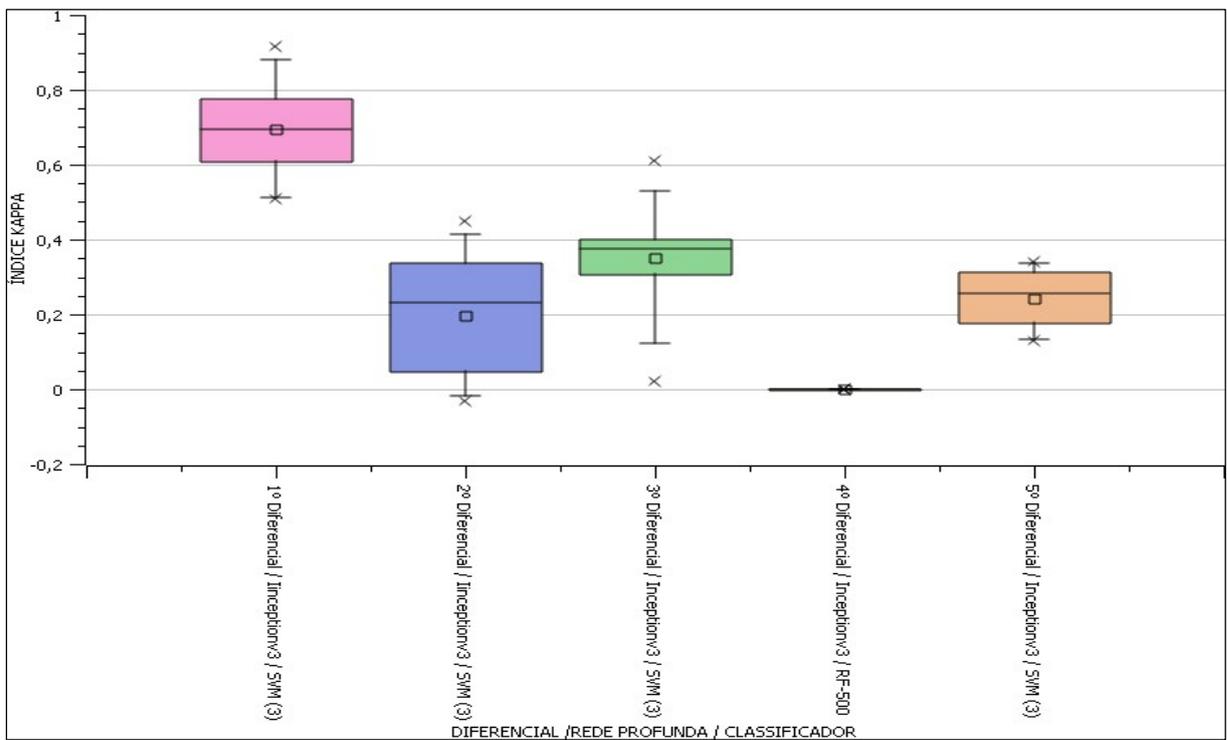
Boxplot mostrando a distribuição da acurácia (Figura 18), índice *kappa* (Figura 19), sensibilidade (Figura 20), especificidade (Figura 21) e da área sob a curva ROC (Figura 22), para os resultados obtidos através da classificação das imagens do melhor modelo de rede profunda e classificador de acordo com cada diferencial Tabela 6. Os dados foram processados com extração de atributos pela rede neural profunda Inception v3, e classificadores como *Random Forest* 500 (quarto diferencial) e SVM de grau 3( primeiro, segundo, terceiro e quinto diferencial), executados no modo de teste. Os resultados são baseados nas Tabelas 24, 25, 26, 27 e 28, refletindo o melhor desempenho dos modelos de acordo com cada diferencial.

Figura 18: Boxplot mostrando a distribuição da melhor acurácia de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5.



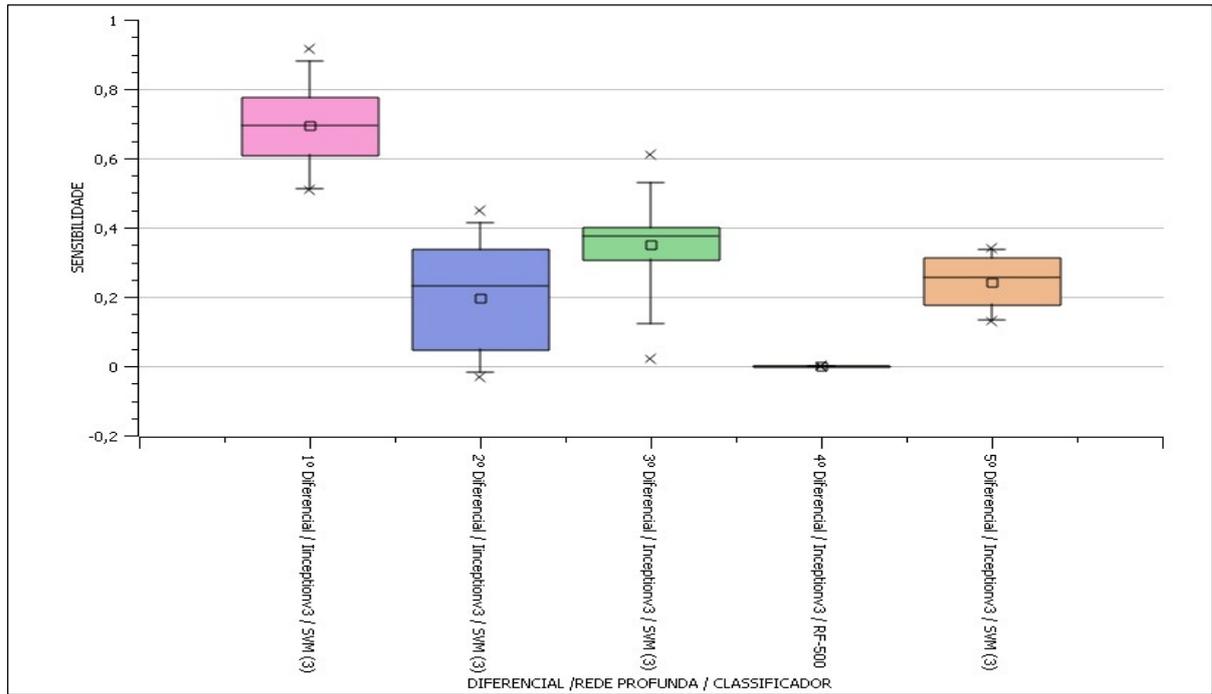
Fonte: Próprio autor (2024).

