



Campus Recife

Departamento Acadêmico dos Cursos Superiores

Curso de Tecnologia em Análise e Desenvolvimento de Sistemas

JAMILLY BRAGA MELO

**ANS2JSON: uma API baseada em IA, para extração de
informações da Avaliação Neurológica de pacientes com
hanseníase.**

Recife

2025

JAMILLY BRAGA MELO

**ANS2JSON: uma API baseada em IA, para extração de informações da
Avaliação Neurológica de pacientes com hanseníase.**

Trabalho de Conclusão de Curso apresentado
como requisito parcial para obtenção do título
de Tecnólogo em Análise e Desenvolvimento de
Sistemas, no Instituto Federal de Educação,
Ciência e Tecnologia de Pernambuco (IFPE),
Campus Recife.

Orientador: Me. Hilson Gomes Vilar de Andrade

Recife

2025

Ficha catalográfica elaborada pela bibliotecária Danielle Castro da Silva CRB4/1457

M528a
2025

Melo, Jamilly Braga

ANS2JSON: uma API baseada em IA, para extração de informações da Avaliação Neurológica de pacientes com hanseníase / Jamilly Braga Melo. --- Recife: A autora, 2025.

70f. il. Color.

Trabalho de Conclusão (Curso Superior Tecnológico em Análise e Desenvolvimento de Sistemas) – Instituto Federal de Pernambuco, Recife, 2025.

Inclui Referências.

Orientador: Prof. Me. Hilson Gomes Vilar de Andrade.

1. Hanseníase. 2. Avaliação Neurológica Simplificada. 3. Inteligência Artificial. 4. Digitalização Automatizada. I. Título. II. Andrade, Hilson Gomes Vilar de (orientador). III. Instituto Federal de Pernambuco.

CDD 658.404 (23 ed.)

JAMILLY BRAGA MELO

**ANS2JSON: uma API baseada em IA, para extração de informações da
Avaliação Neurológica de pacientes com hanseníase.**

Trabalho de Conclusão de Curso
apresentado, como requisito
parcial para obtenção do título de Tecnólogo
em Análise e Desenvolvimento de Sistemas,
do Instituto Federal de Educação, Ciência e
Tecnologia de Pernambuco – Campus Recife.

Aprovado em: 16 de setembro de 2025.

Banca Examinadora:

Prof. Me. Hilson Gomes Vilar de Andrade (Orientador)
IFPE - Campus Recife

Prof. Dr. Ramide Augusto Sales Dantas
IFPE - Campus Recife

Prof^a. Me. Maria Eduarda Ferro de Mello (Externo)
Faculdade Estácio de Sá

Recife

2025

AGRADECIMENTOS

Agradeço, primeiramente, à minha família, que, mesmo sem ter vivido o ensino superior, sempre acreditou no poder da educação e nunca mediu esforços para que eu pudesse trilhar esse caminho.

Ao meu amigo Lucas André, que, desde o instante em que entrou em minha vida, se tornou uma luz em meu caminho, tornando a vida acadêmica mais leve.

Aos professores, que não apenas ensinaram, mas também acreditaram em mim, deixando aprendizados que levarei para além destas páginas.

Ao meu orientador, Hilson Vilar, que me convidou para este projeto e, diante dos desafios, me recordou que “um passo à frente e você já não está no mesmo lugar”, ensinando-me que cada começo é, por si só, um avanço.

E também ao grupo de pesquisa dotLAB Brazil, pelos projetos que promovem o uso da tecnologia no auxílio a pessoas afetadas por doenças tropicais negligenciadas, em particular o projeto Hansen.ai. Agradeço, em especial, à Patrícia Endo, bem como a todos os colegas que colaboraram para a realização deste trabalho.

RESUMO

A Avaliação Neurológica Simplificada (ANS) é uma ferramenta obrigatória e fundamental para o monitoramento de pacientes com hanseníase no Brasil. Contudo, seu formato físico e preenchimento manual representam um grande obstáculo, retendo dados clínicos cruciais em arquivos de papel e impedindo análises longitudinais, pesquisas em larga escala e a formulação de políticas de saúde baseadas em evidências. Este trabalho apresenta o ANS2JSON, um aplicativo inovador que rompe essa barreira através de uma arquitetura híbrida de Inteligência Artificial para a digitalização completa e automática das fichas ANS. A solução enfrenta o desafio de interpretar um formulário complexo, que mescla dados textuais com diagramas visuais, empregando um pipeline de dois estágios: (1) um modelo de visão computacional (YOLOv8 + Random Forest) para detectar e mapear anatômicamente os pontos da avaliação de sensibilidade, e (2) um modelo de Processamento Inteligente de Documentos para extrair os campos textuais. A abordagem demonstrou altíssima eficácia, alcançando 97,5% de precisão na interpretação dos dados visuais e mais de 95% na extração de indicadores clínicos chave. O ANS2JSON, disponibilizado publicamente via web, oferece uma ferramenta robusta para converter um vasto acervo de dados não estruturados em informações digitais, prontas para uso, desbloqueando o potencial de anos de registros clínicos para o avanço no tratamento e na pesquisa da hanseníase. O aplicativo está disponível em: <https://ans2json.dotlabbrasil.com.br>.

Palavras-chaves: Hanseníase; Inteligência Artificial; Digitalização Automatizada; Visão Computacional; Avaliação Neurológica Simplificada.

ABSTRACT

The Simplified Neurological Assessment (SNA) is a mandatory and essential tool for monitoring patients with leprosy in Brazil. However, its physical format and manual completion represent a significant obstacle, as they confine crucial clinical data to paper records and hinder longitudinal analyses, large-scale research, and evidence-based health policymaking. This study introduces ANS2JSON, an innovative application that overcomes this barrier through a hybrid Artificial Intelligence architecture for the complete and automatic digitization of SNA forms.

The proposed solution addresses the challenge of interpreting a complex form that combines textual data with visual diagrams by employing a two-stage pipeline: (1) a computer vision model (YOLOv8 + Random Forest) to detect and anatomically map sensitivity assessment points, and (2) an Intelligent Document Processing model to extract textual fields. The approach demonstrated outstanding effectiveness, achieving 97.5% precision in visual data interpretation and over 95% accuracy in extracting key clinical indicators.

ANS2JSON, publicly available via the web, provides a robust tool for converting a vast repository of unstructured data into ready-to-use digital information, unlocking the potential of years of clinical records to advance both leprosy treatment and research. The application is accessible at: <https://ans2json.dotlabbrazil.com.br>.

Keywords: Leprosy; Artificial Intelligence; Automated Digitization; Computer Vision; Simplified Neurological Assessment.

LISTA DE FIGURAS

Figura 1	Formulário para Avaliação Neurológica Simplificada	6
Figura 2	Exemplo da Avaliação Sensitiva das mãos, na Avaliação Neurológica Simplificada (ANS)	8
Figura 3	Resumo do funcionamento do YOLO.....	10
Figura 4	Descrição anatômica dos pontos da mão	16
Figura 5	Descrição anatômica dos pontos do pé	16
Figura 6	Interface do Azure Document Intelligence Studio na rotulação da ficha ANS.	20
Figura 7	Matriz de confusão do modelo YOLOv8.....	23
Figura 8	Detecções realizadas pelo modelo YOLOv8 no conjunto de teste.....	24
Figura 9	Matriz de confusão do modelo RF.	26

LISTA DE TABELAS

Tabela 1	Comparação entre diferentes modelos de <i>Machine Learning</i> (ML) para o mapeamento anatômico	18
Tabela 2	Desempenho do modelo YOLOv8 no conjunto de validação.	23
Tabela 3	Desempenho do modelo RF para identificação de pontos anatômicos.	25
Tabela 4	Desempenho do modelo de extração de dados textuais (API Azure).	27

LISTA DE ABREVIATURAS E SIGLAS

ANS	Avaliação Neurológica Simplificada
API	<i>Application Programming Interface</i>
CAAE	Certificado de Apresentação de Apreciação Ética
CBH	Congresso Brasileiro de Hansenologia
CEP	Comitê de Ética em Pesquisa
CNNs	<i>Convolutional Neural Networks</i>
IA	Inteligência Artificial
IDP	<i>Intelligent Document Processing</i>
IOU	<i>Intersection over Union</i>
JSON	<i>JavaScript Object Notation</i>
KNN	<i>K-Nearest Neighbors</i>
ML	<i>Machine Learning</i>
OCR	<i>Optical Character Recognition</i>
OMS	Organização Mundial de Saúde
OPAS	Organização Pan-Americana da Saúde
PQT	Poliquimioterapia
R-CNN	<i>Region-based Convolutional Neural Network</i>
RF	<i>Random Forest</i>
RPA	<i>Robotic Process Automation</i>
SBC	Sociedade Brasileira de Computação

SVM *Support Vector Machine*

UPE Universidade de Pernambuco

YOLO *You Only Look Once*

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Objetivos	2
1.1.1	Objetivo Geral	2
1.1.2	Objetivos Específicos	2
1.2	Organização do trabalho	3
2	FUNDAMENTAÇÃO TEÓRICA	4
2.1	A hanseníase como um problema de saúde pública no Brasil	4
2.2	Avaliação Neurológica Simplificada	5
2.3	Uso de IA para reconhecimento de objetos	7
2.3.1	YOLO	8
2.4	Processamento Inteligente de Documentos	10
2.5	IA na Saúde e a Digitalização de Dados Clínicos	11
3	METODOLOGIA	13
3.1	Preparação dos Dados	13
3.1.1	Pré-processamento dos Dados	13
3.2	Construção do Dataset e Treinamento do Modelo	14
3.2.1	Rotulação das Imagens	14
3.2.2	Treinamento do Modelo	14
3.3	Pós-processamento dos Dados	15
3.3.1	Abordagem Inicial: Algoritmo simples	16
3.3.2	Abordagem Avançada: Mapeamento Anatômico usando ML	17
3.3.2.1	<i>Criação do Dataset de Treinamento</i>	18
3.3.3	<i>Pipeline</i> de Inferência	19
3.4	Extração de Dados Estruturados	19
3.4.1	Treinamento do Modelo de Extração Personalizado	20
3.4.2	Integração e Extração de Dados Completa da Ficha ANS	21
4	RESULTADOS	22
4.1	Desempenho do Modelo de Detecção de Objetos (YOLOv8)	22

4.2	Desempenho do Classificador de Pontos Anatômicos (modelo RF)	24
4.3	Desempenho do Modelo de Extração de Documentos (API Azure)	25
5	CONSIDERAÇÕES FINAIS	28
5.1	Principais Contribuições.....	28
5.2	Limitações e Trabalhos Futuros.....	29
	REFERÊNCIAS	31
	APÊNDICES	32
A	PARECER DO COMITÊ DE ÉTICA EM PESQUISA	33
B	CERTIFICADO DE MELHOR ARTIGO	42
C	SCRIPT DE PRÉ-PROCESSAMENTO DAS REGIÕES DE INTERESSE	44
D	ALGORITMO PARA MAPEAMENTO ANATÔMICO	46
E	CRIAÇÃO DO DATASET DE TREINAMENTO VIA BOOTS-TRAPPING	51
F	INFERÊNCIA	55

1 INTRODUÇÃO

A detecção de objetos é uma das áreas mais importantes da visão computacional, com aplicações que abrangem desde a automação industrial até a área médica. Esta técnica, que envolve a identificação e localização de objetos em imagens ou vídeos, tem sido impulsionada por avanços significativos em algoritmos de aprendizado de máquina, especialmente *Convolutional Neural Networks* (CNNs). Dentre estes algoritmos, o *You Only Look Once* (YOLO) se destaca por sua eficiência na detecção de objetos em tempo real, superando abordagens tradicionais ao realizar a detecção em uma única etapa, o que simplifica o processo e o torna mais rápido. Esta característica torna o YOLO uma ferramenta valiosa, inclusive para usuários com conhecimentos limitados em Inteligência Artificial (IA). Na área médica, o uso da visão computacional para o processamento de imagens e reconhecimento de padrões vem avançando nos últimos anos [Gao et al. 2018], porém poucos trabalhos relatam o uso desta tecnologia para a extração de informações de documentos clínicos.

A hanseníase é uma doença infecciosa crônica causada pelo *Mycobacterium leprae* e que afeta principalmente a pele e os nervos periféricos, podendo levar a incapacidades permanentes se não for tratada precocemente [Talhari et al. 2012]. Apesar dos avanços na medicina e na vigilância epidemiológica, esta doença ainda representa um grande desafio para a saúde pública, especialmente em países endêmicos como o Brasil, que, em 2023, foi responsável por 92% dos novos casos registrados nas Américas [World Health Organization 2023]. A prevenção de incapacidades na hanseníase é um dos pilares estratégicos da Organização Mundial de Saúde (OMS) na Estratégia Global para Hanseníase 2021-2030 [Organization 2021]. No Brasil, a Avaliação Neurológica Simplificada (ANS) é recomendada no protocolo clínico de diretrizes terapêuticas da hanseníase e possui caráter obrigatório. Este instrumento é fundamental para o monitoramento da função neural do paciente, permitindo a identificação de alterações autonômicas, redução da sensibilidade e diminuição da força muscular decorrentes do dano neural.

Apesar de sua relevância clínica e epidemiológica, os dados coletados na ANS ainda são registrados manualmente em formulários físicos, o que dificulta a sistematização e a análise dessas informações. Embora estudos recentes tenham explorado o uso de IA para o diagnóstico da hanseníase por meio da análise de imagens de lesões cutâneas, vigilância

epidemiológica e suporte ao tratamento [Andrade et al. 2025], ainda há uma lacuna na literatura sobre a aplicação desta tecnologia para a digitalização e análise de formulários utilizados no monitoramento longitudinal do tratamento, como a ANS. A ausência de uma abordagem automatizada para a interpretação desses registros limita seu potencial para estudos epidemiológicos e para a previsão da evolução da incapacidade física em pacientes afetados.

Diante desse cenário, surge a necessidade de desenvolver modelos computacionais que possam automatizar a digitalização integral das fichas da ANS, um desafio que envolve tanto a interpretação de dados não-textuais (como o resultado da Avaliação Sensibilidade), quanto a extração de dados textuais e numéricos de um formulário estruturado. Este trabalho tem como objetivo apresentar um aplicativo, baseado no uso de visão computacional e aprendizado de máquina, para extração automática de dados da ANS, convertendo os dados das fichas digitalizadas em dados estruturados, com formato de fácil leitura, armazenamento e processamento.

1.1 Objetivos

1.1.1 Objetivo Geral

Este trabalho tem como objetivo principal desenvolver um sistema automatizado para a digitalização completa e automatizada das fichas de Avaliação Neurológica Simplificada, utilizando técnicas de visão computacional, aprendizado de máquina e processamento inteligente de documentos.

1.1.2 Objetivos Específicos

- Desenvolver um modelo baseado em técnicas de visão computacional e aprendizado de máquina para extração de dados não-textuais do formulário ANS;
- Implementar um modelo de Processamento Inteligente de Documentos (IDP) para a extração dos dados textuais e numéricos dos campos estruturados da ficha ANS.
- Desenvolver uma *Application Programming Interface* (API) que integre os resultados dos modelos de visão computacional e de extração de texto, consolidando todas as informações da ficha ANS em um único arquivo JSON estruturado.

- Desenvolver uma aplicação *web* para validar a solução completa, permitindo que usuários finais submetam imagens das fichas ANS e recebam os dados digitalizados de forma prática e acessível.

