

Gabriel Zambelli Scalabrini
João Victor Texeira Degelo
Johan Su Kwok

Reconhecimento facial para detecção de Doença Renal Crônica

São Paulo, SP

2024

Gabriel Zambelli Scalabrini
João Victor Texeira Degelo
Johan Su Kwok

Reconhecimento facial para detecção de Doença Renal Crônica

Trabalho de conclusão de curso apresentado
ao Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo para obtenção do Título de Engenheiro.

Universidade de São Paulo – USP

Escola Politécnica

Departamento de Engenharia de Computação e Sistemas Digitais (PCS)

Orientador: Profa. Dra. Anarosa Alves Franco Brandão

Coorientadores: Profa. Dra. Daniela Ponce

Profa. Dra. Maristela Carvalho da Costa,

Prof. Rogerio da Hora Passos

São Paulo, SP

2024

Resumo

A Doença Renal Crônica (DRC) afeta um grande número de pessoas em todo o mundo. Considerando apenas o cenário brasileiro, estima-se que mais de 10 milhões de cidadãos brasileiros tenham DRC, com 90 mil em tratamento dialítico. Portanto, a DRC impacta a sociedade de duas maneiras: como problema de saúde pública e como problema econômico. Este trabalho propõe o desenvolvimento de uma solução baseada em inteligência artificial utilizando imagens do paciente para auxiliar no diagnóstico de DRC. O projeto inclui o desenvolvimento de um aplicativo para a coleta de dados, além do desenvolvimento de um sistema de diagnóstico contendo um aplicativo destinado aos profissionais de saúde. Para isso, foram treinados modelos de aprendizado de máquina com informações demográficas, de saúde e imagens faciais e das mãos. A aplicação clínica da solução visa aumentar a eficiência no diagnóstico precoce da DRC, reduzindo o impacto da doença por meio de uma abordagem acessível e não invasiva. Este trabalho está alinhado ao ODS 3 - 3.4.

Palavras-chave: Doença Renal Crônica. Inteligência Artificial. Aprendizado de Máquina. Reconhecimento Facial. Diagnóstico Médico. Aplicativos Móveis.

Abstract

Chronic Kidney Disease (CKD) affects a large number of people worldwide. Considering the Brazilian scenario alone, it is estimated that more than 10 million Brazilian citizens have CKD, with 90 thousand undergoing dialysis treatment. Therefore, CKD impacts society in two ways: as a public health problem and as an economic problem. This work proposes developing a solution based on artificial intelligence using patient images to assist in the diagnosis of CKD. The project includes the development of an application for data collection, in addition to the development of a diagnostic system containing an application aimed at health professionals. To this end, machine learning models were trained with demographic and health information, as well as facial and hand images. The clinical application of the solution aims to increase the efficiency in the early diagnosis of CKD, reducing the impact of the disease through an accessible and non-invasive approach. This work is aligned with SDG 3 - 3.4.

Keywords: Chronic Kidney Disease. Artificial Intelligence. Machine Learning. Facial Recognition. Medical Diagnosis. Mobile Applications.

Lista de ilustrações

Figura 1 – Telas de login do auxiliar	35
Figura 2 – Listagem de pacientes, e criação de novo paciente	36
Figura 3 – Informações e fotos do paciente	37
Figura 4 – Diagrama de Arquitetura para o Sistema de Coleta	38
Figura 5 – Volume de dados coletados diariamente	42
Figura 6 – Pacientes em cada estágio da doença, separado de acordo com o sexo .	43
Figura 7 – Distribuição da idade dos pacientes separada pelo estágio da doença . .	43
Figura 8 – Proporção de pacientes afetados por cada condição de saúde em cada estágio da doença	44
Figura 9 – Número de pacientes participantes separados de acordo com o sexo . .	45
Figura 10 – Distribuição da idade dos participantes da pesquisa	45
Figura 11 – Proporção de pacientes afetados por cada condição de saúde	45
Figura 12 – Listagem de pacientes, e criação de novo paciente	53
Figura 13 – Perfil contendo informações do paciente	54
Figura 14 – Criação de novo diagnóstico	55
Figura 15 – Informações do diagnóstico do paciente	55
Figura 16 – Diagrama de Arquitetura	56
Figura 17 – Curva ROC para algoritmos XGBoost no conjunto de treino	60
Figura 18 – Curva ROC para algoritmos XGBoost no conjunto de validação	60
Figura 19 – Curva ROC para Redes Neurais no conjunto de treino	61
Figura 20 – Curva ROC para Redes Neurais no conjunto de validação	62
Figura 21 – Curva ROC para algoritmos selecionados no conjunto de teste	63

Sumário

1	INTRODUÇÃO	11
1.1	Definição	11
1.2	Estágios da DRC	11
1.3	Sintomas	12
1.4	Motivação	13
1.5	Objetivo	14
1.6	Justificativa	14
1.7	Organização do Trabalho	14
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Inteligência Artificial	17
2.1.1	Aprendizado Supervisionado	18
2.1.2	Classificação binária	19
2.1.3	Receiver Operating Characteristics	19
2.2	Revisão da Literatura para DRC	20
2.2.1	Análise de subtipos	20
2.2.2	Análise dos olhos	21
2.2.3	Análise das imagens de ultrassom dos rins	21
2.3	Revisão da Literatura de Reconhecimento Facial para diagnóstico de doenças	22
2.3.1	Doença Arterial Coronariana	23
2.3.2	Acromegalia	23
2.3.3	Síndrome de Cushing	24
3	MÉTODO DO TRABALHO	25
3.1	Submissão do projeto ao conselho de Ética da Plataforma Brasil	25
3.2	Desenvolvimento de Aplicativo para Coleta de Dados	25
3.3	Coleta dos dados	25
3.4	Tratamento e análise dos dados coletados	26
3.5	Seleção de algoritmo de visão computacional	26
3.6	Desenvolvimento de aplicativo para disponibilizar o modelo desenvolvido	26
4	DESENVOLVIMENTO DO APLICATIVO DE COLETA	27
4.1	Especificação Técnica	27
4.1.1	Requisitos Funcionais	27

4.1.2	Requisitos Não-Funcionais	27
4.1.3	Descrição dos Casos de Uso	28
4.1.3.1	Descrição de Criar um usuário	28
4.1.3.2	Descrição de Realizar login	29
4.1.3.3	Descrição de Consultar paciente	29
4.1.3.4	Descrição de Criar um Paciente	30
4.1.3.5	Descrição de Editar Paciente	31
4.1.3.6	Descrição de Deletar paciente	32
4.2	Tecnologias Utilizadas	33
4.2.1	React Native	33
4.2.2	Expo Go	33
4.2.3	C#	33
4.2.4	.NET 8	34
4.2.5	AWS Lambda	34
4.2.6	AWS S3	34
4.2.7	AWS DynamoDB	35
4.3	Projeto e Implementação	35
4.3.1	Telas do Aplicativo	35
4.3.2	Diagrama de Arquitetura	37
4.3.3	Decisões de projeto	38
4.3.3.1	Aplicativo Mobile	38
4.3.3.2	API - Application Programming Interface	38
4.3.3.3	Banco de dados	39
4.3.4	Privacidade e Segurança dos dados	39
4.4	Teste do Aplicativo	40
4.5	Implantação	41
4.6	Resultados e Análises dos dados	42
5	DESENVOLVIMENTO DO SISTEMA DE DIAGNÓSTICO	47
5.1	Especificação Técnica	47
5.1.1	Requisitos Funcionais	47
5.1.2	Requisitos Não-Funcionais	47
5.1.3	Descrição dos Casos de Uso	48
5.1.3.1	Descrição de Criar um Paciente	48
5.1.3.2	Descrição de Editar Paciente	49
5.1.3.3	Descrição de Criar Diagnóstico	50
5.1.3.4	Descrição de Deletar Diagnóstico	51
5.2	Tecnologias Utilizadas	52
5.2.1	AWS Sagemaker	52
5.2.2	Python	52

5.2.3	MediaPipe	52
5.2.4	Scikit-learn	52
5.2.5	TensorFlow	53
5.3	Projeto e Implementação	53
5.3.1	Telas do Aplicativo	53
5.3.2	Diagrama de Arquitetura	56
5.3.3	Aplicativo de Diagnóstico	56
5.4	Tratamento e limpeza dos dados	57
5.4.1	Seleção e treino de modelos	57
5.5	Avaliação e Testes do Sistema	59
5.5.1	Abordagens com algoritmo XGBoost	59
5.5.2	Abordagens utilizando Redes Neurais	61
5.5.3	Avaliação dos modelos selecionados no conjunto de teste	62
5.5.4	Testes aplicativo de diagnóstico	63
6	CONSIDERAÇÕES FINAIS	65
6.1	Conclusões do Projeto de Formatura	65
6.2	Contribuições	65
6.3	Perspectivas de Continuidade	66
	REFERÊNCIAS	69
	APÊNDICES	71
	APÊNDICE A – BROCHURA DA PESQUISA SUBMETIDO PARA A PLATAFORMA BRASIL	73
	APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ES- CLARECIDO	85

1 Introdução

1.1 Definição

A doença renal crônica (DRC) é uma condição de longa duração que afeta os rins, órgãos que são responsáveis pela filtragem do sangue, removendo resíduos metabólicos, impurezas e regulando o equilíbrio de líquidos e pH do corpo. A DRC é caracterizada pela perda gradual e irreversível da função renal ao longo do tempo. Estima-se que aproximadamente dez milhões de brasileiros tenham a doença, dos quais 90 mil estão em diálise (estágio terminal da doença) ([MINISTÉRIO DA SAÚDE, 2019](#)), e os gastos com o tratamento da DRC, e de outras doenças relacionadas, correspondem a 12,97% das despesas em saúde ([ALCALDE, 2018](#)).

Os rins são os órgãos responsáveis pela remoção de resíduos metabólicos, hormônios e substâncias tóxicas do corpo, os rins das pessoas acometidas pela DRC não conseguem realizar adequadamente essa função. Os estágios da doença renal crônica são classificados de acordo com a taxa de filtração glomerular (TFG), que indica a capacidade dos rins de filtrar o sangue. Sintomas iniciais podem passar despercebidos, mas, conforme a DRC progride, podem surgir sintomas como fadiga, edema, pressão arterial elevada e alterações na frequência urinária. Complicações graves podem ocorrer, incluindo insuficiência renal, que exige tratamentos como diálise ou transplante renal ([MINISTÉRIO DA SAÚDE, 2022](#)).

1.2 Estágios da DRC

O primeiro estágio da DRC ocorre quando a TFG é superior a 90, o que indica que os rins estão funcionando adequadamente, porém apresentam sinais de danos leves. Alguns sintomas possíveis são pressão alta, presença de proteína na urina e inchaço nas pernas.

O segundo estágio ocorre quando o nível da TFG está entre 60 e 89, o que representa que os rins ainda estão funcionando, porém com sinais claros de danos. Os sintomas ainda são leves, e alguns pacientes podem não apresentá-los, caracterizando-se pelo acúmulo de fluidos, principalmente nas mãos e pés.

No terceiro estágio, quando o TFG está entre 30 e 59, os rins já apresentam danos sérios. Além dos sintomas dos primeiros 2 estágios, pode ser observada a anemia, e distúrbio mineral ósseo, caracterizado pela não-filtragem de minerais responsáveis pela manutenção óssea, como cálcio e fósforo, o que pode afetar vasos sanguíneos e o coração.

O quarto estágio é o último antes da falência renal, e ocorre quando os níveis de

TFG estão entre 15 e 29, e diversos sintomas podem ocorrer, como icterícia, fadiga, inchaços, náusea, câimbra ou problemas cardíacos. O tratamento pode consistir na hemodiálise, em que o sangue circula por uma máquina que filtra e remove resíduos, ou diálise peritoneal, em que a diálise é feita através do abdômen, e pode ser realizada em casa.

Por fim, o quinto estágio é caracterizado pela falência renal, quando a TFG é inferior a 15, indicando que os rins não funcionam mais. Neste caso, os resíduos metabólicos se acumulam no sangue e no corpo, gerando mudanças de coloração, ausência de urina, dificuldades respiratórias, e podendo levar ao óbito. Os tratamentos são a hemodiálise e o transplante renal.

O diagnóstico precoce é crucial para retardar a progressão da doença e gerenciar os sintomas. Medidas preventivas, como controle da pressão arterial, monitoramento dos níveis de glicose e adoção de um estilo de vida saudável, são fundamentais na gestão da doença renal crônica (BASTOS RACHEL BREGMAN, 2010).

1.3 Sintomas

A presença prolongada dos resíduos metabólicos no sangue gera alterações em diversos órgãos do corpo, e também na pele e face (MALKINA, 2023; LUPI et al., 2011). Dentre essas alterações, pode-se citar:

- Palidez (SHAIKH et al., 2023): Os rins dos pacientes, por não funcionarem corretamente, não conseguem produzir eritropoietina, hormônio responsável pelo estímulo à produção de glóbulos vermelhos, de maneira suficiente, o que leva à anemia. A maioria dos pacientes desenvolve anemia, que pode ocorrer no início da doença e piorar com o tempo.
- Mudanças de coloração da pele (AAD, 2019): Além da palidez, o acúmulo de toxinas no corpo, devido à sua não-filtração pelos rins, pode levar à diversas mudanças na coloração da pele, como manchas escuras, tonalidade cinza, cistos e manchas que se assemelha à espinhas, e pele amarelada, com inchaços e linhas profundas.
- Icterícia (LEVINE, 2022): A partir do estágio 4 da DRC, o paciente pode apresentar icterícia, incluindo o amarelamento do branco dos olhos.
- Inchaço do rosto (AAD, 2019): Os rins removem o excesso de fluido e sal do corpo. Quando eles não conseguem mais fazer isso, os fluidos e o sal se acumulam no corpo, que é responsável pelo inchaço.
- Xerose (AAD, 2019): Em casos nos quais a doença renal crônica está em estado terminal, necessitando de diálise e transplante de rim, é comum que a pele do paciente se torne muito seca, desenvolvendo rachaduras e escamas

- Coceira aguda (AAD, 2019): Esse é um sintoma comum da doença renal crônica avançada. A coceira pode variar de irritante a incapacitante, podendo afetar toda ou a maior parte da pele o tempo todo. É comum os pacientes apresentarem marcas de arranhões, feridas, ou inchaços.

Além desses sintomas na face, esta doença também pode afetar a aparência das mãos e dos pés (AAD, 2019). Além do inchaço e da presença de bolhas, pessoas com doença renal avançada podem desenvolver:

- Coloração branca na parte superior das unhas, e coloração normal ou marrom avermelhada na parte inferior (unhas de Lindsay ou meio-e-meio).
- Unhas pálidas.
- Faixas brancas percorrendo uma ou mais unhas (unhas de Muehrcke).

1.4 Motivação

As doenças não apenas afetam internamente, mas também se refletem nas características faciais. Essas características podem ser indicadores de diagnóstico, especialmente para condições endócrinas, metabólicas, genéticas e neuromusculares. Embora o desenvolvimento da tecnologia de reconhecimento facial tenha sido longo, sua aplicação na medicina clínica tem crescido rapidamente apenas nos últimos dez anos (QIANG et al., 2022).

Entre os trabalhos existentes na área, foram propostas abordagens para algumas doenças, entre elas podemos citar:

- Doenças Endócrinas e Metabólicas: Acromegalia (KONG et al., 2018), Síndrome de Cushing (WEI et al., 2019)
- Doenças Genéticas: Síndrome de Down (QIN et al., 2020), Síndrome de Turner (PAN et al., 2020)
- Doenças Neuromusculares: Paralisia Facial (STOREY et al., 2019)

Os projetos estudados apresentam uma ampla gama de métodos e algoritmos para realizar o diagnóstico através de reconhecimento facial. No entanto, não há nenhuma abordagem na literatura para a Doença Renal Crônica, que é a enfermidade que será estudada nesse projeto. Dessa forma, é muito importante identificar e realizar uma análise profunda de quais são as abordagens que mais se assemelham à doença estudada. Nesse contexto, as doenças endócrinas e metabólicas (Acromegalia e Síndrome de Cushing) são as que mais se aproximam do nosso caso de estudo.