Figura 19: Boxplot mostrando a distribuição do índice kappa de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5.



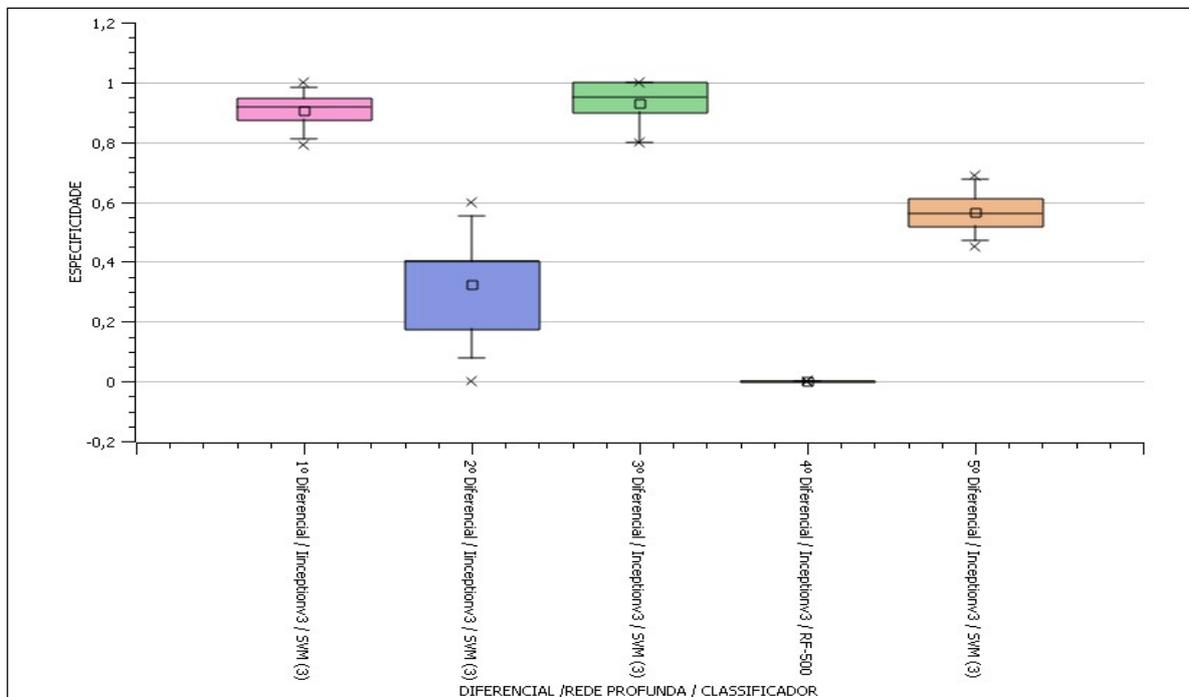
Fonte: Próprio autor (2024).

Figura 20: Boxplot mostrando a distribuição da sensibilidade de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5.



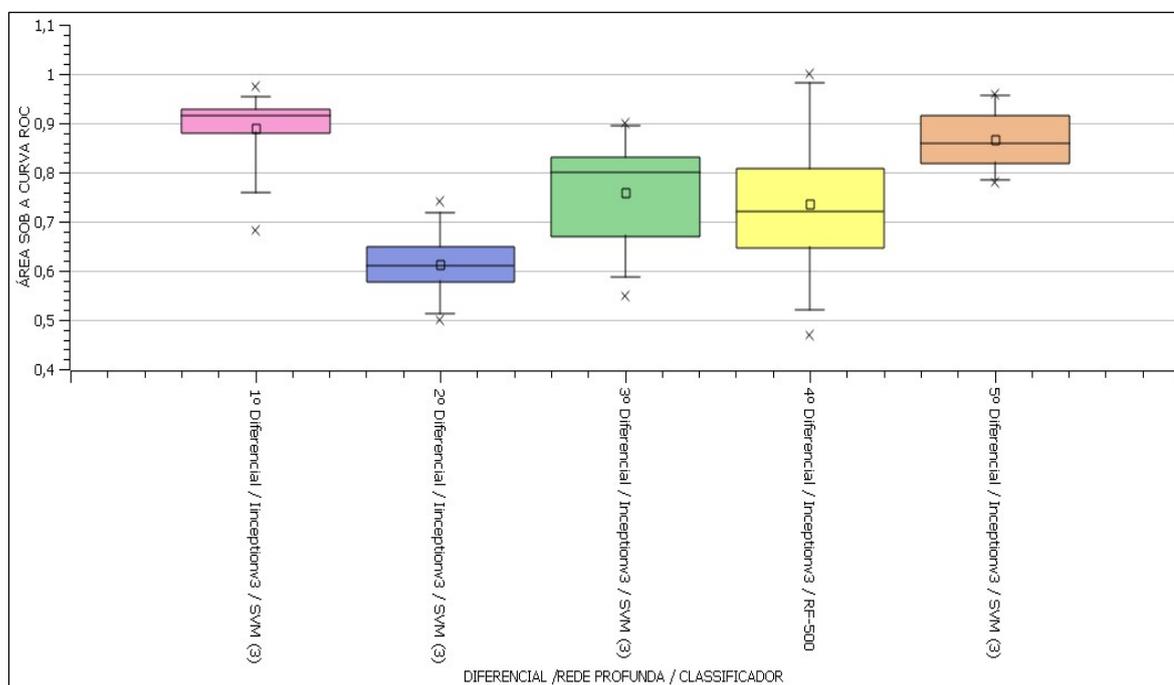
Fonte: Próprio autor (2024)

Figura 21: Boxplot mostrando a distribuição da especificidade de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5.



Fonte: Próprio autor (2024).

Figura 22: Boxplot mostrando a distribuição da área sob a curva ROC de acordo com o melhor classificador considerando os diferenciais 1,2,3,4, e 5.



Fonte: Próprio autor (2024).

## 5.5 MATRIZ DE CONFUSÃO

Os modelos de rede neurais profundas e classificadores foram treinados nos conjuntos de dados especificados nas Tabelas 9 a 23 e testados nos conjuntos das Tabelas 24 a 28. Durante a análise, cada diferencial destacado na Tabela 6 foi cuidadosamente considerado. Matrizes de confusão foram geradas a partir dos resultados, proporcionando uma avaliação detalhada do desempenho dos modelos em relação a cada um desses diferenciais.

A matriz de confusão ilustra as previsões do modelo (nas linhas) em comparação com as classes reais (nas colunas) para cada categoria de doença dermatológica. Notavelmente, os modelos SVM (Kernel de Grau 3) no primeiro, segundo, terceiro e quinto diferencial, com extração de atributos pela Rede Neural Inception V3, representam o melhor conjunto de resultados. Além disso, o *Random Forest* de 500 árvores e a rede neural Inception v3 destacam-se como o melhor conjunto no quarto diferencial.