1.2 Organização do trabalho

Este trabalho está organizado em cinco capítulos. O Capítulo 2 apresenta o referencial teórico, onde serão abordados os principais assuntos discutidos; no Capítulo 3 descreve-se a metodologia, referindo-se a como foram executados os principais assuntos abordados no referencial teórico; e, por fim, nos Capítulos 4 e 5 são apresentados os resultados e as conclusões finais deste trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 A hanseníase como um problema de saúde pública no Brasil

A transmissão da hanseníase ocorre predominantemente por via respiratória, por meio da inalação de gotículas e secreções das vias aéreas superiores eliminadas por pacientes bacilíferos não submetidos à terapia específica, o que exige contato interpessoal próximo e prolongado [Rodrigues e Lockwood 2011]. Trata-se de uma patologia infecciosa de caráter crônico e evolução lenta, causada pelo bacilo *Mycobacterium leprae*, microorganismo que apresenta tropismo acentuado pelo sistema nervoso periférico e pela pele. Esse tropismo resulta em manifestações clínicas como lesões cutâneas e neuropatias periféricas que, quando não diagnosticadas e tratadas precocemente, podem evoluir para incapacidades físicas permanentes e deformidades [Talhari et al. 2012].

Historicamente, a hanseníase figura entre as doenças mais antigas documentadas pela humanidade, com registros que datam de civilizações milenares [Talhari et al. 2012]. No contexto brasileiro, a enfermidade representa um problema persistente e relevante de saúde pública. Segundo o último Boletim Epidemiológico do Ministério da Saúde, publicado em janeiro de 2025, foram notificados 247.139 casos de hanseníase no Brasil entre 2014 e 2023 [MS 2025], evidenciando a persistência da doença no país e a necessidade de fortalecimento e aprimoramento das estratégias de controle e eliminação.

Neste contexto, embora apresente caráter crônico e potencial incapacitante, a hanseníase é uma doença curável quando diagnosticada e tratada adequadamente. O tratamento padrão-ouro consiste na Poliquimioterapia (PQT), um esquema terapêutico que combina diferentes fármacos e cuja duração varia entre 6 e 12 meses, dependendo da classificação operacional do caso (paucibacilar ou multibacilar), conforme diretrizes do Ministério da Saúde [Transmissíveis 2016].

A classificação dos pacientes com hanseníase é realizada com base em critérios clínicos e laboratoriais, sendo utilizada, no contexto assistencial, a classificação operacional da doença, que considera o número de lesões cutâneas, a presença ou ausência de bacilos em exames complementares (baciloscopia ou histopatologia) e o grau de acometimento neural. Essa classificação permite a estratificação dos casos em paucibacilares e multibacilares, orientando a conduta terapêutica e o tempo de tratamento.

Além da classificação operacional, o monitoramento neurológico dos pacientes é

fundamental para a avaliação do impacto da doença sobre a função neural. Nesse cenário, a Organização Mundial de Saúde (OMS) estabelece a classificação do grau de incapacidade física, utilizada para mensurar o comprometimento funcional decorrente do dano neural, a qual é definida da seguinte forma:

- Grau 0 (ausência de comprometimento funcional neural);
- Grau 1 (diminuição ou perda de sensibilidade protetora em olhos, mãos e/ou pés, sem danos visíveis);
- Grau 2 (presença de lesões tróficas, traumáticas e/ou deformidades em olhos, mãos e/ou pés decorrentes do dano neural).

O diagnóstico precoce é um pilar fundamental na abordagem da hanseníase, pois permite a instituição oportuna do tratamento, minimizando o risco de desenvolvimento de danos neurais irreversíveis e incapacidades permanentes, que impactam severamente a qualidade de vida dos acometidos. Adicionalmente, a detecção e o tratamento precoces são essenciais para a interrupção da cadeia de transmissão do *M. leprae*, contribuindo para a redução da incidência da doença na população [Rodrigues e Lockwood 2011]. Neste contexto, o exame clínico dermatoneurológico minucioso, associado a instrumentos padronizados como a ANS, desempenha um papel central na identificação, estadiamento e acompanhamento dos pacientes.

2.2 Avaliação Neurológica Simplificada

No contexto da atenção à hanseníase no Brasil, a ANS integra de forma obrigatória o protocolo clínico e as diretrizes terapêuticas vigentes, sendo uma ferramenta essencial para a avaliação sistemática da função neural dos pacientes. Por meio desse instrumento, é possível identificar precocemente alterações autonômicas, bem como perdas de sensibilidade e redução da força muscular associadas ao comprometimento neural. Os resultados da ANS viabilizam o seguimento clínico do paciente ao longo do tratamento, auxiliam no manejo das neurites e apoiam a tomada de decisão pelos profissionais de saúde. Além disso, esses dados são relevantes para o planejamento e a avaliação de políticas públicas direcionadas ao controle da hanseníase. A Figura 1 ilustra o modelo atualmente adotado do formulário de ANS no Brasil.

INSAISSE
136

ANÁLISE CLÍNICA
MATER

Monitoramento da Capacidade Física em Pacientes com Lesão Medular Transversal
Estatísticas de Desempenho e Avaliação Clínica em Pacientes com Lesão Medular Transversal
Coordenação-Geral de Vigilância em Doenças em Eliminação

Fórmula para a Avaliação Clínica Simplificada e Classificação do Grau de Incapacidade Física em Hanseníase

Nome: _____		Sexo: ☐ M: ☐ F: ☐
Ocupação: _____		Data Nascimento: ____ / ____ / ____
Município: _____		UF: _____
Classificação Operacional: PB: ☐ MB: ☐		
Data início PQT-U: ____ / ____ / ____		Data Alta PQT-U: ____ / ____ / ____

FACE		S1			S2			S3			S4		
Nariz		D	E	D	E	D	E	D	E	D	E	D	E
Queixas:													
Ressecamento		(S/N)											
Furúnculo		(S/N)											
Perfuração de septo		(S/N)											
Olhos		D	E	D	E	D	E	D	E	D	E	D	E
Queixas:													
Diminuição da sensibilidade da córnea		(S/N)											
Diminuição de força muscular das pálpebras superiores		(S/N)											
Febra olhos sem força		(Fenda)											
Febra olhos com força		"mm"/"cu"/"gr"											
Triângulo		(S/N)											
Ectropião		(S/N)											
Opacidade corneana		(S/N)											
Acuidade visual		(Menor que normal)											

Legenda: Sim = S; Não = N; Em caso de fenda anotar em milímetros (mm), em caso de ausência de fenda anotar 0 (zero).
Acuidade visual: se usar óculos para longe, anotar durante o exame. Utilizar a tabela de continuação "A" distorcida a 3 metros para medição da acuidade visual.

MEMBRAS SUPERIORES		S1	/	S2	/	S3	/	S4	/	OP	/	/	
Queixas:													
PALPAÇÃO DE NERVOS		D	E	D	E	D	E	D	E	D	E	D	E
Radial													
Ulnar													
Mediano													

Legenda: Normal = N; Exposto = E; Der = D; Choque = C

AVALIAÇÃO DEFORÇA		D	E	D	E	D	E	D	E	D	E	D	E
Elevar o punho / Extensão punho (nervo radial)													
Abair dedo mínimo / Abdução do 5º dedo (nervo ulnar)													
Elevar o polegar / Abdução do polegar (nervo mediano)													

Legenda: Forte = 4, Resistência Parcial = 3, Movimento completo = 2, Contração = 1, Paralisado = 0. OU
 Forte = 4, Resistência = 3, Movimento = 2, Paralisado = 0

INSPEÇÃO E AVALIAÇÃO SENSITIVA*

S1			S2			S3			S4		
D	E		D	E		D	E		D	E	

Legenda: Seguir as cores dos monitores conforme instruções do fabricante.
 Cor vermelha = R; Cor verde = G; Borbulhão = B; Pontos brancos = P; Linhas tracejadas = L

MEMBROS INFERIORES		28		/		29		/		30		/		31		/		32		/		33		/		34		/		35		/		36		/		37		/		38		/		39		/		40		/		41		/		42		/		43		/		44		/		45		/		46		/		47		/		48		/		49		/		50		/		51		/		52		/		53		/		54		/		55		/		56		/		57		/		58		/		59		/		60		/		61		/		62		/		63		/		64		/		65		/		66		/		67		/		68		/		69		/		70		/		71		/		72		/		73		/		74		/		75		/		76		/		77		/		78		/		79		/		80		/		81		/		82		/		83		/		84		/		85		/		86		/		87		/		88		/		89		/		90		/		91		/		92		/		93		/		94		/		95		/		96		/		97		/		98		/		99		/		100		/		101		/		102		/		103		/		104		/		105		/		106		/		107		/		108		/		109		/		110		/		111		/		112		/		113		/		114		/		115		/		116		/		117		/		118		/		119		/		120		/		121		/		122		/		123		/		124		/		125		/		126		/		127		/		128		/		129		/		130		/		131		/		132		/		133		/		134		/		135		/		136		/		137		/		138		/		139		/		140		/		141		/		142		/		143		/		144		/		145		/		146		/		147		/		148		/		149		/		150		/		151		/		152		/		153		/		154		/		155		/		156		/		157		/		158		/		159		/		160		/		161		/		162		/		163		/		164		/		165		/		166		/		167		/		168		/		169		/		170		/		171		/		172		/		173		/		174		/		175		/		176		/		177		/		178		/		179		/		180		/		181		/		182		/		183		/		184		/		185		/		186		/		187		/		188		/		189		/		190		/		191		/		192		/		193		/		194		/		195		/		196		/		197		/		198		/		199		/		200		/		201		/		202		/		203		/		204		/		205		/		206		/		207		/		208		/		209		/		210		/		211		/		212		/		213		/		214		/		215		/	
--------------------	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--	-----	--	---	--

O exame ANS compreende um teste de avaliação sensorial, no qual é realizada avaliação da força muscular de grupos específicos (mãos, pés e olhos), a palpação de troncos nervosos periféricos (ulnar, mediano, radial, fibular comum, tibial posterior) e, de forma crucial, a pesquisa da sensibilidade cutânea [Saúde 2017]. Os dados coletados permitem classificar o grau de comprometimento neural e orientar as condutas terapêuticas e de reabilitação.

6

monofilamento e, conseqüentemente, a um nível de sensibilidade cutânea, em ordem, do maior para o menor:

- Verde (0,07 g);
- Azul (0,2 g);
- Violeta (2,0 g);
- Vermelho (4,0 g);
- Laranja (10,0 g);
- Rosa (300 g);
- Preto (Ausência de resposta ao rosa).

Os valores exatos e as cores podem ter pequenas variações dependendo do fabricante do estesiômetro e das diretrizes locais. A referência principal deve ser o manual que acompanha o kit e as diretrizes oficiais [Saúde 2017].

Tradicionalmente, esses resultados são desenhados manualmente em formulários impressos, representando os pontos testados nas mãos e nos pés esquematicamente, conforme apresentado na Figura 2. Esse processo manual impõe desafios significativos relacionados ao tempo de execução do exame, ao armazenamento físico das fichas, à recuperação das informações e à análise de séries históricas de dados. Além disso, dificulta a integração dessas informações a sistemas eletrônicos de saúde e limita o uso dos dados para análises clínicas e epidemiológicas em larga escala [Lastória e Macharelli 2014].

2.3 Uso de IA para reconhecimento de objetos

A IA é um ramo amplo e interdisciplinar da ciência da computação cujo objetivo é desenvolver sistemas e agentes computacionais capazes de exibir comportamentos que, se realizados por humanos, seriam considerados inteligentes [Russell e Norvig 2016]. Isso engloba uma vasta gama de capacidades, como raciocínio lógico, aprendizado, planejamento, percepção, compreensão de linguagem natural e resolução de problemas complexos.

Dentro do espectro da IA, encontra-se o Aprendizado de Máquina (Machine Learning, ML), um subcampo focado no desenvolvimento de algoritmos que permitem aos

Figura 2: Exemplo da Avaliação Sensitiva das mãos, na ANS

INSPEÇÃO E AVALIAÇÃO SENSITIVA							
1ª 07/06/2023		2ª 29/06/2023		3ª 15/12/2023		4ª 12/04/2024	
D	E	D	E	D	E	D	E
Legenda: Monofilamentos = Seguir cores OU Caneta (2g), Sente = ✓ Não sente= X Garra móvel = M, Garra rígida = R, Reabsorção = Lesões tróficas = □ Lesões traumáticas = Δ							

Fonte: Ficha ANS digitalizada pela autora (2025)

sistemas computacionais aprender padrões e tomar decisões a partir de dados, sem serem explicitamente programados para cada tarefa específica [Bishop 2006]. Em vez de seguir instruções rígidas, os modelos de ML ajustam seus parâmetros internos com base na exposição a conjuntos de dados de treinamento. As abordagens de ML são comumente categorizadas em: Aprendizado Supervisionado (onde o modelo aprende a partir de dados rotulados, mapeando entradas para saídas conhecidas), Aprendizado Não Supervisionado (onde o modelo identifica padrões e estruturas em dados não rotulados) e Aprendizado por Reforço (onde o agente aprende a tomar sequências de decisões interagindo com um ambiente para maximizar uma recompensa) [Goodfellow, Bengio e Courville 2016].

2.3.1 YOLO

YOLO é uma família de algoritmos de detecção de objetos em tempo real, pertencente ao domínio da visão computacional e que emprega redes neurais profundas, tipicamente CNNs. Diferentemente de abordagens anteriores que aplicavam detectores em múltiplas localizações e escalas da imagem (e.g., *Region-based Convolutional Neural Network* (R-CNN) e suas variantes), o YOLO trata a detecção de objetos como um problema de regressão [Redmon et al. 2016].

A principal inovação do YOLO é sua abordagem unificada e de passagem única (*single-pass*). O funcionamento do YOLO, ilustrado na Figura 3, pode ser resumido da seguinte forma:

1. **Divisão da Imagem em Grade:** A imagem de entrada é dividida em uma grade $S \times S$. Cada célula dessa grade é responsável por detectar objetos cujo centro caia

dentro dela.

2. **Predição por Célula:** Para cada célula da grade, o modelo YOLO prediz simultaneamente:

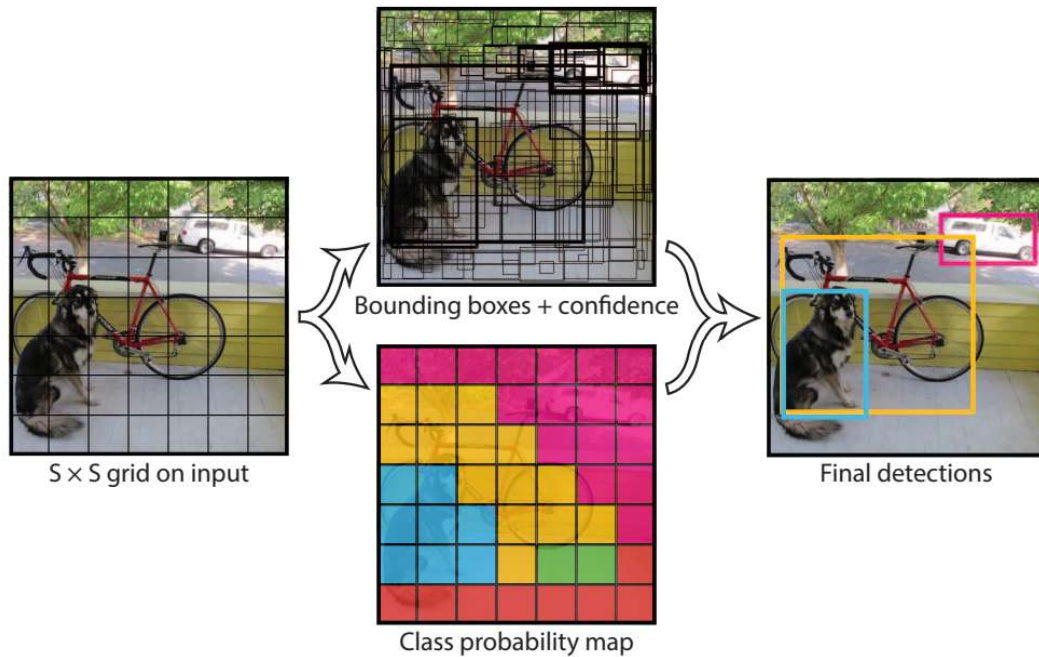
- **Caixas Delimitadoras (*Bounding Boxes*):** Um número B de caixas delimitadoras potenciais. Cada predição de caixa inclui 5 valores: as coordenadas do centro da caixa (bx , by), a largura (bw) e a altura (bh), relativas às dimensões da célula da grade, e uma pontuação de confiança (*confidence score*).
- **Pontuação de Confiança:** Esta pontuação reflete quão confiante o modelo está de que a caixa contém um objeto e quão precisa é a predição da caixa. É calculada como $P(\text{Objeto}) * \text{IOU}(\text{pred}, \text{verdade})$, onde $P(\text{Objeto})$ é a probabilidade de haver um objeto na célula e *Intersection over Union* (IOU) é a interseção sobre a união entre a caixa prevista e a caixa real (ground truth).
- **Probabilidades de Classe:** Para cada célula, o modelo também prediz as probabilidades condicionais de classe, $P(\text{Classe}_i \mid \text{Objeto})$, para cada uma das C classes de objetos que o modelo foi treinado para detectar.