1.5 Objetivo

O presente trabalho tem como objetivo desenvolver uma solução tecnológica integrada para aprimorar o registro, a organização e a análise de dados clínicos no contexto da nefrologia. Para isso, propõe-se a criação de um aplicativo capaz de facilitar o armazenamento e a gestão de informações clínicas e imagens dos pacientes, assegurando a integridade e a segurança dos dados. Paralelamente, busca-se projetar e treinar um modelo preditivo utilizando os dados coletados, com o propósito de oferecer suporte à tomada de decisão médica. Esse modelo visa identificar de forma precoce a Doença Renal Crônica (DRC), auxiliando nefrologistas no diagnóstico e na definição de estratégias de tratamento, com vistas a melhorar o prognóstico e a qualidade de vida dos pacientes.

1.6 Justificativa

O trabalho é importante dado o contexto atual da tendência de crescimento de inteligência artificial e suas aplicações, por diversos motivos. Ela preenche uma lacuna na pesquisa médica por não existir até o presente momento nenhum estudo realizado para o diagnóstico de DRC utilizando técnicas de inteligência artificial de reconhecimento facial, por meio de uma técnica acessível e não invasiva. Além disso, existe o potencial de identificar marcadores de risco no rosto não documentados previamente pela comunidade médica¹. Essa abordagem inovadora tem o potencial de melhorar significativamente o processo de detecção e auxílio ao diagnóstico, principalmente para locais com incidência maior de DRC ou com escassez de médicos especialistas.

1.7 Organização do Trabalho

Para uma melhor organização do trabalho, foi realizado uma divisão dele através de diversos capítulos. O capítulo 1 apresenta diversos conceitos importantes sobre a Doença Renal Crônica, com sua definição 1.1, seus estágios 1.2, seus principais indicadores nas faces dos pacientes 1.3, a motivação desse projeto 1.4, o objetivo dele 1.5 e sua justificativa 1.6.

No capítulo 2, desenvolve-se a fundamentação teórica do projeto, abordando de forma abrangente os conceitos fundamentais de Inteligência Artificial 2.1. O capítulo promove uma revisão crítica da literatura sobre o uso de reconhecimento facial na detecção de doenças, apresentando uma análise detalhada das diferentes abordagens metodológicas propostas para cada condição médica investigada. A revisão bibliográfica estrutura-se em

¹ "Second, our deep learning algorithm may have extracted some features that are strongly associated with CAD but are beyond human perception or understanding."(LIN et al., 2020)

duas vertentes principais: revisões sistemáticas da literatura 2.2 e estudos de caso focados em diagnósticos de doenças específicas 2.3.

O capítulo 3, por sua vez, traz a metodologia a ser utilizada ao longo do trabalho, citando em detalhes as fases de desenvolvimento do projeto, de submissão do projeto ao Conselho de Ética 3.1, desenvolvimento do aplicativo de coleta e a coleta em si 3.2 e 3.3, a análise dos dados 3.4, o desenvolvimento do modelo 3.5 e aplicativo de diagnóstico 3.6.

O capítulo 4 aborda o desenvolvimento do aplicativo de coleta, detalhando a especificação técnica 4.1 do projeto, com os requisitos funcionais 4.1.1 e não-funcionais 4.1.2, e a descrição dos casos de uso 4.1.3. São apresentadas as tecnologias utilizadas 4.2 e informações sobre a implementação do aplicativo 4.3: telas do sistema 4.3.1, diagrama de arquitetura 4.3.2, as principais decisões de projeto 4.3.3 e questões relacionadas à privacidade e segurança dos dados 4.3.4. Por fim, são descritos os testes realizados 4.4, o processo de implantação 4.5 e uma análise dos resultados obtidos 4.6.

O capítulo 5 trata do desenvolvimento do sistema de diagnóstico, que integra o modelo de inteligência artificial e o aplicativo de diagnóstico. Ele traz a especificação técnica 5.1, detalhando os requisitos funcionais 5.1.1, não funcionais 5.1.2, e casos de uso 5.1.3, além de descrever as tecnologias empregadas 5.2. A seção de projeto e implementação 5.3 explora as telas do aplicativo 5.3.1, o diagrama de arquitetura 5.3.2 e os detalhes da construção do aplicativo de diagnóstico 5.3.3. Finalmente, ele aborda o tratamento e limpeza dos dados 5.4 e uma avaliação detalhada 5.5 dos modelos treinados.

Por fim, são feitas as considerações finais do trabalho no capítulo 6, as conclusões finais do projeto, suas contribuições para o mundo 6.2 e perspectivas de continuidade 6.3.

2 Fundamentação Teórica

2.1 Inteligência Artificial

A Inteligência Artificial (IA) representa um campo multidisciplinar que busca desenvolver agentes capazes de realizar tarefas que tradicionalmente requerem inteligência humana. De acordo com (NORVIG; RUSSELL, 2013), a IA pode ser definida através de quatro abordagens fundamentais que descrevem diferentes perspectivas de desenvolvimento e compreensão dos sistemas inteligentes. Essas abordagens compreendem: sistemas que pensam como humanos, focando na simulação dos processos cognitivos mentais; sistemas que agem como humanos, priorizando comportamentos que mimetizam a inteligência humana; sistemas que pensam racionalmente, baseados em modelos lógicos e algorítmicos de raciocínio; e sistemas que agem racionalmente, orientados pela busca da otimização e eficiência decisória.

Na perspectiva moderna, a IA concentra-se principalmente no desenvolvimento de agentes racionais - entidades computacionais sofisticadas que podem perceber seu ambiente através de sensores complexos e atuar sobre esse ambiente através de atuadores precisos, sempre buscando maximizar seu desempenho esperado. Esses agentes são projetados para tomar decisões inteligentes, adaptando-se dinamicamente a diferentes contextos e aprendendo continuamente com suas interações e experiências. A racionalidade, neste contexto, não se limita apenas à execução de tarefas, mas envolve uma capacidade sofisticada de análise, inferência e tomada de decisão que considera múltiplas variáveis e possíveis cenários.

Os avanços contemporâneos em IA têm demonstrado uma capacidade crescente de desenvolver sistemas que não apenas replicam comportamentos humanos, mas que são capazes de superar limitações cognitivas humanas em domínios específicos, como processamento de grandes volumes de dados, reconhecimento de padrões complexos e tomada de decisões em ambientes de alta incerteza.

Dentre as diversas abordagens e técnicas desenvolvidas no campo da Inteligência Artificial, o aprendizado de máquina emerge como um dos domains mais promissores e transformadores, representando uma evolução significativa na capacidade dos sistemas computacionais de adquirir conhecimento e desenvolver inteligência adaptativa. Se os primeiros conceitos de IA se concentravam em criar sistemas capazes de simular processos cognitivos humanos, o aprendizado de máquina inaugura uma nova perspectiva, onde os sistemas não apenas replicam comportamentos, mas são capazes de aprender autonomamente a partir de dados, identificar padrões complexos e realizar generalizações preditivas.

Neste contexto, o aprendizado supervisionado se destaca como uma das modalidades fundamentais, oferecendo um arcabouço metodológico robusto para o desenvolvimento de modelos computacionais capazes de transformar dados históricos em conhecimento preditivo, bridging o gap entre a concepção teórica de inteligência artificial e suas aplicações práticas mais impactantes.

2.1.1 Aprendizado Supervisionado

O aprendizado supervisionado constitui uma das técnicas fundamentais e mais estabelecidas no campo do aprendizado de máquina, representando um paradigma no qual sistemas computacionais desenvolvem capacidades preditivas a partir de conjuntos de dados previamente rotulados e estruturados. Neste modelo de aprendizagem, o algoritmo recebe um conjunto de treinamento composto por pares de entrada-saída, onde cada instância de entrada está precisamente associada à sua respectiva saída desejada ou classe predeterminada, permitindo que o sistema aprenda padrões e relações subjacentes que possibilitam generalizações para novos dados não vistos.

Esta abordagem se mostra particularmente eficaz em cenários onde existe um conjunto significativo de dados históricos rotulados, especialmente quando a relação entre entradas e saídas apresenta consistência e o problema pode ser rigorosamente definido através de exemplos práticos. O processo de aprendizado supervisionado envolve essencialmente duas etapas principais: a fase de treinamento, na qual o modelo aprende a mapear entradas para saídas corretas, e a fase de inferência, onde o modelo já treinado pode fazer previsões ou classificações para novos conjuntos de dados.

Os algoritmos de aprendizado supervisionado podem ser categorizados em dois tipos principais: algoritmos de classificação e algoritmos de regressão. Nos problemas de classificação, o objetivo é prever uma classe ou categoria discreta, como identificar se uma imagem contém um gato ou um cachorro, ou diagnosticar uma condição médica específica. Já nos problemas de regressão, o foco está em prever um valor contínuo, como estimar preços de imóveis, prever temperaturas futuras ou calcular o tempo de vida útil de um equipamento.

Entre os algoritmos supervisionados mais prominentes estão as Árvores de Decisão, Regressão Logística, Máquinas de Vetores de Suporte (SVM), Redes Neurais Feedforward, Florestas Aleatórias e XGBoost, cada um com características e adequações específicas para diferentes tipos de problemas e estruturas de dados. A escolha do algoritmo mais apropriado depende de múltiplos fatores, como a natureza dos dados, a complexidade do problema, o tamanho do conjunto de treinamento e os requisitos computacionais disponíveis.

2.1.2 Classificação binária

A classificação binária representa um paradigma fundamental no aprendizado supervisionado, caracterizado pela tarefa de categorizar instâncias em uma de duas classes mutuamente exclusivas e exaustivas. Este modelo computacional de decisão simplificada possui aplicações ubíquas e críticas em diversos domínios científicos e tecnológicos, transcendendo múltiplos setores como saúde, finanças, segurança digital, análise de risco e processamento inteligente de informações.

Métricas de avaliação desempenham um papel crucial na validação e seleção de modelos de classificação binária. Métricas como acurácia, precisão, recall, F1-score e a curva ROC (Receiver Operating Characteristic) fornecem insights abrangentes sobre o desempenho preditivo, permitindo uma caracterização multidimensional da qualidade do classificador.

2.1.3 Receiver Operating Characteristics

A curva Característica de Operação do Receptor, em inglês ROC curve (Receiver Operating Characteristic), é uma ferramenta gráfica poderosa que representa a relação entre a sensibilidade e a especificidade de um classificador binário em diferentes pontos de corte. A sensibilidade, também conhecida como taxa de verdadeiros positivos (TVP), representa a proporção de casos positivos corretamente identificados pelo modelo. Por outro lado, a especificidade representa a proporção de casos negativos corretamente identificados. Na curva ROC, analisa-se a TVP no eixo y contra a taxa de falsos positivos (TFP = 1 - especificidade) no eixo x.

O gráfico resultante forma uma curva que se estende do ponto (0,0) ao ponto (1,1). Um classificador perfeito produziria uma curva que passa pelo ponto (0,1), indicando 100% de sensibilidade e 100% de especificidade. Na prática, a maioria dos classificadores produz curvas que se arqueiam entre estes extremos. A diagonal do gráfico ($y = x$) representa o desempenho de um classificador aleatório.

A área sob a curva ROC (AUC-ROC) fornece uma medida agregada do desempenho do classificador. Um valor de AUC de 1.0 representa um classificador perfeito, enquanto um valor de 0.5 indica um classificador que não é melhor que o acaso. Valores típicos em aplicações reais variam entre 0.5 e 1.0, onde valores mais altos indicam melhor desempenho.

A curva ROC é especialmente útil em situações onde as classes são desbalanceadas, pois sua interpretação não é afetada pela proporção entre as classes. Além disso, diferentes pontos na curva representam diferentes compromissos entre sensibilidade e especificidade, permitindo que os usuários escolham um ponto de operação que melhor atenda às necessidades específicas de sua aplicação.

Por exemplo, em um contexto médico de diagnóstico de doenças graves, pode-se

preferir um ponto de operação com alta sensibilidade (mesmo que isso resulte em mais falsos positivos) para minimizar o risco de não detectar casos positivos reais. Por outro lado, em um sistema de detecção de fraudes bancárias, pode-se optar por um ponto com maior especificidade para reduzir o número de alertas falsos que precisariam ser investigados manualmente.

A construção da curva ROC envolve variar o limiar de classificação do modelo e calcular os pares de TVP e TFP correspondentes. Para classificadores que produzem probabilidades ou scores contínuos, cada valor possível de limiar gera um ponto na curva. A suavidade da curva resultante geralmente depende do número de instâncias no conjunto de teste e da granularidade dos scores produzidos pelo modelo.

A análise ROC também permite a comparação direta entre diferentes modelos: aquele com maior AUC geralmente é considerado superior. No entanto, é importante notar que em algumas aplicações, pode-se estar mais interessado no desempenho do modelo em regiões específicas da curva ROC, em vez do desempenho global medido pela AUC.

2.2 Revisão da Literatura para DRC

Para um melhor entendimento sobre o contexto do projeto e as possíveis abordagens, foi realizado um estudo na literatura sobre o uso de inteligência artificial no contexto de DRC. Inicialmente, foram analisados três artigos que usam diferentes abordagens para diagnosticar ou classificar a doença renal crônica.

2.2.1 Análise de subtipos

O artigo "Identifying subtypes of chronic kidney disease with machine learning: development, internal validation and prognostic validation using linked electronic health records in 350,067 individuals" (DASHTBAN et al., 2023) explora uma nova classificação para a doença renal crônica (DRC) com base em aprendizado de máquina (ML), buscando melhorar a previsão de resultados e intervenções. A pesquisa analisou o prontuário e exames eletrônicos de 350.067 indivíduos com DRC (incidente e prevalente) e utilizou sete métodos de ML para identificar cinco subtipos: 1) início precoce, 2) início tardio, 3) câncer, 4) metabólico e 5) cardiometabólico.

Os subtipos mostraram diferenças significativas em mortalidade, internações hospitalares e desenvolvimento de novas doenças crônicas. O subtipo cardiometabólico apresentou maior risco de mortalidade e internações, enquanto o subtipo de início precoce mostrou menor risco. A análise também revelou uma alta carga de medicação, especialmente nos subtipos cardiometabólico e metabólico.

Apesar de não tratar do diagnóstico da DRC em si, este estudo enfatiza a necessidade

de validação externa dos subtipos e sugere que essa nova classificação pode informar diretrizes clínicas e intervenções personalizadas. Destaca-se que as classificações tradicionais de DRC não capturam a complexidade da doença, seus fatores de risco e os impactos dos tratamentos.

2.2.2 Análise dos olhos

O artigo "An efficient machine learning approach to nephrology through iris recognition" (DIVYA et al., 2021) explora o uso da iridologia para diagnosticar doenças renais através da análise da íris do olho. A iridologia sugere que alterações na íris, como padrões, formas e cores, refletem o estado dos órgãos internos, permitindo a identificação de condições como inflamação e danos em diferentes partes do corpo. Especificamente, o estudo foca na detecção de problemas renais, como insuficiência renal e diabetes, analisando mudanças nos tecidos e na coloração da íris.

A abordagem utiliza algoritmos de aprendizado de máquina, como o Modelo Oculto de Markov (HMM), para processar imagens da íris e identificar padrões associados a doenças renais. O sistema segmenta e transforma as imagens para detectar anomalias relacionadas aos rins, analisando os nervos conectados aos olhos e aos órgãos internos. O estudo avaliou 100 imagens de íris, classificando-as entre casos de diabetes, cálculos renais, insuficiência renal e olhos normais. Com uma precisão aproximada de 80%, o sistema mostrou ser promissor para detectar problemas renais com base nas imagens da íris.