### 5.5.1 Primeiro diferencial

A Figura 23 apresenta a matriz de confusão referente ao treino do melhor classificador e da melhor rede neural profunda do primeiro diferencial. Os detalhes dessa matriz estão disponíveis na Tabela 9, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado na Figura 13 a 17 .

Figura 23: Matriz de confusão referente ao primeiro diferencial no modo treino.

	a	b	c	d			
a	463	32	21	0	I	a=0	Câncer de pele
b	44	435	42	0	I	b=1	Hanseníase
c	22	38	460	0	I	c=2	Psoríase
d	0	0	0	509	I	d=3	Artrite psoriásica

Fonte: Próprio autor (2024).

Na categoria de Câncer de pele (a), o modelo realizou corretamente 463 previsões, enquanto 32 amostras foram erroneamente classificadas como Hanseníase (b) e 21 como Psoríases (c); não houve nenhuma classificação equivocada como Artrite psoriásica (d).

No grupo de Hanseníase (b), o modelo acertou 435 previsões, porém 44 amostras foram incorretamente classificadas como Câncer de pele (a) e 42 como Psoríases (c); novamente, não houve classificações errôneas como Artrite psoriásica (d).

Já na categoria Psoríase (c), o modelo acertou 460 previsões, entretanto, 22 amostras foram classificadas incorretamente como Câncer de pele (a) e 38 como Hanseníases (b); não foram feitas classificações equivocadas como Artrite psoriásica (d). Por fim, na categoria de

Artrite psoriática (d), o modelo acertou todas as 509 previsões, sem nenhuma amostra classificada erroneamente como outra categoria.

Figura 24 apresenta a matriz de confusão referente ao teste do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 24, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 18 a 22.

Figura 24: Matriz de confusão referente ao primeiro diferencial no modo teste.

	a	b	c	d			
a	108	16	5	0	L	a=	Câncer de pele
b	10	109	11	0	I	b=	Hanseníase
c	13	18	77	0	I	c=	Psoríase
d	0	1	1	0	I	d=	Artrite psoriásica

Fonte: Próprio autor (2024).

A análise da matriz de confusão revelou resultados promissores na classificação das diferentes doenças dermatológicas. Na categoria de Câncer de pele (a), o modelo alcançou 108 previsões corretas, enquanto 16 amostras foram erroneamente classificadas como Hanseníase (b), 5 como Psoríase (c), e nenhuma como Artrite psoriásica (d).

Em relação à Hanseníase (b), foram acertadas 109 previsões, com 10 amostras erroneamente classificadas como Câncer de pele (a), 11 como Psoríase (c), e nenhuma como Artrite psoriásica (d).

Para a categoria de Psoríase (c), o modelo obteve 77 previsões corretas, no entanto, 13 amostras foram classificadas erroneamente como Câncer de pele (a), e 18 como Hanseníase (b), sem nenhuma amostra incorretamente classificada como Artrite psoriásica (d).

Por fim, na categoria de Artrite psoriásica (d), o modelo não realizou previsões corretas. uma amostra foi classificada erroneamente como Psoríase (c) e outra como Hanseníase (b).

A artrite psoriásica, por outro lado, apresentou um desafio maior para o modelo, com uma incapacidade de prever corretamente qualquer caso da doença. Isso pode ser atribuído à similaridade das características visuais entre a artrite psoriásica e outras condições dermatológicas e ao número pequeno do conjunto, dificultando sua distinção com base apenas em imagens.

### 5.5.2 Segundo diferencial

Figura 25 apresenta a matriz de confusão referente ao treino do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 12, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras de 13 a 17.

Figura 25: Matriz de confusão referente ao segundo diferencial no modo treino.

	a	b	c			
a	267	24	21	I	a=	Carcinoma basocelular
b	3	309	0	I	b=	Carcinoma células escamosas
c	0	0	310	I	c=	Melanoma

Fonte: Próprio autor (2024)

A análise da matriz de confusão mostrou resultados promissores na classificação das diferentes tipos de câncer de pele. Na categoria de Carcinoma basocelular (a), o modelo obteve 267 previsões corretas, com 24 amostras erroneamente classificadas como Carcinoma de células escamosas (b) e 21 como Melanoma (c).

Em relação ao Carcinoma de células escamosas (b), o modelo acertou 309 previsões, sem nenhuma amostra erroneamente classificada como Carcinoma basocelular (a) ou Melanoma (c).

Para o Melanoma (c), o modelo obteve 310 previsões corretas, sem nenhuma amostra erroneamente classificada como Carcinoma basocelular (a) ou Carcinoma de células escamosas (b).

Figura 26 apresenta a matriz de confusão referente ao teste do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 25, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 18 a 22.

Figura 26: Matriz de confusão referente ao segundo diferencial no modo teste.

	a	b	c			
a	70	4	4	I	a=	Carcinoma basocelular
b	17	4	3	I	b=	carcinoma células escamosas
c	17	3	6	I	c=	Melanoma

Fonte: Próprio autor (2024)

A análise da matriz de confusão expôs resultados na classificação dos diferentes tipos de câncer de pele. Na categoria de Carcinoma basocelular (a), o modelo obteve 70 previsões corretas, com 4 amostras erroneamente classificadas como Carcinoma de células escamosas (b) e 4 como Melanoma (c).

Em relação ao Carcinoma de células escamosas (b), o modelo acertou 4 previsões, com 17 amostras erroneamente classificadas como Carcinoma basocelular (a) e 3 como Melanoma (c).

Para o Melanoma (c), o modelo obteve 6 previsões corretas, com 17 amostras erroneamente classificadas como Carcinoma basocelular (a) e 3 como Carcinoma de células escamosas (b).

### 5.5.3 Terceiro diferencial

SVM (Kernel de Grau 3) com extração de atributos pela Rede Neural Inception V3, Figura 27, apresenta a matriz de confusão referente ao treino do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 15, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras de 13 a 17.

Figura 27: Matriz de confusão referente ao terceiro diferencial no modo treino.

	a	b	c			
a	208	3	9	I	a=	Borderline
b	0	29	156	I	b=	Lepromatous
c	0	15	28	I	c=	Tuberculoid

Fonte: Próprio autor (2024)

A análise da matriz de confusão amostra resultados significativos na classificação das diferentes classes dermatológicas. Para a classe Borderline (a), observamos que 208 amostras foram corretamente classificadas, enquanto 3 foram erroneamente identificadas como Lepromatous (b) e 9 como Tuberculoid (c). Na classe Lepromatous (b), 29 amostras foram corretamente classificadas, e 156 como (c) tuberculoid.