3. **Combinação das Predições:** As pontuações de confiança das caixas e as probabilidades de classe são combinadas para obter pontuações de confiança específicas de classe para cada caixa delimitadora: $P(\text{Classe}_i) * \text{IOU}(\text{pred}, \text{verdade})$. Isso indica a probabilidade de uma classe específica estar presente naquela caixa e quão bem a caixa se ajusta ao objeto.

4. **Supressão de Não-Máximos (*Non-Max Suppression* - NMS):** Como múltiplas células podem detectar o mesmo objeto, resultando em várias caixas sobrepostas, o NMS é aplicado para filtrar as detecções redundantes, mantendo apenas as caixas com maior pontuação de confiança para cada objeto detectado.

A arquitetura subjacente do YOLO é geralmente uma CNN profunda (como Darknet, utilizada nas versões originais, ou backbones mais modernos como CSPDarknet, ResNet, EfficientNet nas versões mais recentes como YOLOv4, YOLOv5, YOLOR, YOLOv7, YOLOv8) que processa a imagem inteira de uma só vez e produz as predições de caixas e classes diretamente a partir das características extraídas [Redmon e Farhadi 2017, Bochkovskiy, Wang e Liao 2020].

Figura 3: Resumo do funcionamento do YOLO



Fonte: Redmon (2016)

No contexto deste trabalho, o modelo YOLO mostra-se particularmente adequado por sua capacidade de detectar múltiplos objetos em uma única imagem de forma eficiente e precisa. Essa característica é compatível com o desafio de identificar automaticamente múltiplas marcações gráficas presentes nos diagramas das mãos e dos pés nos formulários da ANS, os quais contêm diversos pontos de interesse distribuídos espacialmente. Dessa forma, a utilização do YOLO oferece uma base metodológica promissora para a extração automatizada de informações visuais e sua posterior conversão em dados clínicos estruturados.

2.4 Processamento Inteligente de Documentos

Além da análise de diagramas por meio da detecção de objetos, a digitalização completa de um formulário como a ANS exige a extração de texto e dados estruturados a partir de campos de formulário. Esta tarefa é o foco do campo de Processamento Inteligente de Documentos - *Intelligent Document Processing* (IDP), definido como um conjunto de tecnologias capazes de transformar dados não estruturados e semiestruturados em formato estruturado, indo além do *Optical Character Recognition* (OCR) tra-

dicional. O IDP emprega técnicas de IA, como ML e Processamento de Linguagem Natural, para capturar, classificar e extrair dados complexos, que podem ser posteriormente integrados a fluxos de automação por meio de *Robotic Process Automation* (RPA) [Cutting e Cutting-Decelle 2021].

Plataformas de IDP, como o *Azure AI Document Intelligence*, oferecem modelos pré-treinados e a capacidade de construir modelos personalizados (*custom models*) para extrair informações de documentos específicos. Diferente do OCR tradicional, que apenas transcreve texto, os modelos de extração personalizados podem ser treinados para reconhecer e extrair pares chave-valor (e.g., "Nome do Paciente": "João da Silva"), tabelas e campos de seleção. A utilização de um serviço de IDP neste trabalho complementa a análise de visão computacional dos diagramas de sensibilidade, permitindo a digitalização completa do formulário ANS e a extração de dados textuais e numéricos.

2.5 IA na Saúde e a Digitalização de Dados Clínicos

A IA tem se consolidado como uma ferramenta transformadora na área da saúde, oferecendo novas perspectivas para diagnóstico, prognóstico, tratamento e gestão de dados clínicos [Topol 2019]. Algoritmos de ML, em particular os de aprendizado profundo, demonstram capacidade notável na análise de grandes volumes de dados complexos, como imagens médicas (radiografias, tomografias, histopatologia), sinais fisiológicos (eletrocardiograma, eletroencefalograma) e dados genômicos [Esteva et al. 2019].

Nesse contexto, a aplicação da IA pode contribuir também de forma significativa para a digitalização, organização e análise de dados clínicos, permitindo transformar informações tradicionalmente registradas em formatos não estruturados, como formulários em papel e imagens, em bases de dados digitais passíveis de análise automatizada. Esse processo amplia o potencial de uso desses dados tanto na prática clínica quanto na pesquisa em saúde, favorecendo o monitoramento longitudinal de pacientes e a identificação de padrões associados à progressão de doenças crônicas [Gill et al. 2023].

No âmbito das doenças crônicas negligenciadas, como a hanseníase, o uso de ferramentas computacionais baseadas em IA apresenta especial relevância. A possibilidade de integrar dados clínicos, epidemiológicos e funcionais pode auxiliar na identificação precoce de complicações, no acompanhamento da evolução da doença e no suporte à tomada de decisão clínica, contribuindo para intervenções mais oportunas e personalizadas.

Assim, a IA emerge não apenas como uma tecnologia voltada ao diagnóstico automatizado, mas também como um recurso essencial para viabilizar a digitalização e o aproveitamento de dados clínicos historicamente subutilizados, criando novas oportunidades para o fortalecimento da vigilância em saúde, da pesquisa clínica e do planejamento de políticas públicas.

3 METODOLOGIA

Este capítulo apresenta o delineamento metodológico adotado para o desenvolvimento do presente trabalho. A metodologia foi estruturada em etapas sequenciais, que abrangem desde a coleta e o pré-processamento dos dados brutos até a construção, treinamento e validação do modelo de IA para a interpretação de fichas de ANS. As seções a seguir detalham cada uma dessas fases, justificando as escolhas técnicas e descrevendo os procedimentos empregados para atingir os objetivos propostos.

3.1 Preparação dos Dados

A primeira etapa do projeto consistiu na formação de um banco de dados a partir de fichas físicas de Avaliação Neurológica Simplificada (ANS). A coleta de dados foi realizada no Hospital Otávio de Freitas, uma instituição de referência no tratamento de hanseníase no estado de Pernambuco. Após aprovação do Comitê de Ética em Pesquisa (CEP) da Universidade de Pernambuco (UPE), conforme Certificado de Apresentação de Apreciação Ética (CAAE) disponível no Apêndice A, foi concedido acesso ao arquivo físico das fichas, permitindo a digitalização de um lote inicial. Para a fase de prova de conceito e desenvolvimento do modelo, um corpus de 200 fichas foi digitalizado. Este volume foi considerado suficiente para validar a abordagem proposta e treinar a primeira versão do sistema.

3.1.1 Pré-processamento dos Dados

Com as imagens digitalizadas, a etapa seguinte foi o isolamento das Regiões de Interesse, especificamente, as áreas correspondentes a região de avaliação sensitiva das mãos e dos pés. Para automatizar essa tarefa, foi desenvolvido um *script* em linguagem Python. O algoritmo opera com base em um sistema de coordenadas médias, pré-definidas a partir da estrutura padrão das fichas, para recortar de forma precisa as áreas desejadas. A execução deste script resultou na criação de dois subconjuntos de imagens: um contendo exclusivamente os diagramas das mãos e outro, os dos pés, otimizando o processamento nas etapas subsequentes. O *script* desenvolvido para esta finalidade encontra-se detalhado no Apêndice C.

3.2 Construção do Dataset e Treinamento do Modelo

3.2.1 Rotulação das Imagens

A rotulação de imagens é uma etapa fundamental no treinamento de modelos de aprendizado supervisionado, que consiste em identificar e nomear os objetos de interesse para que o modelo possa aprender a reconhecê-los.

Para a anotação das imagens, utilizou-se o *software* LabelImg. Esta ferramenta permite a demarcação de *bounding boxes* (caixas delimitadoras) em torno dos objetos e sua associação a uma classe. O resultado deste processo é a geração de arquivos de anotação em formato .txt, compatíveis com o padrão exigido pelo modelo YOLO.

Para este trabalho, foram definidas as seguintes categorias de classes:

- **Cores de Sensibilidade:** verde, azul, roxo, vermelho, rosa, laranja e preto, correspondendo aos diferentes níveis de sensibilidade registrados na ficha.
- **Partes do Corpo:** md, me, pd e pe (sendo as mãos e pés, direitos ou esquerdos).

A fim de mitigar a detecção de falsos positivos em fichas não preenchidas e evitar que dados nulos fossem processados, a estratégia de rotulação foi aplicada exclusivamente aos esquemas de mãos e pés que continham preenchimento colorido, indicativo de uma consulta realizada.

O conjunto de dados rotulado, composto por 130 imagens de pés e 130 de mãos, foi então dividido aleatoriamente em subconjuntos de treino e validação, seguindo a proporção de 70% e 30%, respectivamente. Esta partição é uma prática padrão que permite treinar o modelo com a maior parte dos dados e avaliar sua capacidade de generalização com dados não vistos.

3.2.2 Treinamento do Modelo

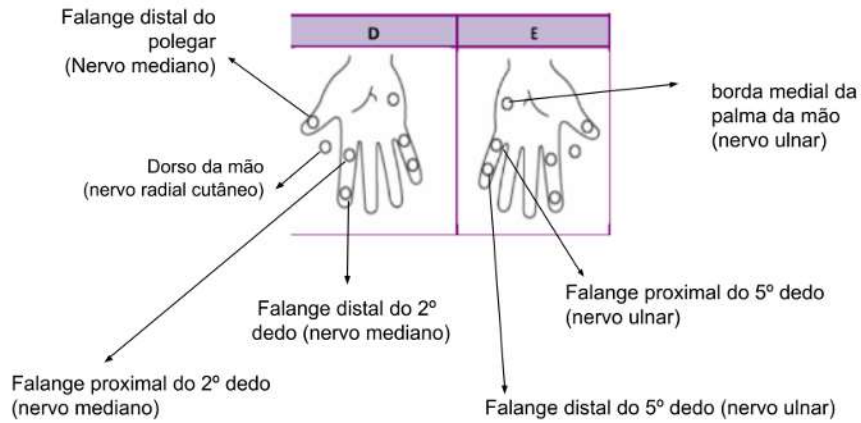
O treinamento do modelo de detecção de objetos foi conduzido no ambiente *Google Colab*, que oferece a infraestrutura de *hardware* necessária para realizar o processamento. Foi empregado o modelo YOLOv8, que é conhecida por seu equilíbrio entre velocidade e precisão. O treinamento foi iniciado utilizando 100 épocas e os hiperparâmetros padrão da arquitetura.

3.3 Pós-processamento dos Dados

A saída nativa do modelo YOLOv8, embora eficaz na detecção de objetos, fornece uma lista de classes detectadas e suas respectivas contagens (e.g., "8 verdes, 8 azuis, 1 pd, 1 pe"), formato que se mostra insuficiente para o objetivo final deste trabalho. O propósito não é apenas quantificar as cores, mas sim associá-las a pontos anatômicos específicos para permitir o monitoramento da evolução ou regressão da sensibilidade do paciente. Portanto, tornou-se imperativo o desenvolvimento de um algoritmo de pós-processamento para realizar o mapeamento anatômico. A partir das informações de saída do YOLO, que incluem o rótulo da classe, as coordenadas da *bounding box* e o grau de confiança da detecção, foi projetado um sistema para interpretar espacialmente esses dados. A lógica fundamental do algoritmo de mapeamento consiste em:

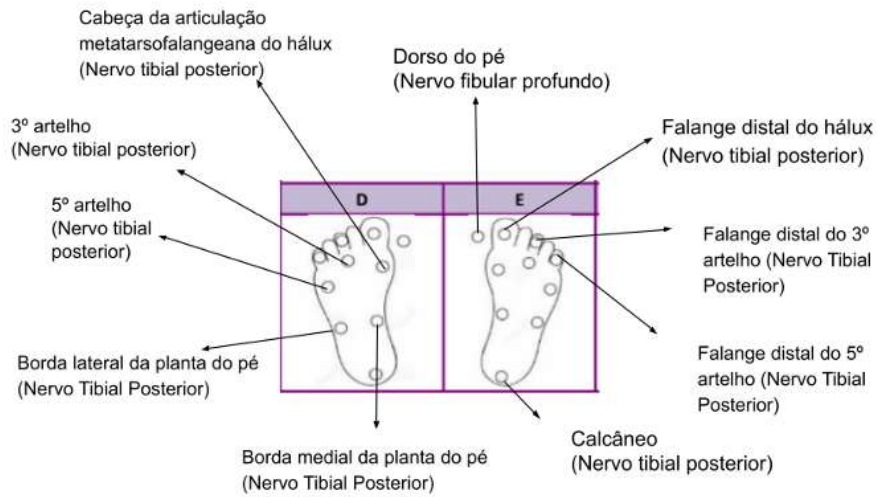
- **Identificação do Membro:** Localizar as bounding boxes dos membros (e.g., mao direita, pe esquerdo).
- **Verificação de Contenção:** Para cada cor detectada, verificar se o centro geométrico de sua bounding box está espacialmente contido dentro da bounding box de um dos membros.
- **Associação Ponto-Cor:** Uma vez que uma cor é associada a um membro, o desafio subsequente é identificar a qual dos múltiplos pontos anatômicos daquele membro ela corresponde.

Figura 4: Descrição anatômica dos pontos da mão



Fonte: Autora (2025)

Figura 5: Descrição anatômica dos pontos do pé



Fonte: Autora (2025)

3.3.1 Abordagem Inicial: Algoritmo simples

A primeira abordagem para resolver a associação ponto-cor foi o desenvolvimento de um algoritmo. Este método baseava-se em um conjunto de regras predefinidas, explorando a disposição espacial esperada dos pontos nos diagramas. As funções ma-

pear_pontos_mao e mapear_pontos, detalhadas no código, implementam essa lógica, ordenando os pontos detectados com base em suas coordenadas (x, y) para inferir suas posições anatômicas. O código-fonte completo que implementa este algoritmo pode ser consultado no Apêndice D.

Contudo, testes preliminares revelaram uma fragilidade significativa nesta abordagem. A alta variabilidade no preenchimento manual das fichas com marcações fora do local exato, sobreposições ou preenchimentos incompletos frequentemente resultava em falhas de mapeamento, comprometendo a robustez do sistema.

3.3.2 Abordagem Avançada: Mapeamento Anatômico usando ML

Para superar as limitações desse método, optou-se por uma solução mais sofisticada e robusta, empregando um modelo de ML para a tarefa de classificação do ponto anatômico. Essa abordagem consiste em uma arquitetura de dois estágios:

1. **Detecção de Objetos (YOLOv8):** O modelo YOLOv8, como descrito anteriormente, detecta todos os objetos de interesse (membros e cores) na imagem.
2. **Classificação do Ponto Anatômico (Random Forest):** Para cada cor detectada dentro de um membro, um segundo modelo, um classificador *Random Forest* (RF), é utilizado para prever o nome do ponto anatômico correspondente.