2.2.3 Análise das imagens de ultrassom dos rins

O estudo "A Deep Learning Approach for Kidney Disease Recognition and Prediction through Image Processing" (KUMAR et al., 2023) propõe um modelo de aprendizado de máquina, especificamente uma rede neural profunda combinada com lógica fuzzy (FDNN), para detecção e prognóstico de DRC. O modelo foi projetado para analisar imagens renais e identificar padrões associados à doença, com uma precisão de 99,23%, superando as abordagens tradicionais. O processo inclui a normalização das imagens renais, extração de características relevantes e posterior classificação com base nos algoritmos de aprendizado profundo.

O modelo FDNN oferece avanços significativos em termos de sensibilidade, precisão e métricas de F-measure em comparação com métodos convencionais de mineração de dados. A análise de dados clínicos e laboratoriais auxilia na identificação precoce dos estágios da DRC, permitindo um diagnóstico mais assertivo. Além disso, a integração de técnicas como redes neurais convolucionais (CNN) para processamento de imagens e a modelagem fuzzy reduzem a dependência de especialistas médicos, automatizando grande parte do processo diagnóstico. O estudo sugere que o uso de modelos híbridos

pode revolucionar a detecção precoce de DRC e outras doenças crônicas, oferecendo uma ferramenta robusta para médicos e sistemas de saúde.

2.3 Revisão da Literatura de Reconhecimento Facial para diagnóstico de doenças

Para um melhor entendimento desse tema, foram estudados dois artigos de revisão de literatura sobre o uso de reconhecimento facial para o diagnóstico de doenças. O primeiro deles “Review on Facial-Recognition-Based Applications in Disease Diagnosis” (QIANG et al., 2022) realizou uma análise da literatura para identificar os principais estudos na área de reconhecimento facial aplicados ao diagnóstico de doenças. Este estudo destacou práticas recomendadas para a captura de fotos, incluindo padronização da expressão facial dos pacientes, posicionamento da câmera e condições ambientais. Além disso, foram apresentados os principais algoritmos utilizados no diagnóstico e as circunstâncias em que cada um deles é mais apropriado.

O segundo artigo “Facial Recognition Intensity in Disease Diagnosis Using Automatic Facial Recognition” (WU et al., 2021) realizou um compilado de 141 estudos que utilizaram o reconhecimento facial para o diagnóstico de doenças, e uma análise estatística sobre os fatores que influenciam a eficiência dos modelos de inteligência artificial.

Os autores deste estudo definiram o conceito de Facial Recognition Intensity (FRI), um indicador que descreve a diferença entre as feições faciais da doença estudada e o grupo de controle. O FRI é calculado pela fórmula $N_f * P_{max}$, onde N_f representa o número de fenótipos faciais relevantes para a doença, ou seja, o número de características faciais que são afetadas pela doença, e P_{max} representa a penetrância máxima destas características, isto é, o máximo percentual de indivíduos acometidos pela doença que apresentam esta característica. Por exemplo, uma doença como Síndrome de Down apresenta ao todo 9 fenótipos, e uma penetrância máxima de 100%, o que resulta em um FRI de 9.

O estudo em questão identificou que há uma forte correlação entre o FRI e a eficiência do modelo. Entretanto, outros fatores como resolução da imagem utilizada no treino, tamanho do dataset utilizado para treino e algoritmo de inteligência artificial utilizado, não apresentam uma correlação tão significativa. Além disso, para alcançar uma eficiência maior no diagnóstico de doenças que apresentam um FRI baixo, o tamanho do dataset utilizado para treino deve ser significativamente maior do que para doenças com FRI elevado. Estas informações são de extrema relevância, visto que impactam como a abordagem para este projeto deve ser realizada.

2.3.1 Doença Arterial Coronariana

Outro artigo analisado foi o “Feasibility of using deep learning to detect coronary artery disease based on facial photo” (LIN et al., 2020), que analisa a utilização de um algoritmo de reconhecimento facial baseado em deep learning para a detecção da doença arterial coronariana (CAD).

Além do treinamento do modelo de inteligência artificial para realizar o diagnóstico, os autores também realizaram diversas análises sobre quais eram os fatores que mais impactavam a eficiência deste. A primeira análise comparava o diagnóstico quando eram fornecidas ao modelo, além das fotos faciais, informações sobre os pacientes tais como idade, sexo, sedentarismo, tempo de trabalho, utilização de álcool ou cigarro, entre outros, e quando somente eram fornecidas ao modelo as fotos faciais. O resultado observado era que, a presença de informações adicionais sobre os pacientes não gerava uma diferença significativa na eficiência do modelo. Entretanto é importante citar que esta conclusão pode variar conforme a doença estudada.

A segunda avaliação realizada foi sobre quais eram os fenótipos faciais mais determinantes para o diagnóstico. Para isso, foi comparada a eficiência do diagnóstico utilizando as fotos sem alteração, e utilizando as fotos com determinadas partes do rosto sendo cobertas, como olhos, nariz, ou região bucal. Com base nos resultados, foi construído um mapa de calor do rosto, indicando quais as características mais importantes na eficiência do diagnóstico.

O presente projeto pretende seguir os passos deste último artigo, realizando análises sobre quais as informações mais impactantes para o modelo durante o processo de diagnóstico utilizando imagens e dados de saúde dos pacientes.

2.3.2 Acromegalia

Outro estudo relevante é “Automatic Detection of Acromegaly From Facial Photographs Using Machine Learning Methods” (KONG et al., 2018). Nele, foram empregados diversos algoritmos populares de aprendizado de máquina para analisar um conjunto de dados composto por 527 pacientes diagnosticados com acromegalia e 596 indivíduos saudáveis como grupo de controle.

Inicialmente, utilizou-se a biblioteca de Python OpenCV para identificar e dimensionar retângulos delimitadores dos rostos, seguido pelo recorte e redimensionamento para as mesmas dimensões. A partir dos rostos detectados, foram extraídas as localizações de pontos de referência faciais, potenciais indicadores clínicos. Para aprimorar o desempenho, adotou-se a técnica de frontalização para sintetizar vistas faciais frontais.

Diversos métodos de aprendizado de máquina foram empregados para identificar de maneira automática a acromegalia com base nas fotografias faciais, pontos de referência

faciais e faces frontais sintetizadas, incluindo LM, KNN, SVM, RT, CNN e EM. A avaliação dos modelos treinados foi realizada em um conjunto de dados separado, no qual metade dos casos foi diagnosticada como acromegalia por meio do teste de supressão do hormônio do crescimento.

Em suma, o artigo propõe um sistema de diagnóstico prático e automático para acromegalia, que pode permitir aos profissionais médicos e pacientes acompanhar de forma proativa as mudanças faciais e detectar a acromegalia com antecedência. Isso, por sua vez, facilita os tratamentos e aumenta a probabilidade de prevenir complicações irreversíveis da secreção excessiva do hormônio do crescimento.

2.3.3 Síndrome de Cushing

Por fim, outro artigo relevante para o projeto é “Deep-Learning Approach to Automatic Identification of Facial Anomalies in Endocrine Disorders” (WEI et al., 2019). Nele, foram coletadas imagens faciais de pacientes com hipercortisolismo (Síndrome de Cushing) e técnicas de augmentation foram utilizadas para ampliar esse dataset com amostras negativas adicionais de bancos de dados públicos. Um modelo com uma rede de aprendizado profundo pré-treinada foi construída para identificar automaticamente a hipersecreção com base em mudanças faciais características. O desempenho dele foi comparado com o de especialistas endócrinos e uma investigação foi realizada para identificar as regiões da face que o modelo considerava mais importante para o diagnóstico dessa doença.

Ao final desse estudo, o desempenho do algoritmo de inteligência artificial superou a dos especialistas endócrinos consultados. Além disso, as regiões de interesse identificadas pela máquina foram principalmente as mesmas que aquelas nas quais os humanos se concentraram. Por fim, o estudo conclui que o modelo de aprendizado profundo aprendeu as características faciais com base apenas em dados rotulados, sem necessidade de possuir conhecimento médico prévio. O modelo final tem o potencial de auxiliar no diagnóstico e acompanhamento da Síndrome de Cushing.

3 Método do trabalho

Para o melhor andamento do projeto, foi necessário uma divisão dele em diversas fases. Elas estão listadas a seguir:

3.1 Submissão do projeto ao conselho de Ética da Plataforma Brasil

Na etapa inicial do projeto, foi elaborada a documentação para submissão ao Comitê de Ética da Plataforma Brasil (Brochura da Pesquisa), visando viabilizar a coleta de dados. Esta documentação inclui a descrição das imagens a serem capturadas dos pacientes, as informações que serão coletadas, o público-alvo da pesquisa e o local onde essas informações serão coletadas. Além disso, também foi elaborado o Termo de Consentimento Livre e Esclarecido que será preenchido pelos participantes da pesquisa. Ambos os documentos estão no apêndice.

3.2 Desenvolvimento de Aplicativo para Coleta de Dados

Para otimizar a coleta de dados, o grupo desenvolveu um aplicativo mobile dedicado a essa finalidade. Os requisitos do aplicativo foram definidos com base nas informações do questionário e nas fotos a serem coletadas. Após essa etapa, o desenvolvimento foi iniciado, com as responsabilidades distribuídas entre o aplicativo e a API. Concluído o desenvolvimento, o sistema foi submetido a testes, com a participação de um grupo selecionado de usuários para avaliar sua funcionalidade e desempenho. Após as correções necessárias, o aplicativo foi disponibilizado aos profissionais de saúde, tornando-se uma ferramenta essencial para a próxima etapa do projeto.

3.3 Coleta dos dados

Após a autorização do Comitê de Ética e do centro de coleta de dados, o levantamento de informações dos pacientes com Doença Renal Crônica teve início. Ela aconteceu no HCFMB-UNESP - Hospital das Clínicas da Faculdade de Medicina de Botucatu da Universidade Estadual Paulista e foram coletadas fotografias faciais e das mãos dos pacientes. Além disso, foi preenchido um questionário de modo a obter informações de saúde dos pacientes. Essa etapa foi realizada com o auxílio de duas técnicas de enfermagem, que receberam acesso ao aplicativo desenvolvido.

3.4 Tratamento e análise dos dados coletados

Sob posse dos dados dos pacientes, o grupo realizou diversas análises acerca deles. Essa etapa permitiu uma compreensão mais aprofundada do perfil dos participantes da pesquisa e possibilitou a identificação de eventuais desequilíbrios nos dados, que poderiam impactar o desempenho do modelo a ser desenvolvido. Informações como idade, sexo e histórico familiar de doenças são muito importantes nessa análise, fornecendo insights valiosos sobre os segmentos da população mais afetados pela Doença Renal Crônica.

3.5 Seleção de algoritmo de visão computacional

Concluída a fase de análise dos dados coletados, iniciou-se o desenvolvimento e a avaliação de diferentes algoritmos de inteligência artificial. Durante essa etapa, foram testados diversos modelos e combinações de algoritmos para identificar aqueles que apresentavam o melhor desempenho na detecção de doença renal crônica. Após a definição do algoritmo final para o diagnóstico, seu desempenho foi avaliado utilizando o grupo de dados reservado exclusivamente para testes, garantindo uma análise robusta de sua eficácia.

3.6 Desenvolvimento de aplicativo para disponibilizar o modelo desenvolvido

Na etapa final, foi desenvolvido um aplicativo mobile direcionado principalmente a médicos nefrologistas, público-alvo do projeto. Por meio de um token fornecido pelo grupo, os usuários podem se cadastrar no aplicativo e utilizar o modelo treinado para calcular a probabilidade de um paciente apresentar Doença Renal Crônica (DRC), com base nos dados clínicos e nas fotos fornecidas pelo paciente.

4 Desenvolvimento do Aplicativo de Coleta

O aplicativo de coleta de dados otimizou o gerenciamento do grande volume de informações, reduzindo significativamente os erros comuns em processos manuais. Além disso, a solução acelerou o ciclo de desenvolvimento ao viabilizar uma integração eficiente e fluida com a pipeline do aplicativo de diagnóstico.

4.1 Especificação Técnica

4.1.1 Requisitos Funcionais

Pode-se elencar os seguintes requisitos funcionais do sistema:

1. O sistema deve permitir o cadastro de novos auxiliares de coleta que apresentem um token de criação válido;
2. O auxiliar de coleta cadastrado consegue realizar login no aplicativo utilizando suas credenciais;
3. O sistema deve permitir que os auxiliares de coleta cadastrem novos pacientes para salvar seus dados e suas fotos no banco de dados;
4. O sistema deve permitir que os auxiliares de coleta acessem e alterem os dados dos pacientes cadastrados por eles;
5. O sistema deve permitir que os auxiliares de coleta apaguem todos os dados dos pacientes cadastrados por eles.

4.1.2 Requisitos Não-Funcionais

Entre os requisitos não-funcionais do sistema, pode-se citar:

1. Security - devem ser adotadas políticas de segurança para proteger o sistema e também os dados sensíveis dos pacientes, garantindo total privacidade;
2. Disponibilidade - sistema deve apresentar disponibilidade de 99,9
3. Baixo tempo de resposta - sistema deve apresentar tempo de resposta médio abaixo de 1 segundo.

4.1.3 Descrição dos Casos de Uso

Ao todo, o aplicativo de coleta possui 6 casos de uso:

1. Criar usuário (4.1.3.1)
2. Realizar login (4.1.3.2)
3. Consultar paciente (4.1.3.3)
4. Cadastrar paciente (4.1.3.4)
5. Editar paciente (4.1.3.5)
6. Deletar Paciente (4.1.3.6)

O diagrama de casos de uso não foi incluído neste relatório, pois seria redundante, dado que há apenas um ator responsável por todos os casos mencionados.

4.1.3.1 Descrição de Criar um usuário

Nome: Criar um usuário.

Descrição: Cria um usuário para acessar o aplicativo.

Evento Iniciador: Acionamento do botão "Primeiro Acesso", na tela inicial do aplicativo.

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Aplicativo baixado no dispositivo móvel.

Sequência de Eventos:

1. Sistema solicita o preenchimento dos dados cadastrais do usuário: nome completo, email, celular, senha, e token de criação.
2. Ator fornece os dados solicitados.
3. Ator clica no botão "Criar Usuário".
4. Sistema valida os dados fornecidos.
5. Sistema cria o usuário, e mostra a tela de login.
6. Fim do Caso de Uso.

Pós-Condição: Usuário criado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Dado inserido é inválido (passo 4)
 - a) Sistema informa quais dados estão incorretos e retorna ao passo 1.
2. Ator cancela o cadastro (passo 2)
 - a) Sistema cancela o cadastro e encerra o caso de uso.

4.1.3.2 Descrição de Realizar login

Nome: Realizar login.

Descrição: Realizar login para acessar o sistema.

Evento Iniciador: Abrir aplicativo.

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Ator possui um usuário cadastrado no sistema.

Sequência de Eventos:

1. Sistema solicita o preenchimento das credenciais do usuário: email e senha.
2. Ator fornece os dados solicitados.
3. Ator clica no botão "Fazer Login".
4. Sistema valida os dados fornecidos.
5. Sistema redireciona o ator para a tela de listagem dos pacientes cadastrados.
6. Fim do Caso de Uso.

Pós-Condição: Ator autenticado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Credenciais inseridas são inválidas (passo 4)
 - a) Sistema informa que credenciais inseridas são inválidas e retorna ao passo 1.

4.1.3.3 Descrição de Consultar paciente

Nome: Consultar paciente.

Descrição: Consultar paciente para visualizar seus dados e fotos cadastrados.

Evento Iniciador: Ator seleciona barra de "Pesquisar".

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator escreve nome do paciente que deseja consultar.
2. Sistema realiza uma filtragem dos pacientes que apresentam o nome digitado.
3. Ator clica no paciente que deseja consultar.
4. Ator é redirecionado para a tela de perfil do paciente.
5. Fim do Caso de Uso.

Pós-Condição: Ator na tela de perfil do paciente.

Fluxos alternativos: Não Contém.

Exceções: Não Contém

4.1.3.4 Descrição de Criar um Paciente

Nome: Criar um paciente.

Descrição: Criar um novo paciente, cadastrando seus dados e fotos.