Por fim, na classe Tuberculoid (c), todas as 28 amostras foram corretamente identificadas, porém 15 foram erroneamente classificadas como lepromatous (b). Esses resultados destacam a importância da matriz de confusão na avaliação do desempenho do modelo de classificação, fornecendo informações detalhadas sobre sua capacidade de distinguir entre as diferentes classes.

Figura 28 apresenta a matriz de confusão referente ao teste do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 26, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 18 a 22.

Figura 28: Matriz de confusão referente ao terceiro diferencial no modo teste.

	a	b	c			
a	11	10	8	I	a=	Borderline
b	4	41	10	I	b=	Lepromatous
c	3	18	25	I	c=	tuberculoid

Fonte: Próprio autor (2024)

A análise da matriz de confusão proporciona informações essenciais sobre a capacidade do modelo em classificar adequadamente as diversas classes dermatológicas. Neste contexto, examinamos os resultados relativos às classes Borderline (a), Lepromatous (b) e Tuberculoid (c): Para a classe Borderline (a), foram corretamente identificadas 11 amostras, enquanto 10 amostras foram erroneamente classificadas como Lepromatous (b) e 8 como Tuberculoid (c). Na classe Lepromatous (b), todas as 41 amostras foram corretamente classificadas, entretanto, 4 amostras foram erroneamente identificadas como Borderline (a) e 10 como Tuberculoid (c). Por fim, na classe Tuberculoid (c), 25 amostras foram corretamente identificadas, mas 3 foram erroneamente classificadas como Borderline (a) e 18 como Lepromatous (b).

#### 5.5.4 Quarto diferencial

Classificador *Random Forest* de 500 com extração de atributos pela Rede Neural Inception V3. Figura 29 apresenta a matriz de confusão referente ao treino do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 18, onde são for-

necidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 13 a 17.

Figura 29: Matriz de confusão referente ao quarto diferencial no modo treino.

	a	b	c	d	e	f	G			
a	154	7	11	1	1	11	3	I	a=	Placas
b	0	187	0	0	0	0	0	I	b=	Eritrodérmica
c	0	0	188	0	0	0	0	I	c=	Gutata
d	0	0	0	186	0	0	0	I	d=	Invertida
e	0	0	0	0	188	0	0	I	e=	Palmar e plantar
f	0	0	0	0	0	186	0	I	f=	Pustular
g	0	0	0	0	0	0	184	I	g=	Ungueal

Fonte: Próprio autor (2024).

A análise minuciosa da matriz de confusão proporciona insights valiosos sobre a capacidade do modelo em classificar as diversas manifestações clínicas da psoríase. Podemos compreender melhor o desempenho do modelo observando os resultados para classe Placas (a), 154 amostras foram corretamente identificadas, enquanto ocorreram 7 classificações incorretas como Eritrodérmica (b), 11 como Gutata (c), 1 como Invertida (d), 1 como Palmar e plantar (e), 11 como Pustular (f) e 3 como Ungueal (g).

Na classe Eritrodérmica (b), todas as 187 amostras foram corretamente classificadas, sem classificações incorretas para as outras classes. Para a classe Gutata (c), todas as 188 amostras foram corretamente classificadas, sem classificações incorretas para as outras classes. Na classe Invertida (d), todas as 186 amostras foram corretamente classificadas, sem classificações incorretas para as outras classes. Para a classe Palmar e plantar (e), todas as 188 amostras foram corretamente classificadas, sem classificações incorretas para as outras classes. Na classe Pustular (f), todas as 186 amostras foram corretamente classificadas, sem classificações incorretas para as outras classes. Para a classe Ungueal (g), todas as 184 amostras foram corretas.

A classe Placas (a) apresentou alguns desafios para o modelo, com classificações incorretas em várias outras classes. Em contraste, as classes Eritrodérmica, Gutata, Invertida, Palmar e plantar, Pustular e Ungueal foram classificadas com 100% de precisão. embora o modelo seja altamente preciso para a maioria das classes, há espaço para melhorias na classificação da classe Placas.

Figura 30 apresenta a matriz de confusão referente ao teste do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 27, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 18 a 22.

Figura 30: Matriz de confusão referente ao quarto diferencial no modo teste.

	a	b	c	d	e	f	g			
a	47	0	0	0	0	0	0	I	a=	Placas
b	8	0	0	0	0	0	0	I	b=	Eritrodérmica
c	14	0	0	0	0	0	0	I	c=	Gutata
d	4	0	0	0	0	0	0	I	d=	Invertida
e	5	0	0	0	0	0	0	I	e=	Palmar e plantar
f	17	0	0	0	0	0	0	I	f=	Pustular
g	8	0	0	0	0	0	00	I	g=	Ungueal

Fonte: Próprio autor (2024).

A análise da matriz de confusão revelou que todas as amostras foram classificadas como Placas (a). Não foram identificadas amostras em outras classes. A falta de amostras em algumas classes pode resultar em um viés do modelo em favor das classes com mais amostras, levando a uma classificação incorreta das amostras das classes menos representadas. Isso ressalta a importância de ter conjuntos de dados bem equilibrados para treinar modelos de aprendizado de máquina, especialmente em tarefas de classificação multiclasse.

### 5.5.5 Quinto diferencial

SVM (Kernel de Grau 3) com extração de atributos pela Rede Neural Inception V3, Figura 31, apresenta a matriz de confusão referente ao treino do melhor classificador e da rede neural profunda. Os detalhes dessa matriz estão disponíveis na Tabela 21, onde são fornecidos os valores para as diferentes classes. Além disso, o Gráfico de Boxplot relacionado ao modo de treino pode ser observado nas Figuras 13 a 17.

Figura 31: Matriz de confusão referente ao quinto diferencial no modo treino.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	
a	245	14	17	3	6	12	4	0	1	0	0	0	2	0	I
b	0	311	0	0	0	1	0	0	0	0	0	0	0	0	I
c	0	0	310	0	0	0	0	0	0	0	0	0	0	0	I
d	1	1	1	305	2	0	0	0	0	0	0	0	0	0	I
e	8	0	0	12	270	12	3	0	0	0	0	1	1	0	I
f	7	1	0	4	5	286	5	0	0	0	1	0	0	0	I
g	3	2	1	5	5	1	278	0	0	0	0	7	2	0	I
h	0	0	0	0	0	0	0	309	0	0	0	0	0	0	I
i	0	0	0	0	0	0	0	0	312	0	0	0	0	0	I
j	0	0	0	0	0	0	0	0	0	309	0	0	0	0	I
k	0	0	0	0	0	0	0	0	0	0	313	0	0	0	I
l	0	0	0	0	0	0	0	0	1	0	0	312	0	0	I
m	0	0	0	0	0	0	0	0	0	0	0	0	313	0	I

N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	312	I
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	-----	---

Fonte : Próprio autor (2024).