A seleção do classificador para este segundo estágio foi baseada em uma análise comparativa para determinar o modelo de melhor desempenho. Cinco modelos foram avaliados: RF, XGBoost, Regressão Logística, *K-Nearest Neighbors* (KNN) e *Support Vector Machine* (SVM), utilizando Acurácia, F1-Score (Macro) e tempo de treinamento como métricas. Os resultados, detalhados na Tabela 1, demonstraram a clara superioridade dos modelos de *ensemble* baseados em árvores (RF e XGBoost), enquanto os demais modelos não se mostraram muito eficientes para a complexidade da tarefa.

Com base nos dados, o RF foi selecionado como o modelo definitivo para a aplicação. A escolha foi justificada por apresentar a melhor combinação de performance e eficiência, obtendo a maior Acurácia (94.68%) e o maior F1-Score (93.76%), além de um tempo de treinamento inferior ao do seu concorrente mais próximo, o XGBoost.

Tabela 1: Comparação entre diferentes modelos de ML para o mapeamento anatômico

Modelo	Acurácia	F1-Score (Macro)	Tempo de Treino (s)
RF	0.9468	0.9376	1.37
XGBoost	0.9365	0.9239	1.95
Regressão Logística	0.7787	0.8093	16.92
KNN	0.5619	0.4953	0.23
SVM	0.2085	0.0677	4.41

Fonte: Autora (2025)

3.3.2.1 Criação do Dataset de Treinamento

Um dos principais desafios foi a criação de um dataset rotulado para treinar o modelo RF. A solução adotada foi um processo no qual o algoritmo inicial (detalhado no Apêndice D) foi aplicado a um subconjunto de imagens consideradas "perfeitas" (ou seja, preenchidas de forma clara e previsível). Para cada ponto que o algoritmo conseguiu mapear corretamente, foi executado um processo de engenharia de features, extraindo um conjunto de características descritivas da relação espacial entre o ponto e o membro. As principais features calculadas foram:

- **Coordenadas Relativas:** Posição (x, y) do centro do ponto de cor, normalizada em relação às dimensões da *bounding box* do membro.
- **Distância Relativa:** Distância euclidiana entre o centro do ponto de cor e o centro do membro.
- **Área Relativa:** Proporção da área do ponto de cor em relação à área total do membro.
- **Razão de Aspecto do Membro:** Proporção entre a largura e a altura da *bounding box* do membro.
- **Confianças de Detecção:** Scores de confiança fornecidos pelo YOLO para a detecção do ponto e do membro.

Este processo gerou um *dataset* tabular onde cada linha representava um ponto de cor mapeado com seu respectivo vetor de features e o rótulo do ponto anatômico. O *script* desenvolvido para a geração deste dataset encontra-se no Apêndice E. Após a criação do

dataset, o modelo RF foi treinado para aprender a associar o vetor de features ao seu rótulo anatômico correspondente.

3.3.3 Pipeline de Inferência

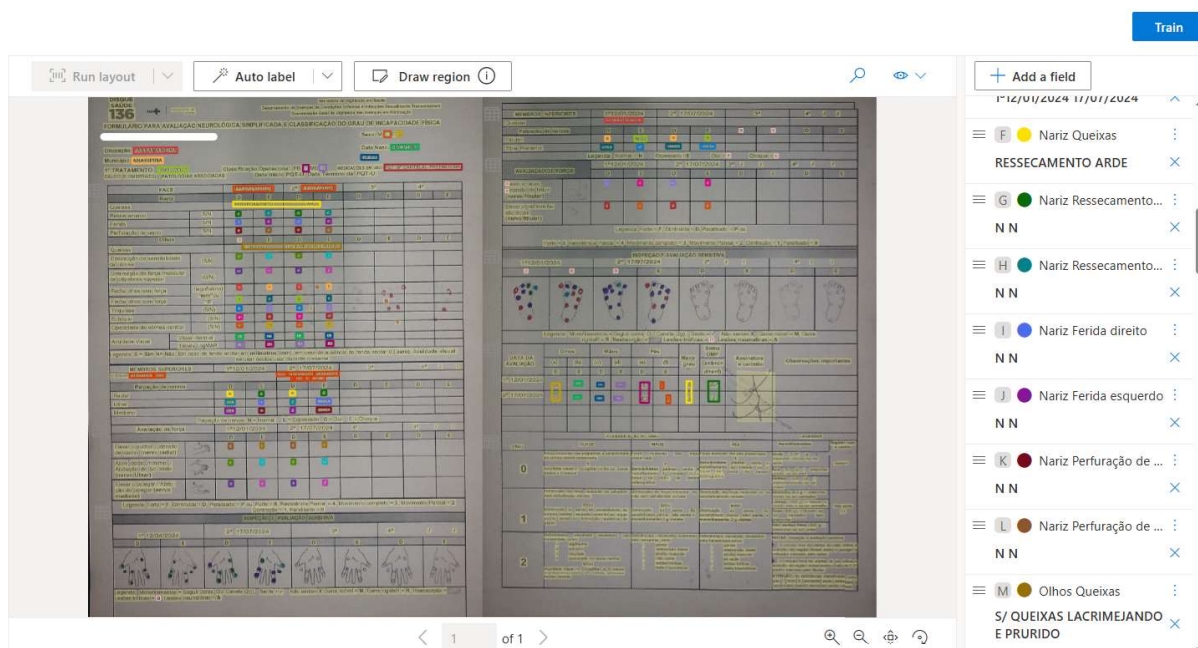
O *pipeline* de processamento orquestra a interação entre os dois modelos de ML: O RF e YOLO. O código-fonte desta função, responsável por interpretar os campos visuais da avaliação sensitiva da ficha ANS, é apresentado em detalhe no Apêndice F. Ao receber uma imagem, o sistema executa a seguinte sequência:

1. O modelo YOLOv8 detecta todas as instâncias de membros e pontos de cor;
2. O sistema associa cada ponto de cor detectado a um membro com base em sua localização espacial;
3. Para cada ponto de cor associado, o sistema calcula um vetor de *features*, que inclui características espaciais (posição relativa, distância ao centro) e de confiança da detecção;
4. O vetor de *features* é fornecido ao modelo RF, que prevê o rótulo do ponto anatômico correspondente;
5. Finalmente, o sistema constrói uma saída estruturada preenchendo um gabarito de todos os pontos anatômicos esperados para aquele membro. Os pontos detectados são preenchidos com a cor identificada, e os demais são marcados como "Não identificado", garantindo um resultado completo e padronizado.

3.4 Extração de Dados Estruturados

Para realizar a digitalização completa da ficha ANS, incluindo dados demográficos do paciente, datas e resultados dos testes motores, foi adotada uma segunda frente metodológica utilizando a plataforma *Azure AI Document Intelligence*. Foi empregado o recurso de construção de modelos de extração personalizados (*custom extraction models*) para automatizar a leitura dos campos do formulário.

Figura 6: Interface do Azure Document Intelligence Studio na rotulação da ficha ANS.



Fonte: Autora (2025)

3.4.1 Treinamento do Modelo de Extração Personalizado

O processo de criação do modelo de extração seguiu as etapas abaixo:

1. **Criação do Dataset de Treinamento:** Um conjunto de 30 fichas ANS foi selecionado para compor o *dataset* de treinamento;
2. **Rotulação dos Campos:** Utilizando a interface do *Document Intelligence Studio*, cada uma das 30 imagens foi manualmente rotulada. Foram definidos 71 campos-chave (ou *labels*) para extração, abrangendo todas as seções relevantes do formulário, como "Dados do Paciente" (sexo, data de nascimento, medicações), "Avaliação Motora" (graus de força muscular) e "Data da Avaliação", e anonimizando o nome do paciente. A Figura 6 ilustra a interface de rotulação;
3. **Treinamento e Validação do Modelo:** Após a rotulação, o modelo foi treinado na plataforma da Azure. O serviço utiliza esses exemplos para aprender a estrutura visual e semântica do formulário, tornando-se capaz de generalizar e extrair os mesmos campos de fichas não vistas anteriormente.

3.4.2 Integração e Extração de Dados Completa da Ficha ANS

A integração dos dois sistemas desenvolvidos: o *pipeline* de Visão Computacional para detecção das áreas de sensibilidade e o modelo de extração de dados textuais baseado no *Azure Form Recognizer*, possibilitou a construção de um fluxo de trabalho automatizado para a digitalização completa da ficha de ANS.

Diante da submissão de uma nova imagem da ficha ANS, o sistema executa ambos os pipelines em paralelo. O resultado é exibido em um objeto *JavaScript Object Notation* (JSON) contendo, de forma estruturada, tanto os campos visuais da avaliação sensitiva quanto os dados textuais e numéricos do formulário. Dessa forma, obtém-se um registro digital unificado e completo da avaliação neurológica do paciente.

A etapa final do processo consistiu na integração desses módulos em um sistema de inferência coeso, garantindo interoperabilidade e escalabilidade. Para assegurar disponibilidade e capacidade de processamento, o *backend* foi implantado em uma instância *Amazon EC2* (AWS). O *frontend*, por sua vez, foi disponibilizado em ambiente *web*, acessível por meio do endereço <https://ans2json.dotlabbrazil.com.br/>, viabilizando sua aplicação prática. O código-fonte completo encontra-se em repositório público no GitHub¹, a fim de promover transparência e possibilitar a replicação da solução.

¹<https://github.com/jamilly-b/ANS2JSON>

4 RESULTADOS

Este capítulo apresenta os resultados obtidos a partir da aplicação da metodologia descrita anteriormente. A avaliação do sistema foi dividida em etapas, analisando o desempenho de cada componente do *pipeline* de processamento: o modelo de detecção de objetos (YOLOv8), o modelo de classificação de pontos anatômicos (RF) e, por fim, o sistema integrado. Os resultados são apresentados por meio de métricas quantitativas e análises qualitativas de exemplos práticos.

4.1 Desempenho do Modelo de Detecção de Objetos (YOLOv8)

A primeira etapa do *pipeline* consistiu na detecção de objetos de interesse (as cores de sensibilidade e os membros (mãos e pés)), utilizando o modelo YOLOv8. Os resultados detalhados são apresentados na Tabela 2. O modelo alcançou um desempenho geral excelente, com uma média de precisão (*Precision*) de 0.975, uma média de revocação (*Recall*) de 0.966 e uma média de precisão média (mAP@.5) de 0.981.

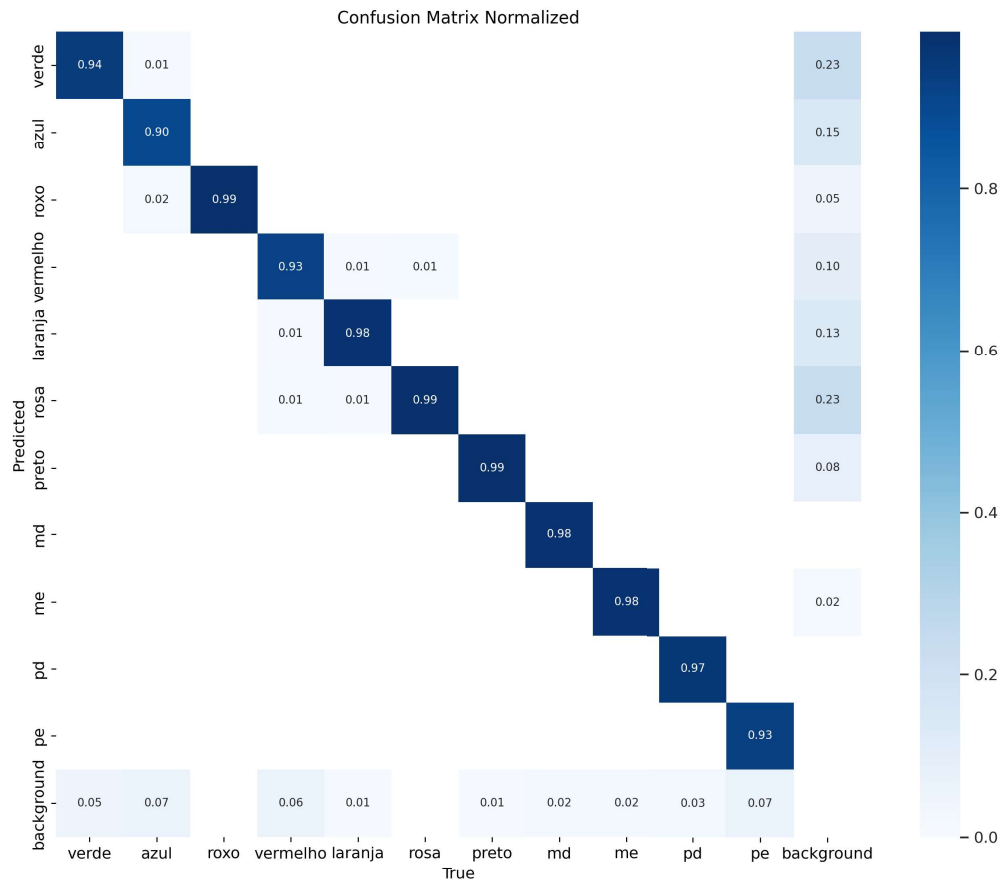
Analizando as classes individualmente, observa-se que as categorias de membros (md, me, pd, pe) apresentaram valores de mAP@.5 excepcionalmente altos, todos acima de 0.98, o que demonstra a capacidade robusta do modelo em localizar corretamente as regiões de interesse primárias. As classes de cores também obtiveram um desempenho sólido, com a maioria apresentando mAP@.5 superior a 0.96. Esses resultados indicam que a primeira fase do sistema é altamente confiável, fornecendo detecções precisas para a etapa subsequente de classificação. Para uma análise visual mais detalhada, a Figura 7 exibe a matriz de confusão gerada durante a validação, e a Figura 8 ilustra exemplos de detecções bem-sucedidas em imagens do conjunto de teste.

Tabela 2: Desempenho do modelo YOLOv8 no conjunto de validação.

Classe	Instâncias	Precision (P)	Recall (R)	mAP@.5
Verde	559	0.963	0.943	0.962
Azul	335	0.974	0.911	0.944
Roxo	423	0.977	0.994	0.989
Vermelho	178	0.949	0.935	0.965
Laranja	156	0.936	0.974	0.984
Rosa	350	0.963	0.997	0.994
Preto	492	0.984	0.992	0.993
Mão Direita (md)	95	0.993	0.979	0.985
Mão Esquerda (me)	95	0.982	0.979	0.982
Pé Direito (pd)	59	0.999	0.983	0.993
Pé Esquerdo (pe)	59	1.000	0.943	0.995
Média Geral (all)	2801	0.975	0.966	0.981

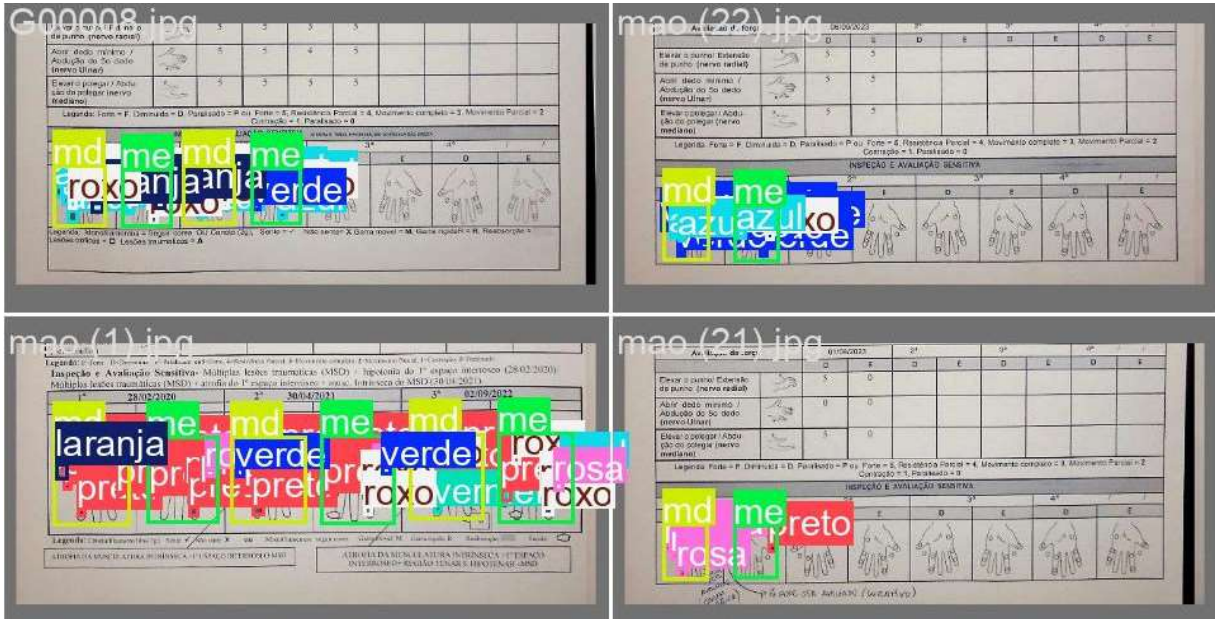
Fonte: Autora (2025)

Figura 7: Matriz de confusão do modelo YOLOv8.