Evento Iniciador: Acionar botão "+" na tela de listagem de pacientes.

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Sistema solicita o preenchimento dos dados do paciente (nome, data de nascimento, sexo, se possui hipertensão ou diabetes e o nível de creatinina).
2. Ator fornece os dados solicitados.
3. Ator clica no botão "Criar paciente".
4. Sistema valida os dados fornecidos.
5. Sistema cria o paciente e redireciona o ator para a tela de perfil do paciente criado.
6. Ator clica no botão "Fotos".
7. Sistema redireciona o ator à seção de envio de fotos.

8. Sistema solicita o envio das quatro fotos do paciente (facial frontal, facial em perfil, facial superior e foto das mãos)
9. Ator submete as fotos
10. Sistema valida se todas as fotos foram fornecidas
11. Fim do Caso de Uso.

Pós-Condição: Paciente cadastrado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Credenciais inseridas são inválidas (passo 4)
 - a) Sistema informa que dados inseridos são inválidos e retorna ao passo 1.
2. Ator cancela o cadastro (passo 3)
 - a) Sistema cancela o cadastro e encerra o caso de uso.

4.1.3.5 Descrição de Editar Paciente

Nome: Editar Paciente.

Descrição: Editar dados e fotos do paciente.

Evento Iniciador: Entrar no perfil do paciente.

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator clica no botão com ícone de lápis.
2. Sistema redireciona o ator para a tela de edição de paciente.
3. Sistema informa os dados atuais do paciente, e solicita o preenchimento dos novos dados (nome, data de nascimento, sexo, se possui hipertensão ou diabetes e o nível de creatinina).
4. Ator altera os dados do paciente.
5. Ator clica no botão "Editar paciente".
6. Sistema valida os dados fornecidos.

7. Sistema edita o paciente e redireciona o ator para a tela de perfil do paciente.
8. Ator clica no botão "Fotos".
9. Sistema redireciona o ator à seção de envio de fotos.
10. Ator exclui as fotos desejadas clicando nas lixeiras correspondentes
11. Sistema exclui as fotos selecionadas pelo Ator
12. Sistema solicita o envio das fotos excluídas
13. Ator submete as fotos solicitadas
14. Sistema valida se todas as fotos foram fornecidas
15. Fim do Caso de Uso.

Pós-Condição: Paciente atualizado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Dados inseridos são inválidos (passo 5)
 - a) Sistema informa que dados inseridos são inválidos e retorna ao passo 2.
2. Ator cancela o cadastro (passo 3)
 - a) Sistema cancela a edição e encerra o caso de uso.

4.1.3.6 Descrição de Deletar paciente

Nome: Deletar paciente

Descrição: Deletar todos os dados para o paciente

Evento Iniciador: Entrar no perfil do paciente.

Ator: Médicos e enfermeiros responsáveis pela coleta.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator clica no botão com ícone de lixeira.
2. Sistema solicita confirmação da deleção.
3. Ator clica no botão "Confirmar".

4. Sistema deleta todos os dados para esse paciente.
5. Sistema redireciona o ator para tela de listagem de pacientes.
6. Fim do Caso de Uso.

Pós-Condição: Paciente deletado do sistema.

Fluxos alternativos: Não contém.

Exceções:

1. Ator cancela a deleção (passo 3)
 - a) Sistema cancela a deleção e encerra o caso de uso.

4.2 Tecnologias Utilizadas

4.2.1 React Native

Para o desenvolvimento do aplicativo, foi escolhido o framework React Native. Criado pelo Facebook, React Native permite a construção de interfaces nativas utilizando JavaScript, aproveitando componentes nativos para garantir um desempenho próximo ao de aplicativos nativos para Android e iOS. Sua simplicidade no desenvolvimento e a compatibilidade com ambas as plataformas tornaram-no a escolha ideal, oferecendo uma solução eficiente e de fácil manutenção para o projeto.

4.2.2 Expo Go

Foi também utilizado o Expo Go para o desenvolvimento do aplicativo, um ambiente de execução que simplifica o processo de desenvolvimento e teste de aplicativos criados com React Native. Ele oferece uma série de ferramentas e bibliotecas pré-configuradas, permitindo aos desenvolvedores visualizar, depurar e testar os aplicativos diretamente em dispositivos móveis ou emuladores, sem a necessidade de configurações complexas de ambiente nativo. Essa facilidade de uso torna o processo de desenvolvimento mais ágil e eficiente.

4.2.3 C#

Para o backend do aplicativo, foi escolhida a linguagem C#. C# é uma linguagem moderna, orientada a objetos e com tipagem forte e estática, desenvolvida pela Microsoft. Amplamente utilizada em aplicativos empresariais, jogos e sistemas embarcados, C# se destaca por sua versatilidade, alto desempenho e suporte a múltiplas plataformas, incluindo

Windows, web e dispositivos móveis. Seus principais benefícios incluem a tipagem forte, que oferece maior segurança ao detectar erros em tempo de compilação, e recursos avançados como LINQ e `async/await`, que facilitam o desenvolvimento de soluções escaláveis, seguras e de alta qualidade.

4.2.4 .NET 8

O .NET 8, a versão mais recente da plataforma de desenvolvimento unificada da Microsoft, complementa a escolha do C# para o backend do aplicativo, proporcionando uma base robusta para a criação de aplicações modernas, escaláveis e de alto desempenho. Para Web APIs, o .NET 8 oferece suporte avançado a tecnologias como Minimal API, autenticação com JWT e OAuth, e integração com OpenAPI/Swagger para documentação automática, além de otimizações significativas de desempenho. Sua compatibilidade com ambientes de contêiner e integração com serviços em nuvem reforçam sua flexibilidade, tornando-o uma escolha ideal para projetos que exigem confiabilidade, segurança e escalabilidade, alinhando-se perfeitamente às necessidades do desenvolvimento backend do aplicativo.

4.2.5 AWS Lambda

Para a hospedagem do backend, foi escolhida a AWS Lambda, um serviço de computação sem servidor da Amazon Web Services (AWS). Este serviço permite a execução de código em resposta a eventos, gerenciando automaticamente os recursos de computação necessários, o que elimina a complexidade de gerenciar servidores. A solução oferece alta flexibilidade e escalabilidade, facilitando a automação de tarefas, criação de APIs e processamento de dados em grande escala, sem a necessidade de infraestrutura dedicada. Além de sua facilidade de manutenção e escalabilidade, o AWS Lambda se destaca pelo baixo custo, tornando-se uma escolha ideal para atender ao crescimento e à demanda do aplicativo de forma eficiente e econômica.

4.2.6 AWS S3

Para o armazenamento em nuvem das fotos coletadas dos pacientes, foi escolhido o Amazon S3 (Simple Storage Service). O S3 é uma solução de armazenamento de objetos da AWS, altamente escalável e durável, ideal para armazenar dados de diversos tipos, como imagens, vídeos e arquivos de grandes volumes. Com recursos de segurança avançados, alta disponibilidade e integração nativa com outros serviços da AWS, o S3 oferece uma solução confiável e eficiente, tornando-se a escolha perfeita para atender às necessidades de armazenamento do aplicativo.

4.2.7 AWS DynamoDB

Por fim, para o armazenamento dos dados clínicos dos pacientes, foi escolhido o AWS DynamoDB, um banco de dados NoSQL gerenciado, projetado para alto desempenho, escalabilidade e baixa latência. Ideal para aplicativos que demandam grande volume de leituras e gravações, como sistemas de e-commerce, jogos e análises em tempo real, o DynamoDB oferece suporte a indexação, consultas flexíveis e integração com outros serviços da AWS, como Lambda e S3. Essa combinação torna o DynamoDB uma solução robusta e eficiente, permitindo uma gestão ágil e escalável dos dados clínicos na nuvem.

4.3 Projeto e Implementação

4.3.1 Telas do Aplicativo

As figuras a seguir apresentam as telas do aplicativo, juntamente com suas respectivas funcionalidades. A primeira delas é a tela de login, mostrada na Figura 1, onde o técnico de enfermagem pode inserir seus dados para acessar o sistema.

Caso o usuário ainda não tenha uma conta, é necessário acessar a tela de primeiro acesso, onde, utilizando o token fornecido pelo grupo, será possível cadastrar um novo usuário no aplicativo.

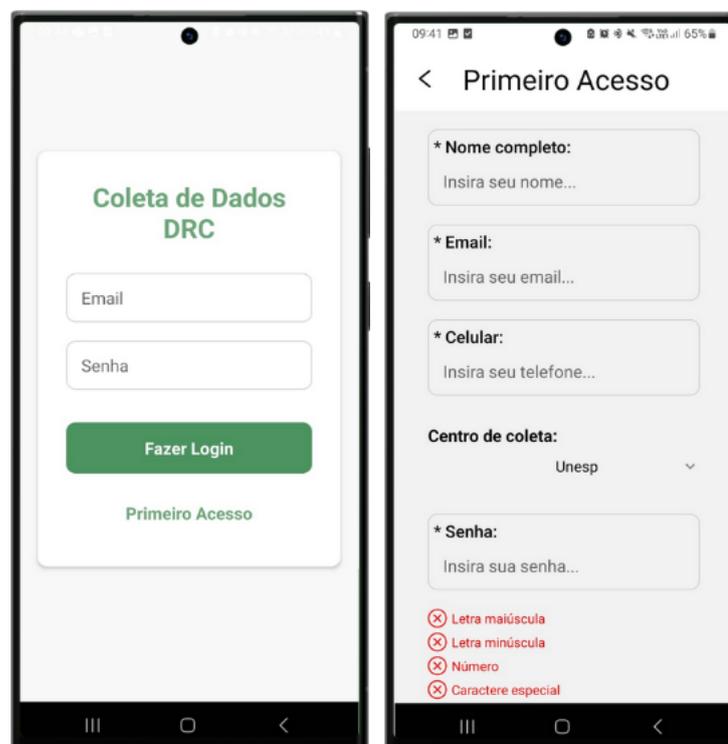


Figura 1 – Telas de login do auxiliar

Após o login bem sucedido no aplicativo, o auxiliar de coleta é redirecionado para

a seguinte tela da Figura 2, listando todos os pacientes que já foram cadastrados.

Para cadastrar novos pacientes é necessário informar seus dados clínicos, como data de nascimento, sexo, e presença de comorbidades como diabetes e hipertensão. Além disso, para saber se o paciente possui a Doença Renal Crônica, em qual estágio está, é necessário informar o nível de creatinina, medido em mg/dL.

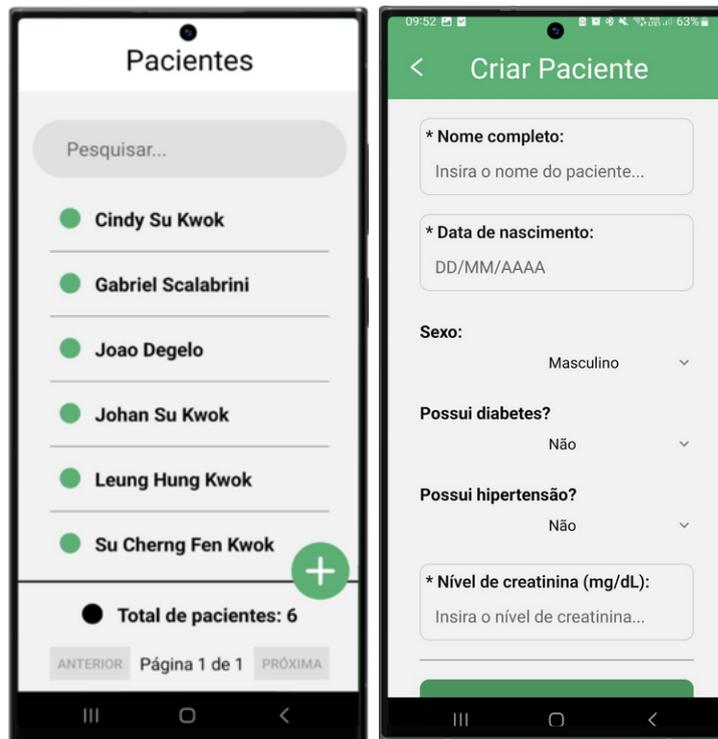


Figura 2 – Listagem de pacientes, e criação de novo paciente

Após preencher os dados ele é redirecionado para a tela da Figura 3 contendo as informações do paciente criado. Nesta tela, é possível editar os dados informados do paciente, deletá-lo, e também realizar o envio das fotos faciais e das mãos

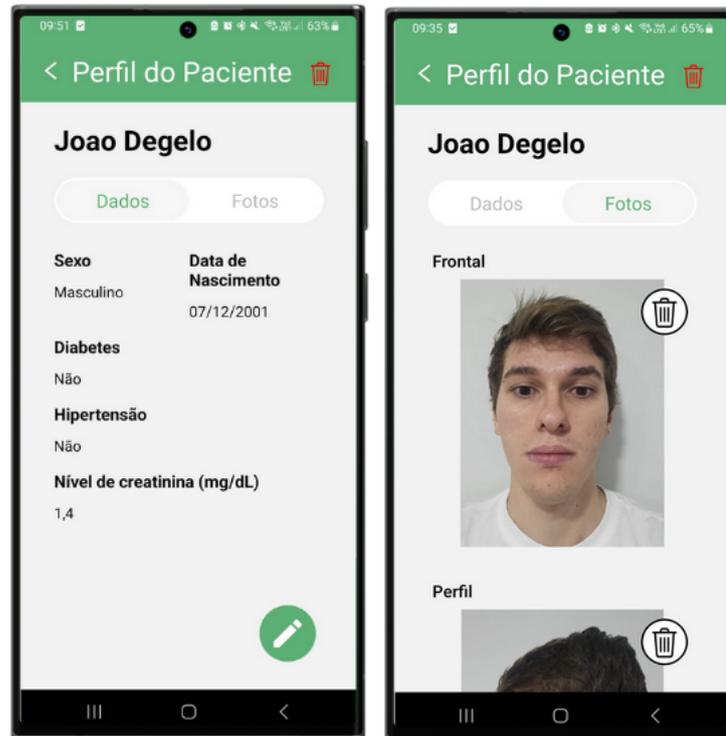


Figura 3 – Informações e fotos do paciente

4.3.2 Diagrama de Arquitetura

O diagrama de arquitetura para o sistema de coleta é apresentado na Figura 4. Desenvolvido para uso dos auxiliares durante a coleta de dados dos pacientes, o aplicativo viabiliza a construção do modelo de Reconhecimento Facial.

O banco de dados armazena três categorias principais de informações: os dados de login do profissional de saúde, os dados cadastrais dos pacientes, e o relacionamento entre os médicos e os pacientes.

O sistema implementa um protocolo robusto de segurança para proteger a comunicação entre o aplicativo e a API. A autenticação é realizada através de tokens JWT (JSON Web Tokens), enquanto todas as requisições são criptografadas via HTTPS, garantindo a confidencialidade e integridade dos dados transmitidos.