Figura 32: Legenda da figura 31.

a= Carcinoma basocelular	i= Gutata
b= Carcinoma células escamosas	j= Invertida
c= Melanoma	k= Palmar e plantar
d= Borderline	l= Pustular
e= Lepromatous	m= Ungueal
f= Tuberculoid	n= Artrite psoriásica
g= Placas	
h= Eritrodérmica	

Fonte : Próprio autor (2024)

Na classe Carcinoma basocelular (a), foram corretamente classificadas 245 amostras, enquanto houve 14 amostras erroneamente classificadas como Carcinoma células escamosas (b), 17 como Melanoma (c), 3 como Borderline (d), 6 como Lepromatous (e), 12 como Tuberculoid (f), 4 como Placas (g), 1 como Gutata (i), e 2 como Ungueal (m).

Na classe Carcinoma células escamosas (b), 311 amostras foram corretamente identificadas mas houve 1 como *Tuberculoid* (f).

Na classe Melanoma (c), 310 amostras foram corretamente identificadas ,

Na classe Borderline (d), 305 amostras foram corretamente identificadas, mas houve 1 amostra erroneamente classificada como Carcinoma basocelular (a), 1 como Carcinoma células escamosas (b) , 1 como Melanoma (c) e 2 como *Lepromatous* (e).

Na classe Lepromatous (e), 270 amostras foram corretamente identificadas, mas houve 8 amostras erroneamente classificadas como Carcinoma basocelular (a) e 12 como Tuberculoid (f).

Na classe *Tuberculoid* (f), 286 amostras foram corretamente identificadas, mas houve 7 amostras erroneamente classificadas como Carcinoma basocelular (a), 1 como Carcinoma células escamosas (b), 4 como *Borderline* (d) e 5 como *Lepromatous* (e).

Na classe Placas (g), 278 amostras foram corretamente identificadas, mas houve 3 amostras erroneamente classificadas como Carcinoma basocelular (a), 2 como Carcinoma células



k	1	0	0	0	0	0	3	0	0	0	0	0	1	0	I
l	1	0	0	3	1	0	8	0	1	0	0	3	0	0	I
m	0	0	0	0	0	0	6	0	0	0	0	0	2	0	I
n	0	0	0	0	0	1	0	0	0	0	0	0	1	0	I

Fonte: Próprio autor (2024).

Figura 34: Legenda da figura 33.

a= Carcinoma basocelular	i= Gutata
b= Carcinoma células escamosas	j= Invertida
c= Melanoma	k= Palmar e plantar
d= Borderline	l= Pustular
e= Lepromatous	m= Ungueal
f= Tuberculoid	n= Artrite psoriásica
g= Placas	
h= Eritrodérmica	

Fonte : Próprio autor (2024).

Na classe (a) - Carcinoma basocelular, foram corretamente classificadas 58 amostras como Carcinoma basocelular. No entanto, houve classificações incorretas, com 5 amostras classificadas como Carcinoma células escamosas, 4 como Melanoma, 1 como Borderline, 1 como Lepromatous, 2 como Tuberculoid, 5 como Placas, 1 como Eritrodérmica, 1 como Gutata, 2 como Invertida, 1 como Palmar e plantar, 3 como Pustular, 6 como Ungueal, e 1 como Artrite psoriásica.

Na classe (b) - Carcinoma células escamosas, 5 amostras foram corretamente classificadas como Carcinoma células escamosas. Houve 11 amostras classificadas incorretamente, com 5 como Carcinoma basocelular, 1 como Melanoma, 3 como Borderline, 3 como Lepromatous, 2 como Tuberculoid, 2 como Placas, 1 como Eritrodérmica, 1 como Gutata, 2 como Pustular, e 1 como Ungueal.

Na classe (c) - Melanoma, 4 amostras foram corretamente classificadas como Melanoma. No entanto, houve 17 amostras classificadas incorretamente, com 17 como Carcinoma baso-

celular, 2 como Borderline, 2 como Lepromatous, 3 como Tuberculoid, 3 como Placas, 1 como Eritrodérmica, 1 como Gutata, 1 como Invertida, e 3 como Pustular.

Na classe (d) - Borderline, 6 amostras foram corretamente classificadas como Borderline. Houve 5 amostras classificadas incorretamente, com 5 como Carcinoma basocelular, 1 como Carcinoma células escamosas, 7 como Lepromatous, 6 como Tuberculoid, 1 como Placas, 1 como Eritrodérmica, 5 como Gutata, 2 como Invertida, e 3 como Pustular.

Na classe (e) - Lepromatous, 31 amostras foram corretamente classificadas como Lepromatous. Houve 4 amostras classificadas incorretamente, com 4 como Carcinoma basocelular, 2 como Carcinoma células escamosas, 7 como Borderline, 5 como Tuberculoid, 5 como Placas, 1 como Eritrodérmica, 1 como Gutata, 1 como Palmar e plantar, 3 como Pustular, e 1 como Ungueal.

Na classe (f) - Tuberculoid, 28 amostras foram corretamente classificadas como Tuberculoid. Houve 3 amostras classificadas incorretamente, com 3 como Carcinoma basocelular, 11 como Lepromatous, 3 como Placas, 1 como Eritrodérmica, 1 como Gutata, 2 como Borderline, 1 como Pustular, e 1 como Ungueal.

Na classe (g) - Placas, 22 amostras foram corretamente classificadas como Placas. Houve 8 amostras classificadas incorretamente, com 8 como Carcinoma basocelular, 2 como Melanoma, 1 como Borderline, 9 como Lepromatous, 5 como Tuberculoid, 1 como Eritrodérmica, 1 como Gutata, 1 como Invertida, e 1 como Pustular.

Na classe (h) - Eritrodérmica, 2 amostras foram corretamente classificadas como Eritrodérmica. Houve 1 amostra classificada incorretamente como Carcinoma basocelular, 1 como Borderline, 1 como Lepromatous, 1 como Placas, 3 como Gutata, 2 como Pustular, e 1 como Artrite psoriásica.