Fonte: Autora (2025)

Figura 8: Detecções realizadas pelo modelo YOLOv8 no conjunto de teste.



Fonte: Autora (2025)

4.2 Desempenho do Classificador de Pontos Anatômicos (modelo RF)

Uma vez que os objetos são detectados, o modelo RF é empregado para associar cada ponto de cor à sua localização anatômica correta. O desempenho deste modelo foi avaliado com base em sua capacidade de classificar corretamente os 1559 pontos anatômicos presentes no conjunto de teste. O modelo alcançou uma acurácia geral de 94.6%. A Tabela 3 apresenta o relatório de classificação detalhado, incluindo as métricas de Precision, Recall e F1-Score para cada ponto anatômico.

A média ponderada do F1-Score foi de 0.95, confirmando a alta performance do classificador. Pontos com localizações anatômicas bem distintas e isoladas, como 'Calcâneo (Direito)' e 'Falange Distal 5º Artelho (Direito)', alcançaram um F1-Score perfeito de 1.00. Em contraste, a classe '3º Artelho (Direito)' registrou o F1-Score mais baixo (0.75), indicando uma maior dificuldade de classificação para o modelo, possivelmente devido à sua proximidade com outros pontos e maior variabilidade espacial. A matriz de confusão, apresentada na Figura 9, permite visualizar quais classes foram mais frequentemente confundidas entre si.

Tabela 3: Desempenho do modelo RF para identificação de pontos anatômicos.

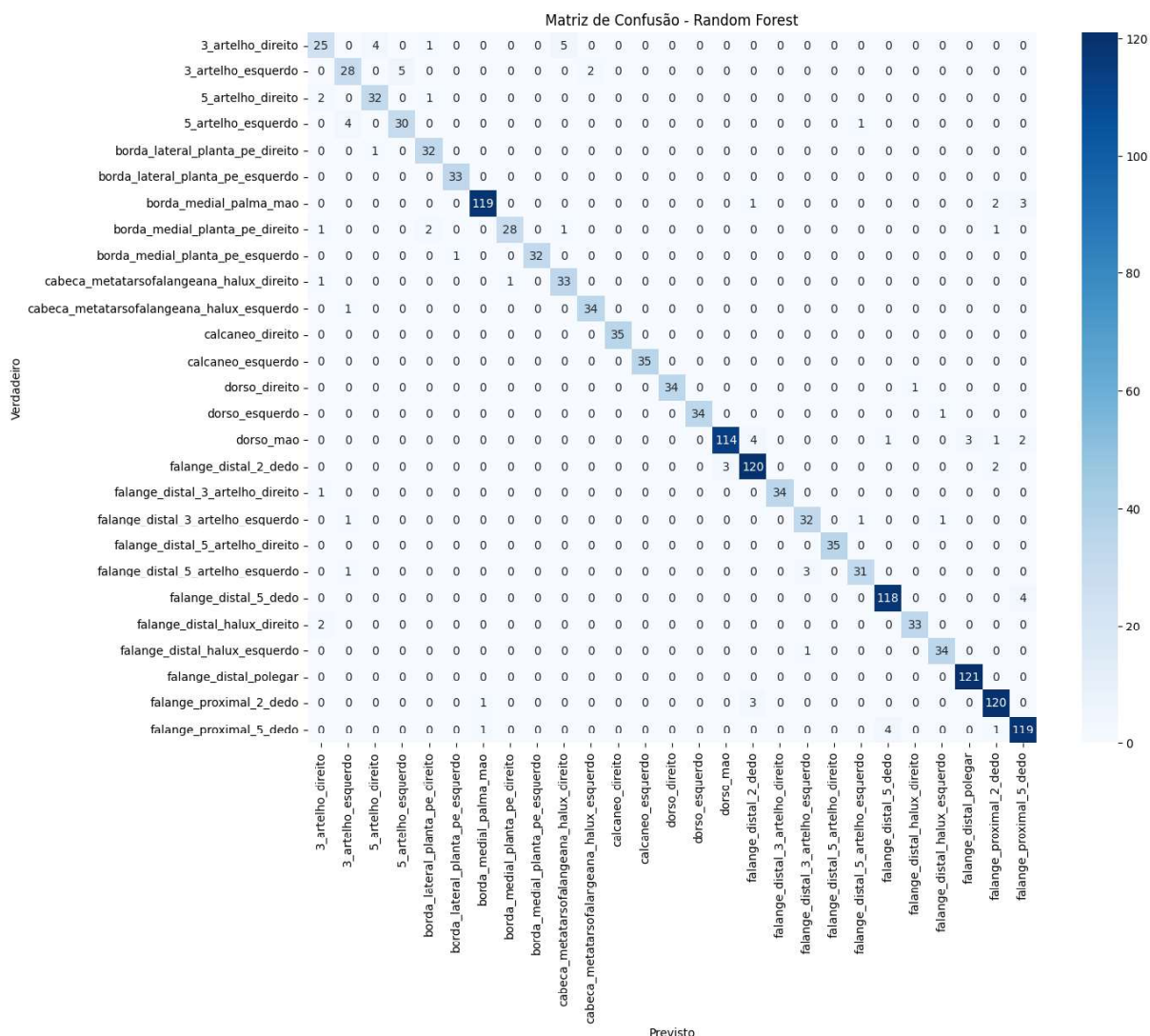
Ponto Anatômico	Precision	Recall	F1-Score	Support
Pontos Anatômicos dos Pés				
3º Artelho (Direito)	0.78	0.71	0.75	35
3º Artelho (Esquerdo)	0.80	0.80	0.80	35
5º Artelho (Direito)	0.86	0.91	0.89	35
5º Artelho (Esquerdo)	0.86	0.86	0.86	35
Borda Lateral da Planta (Direito)	0.89	0.97	0.93	33
Borda Lateral da Planta (Esquerdo)	0.97	1.00	0.99	33
Borda Medial da Planta (Direito)	0.97	0.85	0.90	33
Borda Medial da Planta (Esquerdo)	1.00	0.97	0.98	33
Cabeça Metatarsofalangeana Hálux (Direito)	0.85	0.94	0.89	35
Cabeça Metatarsofalangeana Hálux (Esquerdo)	0.94	0.97	0.96	35
Calcâneo (Direito)	1.00	1.00	1.00	35
Calcâneo (Esquerdo)	1.00	1.00	1.00	35
Dorso do Pé (Direito)	1.00	0.97	0.99	35
Dorso do Pé (Esquerdo)	1.00	0.97	0.99	35
Falange Distal 3º Artelho (Direito)	1.00	0.97	0.99	35
Falange Distal 3º Artelho (Esquerdo)	0.89	0.91	0.90	35
Falange Distal 5º Artelho (Direito)	1.00	1.00	1.00	35
Falange Distal 5º Artelho (Esquerdo)	0.94	0.89	0.91	35
Falange Distal Hálux (Direito)	0.97	0.94	0.96	35
Falange Distal Hálux (Esquerdo)	0.94	0.97	0.96	35
Pontos Anatômicos das Mãos				
Borda Medial da Palma	0.98	0.95	0.97	125
Dorso da Mão	0.97	0.91	0.94	125
Falange Distal 2º Dedo	0.94	0.96	0.95	125
Falange Distal 5º Dedo	0.96	0.97	0.96	122
Falange Distal Polegar	0.98	1.00	0.99	121
Falange Proximal 2º Dedo	0.94	0.97	0.96	124
Falange Proximal 5º Dedo	0.93	0.95	0.94	125
Média Ponderada (Total)	0.95	0.95	0.95	1559

Fonte: Autora (2025)

4.3 Desempenho do Modelo de Extração de Documentos (API Azure)

A digitalização dos dados textuais e estruturados da ficha ANS foi realizada pelo modelo de extração personalizado, treinado na plataforma *Azure AI Document Intelligence*. A performance do modelo foi avaliada campo a campo, com os resultados de acurácia apresentados na Tabela 4. O modelo demonstrou um desempenho geral robusto, embora heterogêneo, refletindo a complexidade e a variabilidade dos campos de um formulário. Campos com estrutura bem definida e menor variação de preenchimento, como "Data de Nascimento" (97.0%), "Grau de Incapacidade (Mãos, Pés, Olhos)" (acima de

Figura 9: Matriz de confusão do modelo RF.



Fonte: Autora (2025)

93%) e dados de avaliação motora como "Elevar o punho esquerdo" (99.5%), alcançaram altíssimos níveis de acurácia. Este sucesso indica que a ferramenta é altamente confiável para a extração de dados essenciais e padronizados.

Por outro lado, campos opcionais, ou que as informações não são sempre preenchidas, apresentaram uma acurácia significativamente menor. Campos como "Observação sobre o paciente" (18.2%), "Medicações em uso" (33.3%) e datas de tratamento como "Início PQT-U" (36.4%) tiveram os desempenhos mais baixos. Este resultado sugere que o modelo teve dificuldade em generalizar para campos que aparecem em branco em muitas das fichas do conjunto de treinamento, um desafio comum em modelos de extração de documentos. Apesar da variação, a maioria dos campos cruciais para a avaliação neu-

rológica e motora obteve acurácia superior a 80%, validando a eficácia da ferramenta como um componente essencial para a digitalização da ficha ANS.

Tabela 4: Desempenho do modelo de extração de dados textuais (API Azure).

Campo	Acurácia
Campos de Alta Acurácia	
Elevar o punho esquerdo	99.5%
Data de Nascimento	97.0%
Grau (Mãos, Pés, Olhos, Maior)	> 93.9%
Sexo	93.9%
Campos de Acurácia Moderada	
Município	87.9%
Classificação Operacional (PB/MB)	> 81.8%
Fecha os olhos com força	> 78.8%
Data da Consulta	78.8%
Campos de Baixa Acurácia (Desafios)	
Patologias Associadas	45.5%
Primeiro Tratamento	36.4%
Medicações em Uso	33.3%
Observação sobre o paciente	18.2%

Fonte: Autora (2025)

5 CONSIDERAÇÕES FINAIS

Este trabalho se propôs a enfrentar o desafio da subutilização de dados clínicos do tratamento da hanseníase, desenvolvendo um aplicativo *web* completo para a digitalização integral e automatizada das Fichas de ANS. Por meio de uma arquitetura híbrida de IA, o aplicativo demonstrou ser capaz de extrair com alta precisão tanto os dados não-textuais dos diagramas de sensibilidade, quanto os dados textuais e numéricos dos demais campos do formulário, culminando em uma solução ponta-a-ponta implantada em ambiente de nuvem.

Os resultados obtidos indicam um desempenho consistente do *pipeline* de visão computacional na interpretação dos testes de sensibilidade, com mAP@.5 de 0.981 e acurácia de classificação de 94.6% nos experimentos realizados, sugerindo boa capacidade de identificação dos monofilamentos. De forma complementar e crucial, a digitalização dos dados do formulário via a API *Azure AI Document Intelligence* demonstrou alta eficácia justamente nos indicadores clínicos mais importantes. O sistema alcançou acurácia excelente na extração dos resultados finais do Grau de Incapacidade (acima de 93% para todos os membros) e em testes de força motora chave, como "Elevar o punho esquerdo" (99.5%) e "Elevar o pé direito" (97.0%). A avaliação de nervos periféricos, queixas e testes oculares também apresentou um boa performance, com a maioria dos campos superando 85% de acurácia.

Considerando os resultados obtidos, a alta performance observada confirma que o aplicativo é uma ferramenta promissora para a digitalização do formulário de ANS. A implementação do aplicativo na nuvem da Amazon Web Services (AWS) e sua disponibilização através de uma interface *web* não apenas validam a solução proposta, mas também a disponibiliza como uma ferramenta prática e escalável. Ao automatizar um processo anteriormente manual, propenso a erros e baixa eficiência, este trabalho oferece um caminho para otimizar a gestão de dados clínicos, otimizando o tempo dos profissionais de saúde e auxiliando na criação de repositórios de dados robustos para pesquisas futuras no tratamento da hanseníase.

5.1 Principais Contribuições

As principais contribuições deste projeto podem ser resumidas em:

- **Inovação Metodológica:** Proposição de uma arquitetura híbrida de IA que combina detecção de objetos, classificação baseada em *features* e processamento inteligente de documentos para a digitalização completa de um formulário clínico utilizado no tratamento da hanseníase;
- **Validação com dados clínicos reais:** Desenvolvimento e validação dos modelos propostos utilizando dados reais, extraídos de Unidade de Referência no tratamento da hanseníase em Pernambuco, obtendo uma acurácia superior a 94%, garantindo a confiabilidade dos dados extraídos pela solução proposta;
- **Disponibilização da solução proposta para o usuário final:** Implementação e disponibilização do aplicativo para utilização dos usuários finais, através da Internet, em parceria com o Grupo de Pesquisa dotlab Brazil, da UPE, e a AWS;
- **Divulgação Científica:** Os resultados preliminares deste trabalho foram publicados no XXI Simpósio Brasileiro de Sistemas de Informação, promovido pela Sociedade Brasileira de Computação (SBC), onde foi apresentado o artigo *Automatic Information Extraction from Neurological Assessment Records for Leprosy Cases Using AI* melo2025automatic e o resumo "Do papel ao aplicativo: ANS2JSON, inovação digital no armazenamento e processamento das Avaliações Neurológicas Simplificadas", que obteve o primeiro lugar na área de Prevenção de Incapacidades e Reabilitação entre os melhores trabalhos apresentados no 19º Congresso Brasileiro de Hansenologia Congresso Brasileiro de Hansenologia (CBH), conforme Apêndice B;
- **Base para Pesquisas Futuras:** A criação de um método eficaz para a digitalização em massa de fichas ANS abre portas para a formação de repositórios públicos de dados clínicos, essenciais para novos estudos para ajudar no cuidado à hanseníase.

5.2 Limitações e Trabalhos Futuros

Apesar dos resultados promissores, o presente estudo possui algumas limitações que abrem caminhos para futuras investigações. A principal limitação é a dependência de um serviço comercial (API *Azure AI Document Intelligence*) para a extração de dados do formulário, o que pode implicar custos operacionais em uma implementação em larga escala. Com base nisso, os seguintes trabalhos futuros são propostos:

1. **Criação de um Banco de Dados Longitudinal:** Utilizar o sistema desenvolvido para digitalizar um grande volume de fichas históricas, criando um banco de dados que permita acompanhar a evolução neurológica dos pacientes ao longo do tempo.
2. **Desenvolvimento de Modelos Preditivos:** Com o banco de dados consolidado, desenvolver modelos de ML para prever o risco de desenvolvimento de incapacidades físicas, auxiliando em intervenções clínicas precoces e personalizadas.
3. **Exploração de Alternativas *Open-Source*:** Pesquisar e desenvolver uma solução de extração de dados de formulário baseada em tecnologias de código aberto (e.g., modelos de linguagem e visão, como o LayoutLM), visando reduzir os custos e aumentar a autonomia da plataforma.
4. **Integração com Plataformas de Saúde Digital:** Integrar o sistema de digitalização a um aplicativo móvel que possibilite o preenchimento nativo-digital da ficha ANS. Essa fusão criaria um ecossistema de dados unificado, onde registros históricos (digitalizados a partir de fichas físicas) e avaliações correntes (inseridas digitalmente) coexistiriam, formando um histórico longitudinal completo para o monitoramento do paciente.