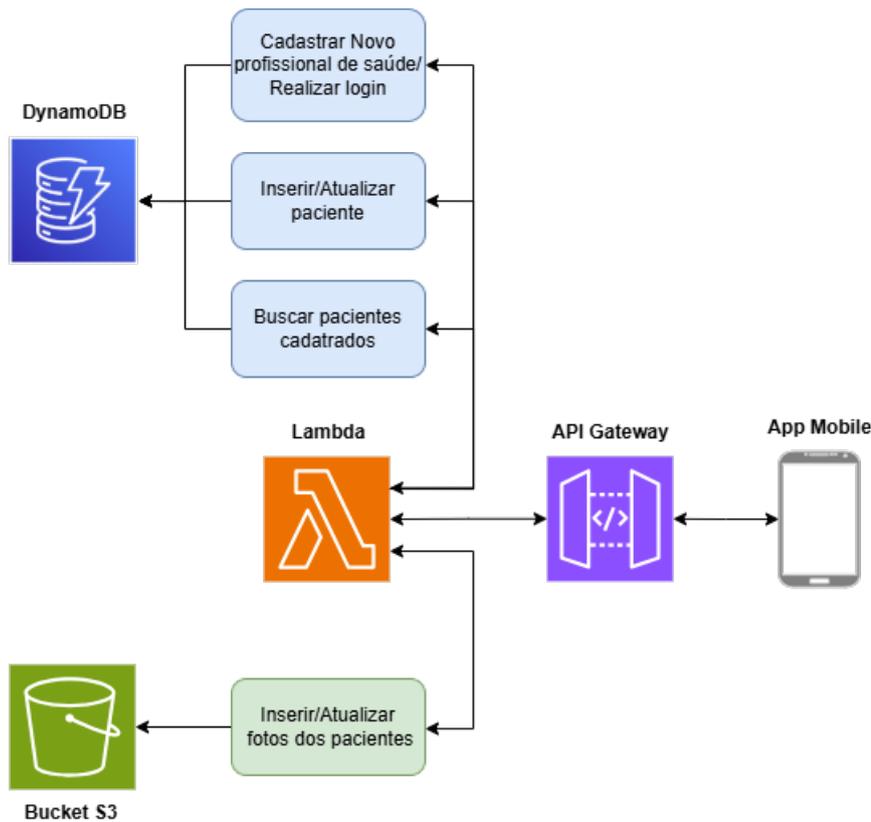


Figura 4 – Diagrama de Arquitetura para o Sistema de Coleta

4.3.3 Decisões de projeto

4.3.3.1 Aplicativo Mobile

Para o desenvolvimento do aplicativo, optou-se pela utilização do framework Expo Go em conjunto com React Native e Typescript. Essa combinação oferece maior flexibilidade e praticidade, permitindo o desenvolvimento simultâneo para as plataformas Android e iOS de forma mais ágil e eficiente. O Expo Go facilita a implementação de funcionalidades nativas sem a necessidade de configuração complexa, enquanto o React Native, com sua base de código unificada, proporciona uma experiência de desenvolvimento mais integrada e consistente, acelerando o ciclo de testes e implantação. Essa escolha foi estratégica para otimizar o tempo e os recursos, mantendo a qualidade e a usabilidade do aplicativo em ambas as plataformas.

4.3.3.2 API - Application Programming Interface

Durante desenvolvimento das API que é utilizada para se comunicar com o aplicativo, optou-se por utilizar o .NET 8. Esse framework oferece diversos benefícios, como alta performance, segurança e facilidade de manutenção. No que se refere à autenticação, foi adotado o uso de Bearer Token com JWT (JSON Web Token). Com esse método, cada requisição à API deve ser acompanhada de um token válido, que é gerado e fornecido ao

usuário após o processo de login. Essa abordagem garante maior segurança nas comunicações, assegurando que apenas usuários autenticados possam acessar os recursos do sistema.

Já para o deploy da API, decidiu-se pela utilização da infraestrutura da Amazon Web Services (AWS), com o serviço AWS Lambda. Essa escolha foi feita devido aos benefícios de custo e escalabilidade proporcionados pela arquitetura serverless. No AWS Lambda, os custos são baseados no número de requisições e no tempo de execução das funções, o que garante uma infraestrutura economicamente eficiente, especialmente em ambientes com demandas variáveis. Além disso, o AWS Lambda oferece baixa latência e alta disponibilidade, automaticamente escalando conforme o volume de requisições aumenta, sem a necessidade de gerenciamento manual de servidores. Essa abordagem não só otimiza os recursos financeiros, mas também garante a continuidade do serviço com mínima interrupção, proporcionando uma experiência mais fluida para os usuários.

4.3.3.3 Banco de dados

O sistema desenvolvido lida com dois tipos principais de dados: os dados do questionário dos pacientes e suas respectivas fotos. Para otimizar o armazenamento e o gerenciamento desses dados, optou-se por utilizar dois serviços distintos da infraestrutura da AWS, cada um adequado ao tipo de informação.

Para armazenar o questionário, escolhemos o banco de dados não relacional DynamoDB, que oferece alta velocidade de leitura e escrita, além de ser extremamente escalável e disponível. O DynamoDB é ideal para lidar com grandes volumes de dados estruturados de forma flexível, sem a necessidade de esquemas rígidos, garantindo o desempenho em cenários de grande demanda.

Já para o armazenamento das imagens dos pacientes, a opção mais eficiente foi o Simple Storage Service (S3), um serviço altamente confiável para a persistência de arquivos na nuvem. O S3 não só permite a gestão de grandes quantidades de dados binários, como também oferece vantagens como redundância automática, alta durabilidade e controle de acessos refinado, garantindo a segurança e integridade dos arquivos.

4.3.4 Privacidade e Segurança dos dados

O projeto envolve a coleta e o processamento de dados sensíveis, incluindo imagens (face e mãos) e informações médicas dos pacientes, exigindo uma arquitetura de segurança robusta e em conformidade com padrões regulatórios. Para garantir a proteção desses dados, foi implementado um sistema de autenticação robusto na API, que restringe o acesso exclusivamente a profissionais de saúde e auxiliares devidamente autorizados, permitindo que acessem apenas os dados dos pacientes sob sua responsabilidade.

A política de senhas adota critérios rigorosos de segurança, exigindo um mínimo de 8 caracteres, com combinação obrigatória de letras maiúsculas e minúsculas, números e caracteres especiais, reduzindo significativamente o risco de ataques de força bruta. Para elevar ainda mais o nível de proteção, especialmente no acesso à infraestrutura AWS que possui privilégios elevados, foi implementada a autenticação em dois fatores (2FA), criando uma barreira adicional contra acessos não autorizados, mesmo em casos de comprometimento de credenciais.

A proteção dos dados é reforçada por meio de criptografia end-to-end, tanto para informações em trânsito quanto em repouso, garantindo sua integridade em todas as etapas do processamento. Este conjunto abrangente de medidas técnicas e procedimentais estabelece um ambiente seguro e confiável para o tratamento das informações sensíveis, garantindo a privacidade e os direitos dos participantes da pesquisa.

4.4 Teste do Aplicativo

A camada de backend, desenvolvida em C#, passou por testes unitários básicos utilizando xUnit para validar a lógica de negócios principal. Os testes de integração verificaram a interação entre diferentes módulos do sistema, incluindo a comunicação com serviços AWS como DynamoDB e S3, garantindo que as operações de CRUD, autenticação e autorização funcionassem conforme esperado.

Para validar a integração completa entre frontend e backend, foram realizados testes end-to-end abrangentes. Estes incluíram verificação das chamadas HTTP, contemplando cenários de sucesso, tratamento de erros e timeout. A persistência de dados foi validada através do fluxo completo, desde a captura dos dados no aplicativo até o armazenamento no DynamoDB ou S3. O processo de gestão de arquivos foi testado considerando upload e download de imagens utilizando o Amazon S3, assim como os mecanismos de autenticação e autorização foram verificados quanto à sua eficácia.

Após a fase inicial de testes automatizados, o aplicativo foi disponibilizado para um grupo controlado de aproximadamente 15 usuários que realizaram testes práticos durante uma semana. Este período de testes com usuários reais foi crucial para validar a usabilidade do sistema em condições reais de uso e identificar potenciais melhorias na experiência do usuário. O feedback coletado durante esta fase foi fundamental para refinar a interface e otimizar os fluxos de trabalho do aplicativo.

O desempenho do sistema também foi objeto de testes específicos. Através de testes de carga, foi possível simular múltiplos usuários simultâneos e verificar a capacidade de resposta do backend. A latência foi medida em diferentes operações, especialmente no upload de imagens e consultas ao banco de dados. O consumo de recursos nas funções Lambda foi monitorado, permitindo ajustes para otimização de custos.

Os resultados combinados dos testes automatizados e do período de testes com usuários foram fundamentais para identificar e corrigir potenciais problemas antes do deploy em produção. Entre as principais melhorias implementadas, destaca-se a otimização do processo de upload de imagens com compressão realizada no aplicativo, a implementação de validações tanto no frontend quanto no backend, o refinamento das políticas de retry na comunicação entre aplicativo e API, além de um tratamento de erros mais sofisticado nas chamadas HTTPS.

A execução contínua destes testes, integrada ao pipeline de CI/CD, garante que novas funcionalidades e correções não introduzam regressões no sistema, mantendo sua estabilidade e confiabilidade. Esta abordagem sistemática de testes, combinada com o feedback direto dos usuários, mostrou-se essencial para manter a qualidade do sistema e proporcionar uma experiência confiável aos usuários finais.

4.5 Implantação

Após a aprovação da pesquisa pelo conselho de ética e pelo hospital parceiro, o aplicativo foi liberado para uso pelas técnicas de enfermagem. Em uma reunião inicial, foram explicados detalhadamente os procedimentos de coleta, abrangendo a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) e apresentando integralmente as funcionalidades do aplicativo desenvolvido. Nesse encontro, foram criadas as contas de acesso personalizadas para cada profissional, possibilitando o cadastro de novos pacientes, e na semana seguinte foi iniciado o processo de coleta de dados.

Durante a utilização do aplicativo, foram necessárias duas correções fundamentais para aprimorar sua funcionalidade. A primeira correção estava relacionada à validação do campo de nível de creatinina. Inicialmente, baseando-se em pesquisas que indicavam valores normais entre 0,6 e 1,3 mg/dL, o sistema possuía um limite máximo de 10,0 mg/dL para esse campo. No entanto, após algumas semanas de uso, as técnicas de enfermagem relataram que pacientes em pré-diálise podem apresentar níveis significativamente elevados de creatinina. Considerando essa importante observação, foi realizado um ajuste no limite superior do campo, permitindo o registro de valores mais amplos e precisos.

O segundo ponto de aprimoramento surgiu em decorrência de uma sobrecarga do aplicativo, ocasionada pelo elevado número de componentes renderizados simultaneamente. Aproximadamente um mês após o início da coleta, quando já haviam sido cadastrados cerca de 200 pacientes, o aplicativo começou a apresentar travamentos que impactavam diretamente a experiência dos usuários. Para solucionar esse problema de desempenho, foi implementada uma estratégia de paginação no frontend, reduzindo significativamente o número de componentes renderizados no carregamento da página e garantindo maior fluidez na utilização do sistema.

De modo geral, o aplicativo demonstrou-se uma ferramenta essencial para a coleta de dados, proporcionando uma otimização significativa dos procedimentos realizados pelas técnicas de enfermagem. Além disso, o sistema assegurou maior privacidade aos pacientes e garantiu segurança e organização eficiente das informações, mesmo com o cadastro diário de um grande volume de dados sensíveis. A figura 5 ilustra graficamente o volume de pacientes cadastrados diariamente durante o período da coleta.

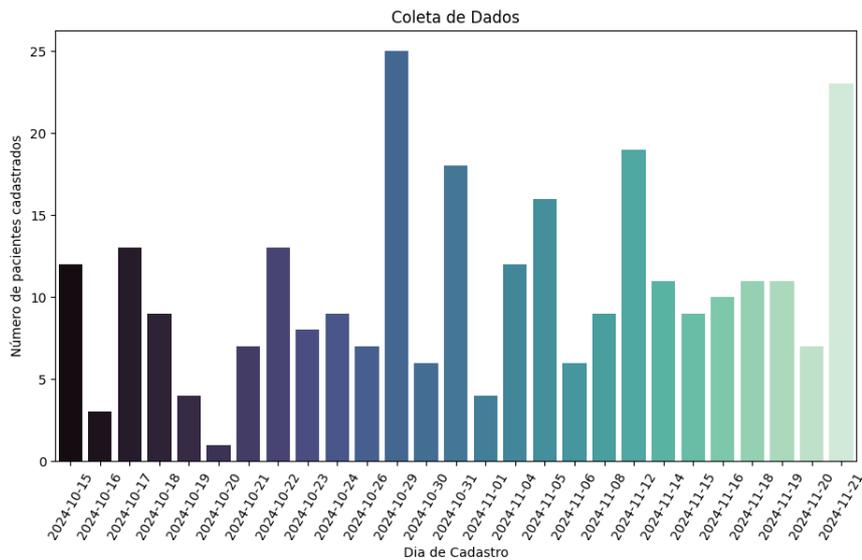


Figura 5 – Volume de dados coletados diariamente

O período analisado da coleta teve duração aproximada de um mês e meio e contou com a participação de 283 pacientes. A coleta continua sendo realizada e pretende-se atingir o número de 500 pacientes cadastrados, ampliando ainda mais o escopo e a representatividade da pesquisa.

4.6 Resultados e Análises dos dados

De maneira geral, o aplicativo atendeu integralmente às especificações técnicas propostas. Todos os requisitos funcionais 4.1.1 foram implementados, possibilitando o uso completo das funcionalidades do sistema. O requisito não funcional de Security foi alcançado por meio de medidas robustas de segurança do sistema e proteção das informações armazenadas.

Além disso, a arquitetura em nuvem projetada permitiu atingir com êxito os requisitos não funcionais 4.1.2 de disponibilidade e baixo tempo de resposta, garantindo uma infraestrutura tecnológica eficiente e de alto desempenho.

Com relação aos resultados da coleta de dados com o aplicativo, foram realizadas algumas análises sobre as informações, de modo a entender melhor o público atingido pela pesquisa. A figura 6 demonstra a distribuição dos participantes separadas de acordo com o

sexo e o estágio da doença. De modo geral, a pesquisa teve uma participação equilibrada entre ambos os sexos. No entanto, é notável que, para o grupo de controle e para os pacientes no Estágio 2 da DRC há uma presença maior de pessoas do sexo masculino.

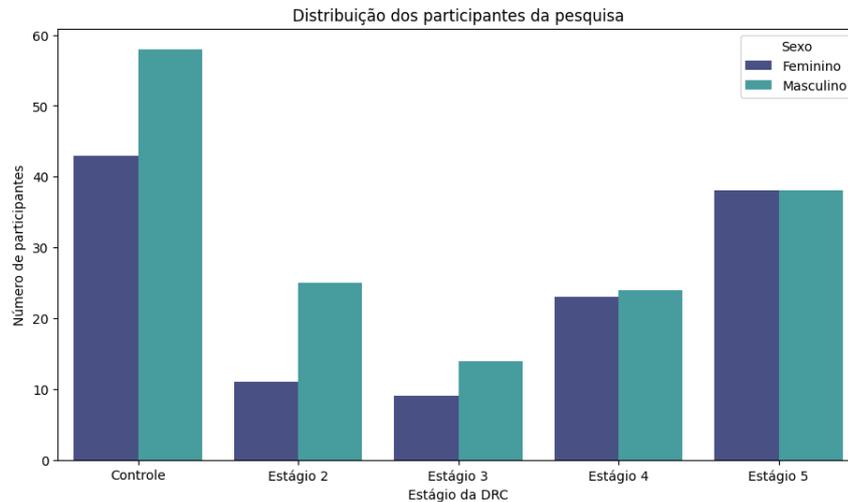


Figura 6 – Pacientes em cada estágio da doença, separado de acordo com o sexo

Realizando uma análise mais aprofundada com relação à idade dos participantes da pesquisa. A figura 7 demonstra graficamente a distribuição da idade dos pacientes em cada estágio da doença. É possível notar que pacientes com Doença Renal Crônica apresentam uma idade média levemente superior quando comparada a os pacientes que compõem o grupo controle.