Na classe (i) - Gutata, 1 amostra foi corretamente classificada como Gutata. Houve 4 amostras classificadas incorretamente, com 4 como Carcinoma basocelular, 5 como Lepromatous, 2 como Placas, 2 como Eritrodérmica, 5 como Pustular, 6 como Ungueal, e 1 como Artrite psoriásica.

Na classe (j) - Invertida, não houve amostras corretamente classificadas como Invertida. Houve 2 amostras classificadas incorretamente, com 2 como Carcinoma basocelular, 2 como Borderline, e 1 como Ungueal.

Na classe (k) - Palmar e plantar, 3 amostras foram corretamente classificadas como Palmar e plantar. Houve 1 amostra classificada incorretamente como Carcinoma basocelular, 3 como Pustular, e 1 como Ungueal.

Na classe (l) - Pustular, 3 amostras foram corretamente classificadas como Pustular. Houve 1 amostra classificada incorretamente como Carcinoma basocelular, 3 como Borderline, 1 como Lepromatous, 8 como Placas, e 1 como gutata.

Na classe (m) - Ungueal, 2 amostras foram corretamente classificadas como Ungueal. Houve 6 amostras classificadas incorretamente.

Na classe (n) - Artrite psoriásica, não houve amostras corretamente classificadas como Artrite psoriásica.

Observa-se que os modelos treinados apresentaram um desempenho satisfatório durante a fase de treinamento, com métricas de avaliação robustas. No entanto, ao avaliar esses modelos no conjunto de teste, encontrou-se uma queda significativa na performance. Avalia-se que a distribuição dos dados entre treinamento e teste é diferente, o que afetou a capacidade de generalização dos modelos. Além disso, o tamanho limitado do conjunto de teste contribuiu para uma avaliação menos confiável dos modelos.

A diferença na distribuição dos dados entre treinamento e teste pode comprometer a capacidade de generalização dos modelos de aprendizado de máquina, pois eles podem não ser capazes de lidar com novos padrões presentes nos dados de teste. Além disso, o tamanho limitado do conjunto de teste pode resultar em estimativas imprecisas da performance do modelo. Esses resultados destacam a importância de uma seleção cuidadosa e preparação dos conjuntos de treinamento e teste para garantir uma avaliação precisa do modelo.

## 6 CONCLUSÃO

Este estudo oferece uma estrutura promissora para o diagnóstico precoce de doenças dermatológicas, com potenciais benefícios práticos na prática clínica. Ao implementar ferramentas de inteligência artificial, é possível acelerar o diagnóstico, reduzir complicações e melhorar a gestão dos recursos de saúde.

No futuro, a pesquisa pode se concentrar em expandir o uso de algoritmos de aprendizado profundo, explorar novos conjuntos de dados e validar essas ferramentas em diferentes contextos clínicos.

Com base nas análises realizadas, fica evidente que o modelo Inception V3 demonstrou um desempenho superior em comparação com os modelos VGG16 e ResNet-50. Essa superioridade pode ser atribuída a diversas vantagens específicas do Inception V3. Sua arquitetura complexa e sofisticada, com módulos de Inception capazes de capturar características em di-

ferentes escalas e níveis de abstração, mostrou-se particularmente adequada para os conjuntos de dados e tarefas específicas em questão.

Além disso, a extração de atributos do Inception V3 resultou na captura de características mais relevantes e discriminativas nos dados de entrada, o que facilitou previsões mais precisas pelo classificador SVM de kernel de grau 3 seguida do Random Forest de 500. A incorporação de técnicas de regularização, como dropout e normalização em lote, na arquitetura do Inception V3 também contribuiu para sua capacidade de generalizar melhor para dados não vistos e seu desempenho superior em comparação com outros modelos.

O pré-treinamento do Inception V3 no conjunto de dados ImageNet desempenhou um papel crucial, proporcionando uma inicialização mais eficaz dos pesos do modelo e permitindo que ele capturasse representações úteis das características dos dados. Além disso, a otimização eficaz dos hiperparâmetros do Inception V3 durante o treinamento contribuiu para seu desempenho superior em comparação com os modelos concorrentes.

Os resultados obtidos também destacam a influência da distribuição diferenciada dos dados entre treinamento e teste, juntamente com o tamanho limitado do conjunto de teste, na generalização dos modelos de aprendizado de máquina em dados médicos. Esses fatores impactaram negativamente a capacidade dos modelos de generalizar para novos dados, enfatizando a importância de considerá-los ao avaliar a performance de modelos de predição em contextos clínicos.

Em suma, os resultados deste estudo reforçam a importância da seleção cuidadosa do modelo e da arquitetura ao abordar problemas de classificação em conjuntos de dados complexos. A combinação da arquitetura Inception V3 com o classificador SVM de kernel de grau 3 mostrou-se eficaz na detecção e classificação de doenças dermatológicas em imagens médicas, demonstrando seu potencial para aplicações clínicas no diagnóstico precoce e preciso de uma variedade de condições dermatológicas.

Recomenda-se que futuras pesquisas explorem métodos para mitigar o impacto da diferença na distribuição dos dados e do tamanho limitado do conjunto de teste. Em última análise, abordar esses desafios é crucial para desenvolver modelos mais robustos e eficazes para a tomada de decisões clínicas. Esperamos que este estudo forneça insights valiosos e inspire futuras pesquisas no campo da aprendizagem de máquina aplicada à medicina.

## 7 REFERÊNCIAS

ALALI-NAJAFABADI, F. et al. Application of Information Theoretic Feature Selection and Machine Learning Methods for the Development of Genetic Risk Prediction Models. *Scientific Reports*, v. 11, n. 1, p. 1-15, 2 dez. 2021. DOI: 10.1038/s41598-021-00854-x.

AGGARWAL, S. L. P. Data Augmentation in Dermatology Image Recognition Using Machine Learning. *Skin Research and Technology*, v. 25, n. 6, p. 815-820, nov. 2019. DOI: 10.1111/srt.12726.

AHMAD, A. et al. Psoriatic Arthritis. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing, 2021. Disponível em: . Acesso em: 18 ago. 2023.

AIJAZ, A. et al. Psoriasis: An updated overview and recent advancements in treatment. *Asian Journal of Pharmaceutical Sciences*, v. 17, n. 4, p. 464-481, 2022.

ALZUBAIDI, L. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, v. 8, n. 1, art. 53, 31 mar. 2021. DOI: 10.1186/s40537-021-00422-8.

AGÊNCIA FIOCRUZ DE NOTÍCIAS. Especial: Doenças Negligenciadas. Rio de Janeiro, 2 ago. 2013. Disponível em: <https://agencia.fiocruz.br/doen%C3%A7as-negligenciadas>. Acesso em: 18 ago. 2023.