Em suma, este trabalho não apenas alcançou seus objetivos técnicos, mas também estabeleceu uma base sólida para a transformação digital no monitoramento da hanseníase, com potencial para impactar positivamente a pesquisa, a prática clínica e a vida dos pacientes.

REFERÊNCIAS

- [Andrade et al. 2025]ANDRADE, H. G. V. de et al. On the usage of artificial intelligence in leprosy care: A systematic literature review. *PLOS Computational Biology*, Public Library of Science San Francisco, CA USA, v. 21, n. 6, p. e1012550, 2025.
- [Bell-Krotoski 1990]BELL-KROTOSKI, J. A. Light touch-deep pressure testing using Semmes-Weinstein monofilaments. In: HUNTER, J. M. et al. (Ed.). *Rehabilitation of the Hand: Surgery and Therapy*. 3rd. ed. St. Louis: Mosby, 1990. p. 575–593.
- [Bishop 2006]BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer, 2006.
- [Bochkovskiy, Wang e Liao 2020]BOCHKOVSKIY, A.; WANG, C.-Y.; LIAO, H.-Y. M. *YOLOv4: Optimal Speed and Accuracy of Object Detection*. 2020.
- [Cutting e Cutting-Decelle 2021]CUTTING, G. A.; CUTTING-DECELLE, A.-F. Intelligent document processing—methods and tools in the real world. *arXiv preprint arXiv:2112.14070*, 2021.
- [Esteva et al. 2019]ESTEVA, A. et al. A guide to deep learning in healthcare. *Nature Medicine*, v. 25, n. 1, p. 24–29, 2019.
- [Gao et al. 2018]GAO, J. et al. Computer vision in healthcare applications. *Journal of healthcare engineering*, v. 2018, p. 5157020, 2018.
- [Gill et al. 2023]GILL, A. Y. et al. Revolutionizing Healthcare: Exploring the Role of Artificial Intelligence in Disease Prediction. *Sustainability*, v. 15, n. 15, p. 11702, 2023.
- [Goodfellow, Bengio e Courville 2016]GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>.
- [Lastória e Macharelli 2014]LASTÓRIA, J. C.; MACHARELLI, C. A. Informatização em hansenologia: uma necessidade atual. *Hansenologia Internationalis: hanseníase e outras doenças infecciosas*, v. 39, n. 1, p. 1–2, 2014. Disponível em: <<https://periodicos.saude.sp.gov.br/index.php/hansenologia/article/view/35855>>.
- [MS 2025]MS. *Boletim Epidemiológico - Hanseníase 2025*. 2025. Disponível em: <<https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/especiais/2025/boletim-epidemiologico-de-hanseníase-numero-especial-jan-2025.pdf>>.
- [Organization 2021]ORGANIZATION, W. H. *Towards zero leprosy. Global leprosy (Hansen's Disease) strategy 2021–2030*. 2021. Disponível em: <<https://www.who.int/publications/i/item/9789290228509>>.
- [Redmon et al. 2016]REDMON, J. et al. You Only Look Once: Unified, Real-Time Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 779–788.

- [Redmon e Farhadi 2017]REDMON, J.; FARHADI, A. YOLO9000: Better, Faster, Stronger. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2017. p. 6517–6525.
- [Rodrigues e Lockwood 2011]RODRIGUES, L. C.; LOCKWOOD, D. N. Leprosy now: epidemiology, progress, challenges, and research gaps. *The Lancet Infectious Diseases*, v. 11, n. 6, p. 464–470, 2011.
- [Russell e Norvig 2016]RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3rd. ed. [S.l.]: Pearson, 2016.
- [Saúde 2017]SAÚDE, D. d. A. B. Secretaria de A. *Guia prático sobre a hanseníase*. 2^a. ed. Brasília, 2017. (Cadernos de Atenção Básica, 36). Disponível em: <https://bvsms.saude.gov.br/bvs/publicacoes/guia_pratico_hanseniase_ab36.pdf>.
- [Talhari et al. 2012]TALHARI, S. et al. *Hansenologia*. 4^a. ed. Rio de Janeiro: Di Livros, 2012.
- [Topol 2019]TOPOL, E. J. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. [S.l.]: Basic Books, 2019.
- [Transmissíveis 2016]TRANSMISSÍVEIS, D. de Vigilância das D. *Diretrizes para vigilância, atenção e eliminação da Hanseníase como problema de saúde pública: manual técnico-operacional*. Brasília, 2016. Disponível em: <https://bvsms.saude.gov.br/bvs/publicacoes/diretrizes_eliminacao_hanseniase_manual_tecnico.pdf>.
- [World Health Organization 2023]World Health Organization. Global leprosy (Hansen disease) update, 2022: adapting and moving forward. *Weekly Epidemiological Record*, v. 98, n. 36, p. 389–412, 2023. Disponível em: <<https://iris.who.int/bitstream/handle/10665/372429/WER9836-389-412-eng-fre.pdf>>.

A PARECER DO COMITÊ DE ÉTICA EM PESQUISA

COMPLEXO HOSPITALAR HUOC/PROCAPE



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Hansen.ai: Auxiliando o diagnóstico e a classificação do grau de incapacidade em pacientes com hanseníase por meio de inteligência artificial

Pesquisador: PATRICIA TAKAKO ENDO

Área Temática:

Versão: 4

CAAE: 79763024.8.0000.5192

Instituição Proponente: FUNDACAO UNIVERSIDADE DE PERNAMBUCO

Patrocinador Principal: MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO
MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO

DADOS DO PARECER

Número do Parecer: 6.986.262

Apresentação do Projeto:

A hanseníase é uma doença infectocontagiosa e de evolução crônica, causada pela bactéria *Mycobacterium leprae* e transmitida pelas vias aéreas superiores. Contudo, necessita de um contato prolongado com o portador sem tratamento e alta carga bacilar. Atinge principalmente a pele, as mucosas e os nervos periféricos. Caso não tratada adequadamente, pode levar a deformidades e incapacidades físicas, por vezes irreversíveis (Brasil, 2022; Brasil, 2023). A doença é compreendida como um grave problema de saúde pública devido à sua magnitude, gravidade e relevância social. Nesse sentido, o Ministério da Saúde (MS) trabalha através do Sistema de Informação de Agravos de Notificação (SINAN), com a coleta de indicadores que tem o objetivo de monitorar, acompanhar, transmitir e disseminar os padrões de ocorrência do agravo e suas tendências epidemiológicas. A hanseníase atinge indivíduos de qualquer sexo e faixa etária e, se relaciona a condições econômicas, sociais e ambientais desfavoráveis, ou seja, não afeta de maneira uniforme todas as pessoas, apresenta tendência para uma maior repercussão nas populações mais vulneráveis (Melo; Moraes; Santos, 2017; Brasil, 2019). Apesar da redução global do número de casos em tratamento para níveis abaixo de 1 por 10.000 habitantes, ainda são necessárias ações para interromper a transmissão, detectar ativamente novos casos, tratar adequadamente e prevenir as incapacidades físicas.

As

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

dificuldades no reconhecimento dos seus sinais e sintomas iniciais, frente à semelhança de suas características clínicas com outras doenças dermatoneurológicas são apontadas como principal causa do diagnóstico tardio e, conseqüentemente, do desenvolvimento do potencial incapacitante da patologia, podendo gerar deformidades, por vezes irreversíveis. (Ribeiro Et Al., 2018; Brasil, 2017). O país ocupa o segundo lugar no cenário internacional em número absoluto dos países com maior número de casos novos registrados, ficando atrás apenas da Índia. A região do Nordeste está entre as três primeiras com maior número de casos e, dentre suas unidades federadas, Pernambuco assume a terceira posição em casos novos por ano. O território pernambucano ocupa o segundo lugar do Nordeste e a quinta posição no país quanto aos indicadores de casos detectados em crianças menores de 15 anos, o que reflete uma alta carga bacilar no estado, que compromete a interrupção da cadeia de transmissão (Brasil, 2021a; Melo; Moraes; Santos, 2017; Brasil, 2020a). Considerada uma doença negligenciada e milenar, a hanseníase possui uma relação estreita entre condições de vida dos indivíduos atingidos por ela e o processo de saúde e doença. A influência dos Determinantes Sociais em Saúde (DSS), como desigualdades socioeconômicas e falta de acesso a serviços de qualidade, se torna evidente na distribuição geográfica da hanseníase, que tende a ser mais prevalente em determinadas regiões do país. Na ausência de tratamento oportuno, a hanseníase cursa com o agravamento do acometimento neural e, possivelmente, instalação de incapacidades físicas e/ou deformidades comprometendo a qualidade de vida dos indivíduos e causando estigma e discriminação social. Além do impacto sobre a qualidade de vida, o sofrimento causado pelo estigma e pela discriminação pode afetar as relações sociais, o bem-estar mental e a condição socioeconômica da pessoa afetada pela doença e sua rede de apoio (Brasil, 2022, D'Azevedo et al. 2021). As incapacidades físicas causam redução nas atividades de vida diária e são importantes fatores no que diz respeito à restrição à participação na sociedade de pessoas atingidas pela doença. Destaca-se que aproximadamente 4 milhões de pessoas no mundo apresentam deficiências físicas visíveis causadas pelas complicações da hanseníase (Brasil, 2017). Destaca-se que modelos de inteligência artificial (IA) podem ser ferramentas valiosas na análise, otimização e desenvolvimento de políticas públicas relacionadas ao tratamento de pacientes com hanseníase e à prevenção da incapacidade decorrente da doença. A análise de grandes conjuntos de dados pode identificar áreas de maior risco de contrair a doença, auxiliar na triagem de pacientes e otimizar o direcionamento de recursos. A IA também pode ajudar na criação de modelos preditivos para identificar pacientes com maior probabilidade de desenvolver incapacidades, permitindo a implementação de medidas

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

preventivas personalizadas. Diferente dos trabalhos da literatura atual, que focam em métodos de inteligência artificial principalmente para o diagnóstico da hanseníase por meio da análise de imagens de lesões cutâneas, este projeto planeja entregar como principais produtos (a) uma plataforma para predição de grau de incapacidade, com interface amigável, que possa auxiliar no tratamento de pacientes com hanseníase e (b) um sistema digital de ANS integrado com dispositivo microprocessado para auxiliar nos testes de sensibilidade. Neste sentido, este projeto representa uma iniciativa para apoiar profissionais de saúde na prevenção de incapacidades causadas pela hanseníase, na tomada de decisões clínicas e na formulação de políticas públicas de saúde com base nos DDS. Essa iniciativa está alinhada com a Estratégia Global de Hanseníase 2021-2030 - Rumo à zero hanseníase e a aplicabilidade dessa proposta ao SUS é multifacetada, podendo contribuir significativamente para o aprimoramento dos serviços de saúde, promoção da equidade em saúde e redução das complicações relacionadas à hanseníase, melhorando a qualidade de vida dos pacientes.

A primeira fase busca desenvolver uma plataforma para predição de incapacidades físicas em pessoas afetadas pela hanseníase. Para alcançar os objetivos, serão realizadas as seguintes etapas: Realização de uma revisão sistemática para expandir os conhecimentos sobre técnicas de inteligência artificial, para identificar quais se encaixam melhor no contexto do problema do projeto. Aquisição da base de dados, através de digitalização e tabulação dos dados dos prontuários dos pacientes com hanseníase atendidos no Hospital Otávio de Freitas (HOF). Os dados serão os sociodemográficos, clínicos e da ficha de avaliação neurológica simplificada (anexo) Pré-processamento da base de dados, com o objetivo de realizar uma limpeza nos dados, a fim de remover outliers e identificar atributos com dados faltantes ou irrelevantes para o problema. Análise dos principais DDS, destacando como fatores sociais, econômicos e culturais desempenham um papel crítico na incidência da doença, no acesso aos cuidados de saúde e no sucesso das estratégias de controle. Desenvolvimento do sistema de coleta de dados das Avaliações Neurológicas Simplificadas (ANS), investigando a possibilidade de integração com dispositivos do tipo wearable. Treino e teste de modelos e análise dos resultados, considerando uma metodologia científica rigorosa para definir e comparar diferentes modelos de inteligência artificial, a fim de identificar possíveis vieses, e também avaliar resultados referentes ao aprendizado dos modelos. Refinamento dos modelos com objetivo de melhorar os resultados das métricas obtidas. Desenvolvimento de um protótipo da plataforma com modelos de inteligência artificial para predição de grau de incapacidade com interface gráfica amigável para os profissionais de saúde do SUS. Realização de testes de

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

usabilidade junto aos profissionais de saúde. Para tal, serão utilizados dados do SINAN e DATASUS (anexo), assim como dados sociodemográficos (idade, sexo, raça, escolaridade, município de residência, ocupação) e clínicos dos pacientes em atendimento em uma unidade de referência estadual de Pernambuco, o Hospital Otávio de Freitas (HOF) (forma clínica, número de lesões cutâneas, classificação operacional, número de nervos afetados, baciloscopia, esquema terapêutico inicial). No HOF, será necessária a coleta de dados diretamente nos prontuários dos pacientes cadastrados na unidade, considerando os últimos 10 anos de atendimento, os dados serão os sociodemográficos e clínicos já citados, assim como informações sobre estado reacional, alterações no esquema terapêutico e evoluções na avaliação neurológica simplificada. Para a coleta de dados serão utilizados um formulário contendo as variáveis sociodemográficas e clínicas. Serão realizados experimentos estatísticos e testes de significância para avaliação da acurácia, sensibilidade, precisão, especificidade. Na fase 2, Será realizada a validação da plataforma junto a juízes especialistas em hansenologia, profissionais de saúde e pacientes em atendimento no ambulatório de Dermatologia Sanitária do Hospital Otávio de Freitas. Para a validação com juízes especialistas, serão realizadas as seguintes etapas: Levantamento de pesquisadores e profissionais de saúde que atuam na atenção à saúde de pessoas com hanseníase, a busca será realizada pela plataforma lattes; Envio dos dados para avaliação da plataforma (informações para acesso a plataforma e formulário de avaliação); Será aplicado o índice de validade de conteúdo (IVC) para apresentar a concordância entre os juízes e a estimativa de confiabilidade entre os avaliadores. Para validação com o público-alvo serão realizadas as seguintes etapas: Treinamento da equipe de saúde para utilização da plataforma HANSEN.AI; Pré-teste da plataforma em consultas no ambulatório e utilização da plataforma junto aos pacientes. Critério de Inclusão: População a ser estudada: Participarão do estudo, na fase de validação da plataforma, juízes especialistas da área da saúde; profissionais de saúde do ambulatório de dermatologia e Pessoas atingidas pela hanseníase atendidas no ambulatório de dermatologia do Hospital Otávio de Freitas (HOF). Estes serão recrutados a partir de um convite em participar da pesquisa no período que estão no ambulatório. Serão incluídos os juízes especialistas na área de hansenologia (professores e/ou pesquisadores que desenvolveram projetos voltados para pessoas afetadas pela hanseníase) que alcancem no mínimo cinco pontos na avaliação de Fehring (1994). Serão incluídos os profissionais de saúde (médicos, enfermeiros, terapeutas ocupacionais e fisioterapeutas) que atuam no atendimento clínico as pessoas afetadas pela hanseníase no ambulatório de dermatologia do HOF e que estejam realizando atendimento aos pacientes no

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

período da coleta de dados. Serão incluídas as pessoas afetadas pela hanseníase com idade igual ou superior a 18 anos, ambos os sexos; em acompanhamento no ambulatório no primeiro e último trimestre do tratamento com a PQT. Critério de Exclusão: Serão excluídos os juízes especialistas aposentados há mais de 5 anos da área de hansenologia, a fim de garantir a atualização do conhecimento e experiência na área. Assim como, os indivíduos que não estejam ativamente envolvidos em atividades relacionadas à hansenologia nos últimos 5 anos, como pesquisa, ensino ou assistência clínica. Serão excluídos os Profissionais de Saúde que estejam em licença ou afastamento do trabalho durante o período da coleta de dados; ou que tenham sido transferidos para outras funções dentro do hospital ou ambulatório, não estando mais diretamente envolvidos no atendimento aos pacientes com hanseníase. Serão excluídas as pessoas afetadas pela hanseníase com comorbidades com uso de medicamentos que possam interferir no tratamento com PQT; Pessoas em reação hansênica tipo I ou II; e abandono do tratamento

Objetivo da Pesquisa:

Objetivo Primário:

Desenvolver a plataforma hansen.ai para classificar pacientes diagnosticados com hanseníase quanto ao grau de incapacidade que podem desenvolver e possivelmente prever complicações decorrentes da doença, utilizando modelos de inteligência artificial.