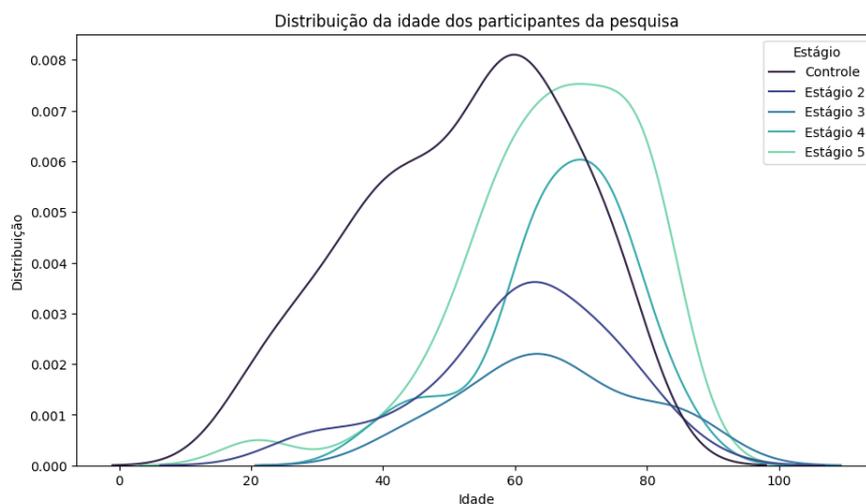


Figura 7 – Distribuição da idade dos pacientes separada pelo estágio da doença

Além dos dados demográficos básicos (sexo e idade), foram analisadas duas comorbidades fundamentais no desenvolvimento da Doença Renal Crônica: diabetes e hipertensão. A figura 8 ilustra a distribuição proporcional dessas condições em cada estágio da doença.

A análise do gráfico evidencia não apenas a estreita relação entre estas comorbidades e a DRC, mas também sua influência na progressão dos diferentes estágios da doença.

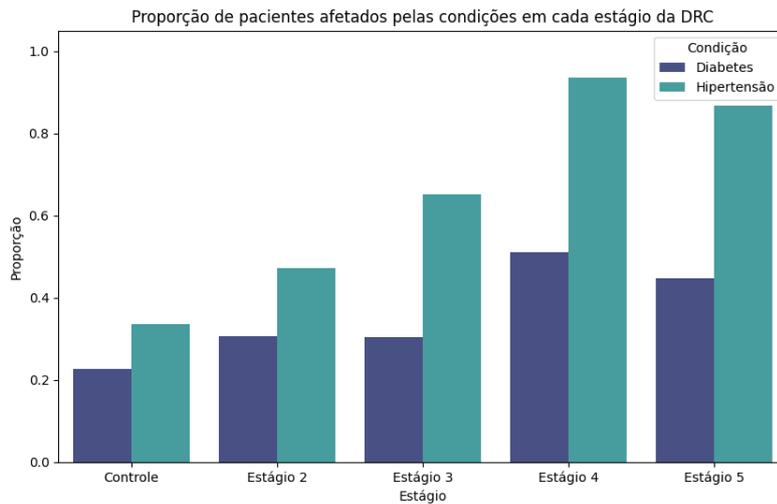


Figura 8 – Proporção de pacientes afetados por cada condição de saúde em cada estágio da doença

No treinamento dos modelos de Inteligência Artificial baseados em aprendizado supervisionado, o problema foi estruturado em duas categorias: pacientes sem DRC (grupo controle) e pacientes com DRC (estágios 2, 3, 4 e 5). As figuras 9, 10 e 11 apresentam a análise da distribuição dos dados em relação à variável target (presença ou ausência de DRC).

As análises realizadas proporcionam insights fundamentais sobre o perfil epidemiológico da Doença Renal Crônica (DRC), revelando nuances importantes para o desenvolvimento do modelo de Inteligência Artificial. Destacam-se, especialmente, os desafios relacionados ao desbalanceamento dos dados, que podem exigir estratégias metodológicas específicas durante o treinamento do modelo. Tais resultados configuraram-se como elementos cruciais para o planejamento e a orientação das etapas subsequentes do projeto, permitindo uma abordagem mais assertiva e fundamentada cientificamente.

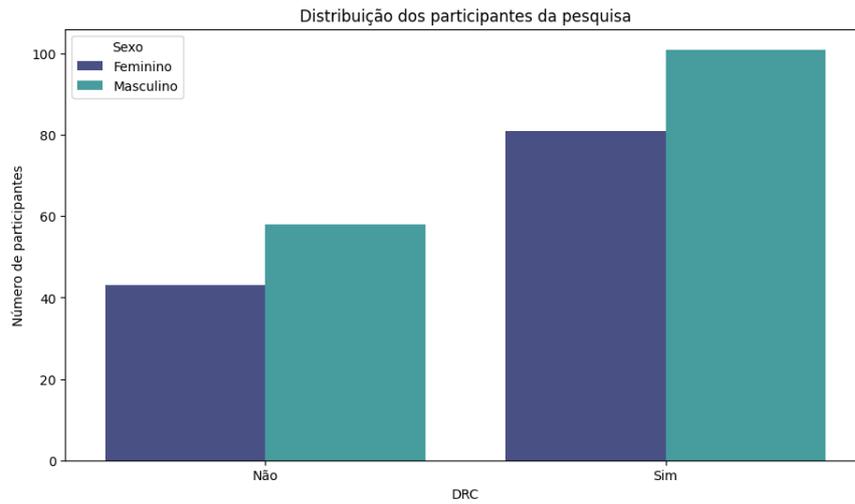


Figura 9 – Número de pacientes participantes separados de acordo com o sexo

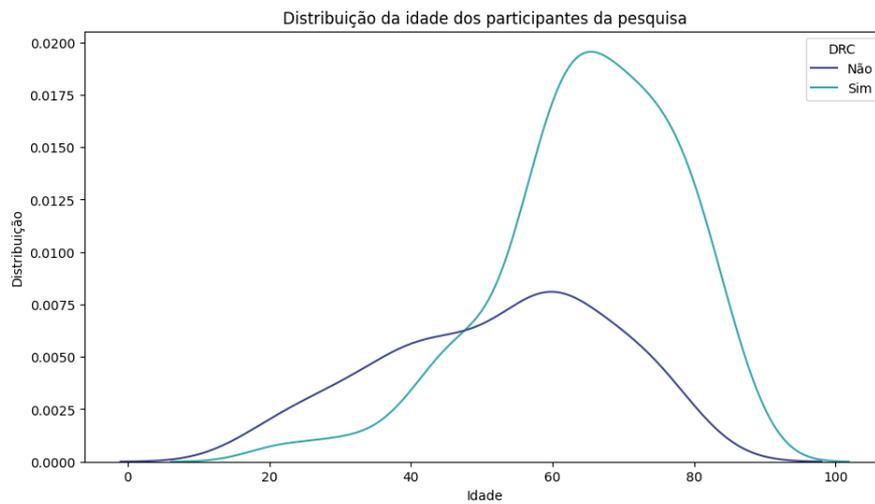


Figura 10 – Distribuição da idade dos participantes da pesquisa

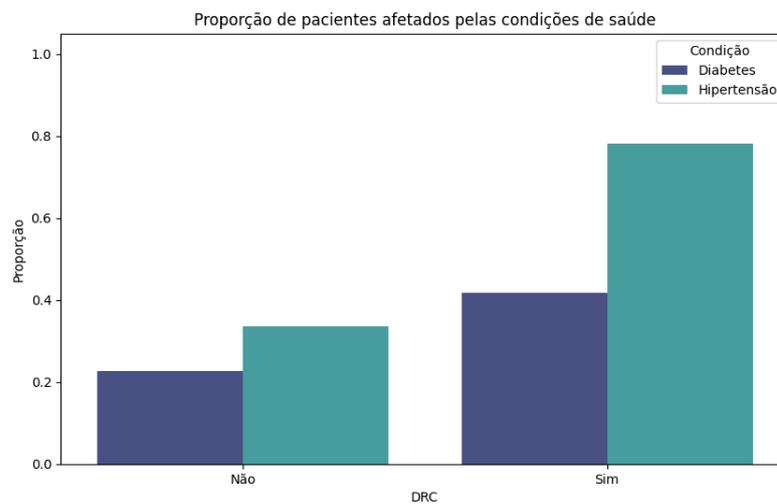


Figura 11 – Proporção de pacientes afetados por cada condição de saúde

5 Desenvolvimento do Sistema de Diagnóstico

Nesta etapa do projeto, os dados coletados foram utilizados para treinar um modelo de Inteligência Artificial destinado a auxiliar no diagnóstico de Doença Renal Crônica. Foi desenvolvido um novo aplicativo que incorporou o modelo treinado, possibilitando a realização de diagnósticos a partir dos dados e imagens médicas fornecidas pelos profissionais de saúde.

5.1 Especificação Técnica

5.1.1 Requisitos Funcionais

Pode-se elencar os seguintes requisitos funcionais do sistema:

1. O sistema deve permitir o cadastro de novos profissionais da saúde com um login e senha;
2. O profissional da saúde cadastrado consegue realizar login no aplicativo utilizando suas credenciais;
3. O sistema deve permitir que os profissionais da saúde cadastrem novos pacientes;
4. O sistema deve permitir que os profissionais da saúde adicione as imagens do paciente;
5. O sistema deve permitir que os profissionais da saúde obtenham novos diagnósticos a partir dos dados cadastrados do paciente.
6. O sistema deve permitir que os profissionais da saúde acessem o histórico dos diagnósticos e dos dados dos pacientes que foram cadastrados por eles;
7. O sistema deve permitir que os profissionais da saúde apaguem todos os dados armazenados no sistema dos pacientes cadastrados por eles.

5.1.2 Requisitos Não-Funcionais

Entre os requisitos não-funcionais do sistema, pode-se citar:

- Security - devem ser adotadas políticas de segurança para proteger o sistema e também os dados sensíveis dos pacientes, garantindo total privacidade;

- Disponibilidade - sistema deve apresentar disponibilidade de 99,9
- Baixo tempo de resposta - sistema deve apresentar tempo de resposta médio abaixo de 1 segundo.
- Confiabilidade - o sistema deve passar por um rigoroso processo de validação e testes para garantir que o diagnóstico seja consistente e confiável;

5.1.3 Descrição dos Casos de Uso

O aplicativo de diagnóstico abrange no total 7 casos de uso:

1. Criar usuário: idêntico ao caso de uso Criar usuário do aplicativo de coleta (4.1.3.1)
2. Realizar login: idêntico ao Realizar Login do aplicativo de coleta (4.1.3.2)
3. Consultar paciente: idêntico ao Consultar Paciente do aplicativo de coleta (4.1.3.3)
4. Cadastrar paciente (5.1.3.1)
5. Editar paciente (5.1.3.2)
6. Deletar Paciente: idêntico ao Deletar Paciente do aplicativo de coleta (4.1.3.6).
7. Criar Diagnóstico (5.1.3.3)
8. Deletar Diagnóstico (5.1.3.4)

O diagrama de casos de uso do aplicativo de diagnóstico, tal como o do aplicativo de coleta, não foi incluído neste relatório, pois seria redundante, uma vez que há um único ator responsável por todos os casos citados.

5.1.3.1 Descrição de Criar um Paciente

Nome: Criar um paciente.

Descrição: Criar um novo paciente, cadastrando seus dados clínicos.

Evento Iniciador: Acionar botão "+" na tela de listagem de pacientes.

Ator: Médicos nefrologistas.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Sistema solicita o preenchimento dos dados do paciente (nome, data de nascimento, sexo e se possui hipertensão ou diabetes).

2. Ator fornece os dados solicitados.
3. Ator clica no botão "Criar paciente".
4. Sistema valida os dados fornecidos.
5. Sistema cria o paciente e redireciona o ator para a tela de perfil do paciente criado.
6. Fim do Caso de Uso.

Pós-Condição: Paciente cadastrado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Credenciais inseridas são inválidas (passo 4)
 - a) Sistema informa que dados inseridos são inválidos e retorna ao passo 1.
2. Ator cancela o cadastro (passo 3)
 - a) Sistema cancela o cadastro e encerra o caso de uso.

5.1.3.2 Descrição de Editar Paciente

Nome: Editar Paciente.

Descrição: Editar dados clínicos do paciente.

Evento Iniciador: Entrar no perfil do paciente.

Ator: Médicos nefrologistas.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator clica no botão com ícone de lápis.
2. Sistema redireciona o ator para a tela de edição de paciente.
3. Sistema informa os dados atuais do paciente, e solicita o preenchimento dos novos dados (nome, data de nascimento, sexo e se possui hipertensão ou diabetes).
4. Ator altera os dados do paciente.
5. Ator clica no botão "Editar paciente".
6. Sistema valida os dados fornecidos.
7. Sistema edita o paciente e redireciona o ator para a tela de perfil do paciente.

8. Fim do Caso de Uso.

Pós-Condição: Paciente atualizado no sistema.

Fluxos alternativos: Não Contém.

Exceções:

1. Dados inseridos são inválidos (passo 5)
 - a) Sistema informa que dados inseridos são inválidos e retorna ao passo 2.
2. Ator cancela o cadastro (passo 3)
 - a) Sistema cancela a edição e encerra o caso de uso.

5.1.3.3 Descrição de Criar Diagnóstico

Nome: Criar um Diagnóstico

Descrição: Cria um novo diagnóstico para um paciente

Evento Iniciador: Entrar no perfil do paciente.

Ator: Médicos nefrologistas.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator clica no botão "Histórico".
2. Sistema redireciona o ator à seção de histórico de diagnósticos.
3. Ator clica no botão "+"
4. Sistema redireciona o ator à tela de criação de diagnóstico.
5. Sistema solicita o envio das quatro fotos do paciente (facial frontal, facial em perfil, facial superior, e foto das mãos).
6. Ator submete as fotos e clica no botão "Criar Diagnóstico"
7. Sistema valida se todas as fotos foram fornecidas.
8. Sistema cria o diagnóstico e redireciona o ator para a tela de histórico de diagnósticos.
9. Fim do Caso de Uso.

Pós-Condição: Diagnóstico criado.

Fluxos alternativos: Não contém.

Exceções:

1. Alguma foto não foi fornecida (passo 7)
 - a) Sistema informa que dados inseridos são inválidos e retorna ao passo 5.
2. Ator cancela a criação (passo 6)
 - a) Sistema cancela a criação e encerra o caso de uso.

5.1.3.4 Descrição de Deletar Diagnóstico

Nome: Deletar diagnóstico

Descrição: Deleta um diagnóstico de um paciente

Evento Iniciador: Entrar no perfil do paciente.

Ator: Médicos nefrologistas.

Pré-condição: Ator autenticado no sistema.

Sequência de Eventos:

1. Ator clica no botão "Histórico".
2. Sistema redireciona o ator à seção de histórico de diagnósticos.
3. Ator clica no diagnóstico que deseja deletar.
4. Sistema redireciona o ator para a tela de dados do diagnóstico.
5. Ator clica no botão com ícone de lixeira.
6. Sistema solicita confirmação da deleção.
7. Ator clica no botão "Confirmar".
8. Sistema deleta o diagnóstico do paciente.
9. Sistema redireciona o ator para a tela de histórico de diagnósticos.
10. Fim do Caso de Uso.

Pós-Condição: Diagnóstico deletado do sistema.

Fluxos alternativos: Não contém.

Exceções:

1. Ator cancela a deleção (passo 7)

- a) Sistema cancela a deleção e encerra o caso de uso.

5.2 Tecnologias Utilizadas

5.2.1 AWS Sagemaker

A plataforma AWS SageMaker é uma solução abrangente de machine learning em nuvem, desenvolvida pela Amazon para simplificar e acelerar o processo de criação, treinamento e implantação de modelos de Inteligência Artificial. Ela oferece ferramentas integradas que permitem aos pesquisadores e desenvolvedores preparar dados, treinar modelos em larga escala, realizar testes comparativos e implantar soluções de forma eficiente e segura, com recursos de escalabilidade e gerenciamento automatizado da infraestrutura computacional.

5.2.2 Python

Python é uma linguagem de programação de alto nível, caracterizada por sua sintaxe clara e legível, amplamente adotada no campo da ciência de dados e machine learning. Seu ecossistema rico em bibliotecas especializadas, como NumPy, Pandas e SciPy, fornece suporte robusto para manipulação de dados, análise estatística e desenvolvimento de algoritmos complexos, e bibliotecas como TensorFlow e Scikit-learn tornam-a uma escolha preferencial para projetos que envolvem Inteligência Artificial e aprendizado de máquina.