AMERICAN CANCER SOCIETY. *Cancer A-Z. Skin cancer*. Atlanta: ACS, 2019. Disponível em: <https://www.cancer.org/cancer/skin-cancer.html>. Acesso em: 10/04/2023.

ASTORINO, Annabella et al. Melanoma detection by means of multiple instance learning. *Interdisciplinary Sciences: Computational Life Sciences*, Springer, v. 12, n. 1, p. 24–31, 2020.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde, Departamento de Doenças de Condições Crônicas e Infecções Sexualmente Transmissíveis. Protocolo Clínico e Diretrizes Terapêuticas da Hanseníase [recurso eletrônico] / Ministério da Saúde, Secretaria de Ciência, Tecnologia, Inovação e Insumos Estratégicos em Saúde, Secretaria de Vigilância em Saúde. – Brasília : Ministério da Saúde, 2022. 152 p. : il.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. Guia prático sobre a hanseníase [recurso eletrônico]. Brasília: Ministério da Saúde, 2017. 68 p. : il.

BAWEJA, Harjatin Singh; PARHAR, Tanvir. Leprosy lesion recognition using convolutional neural networks. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERNETICS (ICMLC), 2016, p. 141-145.

BUZUTI, Lucas Fontes; CARLOS, Eduardo Thomaz. Reconhecimento de faces por aprendizado profundo: dropout em autoencoders. 2018. Disponível em [https://fei.edu.br/sites/sicfei/2018/eletrica/SICFEI\\_2018\\_paper\\_68.pdf](https://fei.edu.br/sites/sicfei/2018/eletrica/SICFEI_2018_paper_68.pdf). Acesso em: 17 mar. 2024.

CASANOVA, J. M., M. Baradad, X. Soria, R. M. Martí. “www.dermatoweb.net. Una web docente para el aprendizaje de la Dermatología en el pregrado\*”. *Actas Dermo-Sifiliográficas*, v. 100, n. 10, p. 866–874, 1o de dezembro de 2009. [https://doi.org/10.1016/S0001-7310\(09\)72915-1](https://doi.org/10.1016/S0001-7310(09)72915-1).

CHEN, Kou-Huang et al. Leprosy: A review of epidemiology, clinical diagnosis, and management. *Journal of Tropical Medicine*, Hindawi, v. 2022, 2022.

COPPIN, B. (2010). *Artificial Intelligence Illuminated*. Jones & Bartlett Publishers.

DE SOUZA, M. L. M. et al. Leprosy Screening Based on Artificial Intelligence: Development of a Cross-Platform App. *JMIR mHealth and uHealth*, v. 9, n. 4, p. e23718, 7 abr. 2021.

DASH, Manoranjan et al. PsLSNet: Automated Psoriasis Skin Lesion Segmentation Using Modified U-Net-Based Fully Convolutional Network. *Biomedical Signal Processing and Control*, v. 52, p. 226-237, 1 de julho de 2019. Disponível em: <https://doi.org/10.1016/j.bspc.2019.04.002>.

DE GOMA, Joel Casuayan; DEVARAJ, Madhavi. Recognizing Common Skin Diseases in the Philippines Using Image Processing and Machine Learning Classification. In: *2020 3rd International Conference on Computing and Big Data*, 2020. p. 68-72.

DE SOUZA, MLM et al. Leprosy Screening Based on Artificial Intelligence: Development of a Cross-Platform App. *JMIR Mhealth Uhealth*, v. 9, n. 4, p. e23718, 7 de abril de 2021. DOI: 10.2196/23718. PMID: 33825685; PMCID: PMC8060869.

DENG, J. et al. ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Anais... Em: *2009 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*. jun. 2009. Disponível em: <<https://ieeexplore.ieee.org/document/5206848>>. Acesso em: 6 mar. 2023.

DERMATOWEB. Acessado em: 3 jul. 2023. Disponível em: <http://dermatoweb.udl.es/.e18438>.

ELREEDY, Dina; ATIYA, Amir F. A Comprehensive Analysis of Synthetic Minority Over-sampling Technique (SMOTE) for handling class imbalance. *Information Sciences*, v. 505, p. 32–64, 1 dez. 2019. Disponível em: <https://doi.org/10.1016/j.ins.2019.07.070>. Acesso em: 17 mar. 2024.

ESTEVA, Andre et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, Nature Publishing Group, v. 542, n. 7639, p. 115–118, 2017.

FERREIRA, Janderson, Allan Feitosa, Juliana Gomes, Abel Silva-Filho, e Wellington Dos Santos. Evaluation of molecular docking by deep learning and Random Forests: A hybrid approach based on pseudo-convolutions, 2023. <https://doi.org/10.21203/rs.3.rs-3637282/v1>.

FOGEL, Alexander L.; KVEDAR, Joseph C. Artificial intelligence powers digital medicine. *NPJ Digital Medicine*, Nature Publishing Group, v. 1, n. 1, p. 1–4, 2018.

GEORGE, Yasmeen; ALDEEN, Mohammad; GARNAVI, Rahil. Automatic Scale Severity Assessment Method in Psoriasis Skin Images Using Local Descriptors. *IEEE Journal of Biomedical and Health Informatics*, v. 24, n. 2, p. 577–585, fevereiro de 2020. DOI: 10.1109/JBHI.2019.2910883.

HE, Jianxing et al. The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, v. 25, n. 1, p. 30–36, 2019.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, v. 60, n. 6, p. 84–90, 24 maio 2017. DOI: 10.1145/3065386.

INSTITUTO NACIONAL DO CÂNCER. ABC do Câncer: Abordagens Básicas para o Controle do Câncer. Rio de Janeiro: Ministério da Saúde. Instituto Nacional de Câncer José Alencar Gomes da Silva, 2011. ISBN 978-85-7318-188-3.

IOFFE, S.; SZEGEDY, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv*, , 2 mar. 2015. Disponível em: <<http://arxiv.org/abs/1502.03167>>. Acesso em: 17 mar. 2024

JIN, S. et al. Comorbidities and complications of psoriasis: A comprehensive review. *Dermatology and Therapy*, v. 13, n. 1, p. 1-19, 2023.

JOSHI, Amruta et al. Skin disease detection and classification. *International Journal of Advanced Engineering Research and Science (IJAERS)*, v. 6, n. 5, p. 396–400, 2019.

KIM, W. B. et al. (2023). Psoriasis: Epidemiology, clinical features, and quality of life. *Dermatology and Therapy*, 13(1), 21-39.