Objetivo Secundário:

- Realizar uma análise dos registros clínicos e sociodemográficos das pessoas afetadas pela hanseníase para identificar os Determinantes Sociais em Saúde (DDS) que podem influenciar o desenvolvimento da doença e a progressão de incapacidades;- Desenvolver um sistema (web, mobile ou Progressive Web App (PWA)) para auxiliar o profissional de saúde, com atuação nas Unidades de Atenção Básica do SUS, na coleta e no preenchimento da Avaliação Neurológica Simplificada (ANS), a partir de uma interface interativa e amigável, com o objetivo de mitigar erros de preenchimento, bem como garantir a persistência dos dados, facilitando a utilização dos dados pelos modelos preditivos; Investigar a viabilidade da integração de um dispositivo microprocessado (tipo wearable ou similar) para auxiliar o profissional de saúde nos testes de sensibilidade em pontos específicos dos pés e das mãos do paciente, integrando-o com o sistema de ANS a ser desenvolvido;- Desenvolver uma plataforma e avaliar quantitativamente modelos de inteligência artificial que possam prever o grau de incapacidade que os pacientes diagnosticados com hanseníase podem desenvolver ao longo do tratamento. Isso envolve o

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

treinamento dos modelos com base em dados históricos de pacientes e avaliações neurológicas simplificadas.- Validar a plataforma junto aos profissionais de saúde e pacientes em uma unidade de referência para o atendimento de pessoas afetadas pela doença no estado de Pernambuco.

Avaliação dos Riscos e Benefícios:

Riscos:

Os riscos psicológicos envolvidos são do tipo desconforto emocional, ou apreensão decorrente da participação na pesquisa, por meio dos questionários que analisarão a probabilidade de surgimento de incapacidades físicas ao longo do tratamento. Para minimizar estes riscos, a equipe do projeto contará com a colaboração dos profissionais de saúde do HOF, que conta com equipe multidisciplinar, entre eles psicóloga e terapeuta ocupacional. Os riscos envolvem ainda a precisão e confiabilidade das previsões, pois será uma plataforma em desenvolvimento e a precisão das previsões pode ser limitada pela qualidade e quantidade de dados disponíveis. É fundamental realizar estudos rigorosos para avaliar a acurácia e a confiabilidade da plataforma antes de sua implementação na prática clínica. Outro risco é a interpretabilidade dos modelos de inteligência artificial, algo crucial para que os profissionais de saúde compreendam os fatores que influenciam as previsões. Para redução do risco, serão desenvolvidas ferramentas, documentação e treinamentos que auxiliem a interpretação clara das sugestões às decisões da plataforma aos profissionais da saúde. Os riscos em relação ao armazenamento, compartilhamento e análise de dados podem existir em várias camadas. Os principais são: acesso indevido aos dados, interceptação de transmissões digitais e perda/furto de equipamentos de armazenamento. O acesso aos dispositivos do projeto são protegidos com encriptação individual das contas e controle restrito às máquinas. Além disso, o banco de dados local adiciona uma sobrecamada de métodos de encriptação e segurança aos dados. As transmissões de dados se darão por rede encriptada e de dados encriptados, aumentando a confiabilidade e segurança do processo. O furto/roubo de equipamentos tem seu risco mitigado pelo protocolo de segurança e encriptação da máquina e estratégias de controle e apagamento remoto permitidos pelos sistemas operacionais utilizados no projeto. Por fim, há o risco relacionado a validade externa, pois a plataforma precisa ser validada em diferentes populações e contextos para garantir que os resultados sejam generalizáveis. É importante considerar a diversidade de fatores que podem influenciar o desenvolvimento de incapacidades em hanseníase. Esse risco será mitigado através de processos de desenvolvimento iterativos

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

de sistemas com contato técnico frequente com profissionais da área. Assim, a solução proposta neste projeto terá sempre acompanhamento dos reais envolvidos no futuro uso, reduzindo o risco de não haver validade externa a posteriori. Destaca-se ainda que a plataforma precisa ser robusta e segura para evitar falhas, erros de software ou ataques cibernéticos. É fundamental garantir a qualidade do desenvolvimento e a implementação de medidas de segurança adequadas conforme apresentados na seção ¿Armazenamento e guarda dos dados¿ a seguir.

Benefícios:

A plataforma Hansen.AI pode oferecer benefícios, como: melhoria da qualidade do atendimento aos pacientes; Redução dos custos de tratamento; Prevenção de incapacidades físicas e promoção da autonomia e da qualidade de vida dos pacientes.

Comentários e Considerações sobre a Pesquisa:

Da forma agora apresentada o projeto encontra-se em conformidade com as resoluções vigentes (Resolução 466/12, Norma Operacional 001/13 CNS-MS).

Considerações sobre os Termos de apresentação obrigatória:

: Apresentados atendendo o protocolo de pesquisas, em conformidade com as resoluções vigentes (Resolução 466/12, Norma Operacional 001/13 CNS-MS).

Conclusões ou Pendências e Lista de Inadequações:

Diante do exposto, o CEP-HUOC/PROCAPE, de acordo com as atribuições definidas na Resolução CNS nº 466 de 2012 e na Norma Operacional nº 001 de 2013 do CNS, manifesta-se, pela aprovação da proposta ao projeto de pesquisa, não apresentando óbice ético, devendo o pesquisador enviar à Plataforma Brasil, relatório parcial caso durante a pesquisa for demonstrando fatos relevantes e resultados parciais de seu desenvolvimento e um relatório final após o encerramento da pesquisa, totalizando seus resultados .

Considerações Finais a critério do CEP:

Diante do exposto, o CEP-HUOC/PROCAPE, de acordo com as atribuições definidas na Resolução CNS nº 466 de 2012 e na Norma Operacional nº 001 de 2013 do CNS, manifesta-se, pela aprovação da proposta ao projeto de pesquisa, não apresentando óbice ético, devendo o pesquisador enviar à Plataforma Brasil, relatório parcial caso durante a pesquisa for demonstrando fatos relevantes e resultados parciais de seu desenvolvimento e um relatório final após o encerramento da pesquisa, totalizando seus resultados .

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

E-mail: cep_huoc.procape@upe.br

COMPLEXO HOSPITALAR HUOC/PROCAPE



Continuação do Parecer: 6.986.262

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_P ROJETO_2316530.pdf	29/07/2024 15:35:44		Aceito
Projeto Detalhado / Brochura Investigador	3projetodetalhado.docx	29/07/2024 15:35:21	PATRICIA TAKAKO ENDO	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_hanseai.docx	29/07/2024 15:34:17	PATRICIA TAKAKO ENDO	Aceito
Outros	INSTRUMENTOS_HANSENAI.pdf	29/07/2024 15:33:54	PATRICIA TAKAKO ENDO	Aceito
Outros	lattesdanielle.pdf	05/07/2024 16:35:35	PATRICIA TAKAKO ENDO	Aceito
Outros	lattespatricia.pdf	05/07/2024 16:34:31	PATRICIA TAKAKO ENDO	Aceito
Outros	cartadeanuencia_HOF.pdf	10/05/2024 11:08:21	PATRICIA TAKAKO ENDO	Aceito
Outros	TERMO_DE_CONFIDENCIALIDADE_a ssinado.pdf	10/05/2024 10:58:57	PATRICIA TAKAKO ENDO	Aceito
Folha de Rosto	PATRICIA_TAKAKO_ENDO_folha_de_r osto_Fensg_assinado.pdf	10/05/2024 10:51:29	PATRICIA TAKAKO ENDO	Aceito

Situação do Parecer:

Aprovado

Necessita apreciação da CONEP:

Não

RECIFE, 06 de Agosto de 2024

Assinado por:
ROSANA ALMEIDA DE MORAES
(Coordenador(a))

Endereço: Rua Arnóbio Marques, 310

Bairro: Santo Amaro

CEP: 50.100-130

UF: PE

Município: RECIFE

Telefone: (81)3184-1271

Fax: (81)3184-1271

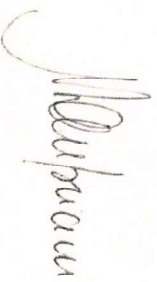
E-mail: cep_huoc.procape@upe.br

B CERTIFICADO DE MELHOR ARTIGO

Certificado

Certificamos que o trabalho INTITULADO “DO PAPEL AO APLICATIVO: ANS2J3ON, INOVAÇÃO DIGITAL NO ARMAZENAMENTO E PROCESSAMENTO DAS AVALIAÇÕES NEUROLÓGICAS SIMPLIFICADAS” de autoria de: “*Jamilly Braga Melo; Kayo Henrique de Carvalho Monteiro; Igor Vitor Teixeira; Aymée Medeiros da Rocha; Patricia Takako Endo; Hillson Gomes Vilar de Andrade*” obteve “**Primeiro Lugar**”, na área **Prevenção de Incapacidades e Reabilitação** entre os melhores trabalhos apresentados no **19º Congresso Brasileiro de Hansenologia, realizado de 25 a 28 de novembro de 2025.**

Foz do Iguaçu/PR, 28 de novembro de 2025.



Marco Andrey C. Frade
Presidente da SBH e do Congresso



Marilda A. Milanez Morgado de Abreu
Diretora Científica da SBH e do Congresso

C SCRIPT DE PRÉ-PROCESSAMENTO DAS REGIÕES DE INTERESSE

```
1 import cv2
2 import os
3 import zipfile
4 from google.colab import files
5
6 # Caminho da pasta com as imagens
7 input_dir = "/content/imagens/dataset_ANS"
8
9 # Criar as pastas de saida
10 hands_output_dir = "/content/hands"
11 feet_output_dir = "/content/feet"
12 os.makedirs(hands_output_dir, exist_ok=True)
13 os.makedirs(feet_output_dir, exist_ok=True)
14
15 # Definir as coordenadas para as areas de maos e pes
16 hands_coords = (0, 1648, 1783, 2450) # Coordenadas (x_min, y_min, x_max
    , y_max)
17 feet_coords = (1740, 544, 3476, 1400) # Coordenadas (x_min, y_min,
    x_max, y_max)
18
19 # Funcao para cortar uma area de interesse
20 def crop_image(image, coords):
21     x_min, y_min, x_max, y_max = coords
22     return image[y_min:y_max, x_min:x_max]
23
24 # Processar todas as imagens e cortar nas areas de interesse
25 for filename in os.listdir(input_dir):
26     if filename.endswith((".png", ".jpg", ".jpeg")):
27         img_path = os.path.join(input_dir, filename)
28         image = cv2.imread(img_path)
29
30         if image is None:
31             print(f"Erro ao carregar {filename}. Pulando...")
32             continue
33
34         # Verificar se a imagem possui o tamanho minimo necessario para
```

```

o corte
35     img_height, img_width = image.shape[:2]
36     if img_width < max(hands_coords[2], feet_coords[2]) or
img_height < max(hands_coords[3], feet_coords[3]):
37         print(f"{filename} eh menor do que o necessario para o corte
. Pulando...")
38         continue
39
40     # Cortar as areas das maos e pes
41     hands_crop = crop_image(image, hands_coords)
42     feet_crop = crop_image(image, feet_coords)
43
44     # Recorta a regio das maos e pes e salva
45     if hands_crop.size != 0:
46         hands_save_path = os.path.join(hands_output_dir, filename)
47         cv2.imwrite(hands_save_path, hands_crop)
48     else:
49         print(f"Recorte das maos vazio para {filename}. Pulando...")
50
51     if feet_crop.size != 0:
52         feet_save_path = os.path.join(feet_output_dir, filename)
53         cv2.imwrite(feet_save_path, feet_crop)
54     else:
55         print(f"Recorte dos pes vazio para {filename}. Pulando...")
56
57     print(f"Processado: {filename}")
58
59 print("Corte e salvamento concluidos!")
60
61 # Compactar as pastas de saida para download
62 !zip -r /content/hands.zip /content/hands
63 !zip -r /content/feet.zip /content/feet
64
65 # Download dos arquivos compactados com os recortes
66 files.download("/content/hands.zip")
67 files.download("/content/feet.zip")

```

Listing C.1: Script de corte das Regiões de Interesse (ROI)

D ALGORITMO PARA MAPEAMENTO ANATÔMICO

```
1 # Funcoes auxiliares
2
3 def get_center_from_box(box):
4     """Calcula o centro de uma bounding box [x1, y1, x2, y2]."""
5     x1, y1, x2, y2 = box
6     return (x1 + x2) / 2, (y1 + y2) / 2
7
8 def is_inside(inner_box_center, container_box):
9     """Verifica se o centro de uma caixa esta dentro de outra."""
10    cx, cy = inner_box_center
11    x1, y1, x2, y2 = container_box
12    return x1 <= cx <= x2 and y1 <= cy <= y2
13
14
15 # Mapeamento de pontos na mao
16
17 def mapear_pontos_mao(pontos_na_mao, tipo_mao):
18     """
19     Mapeia os pontos detectados nas maos em regioes anatomicas.
20     Retorna um dicionario com os nomes mapeados.
21     """
22     pontos_ordenados = sorted(
23         pontos_na_mao,
24         key=lambda p: (p['center'][0], p['center'][1])
25     )
26     pontos_ordenados += [None] * (7 - len(pontos_ordenados))
27
28     posicoes = {}
29     if tipo_mao == "md":
30         posicoes = {
31             "falange_distal_polegar": pontos_ordenados[0],
32             "dorso_mao": pontos_ordenados[1],
33             "falange_distal_2_dedo": pontos_ordenados[2],
34             "falange_proximal_2_dedo": pontos_ordenados[3],
35             "borda_medial_palma_mao": pontos_ordenados[4],
36             "falange_proximal_5_dedo": pontos_ordenados[5],
```



```

37         "falange_distal_5_dedo": pontos_ordenados[6]
38     }
39     elif tipo_mao == "me":
40         posicoes = {
41             "falange_distal_5_dedo": pontos_ordenados[0],
42             "falange_proximal_5_dedo": pontos_ordenados[1],
43             "borda_medial_palma_mao": pontos_ordenados[2],
44             "falange_proximal_2_dedo": pontos_ordenados[3],
45             "falange_distal_2_dedo": pontos_ordenados[4],
46             "dorso_mao": pontos_ordenados[5],
47             "falange_distal_polegar": pontos_ordenados[6]
48         }
49
50     return {nome: ponto for nome, ponto in posicoes.items() if ponto is
51           not None}
52
53 # Mapeamento de pontos no pe
54
55 PONTOS_PE_DIREITO = {
56     'calcaneo': 'calcaneo_direito',
57     'borda_lateral': 'borda_lateral_planta_pe_direito',
58     'borda_medial': 'borda_medial_planta_pe_direito',
59     '5_artelho': '5_artelho_direito',
60     'falange_distal_5_artelho': 'falange_distal_5_artelho_direito',
61     '3_artelho': '3_artelho_direito',
62     'falange_distal_3_artelho': 'falange_distal_3_artelho_direito',
63     'cabeca_metatarsofalangeana': '
64     cabeca_metatarsofalangeana_halux_direito',
65     'falange_distal_halux': 'falange_distal_halux_direito',
66     'dorso': 'dorso_direito'
67 }
68
69 PONTOS_PE_ESQUERDO = {
70     'calcaneo': 'calcaneo_esquerdo',
71     'borda_lateral': 'borda_lateral_planta_pe_esquerdo',
72     'borda_medial': 'borda_medial_planta_pe_esquerdo',
73     '5_artelho': '5_artelho_esquerdo',
74     'falange_distal_5_artelho': 'falange_distal_5_artelho_esquerdo',