5.2.3 MediaPipe

Desenvolvida pelo Google, a biblioteca MediaPipe representa uma solução avançada para processamento de mídias e visão computacional. No contexto deste projeto, suas capacidades de extração de embeddings faciais possibilitam a transformação de imagens de rosto em representações numéricas precisas, capturando características únicas e permitindo análises de reconhecimento e comparação com alta acurácia. A biblioteca se destaca pela eficiência computacional e pela capacidade de trabalhar com diferentes tipos de dados multimodais.

5.2.4 Scikit-learn

Scikit-learn é uma biblioteca fundamental no ecossistema de machine learning em Python, oferecendo uma coleção abrangente de algoritmos para classificação, regressão, agrupamento e redução de dimensionalidade. Sua arquitetura consistente e documentação

extensa facilitam a implementação de técnicas de aprendizado supervisionado e não supervisionado, incluindo ferramentas essenciais para pré-processamento de dados, seleção de características, validação cruzada e avaliação de desempenho de modelos.

5.2.5 TensorFlow

O TensorFlow, desenvolvido pelo Google, é um framework de código aberto projetado para computação numérica e aprendizado de máquina, especialmente voltado para redes neurais profundas. Com integração nativa do Keras, simplifica significativamente o processo de construção e treinamento de modelos, permitindo a criação rápida de arquiteturas complexas. Suas capacidades incluem suporte a computação distribuída, ferramentas de visualização de modelos e flexibilidade para implementar técnicas avançadas como transfer learning e fine-tuning em diversos domínios de aplicação.

5.3 Projeto e Implementação

5.3.1 Telas do Aplicativo

O produto final desse projeto consiste em um aplicativo que será disponibilizado aos profissionais de saúde para diagnóstico de Doença Renal Crônica. A Figura 12 representa a tela que o profissional de saúde é redirecionado após realizar o preenchimento de seu usuário e senha. Nessa tela, ele pode cadastrar novos pacientes ou visualizar os pacientes cadastrados por ele, além de buscar por alguma pessoa específica, da mesma forma que no aplicativo de coleta.

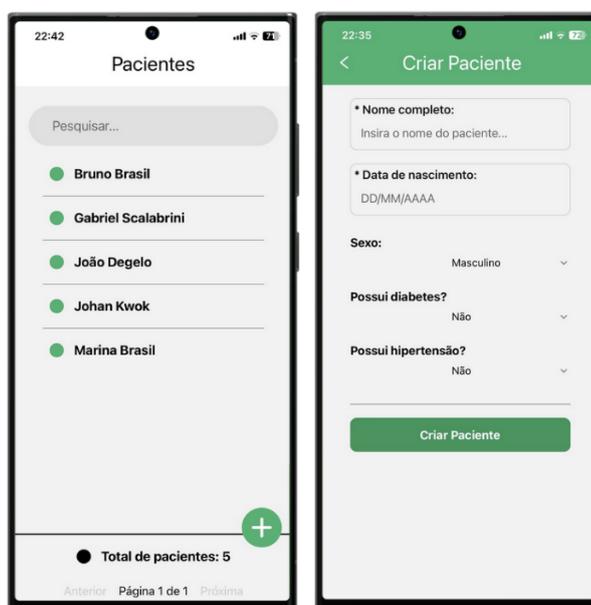


Figura 12 – Listagem de pacientes, e criação de novo paciente

Ao selecionar o paciente que se deseja visualizar, o profissional de saúde é redirecionado para a tela representada na Figura 13 contendo as informações pessoais do paciente e o seu histórico de diagnósticos utilizando o modelo de reconhecimento facial desenvolvido.

O usuário do sistema pode alterar as informações cadastradas para o paciente ou criar um novo diagnóstico com as fotografias. Ao clicar em um dos diagnósticos, é possível visualizar os detalhes dele com as informações que foram utilizadas pelo modelo para calcular a probabilidade exibida.

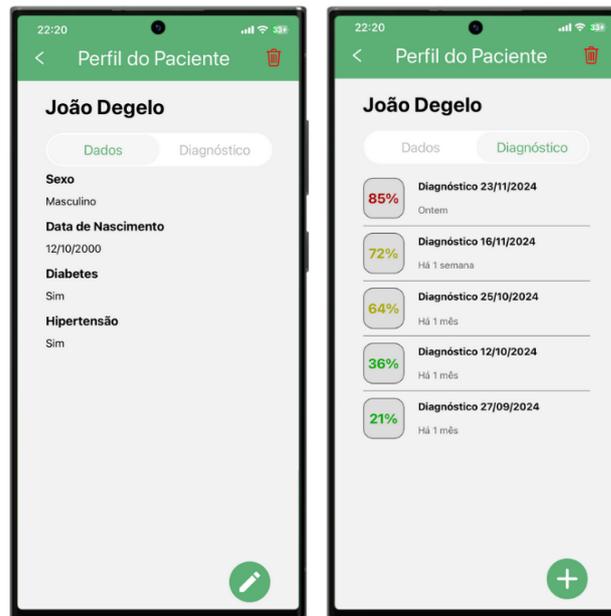


Figura 13 – Perfil contendo informações do paciente

Na tela de criação de diagnóstico, o usuário deve realizar o upload das três fotos faciais, nas posições frontal, perfil e superior, e também da foto das mãos, como mostra a Figura 14. Os dados clínicos utilizados em um diagnóstico são os mesmos do perfil do usuário, e devem ser alterados pela tela correspondente.



Figura 14 – Criação de novo diagnóstico

Ao buscar mais detalhes de um diagnóstico específico na aba de "Histórico", o usuário é redirecionado para as telas representadas na Figura 15. Com essa visão, é possível observar as informações sobre os dados clínicos do paciente assim como as imagens da face e mãos do paciente que foram utilizadas pelo modelo para gerar a probabilidade final.

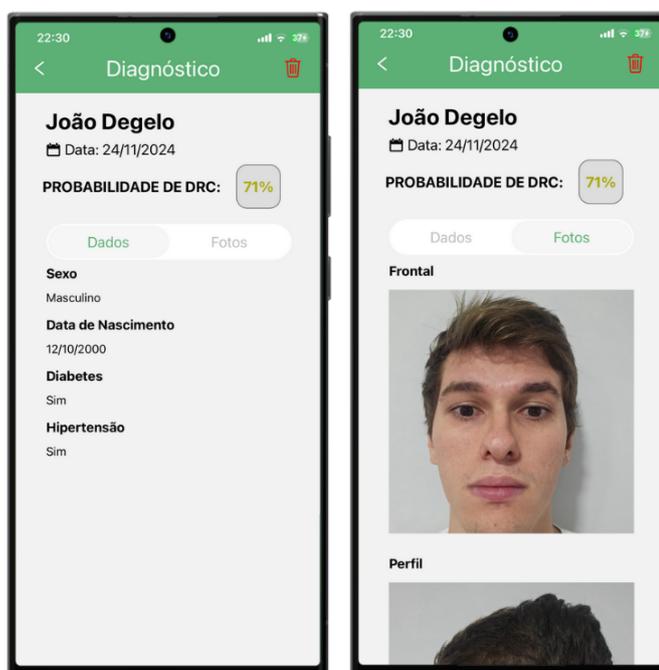


Figura 15 – Informações do diagnóstico do paciente

5.3.2 Diagrama de Arquitetura

A Figura 16 apresenta o diagrama simplificado da arquitetura do sistema de aplicativo de diagnóstico. O aplicativo será utilizado pelos médicos para o cadastro de pacientes e utilização do modelo treinado de Inteligência Artificial. O banco de dados conterá os dados cadastrais dos médicos, a relação entre médicos e pacientes, dados clínicos dos pacientes, e informações sobre cada um dos diagnósticos realizados. O aplicativo se comunica com o banco de dados e o modelo de inteligência artificial por meio de requisições HTTPS realizadas para uma API protegida contra acesso indevido em uma instância Lambda da AWS.

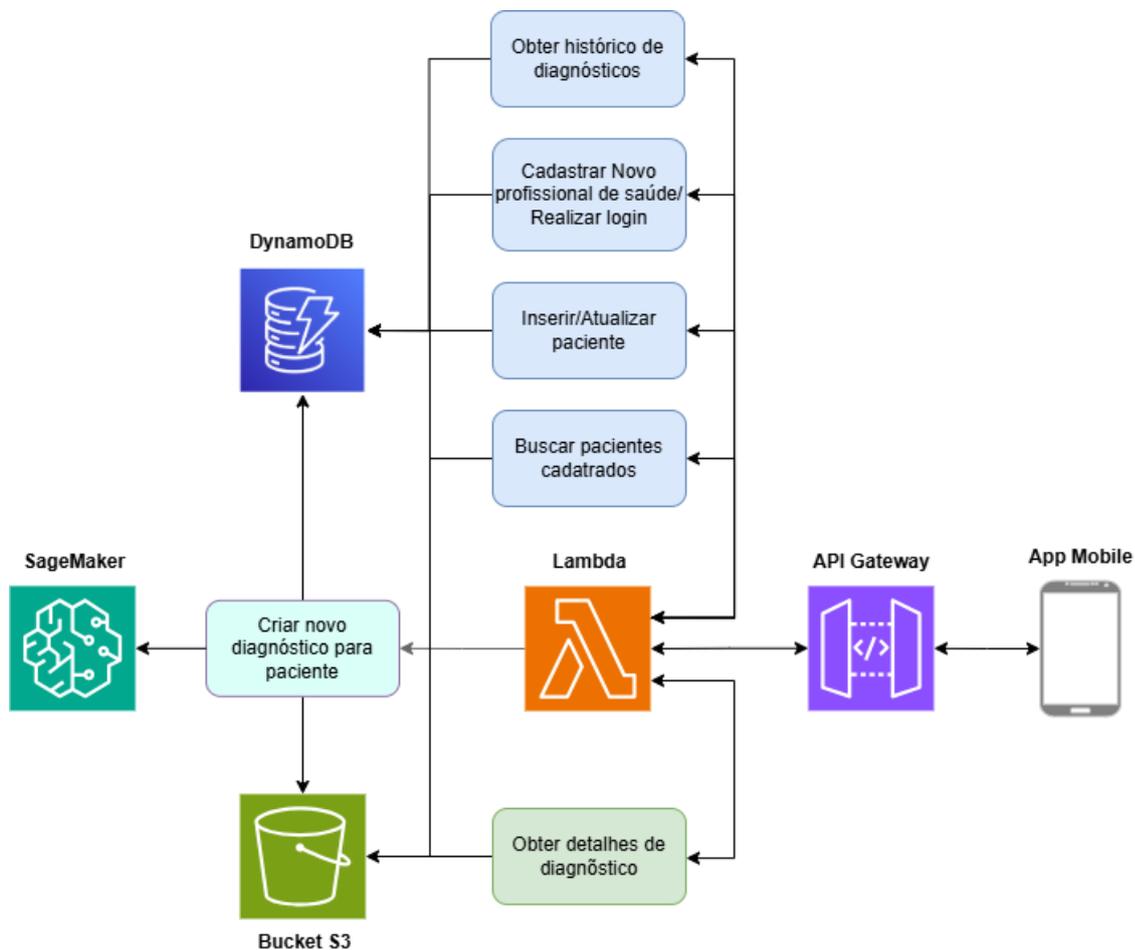


Figura 16 – Diagrama de Arquitetura

5.3.3 Aplicativo de Diagnóstico

O desenvolvimento do aplicativo de diagnóstico de Doença Renal Crônica foi significativamente facilitado pela experiência adquirida no desenvolvimento do aplicativo de coleta de dados. Muitos componentes foram similares ou idênticos aos do aplicativo de coleta, permitindo um alto reaproveitamento de código e uma alocação mais estratégica de recursos no desenvolvimento do modelo de Inteligência Artificial. Foi utilizada a mesma

estrutura do frontend, backend e banco de dados do aplicativo de coleta de dados.

Dado o curto período de tempo para o desenvolvimento do modelo, o aplicativo não foi disponibilizado para a equipe médica no momento.

5.4 Tratamento e limpeza dos dados

Para viabilizar o treinamento utilizando aprendizado supervisionado, foi necessário realizar alguns tratamentos no conjunto de dados coletados. O principal deles foi o cálculo da taxa de filtração glomerular (TFG) a partir do nível de creatinina, utilizando a equação CKD-EPI Creatinine Equation. Com esse valor, foi possível determinar o estágio da doença renal crônica de cada paciente. Além disso, foi realizado o processamento da data de nascimento dos pacientes para calcular suas idades atuais.

Por fim, o dataset passou por uma etapa de limpeza, na qual foram removidos dois registros que não continham as quatro fotos cadastradas pelo aplicativo. Essa inconsistência ocorreu devido a um erro manual durante a coleta dos dados.

5.4.1 Seleção e treino de modelos

Foram testados diversos modelos de Inteligência Artificial para identificar quais abordagens apresentavam o melhor desempenho na tarefa proposta. O processo de treinamento desses modelos envolveu uma divisão do conjunto de dados (244 registros) em subconjuntos de treino (156 registros), validação (39 registros) e teste (49 registros). Os modelos foram inicialmente treinados no conjunto de treino, utilizando o conjunto de validação para seleção daqueles com melhor desempenho. A avaliação final foi realizada no conjunto de teste, garantindo uma análise robusta e imparcial das diferentes abordagens desenvolvidas

O primeiro modelo desenvolvido foi um XGBoost, cuja escolha fundamentou-se em suas características de desempenho em conjuntos de dados pequenos, sendo um algoritmo especialmente robusto para lidar com amostras reduzidas. Diferentemente de modelos de redes neurais que requerem grandes volumes de dados para treinamento, o XGBoost opera eficientemente com datasets limitados, realizando uma regularização intrínseca que previne overfitting.

Esse primeiro modelo utilizou como entrada apenas os dados de saúde dos pacientes (sexo, idade, diabetes e hipertensão). Esse conjunto de dados fornece informações altamente relevantes para o diagnóstico de doença renal crônica. Para otimizar o desempenho do modelo, foi aplicada a técnica RandomizedSearch com validação cruzada, visando encontrar os melhores hiperparâmetros. Essa abordagem realiza uma busca aleatória por combinações de valores, o que reduz o tempo de computação em comparação com a busca exaustiva, ao

mesmo tempo em que aumenta a probabilidade de encontrar uma configuração eficaz e bem ajustada para o modelo.

O segundo modelo desenvolvido utilizou embeddings faciais dos pacientes para extrair features a partir das imagens. Esses embeddings capturam características únicas de cada rosto, permitindo que o modelo identifique padrões relevantes para o diagnóstico. Com essas representações faciais, foi treinado um modelo XGBoost, otimizado por meio de ajuste de hiperparâmetros com RandomizedSearch. A partir dessa abordagem, foram exploradas algumas variantes, incluindo a combinação dos embeddings com os dados de saúde dos pacientes para enriquecer as informações de entrada. Além disso, aplicou-se a técnica de Feature Selection com Recursive Feature Elimination (RFE), que identifica e retém apenas as variáveis mais relevantes, removendo progressivamente aquelas de menor importância a cada rodada de treinamento. Essa técnica reduz a complexidade do modelo e o risco de overfitting, permitindo que ele se concentre nas informações mais úteis para a tarefa, resultando em um desempenho mais eficiente e confiável.

Por fim, o último modelo treinado utilizou uma arquitetura de redes neurais convolucionais (CNNs), com base no modelo pré-treinado EfficientNetB0. Esta é uma arquitetura otimizada para maximizar a precisão de classificação com o menor custo computacional possível, utilizando um escalonamento composto que ajusta a profundidade, largura e resolução da rede de forma eficiente. Optou-se por utilizar modelos pré-treinados com fine tuning devido ao tamanho relativamente pequeno do conjunto de dados disponível para o treinamento. Essa abordagem permite aproveitar o aprendizado adquirido por modelos treinados em grandes bases de dados, ajustando-os de forma eficiente para a tarefa específica, mesmo com uma quantidade limitada de dados.