KUNDAKCI, Nihal; ERDEM, Cengizhan. Leprosy: A great imitator. *Clinics in dermatology*, Elsevier, v. 37, n. 3, p. 200–212, 2019.

LI, Z. et al. Artificial Intelligence in Dermatology Image Analysis: Current Developments and Future Trends. *Journal of Clinical Medicine*, v. 11, n. 22, p. 6826, jan. 2022.

LIU, M. et al. Psoriatic Arthritis: A Review of Disease Presentation, Pathogenesis, and Management. *The Journal of Clinical and Aesthetic Dermatology*, v. 15, n. 1, p. E26-E33, 2022.

LANG, Steven et al. WekaDeeplearning4j: A deep learning package for Weka based on Deeplearning4j. *Knowledge-Based Systems*, Elsevier, v. 178, p. 48–50, 2019.

LOBO, Luiz Carlos. Inteligência artificial, o Futuro da Medicina e a Educação Médica. *Revista Brasileira de Educação Médica*, Rio de Janeiro, v. 42, n. 3, p. 598-600, jul.-set. 2018.

LORENA, A. C.; CARVALHO, A. C. P. L. F. DE. Uma Introdução às Support Vector Machines. *Revista de Informática Teórica e Aplicada*, v. 14, n. 2, p. 43–67, 20 dez. 2007.

LOVE, Thorvardur Jon, Tianxi Cai, e Elizabeth W. Karlson. Validation of Psoriatic Arthritis Diagnoses in Electronic Medical Records Using Natural Language Processing. *Seminars in Arthritis and Rheumatism*, v. 40, n. 5, p. 413–420, abril 2011.

MACHADO, C. K. et al. "Projeto Pele Alerta": prevenção e detecção precoce do câncer de pele direcionado a profissionais de beleza. *Revista Brasileira de Cirurgia Plástica*, v. 36, n. 2, p. 123-135, 2021. DOI: 10.5935/2177-1235.2021RBCP0074.

MITCHELL, T. M. *Machine Learning*. McGraw-Hill, 1997.

NAVARINI, Luca et al. A Machine-Learning Approach to Cardiovascular Risk Prediction in Psoriatic Arthritis. *Rheumatology (Oxford, England)*, v. 59, n. 7, p. 1767–1769, 1 de julho de 2020. DOI: 10.1093/rheumatology/kez677.

WORLD HEALTH ORGANIZATION. *A guide to eliminating leprosy as a public health problem*. 2nd ed. Geneva, 1997.

PEREZ, Fábio; AVILA, Sandra; VALLE, Eduardo. Solo or Ensemble? Choosing a CNN Architecture for Melanoma Classification. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019. p. 2775-2783.

PUIG, L. et al. Generalized Pustular Psoriasis: An Update and Review. *Actas Dermo-Sifilográficas (English Edition)*, v. 114, n. 5, p. 343-354, 2023.

RAJ, A. D.; LONDHE, N. D.; SONAWANE, R. S. Segmentation of psoriasis lesion using image processing techniques: A survey. *Journal of Skin and Stem Cell*, v. 7, n. 1, p. 1-6, 2020. RAY, A.; GUPTA, A.; AL, A. Skin Lesion Classification With Deep Convolutional.

SAÚDE, M. DA. *Guia de Vigilância em Saúde*. 5. ed. Brasília, DF: Ministério da Saúde, 2022.

SANTACROCE, Luigi et al. *Mycobacterium leprae: A historical study on the origins of leprosy and its social stigma*. *Le Infezioni in Medicina*, v. 29, n. 4, p. 623, 2021.

SANTOS, Wellington Pinheiro. *Análise Monoespectral De Imagens Funcionais De Ressonância Magnética Utilizando Mapas Auto-Organizados De Kohonen*. In: *Anais Do 8º Congresso Brasileiro de Redes Neurais*, 1º de janeiro de 2016. Disponível em: [https://www.academia.edu/93494046/An%C3%A1lise\\_Monoespectral\\_De\\_Imagens\\_Funcionais\\_De\\_Resson%C3%A2ncia\\_Magn%C3%A9tica\\_Utilizando\\_Mapas\\_Auto\\_Organizados\\_De\\_Kohonen](https://www.academia.edu/93494046/An%C3%A1lise_Monoespectral_De_Imagens_Funcionais_De_Resson%C3%A2ncia_Magn%C3%A9tica_Utilizando_Mapas_Auto_Organizados_De_Kohonen). Acesso em: 17 mar. 2024.

SANTOS, M. C. (2005). *Introdução ao aprendizado de máquina (Machine Learning)*. Instituto de Ciências Matemáticas e de Computação (ICMC) - Universidade de São Paulo.

SHALU; KAMBOJ, Aman. A color-based approach for melanoma skin cancer detection. In: 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018. p. 508-513.

SHRIVASTAVA, Vimal K. et al. Reliability Analysis of Psoriasis Decision Support System in Principal Component Analysis Framework. *Data & Knowledge Engineering*, v. 106, p. 1-17, 1 de novembro de 2016. DOI: 10.1016/j.datak.2016.09.001.

SILVA, Samuel Freire da. Atlas Dermatológico. 1995. Disponível em: <http://www.atlasdermatologico.com.br/>. Acesso em: 01/03/2024.

SILVA, RODRIGO EMERSON VALENTIM DA. Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens. Monografia (Graduação em Sistemas de Informação) - Universidade Federal do Ceará, Campus Quixadá, 2018.

SILVEIRA, Mariana. Introdução à Aprendizagem Profunda (Deep Learning). *Mentes Artificiais*, 23 de março de 2023. Disponível em: <https://mentesartificiais.com/introducao-a-aprendizagem-profunda-deep-learning/>. Acesso em: (12/12/2023).

SOUZA, Victor et al. Análise Comparativa de Redes Neurais Convolucionais no Reconhecimento de Cenas. 3 de dezembro de 2020. DOI: 10.14210/cotb.v11n1.p419-426.

VELAZQUEZ, Matthew; LEE, Yugyung. Random forest model for feature-based Alzheimer's disease conversion prediction from early mild cognitive impairment subjects. *PLoS ONE*, v. 16, n. 4, 29 de abril de 2021. DOI: 10.1371/journal.pone.0244773.

VILELA, I.F. et al. Impacto da pandemia do vírus COVID-19 nas internações para tratamento de câncer de pele no Brasil. *Revista Brasileira de Cirurgia Plástica*, v. 36, n. 3, p. 303-308, jul.-set. 2021.

YU, K.-H.; BEAM, A. L.; KOHANE, I. S. Artificial intelligence in healthcare. *Nature Biomedical Engineering*, v. 2, n. 10, p. 719-731, out. 2018.