```



```

74     '3_artelho': '3_artelho_esquerdo',
75     'falange_distal_3_artelho': 'falange_distal_3_artelho_esquerdo',
76     'cabeca_metatarsofalangeana': '
cabeca_metatarsofalangeana_halux_esquerdo',
77     'falange_distal_halux': 'falange_distal_halux_esquerdo',
78     'dorso': 'dorso_esquerdo'
79 }
80
81 # Encontra o ponto mais pr ximo abaixo
82
83 def find_nearest_below(reference_point, candidates, used_indices):
84     ref_center = reference_point['center']
85     best_idx = -1
86     min_dist = float('inf')
87
88     for i, candidate in enumerate(candidates):
89         if i in used_indices:
90             continue
91         cand_center = candidate['center']
92         if cand_center[1] > ref_center[1]:
93             dist = ((ref_center[0] - cand_center[0])**2 * 1.5 + (
ref_center[1] - cand_center[1])**2)**0.5
94             if dist < min_dist:
95                 min_dist = dist
96                 best_idx = i
97
98     return (candidates[best_idx], best_idx) if best_idx != -1 else (None
, -1)
99
100 # Mapeia os pontos detectados no pe para seus nomes anatomicos
101
102 def mapear_pontos_pe(pontos_no_pe, tipo_pe):
103
104     if not pontos_no_pe:
105         return {}
106
107     for i, ponto in enumerate(pontos_no_pe):
108         ponto['original_index'] = i
109

```

```

110     candidatos = list(pontos_no_pe)
111     mapeados = {}
112     usados = set()
113     NOMES = PONTOS_PE_DIREITO if tipo_pe == 'pd' else PONTOS_PE_ESQUERDO
114
115     # 1. Calcaneo
116     candidatos.sort(key=lambda p: p['center'][1], reverse=True)
117     for p in candidatos:
118         if p['original_index'] not in usados:
119             mapeados[NOMES['calcaneo']] = p
120             usados.add(p['original_index'])
121             break
122
123     # 2. Linha superior (4 pontos mais altos)
124     candidatos.sort(key=lambda p: p['center'][1])
125     linha_topo = [p for p in candidatos if p['original_index'] not in
126 usados][:4]
127     reverse_order = (tipo_pe == 'pe')
128     linha_topo.sort(key=lambda p: p['center'][0], reverse=reverse_order)
129
130     if len(linha_topo) == 4:
131         mapeados[NOMES['falange_distal_5_artelho']] = linha_topo[0];
132         usados.add(linha_topo[0]['original_index'])
133         mapeados[NOMES['falange_distal_3_artelho']] = linha_topo[1];
134         usados.add(linha_topo[1]['original_index'])
135         mapeados[NOMES['falange_distal_halux']] = linha_topo[2]; usados.
136         add(linha_topo[2]['original_index'])
137         mapeados[NOMES['dorso']] = linha_topo[3]; usados.add(linha_topo
138 [3]['original_index'])
139
140     # 3. Buscar pontos logo abaixo dos dedos
141     pares = [
142         ('falange_distal_5_artelho', '5_artelho'),
143         ('falange_distal_3_artelho', '3_artelho'),
144         ('falange_distal_halux', 'cabeca_metatarsofalangeana')
145     ]
146
147     for ref, alvo in pares:
148         ref_nome = NOMES.get(ref)

```

```

144         if ref_nome in mapeados:
145             ponto_ref = mapeados[ref_nome]
146             prox, idx = find_nearest_below(ponto_ref, pontos_no_pe,
usados)
147             if prox:
148                 mapeados[NOMES[alvo]] = prox
149                 usados.add(idx)
150
151     # 4. Borda medial e lateral (pontos restantes)
152     bordas = [p for i, p in enumerate(pontos_no_pe) if i not in usados]
153     if len(bordas) >= 2:
154         bordas.sort(key=lambda p: p['center'][0])
155         if tipo_pe == 'pd':
156             mapeados[NOMES['borda_lateral']] = bordas[0]
157             mapeados[NOMES['borda_medial']] = bordas[1]
158         else:
159             mapeados[NOMES['borda_medial']] = bordas[0]
160             mapeados[NOMES['borda_lateral']] = bordas[1]
161
162     return mapeados

```

Listing D.1: Algoritmo para mapeamento anatômico

E CRIAÇÃO DO DATASET DE TREINAMENTO VIA BOOTSTRAPPING

```
1
2 from ultralytics import YOLO
3 import pandas as pd
4 from tqdm.notebook import tqdm
5 import cv2
6 import math
7
8 # Carrega o modelo
9 model = YOLO('best.pt')
10
11 # Lista de imagens a serem processadas
12 image_folder = /content/imagens
13 image_files = [f for f in os.listdir(image_folder) if f.lower().endswith
14                 ((''.jpg', '.png', '.jpeg'))]
15
16 # Labels que definem os membros
17 limb_labels = {'md', 'me', 'pd', 'pe'}
18 # Lista para armazenar todos os dados finais
19 final_dataset = []
20
21 # Itera sobre cada imagem com uma barra de progresso
22 for image_name in tqdm(image_files, desc="Processando Imagens"):
23     image_path = os.path.join(image_folder, image_name)
24
25     # Carrega a imagem para obter suas dimensoes
26     img = cv2.imread(image_path)
27     if img is None:
28         print(f"Aviso: Nao foi possivel ler a imagem {image_name}.
29         Pulando.")
30         continue
31
32     img_h, img_w, _ = img.shape
33
34     results = model(img, verbose=False)[0]
35
36     # Separa as deteccoes em membros e pontos
```

```

34     limbs = []
35     points = []
36     for box in results.bboxes.data.tolist():
37         x1, y1, x2, y2, conf, cls_id = box
38         label = results.names[int(cls_id)]
39
40         detection = {
41             'label': label,
42             'box': [x1, y1, x2, y2],
43             'confidence': conf,
44             'center': get_center_from_box([x1, y1, x2, y2])
45         }
46
47         if label in limb_labels:
48             limbs.append(detection)
49         else:
50             points.append(detection)
51
52     # Para cada membro detectado, encontra os pontos e calcula as
53     # features
54     for limb in limbs:
55         limb_label = limb['label']
56         limb_box = limb['box']
57         limb_center = limb['center']
58
59         # Calculo das features do MEMBRO
60         lx1, ly1, lx2, ly2 = limb_box
61         limb_w = lx2 - lx1
62         limb_h = ly2 - ly1
63
64         if limb_w <= 0 or limb_h <= 0: continue # Pula bboxes invalidas
65
66         limb_area = limb_w * limb_h
67         limb_aspect_ratio = limb_w / limb_h
68
69         # Filtra os pontos que estao dentro da bbox do membro atual
70         points_inside_limb = [p for p in points if is_inside(p['center'], limb_box)]

```

```

71     mapped_points = {}
72     if limb_label in ['me', 'md']:
73         mapped_points = mapear_pontos_mao(points_inside_limb,
limb_label)
74     elif limb_label in ['pe', 'pd']:
75         mapped_points = mapear_pontos_pe(points_inside_limb,
limb_label)
76
77     # Adiciona os pontos mapeados e as features ao dataset final
78     for anatomical_name, point_info in mapped_points.items():
79         point_center = point_info['center']
80         px1, py1, px2, py2 = point_info['box']
81         point_w = px2 - px1
82         point_h = py2 - py1
83         point_area = point_w * point_h if point_w > 0 and point_h >
0 else 0
84
85         # 1. Distancia Euclidiana ao Centro do Membro
86         dist_centro = math.sqrt((point_center[0] - limb_center[0])
**2 + (point_center[1] - limb_center[1])**2)
87
88         # 2. Coordenadas Relativas do Ponto (0 a 1)
89         ponto_x_rel = (point_center[0] - lx1) / limb_w
90         ponto_y_rel = (point_center[1] - ly1) / limb_h
91
92         # 3. Area Relativa do Ponto
93         area_rel = point_area / limb_area
94
95         final_dataset.append({
96             # Identificadores
97             'imagem': image_name,
98             'largura_imagem': img_w,
99             'altura_imagem': img_h,
100             'membro_tipo': limb_label,
101             'ponto_anatomico': anatomical_name,
102
103             # Features do Ponto
104             'ponto_centro_x': round(point_center[0], 2),
105             'ponto_centro_y': round(point_center[1], 2),

```

```

106         'cor_detectada': point_info['label'],
107         'distancia_ao_centro_membro': round(dist_centro, 2),
108         'ponto_centro_x_relativo': round(ponto_x_rel, 4),
109         'ponto_centro_y_relativo': round(ponto_y_rel, 4),
110         'area_relativa_ponto': round(area_rel, 6),
111         'aspect_ratio_membro': round(limb_aspect_ratio, 4),
112
113         # Confianças
114         'conf_ponto': round(point_info['confidence'], 4),
115         'conf_membro': round(limb['confidence'], 4)
116     })
117
118 print("\nProcessamento concluído!")
119
120 if final_dataset:
121     # Cria o DataFrame
122     df = pd.DataFrame(final_dataset)

```

Listing E.1: Criação do Dataset de Treinamento via Bootstrapping

F INFERÊNCIA

```
1 import os
2 import math
3 import pandas as pd
4
5 # Nota: Os modelos (YOLO, RF, etc.) e o logger sao passados como
6     argumentos para manter esta funcao independente do ambiente da API.
7
8 def processar_com_ia(image_path, yolo_model, rf_model, label_encoder,
9     training_columns, master_point_list, logger):
10     """
11     Processa uma imagem para detectar membros e pontos de sensibilidade,
12     utilizando
13     um pipeline de dois estagios (YOLOv8 para detecção e Random Forest
14     para classificacao).
15
16     Args:
17         image_path (str): Caminho para o arquivo de imagem.
18         yolo_model: Modelo YOLOv8 pre-carregado.
19         rf_model: Modelo Random Forest pre-carregado.
20         label_encoder: LabelEncoder treinado para os pontos anatomicos.
21         training_columns (list): Lista de colunas de features usadas no
22             treino do RF.
23         master_point_list (dict): Dicionario com a lista de todos os
24             pontos esperados por membro.
25         logger: Instancia do logger para registro de eventos.
26
27     Returns:
28         dict: Um dicionario JSON com os dados da avaliacao sensorial
29             extraídos da imagem.
30     """
31     logger.info(f"Processando imagem: {os.path.basename(image_path)}")
32     json_output = {"imagem_processada": os.path.basename(image_path), "
33     consultas_maos": [], "consultas_pes": []}
34
35     # 1. Deteccao de Objetos com YOLO
36     results = yolo_model(image_path, verbose=False)[0]
```



```

29
30 # 2. Separacao de deteccoes em membros e pontos
31 all_limbs, all_points = [], []
32 limb_labels = {'md', 'me', 'pd', 'pe'}
33 for box in results.bboxes.data.tolist():
34     x1, y1, x2, y2, conf, cls_id = box
35     label = yolo_model.names[int(cls_id)]
36     detection = {
37         'label': label,
38         'box': [x1, y1, x2, y2],
39         'confidence': conf,
40         'center': ((x1 + x2) / 2, (y1 + y2) / 2)
41     }
42     if label in limb_labels:
43         all_limbs.append(detection)
44     else:
45         all_points.append(detection)
46
47     logger.info(f"YOLO detectou {len(all_limbs)} membros e {len(
48         all_points)} pontos de cor.")
49
50 # 3. Agrupamento e processamento por consulta
51 for tipo_grupo, labels_grupo in [("maos", ['md', 'me']), ("pes", ['
52     pd', 'pe'])]:
53     membros_do_tipo = sorted([l for l in all_limbs if l['label'] in
54         labels_grupo], key=lambda m: m['box'][0])
55
56     for i in range(0, len(membros_do_tipo), 2):
57         consulta_id = (i // 2) + 1
58         par_de_membros = membros_do_tipo[i:i+2]
59         consulta_dict = {"consulta_id": consulta_id, "membros": []}
60
61         for membro_atual in par_de_membros:
62             membro_label = membro_atual['label']
63             membro_box = membro_atual['box']
64             lx1, ly1, lx2, ly2 = membro_box
65
66             # Inicia o gabarito com todos os pontos esperados
67             expected_points = master_point_list.get(membro_label,

```

```

[])
65         mapped_points_dict = {point: "Nao identificado" for
point in expected_points}
66
67         # Filtra os pontos que estao dentro do membro atual
68         pontos_no_membro = [p for p in all_points if (lx1 <= p['
center']][0] <= lx2) and (ly1 <= p['center'][1] <= ly2)]
69
70         limb_center = membro_atual['center']
71         limb_w, limb_h = lx2 - lx1, ly2 - ly1
72         if limb_w <= 0 or limb_h <= 0: continue
73         limb_area, limb_aspect_ratio = limb_w * limb_h, limb_w /
limb_h
74
75         # 4. Classificacao de cada ponto com Random Forest
76         for ponto_atual in pontos_no_membro:
77             # 4.1. Engenharia de Features
78             point_center = ponto_atual['center']
79             px1, py1, px2, py2 = ponto_atual['box']
80             point_w, point_h = px2 - px1, py2 - py1
81             point_area = point_w * point_h if point_w > 0 and
point_h > 0 else 0
82             dist_centro = math.sqrt((point_center[0] -
limb_center[0])**2 + (point_center[1] - limb_center[1])**2)
83             ponto_x_rel = (point_center[0] - lx1) / limb_w
84             ponto_y_rel = (point_center[1] - ly1) / limb_h
85             area_rel = point_area / limb_area if limb_area > 0
else 0
86
87             # 4.2. Preparacao dos dados para predicao
88             data_for_prediction = {
89                 'membro_tipo': membro_label, 'cor_detectada':
ponto_atual['label'],
90                 'distancia_ao_centro_membro': dist_centro, '
ponto_centro_x_relativo': ponto_x_rel,
91                 'ponto_centro_y_relativo': ponto_y_rel, '
area_relativa_ponto': area_rel,
92                 'aspect_ratio_membro': limb_aspect_ratio, '
conf_ponto': ponto_atual['confidence'],

```

```

93         'conf_membro': membro_atual['confidence']
94     }
95     df_point = pd.DataFrame([data_for_prediction])
96     df_point_encoded = pd.get_dummies(df_point, columns
= ['membro_tipo', 'cor_detectada'])
97     df_point_reindexed = df_point_encoded.reindex(
columns=training_columns, fill_value=0)
98
99     # 4.3. Predicao
100     prediction_encoded = rf_model.predict(
df_point_reindexed)
101     anatomical_name_predicted = label_encoder.
inverse_transform(prediction_encoded)[0]
102
103     # Atualiza o gabarito com a cor detectada para o
ponto previsto
104     if anatomical_name_predicted in mapped_points_dict:
105         mapped_points_dict[anatomical_name_predicted] =
ponto_atual['label']
106
107     membro_info = {"tipo_membro": membro_label, "
pontos_mapeados": mapped_points_dict}
108     consulta_dict["membros"].append(membro_info)
109
110     if tipo_grupo == "maos":
111         json_output["consultas_maos"].append(consulta_dict)
112     else:
113         json_output["consultas_pes"].append(consulta_dict)
114
115     return json_output

```

Listing F.1: Pipeline de inferência do campo visual da avaliação sensitiva