O fine-tuning foi feito ajustando as camadas finais do modelo para que ele aprendesse características específicas do conjunto de dados disponível, evitando a necessidade de treinar o modelo do zero. Esse modelo final foi projetado para receber como entrada um conjunto de 4 imagens (frontal, perfil, superior e das mãos) combinado com dados de saúde dos pacientes (sexo, idade, diabetes e hipertensão), proporcionando uma análise mais completa.

Foram aplicadas também técnicas de Data Augmentation para enriquecer o conjunto de treinamento, gerando variações nas imagens originais (com uma leve rotação, espelhamento e um pequeno ajuste de brilho) que aumentam a diversidade dos dados e melhoram a capacidade de generalização do modelo. Além disso, foram utilizados métodos de Regularization, como dropout e weight decay, para reduzir o risco de overfitting, regularizando o modelo para que ele mantenha eficiência sem superajustes. Essas estratégias, juntamente com o ajuste fino dos hiperparâmetros, resultaram em um modelo robusto e mais capaz de identificar padrões complexos nas imagens e nos dados de saúde.

5.5 Avaliação e Testes do Sistema

O sistema desenvolvido atendeu plenamente à especificação técnica proposta, demonstrando excelente alinhamento com os requisitos estabelecidos. No âmbito funcional 5.1.1, todos os requisitos foram integralmente implementados, assegurando uma experiência de usabilidade robusta e eficiente.

Quanto aos requisitos não funcionais 5.1.2, destaca-se o cumprimento das medidas de segurança definidas no requisito , com a implementação de múltiplas camadas de proteção. A arquitetura em nuvem adotada mostrou-se fundamental para garantir os requisitos de disponibilidade e desempenho, proporcionando tempos de resposta otimizados, conforme especificado no requisito. Os testes abrangentes realizados no sistema corroboraram o atendimento do requisito de confiabilidade, validando a robustez e a estabilidade da solução desenvolvida.

Para avaliar a performance desses modelos, a principal métrica de avaliação utilizada foi a ROC-AUC (Receiver Operating Characteristic - Area Under the Curve). Esta métrica é particularmente adequada para avaliar o desempenho de modelos de classificação binária, como no caso do diagnóstico de Doença Renal Crônica, onde o resultado pode ser positivo (presença da doença) ou negativo (ausência da doença).

A curva ROC ilustra graficamente a relação entre a taxa de verdadeiros positivos (sensibilidade) e a taxa de falsos positivos (1 - especificidade) para diferentes limiares de classificação, enquanto a AUC representa a área sob esta curva. Um valor de AUC próximo a 1 indica um modelo com excelente capacidade de discriminação, enquanto um valor próximo a 0.5 sugere um desempenho equivalente ao acaso. Esta métrica foi escolhida por sua robustez em cenários de classificação médica, onde é crucial equilibrar a sensibilidade (capacidade de identificar corretamente pacientes doentes) e a especificidade (capacidade de identificar corretamente pacientes saudáveis). Além disso, a ROC-AUC tem a vantagem de ser insensível ao desbalanceamento de classes, uma característica comum em dados médicos, onde frequentemente há uma proporção desigual entre casos positivos e negativos.

5.5.1 Abordagens com algoritmo XGBoost

As figuras 17 e 18 ilustram o desempenho comparativo dos modelos XGBoost nos conjuntos de treino e validação respectivamente. Observa-se uma discrepância significativa entre a performance nos dois conjuntos para todos os algoritmos que utilizam características extraídas de imagens, com exceção do XGBoost treinado exclusivamente com dados clínicos. Esta disparidade é um indicativo claro de overfitting, onde os modelos apresentam excelente capacidade de aprendizado no conjunto de treino, mas falham em generalizar adequadamente para dados não vistos. O fenômeno ocorre principalmente devido à alta dimensionalidade das features extraídas das imagens em contraste com o tamanho limitado

do dataset disponível.

É importante ressaltar que, mesmo com a implementação de diversas técnicas de regularização, como Feature Selection e RFE, o problema de overfitting persistiu nos modelos que utilizavam dados de imagem. Consequentemente, a abordagem que demonstrou maior robustez e melhor capacidade de generalização foi a que empregou exclusivamente dados médicos como entrada, evidenciando que, para este conjunto de dados específico, a simplicidade do modelo favoreceu sua eficácia diagnóstica.

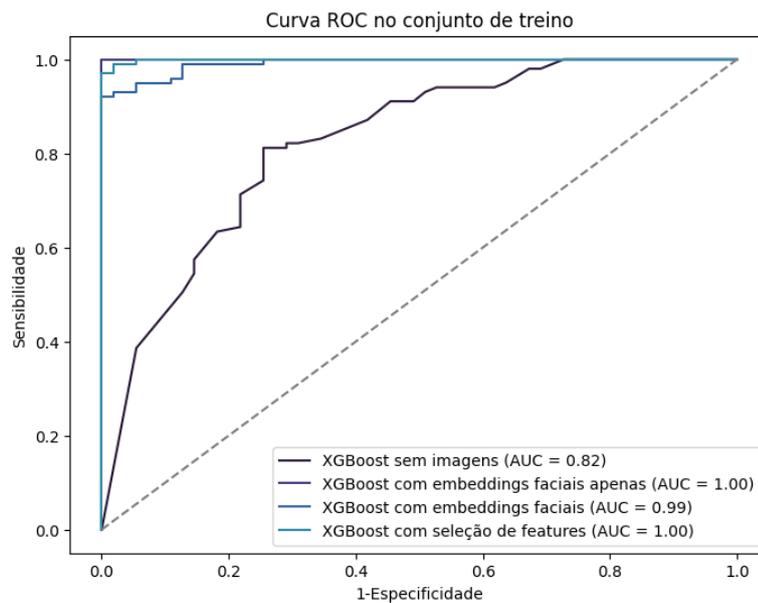


Figura 17 – Curva ROC para algoritmos XGBoost no conjunto de treino

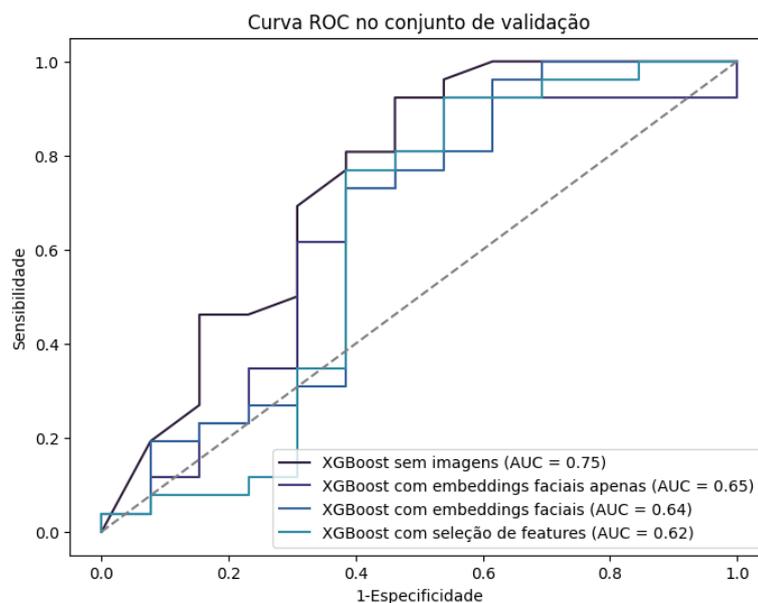


Figura 18 – Curva ROC para algoritmos XGBoost no conjunto de validação

5.5.2 Abordagens utilizando Redes Neurais

Com relação à segunda abordagem, que emprega redes neurais com modelo de imagens pré-treinado, as figuras 19 e 20 apresentam o desempenho desses modelos nos conjuntos de treino e validação respectivamente. A análise dos resultados revela um padrão consistente de generalização, onde todos os modelos treinados demonstraram desempenho comparável entre os conjuntos de treino e validação. Esta estabilidade evidencia a eficácia das técnicas de regularização implementadas, como Data Augmentation e dropout, que foram particularmente bem-sucedidas em controlar o overfitting nesta arquitetura.

Outro aspecto notável dos resultados é a uniformidade no desempenho dos diferentes modelos no conjunto de validação, com destaque para a abordagem que combina imagens das mãos e dados clínicos, que alcançou a melhor performance com uma AUC de 0.85. Esta combinação específica sugere que as características faciais extraídas das imagens das mãos, quando associadas aos dados clínicos, fornecem informações complementares e relevantes para o diagnóstico da doença renal crônica.

Entretanto, é importante ressaltar que dois aspectos fundamentais impossibilitam afirmar a superioridade entre as abordagens: o tamanho reduzido do conjunto de validação e a pequena variação nas performances obtidas (entre 0.85 e 0.80). Esta proximidade nos resultados, combinada com as limitações amostrais, não permite estabelecer uma diferença estatisticamente significativa entre as abordagens, indicando que todas as estratégias implementadas demonstram potencial similar para a tarefa de diagnóstico.

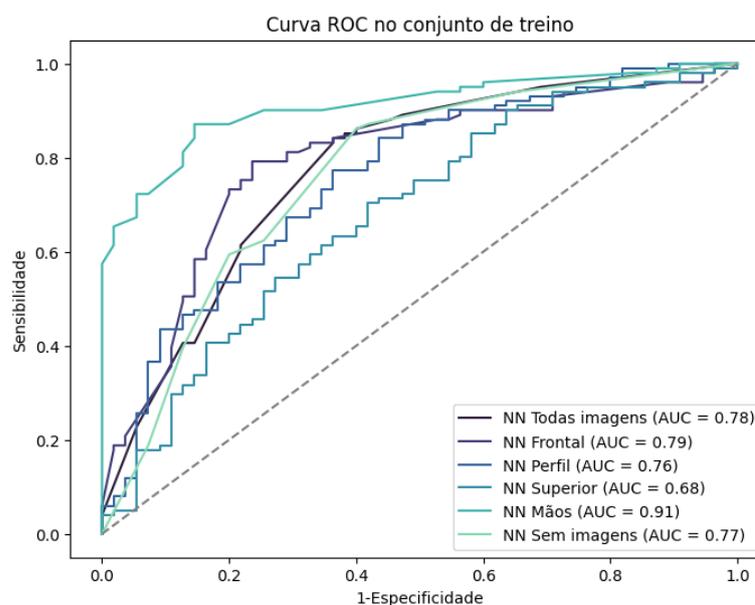


Figura 19 – Curva ROC para Redes Neurais no conjunto de treino

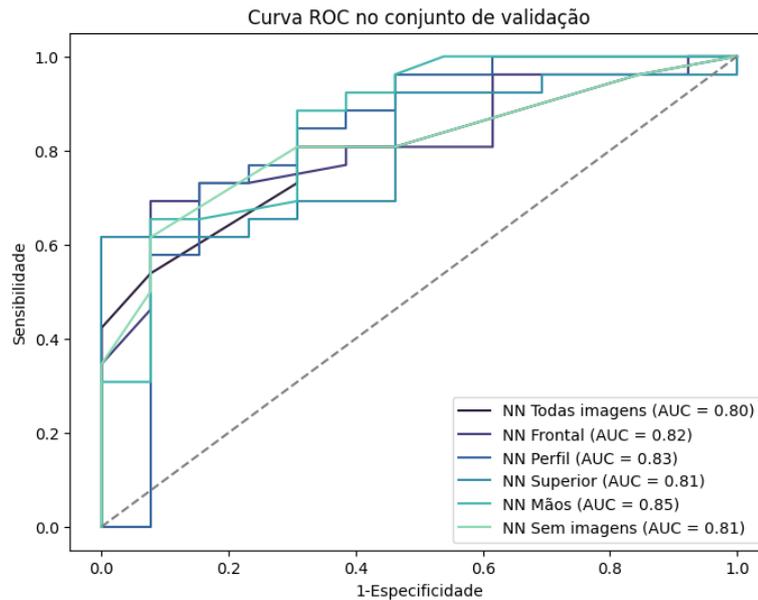


Figura 20 – Curva ROC para Redes Neurais no conjunto de validação

5.5.3 Avaliação dos modelos selecionados no conjunto de teste

Os modelos que demonstraram melhor performance no conjunto de validação foram selecionados para a análise final do algoritmo: XGBoost sem imagens, rede neural com imagens das mãos (NN Mãos) e rede neural com imagens de perfil (NN Perfil). A figura 21 ilustra o desempenho comparativo destes modelos no conjunto de teste. Os resultados indicam que o modelo que combina imagens das mãos com dados clínicos alcançou a melhor performance (AUC=0.77), seguido pelo modelo XGBoost utilizando apenas dados clínicos (AUC=0.73) e pela rede neural com imagens de perfil. Contudo, é fundamental destacar que a diferença de desempenho observada entre os dois melhores modelos não pode ser considerada estatisticamente significativa, devido tanto ao tamanho limitado do conjunto de teste quanto à proximidade entre os valores de AUC obtidos.

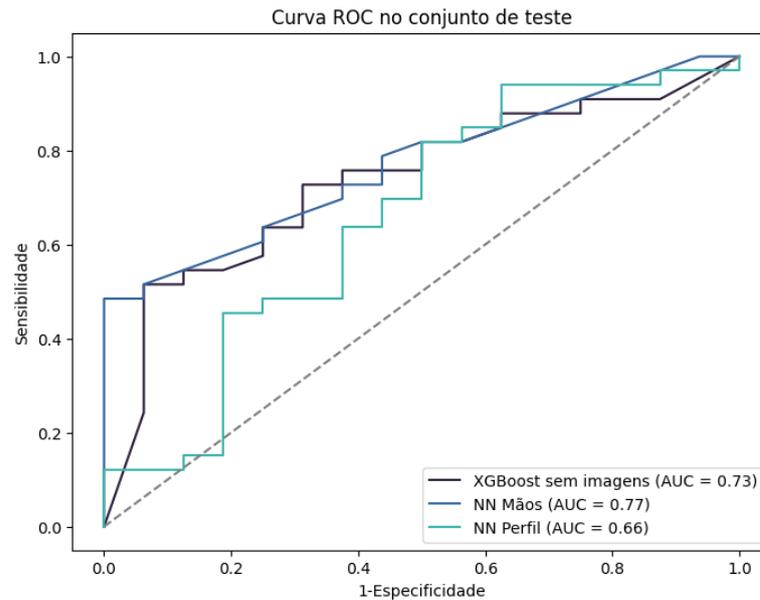


Figura 21 – Curva ROC para algoritmos selecionados no conjunto de teste

5.5.4 Testes aplicativo de diagnóstico

Por conta da significativa similaridade arquitetural entre o aplicativo de coleta e o de diagnóstico, a experiência e os resultados obtidos nos testes anteriores contribuíram substancialmente para otimizar o processo de validação do novo sistema. A reutilização de componentes previamente testados permitiu direcionar os esforços de teste para os novos elementos da arquitetura, com foco especial em duas áreas críticas: a integração da API com o AWS SageMaker e a nova interface de diagnóstico do aplicativo.

Na camada de integração com o AWS SageMaker, foram conduzidos testes extensivos para garantir a robustez e confiabilidade do sistema. Estes incluíram testes de conectividade para verificar a estabilidade da comunicação entre a API e o serviço em nuvem e testes de concorrência para assegurar o correto funcionamento em cenários de múltiplas requisições simultâneas. Adicionalmente, foram realizados testes de resiliência para validar o comportamento do sistema em casos de falha na comunicação ou indisponibilidade temporária do serviço.

Para a nova interface de diagnóstico, foram implementados testes abrangentes de usabilidade e funcionais. Estes envolveram a validação do processo de upload e processamento das imagens, verificação da correta exibição e formatação dos dados clínicos inseridos, e avaliação da apresentação dos resultados do diagnóstico. Foram realizados também testes de responsividade para garantir a adequada adaptação da interface em diferentes dispositivos e resoluções de tela, além de testes de integração end-to-end simulando o fluxo completo de diagnóstico, desde a entrada de dados até a apresentação do resultado final.

6 Considerações Finais

6.1 Conclusões do Projeto de Formatura

O principal objetivo deste projeto foi desenvolver um modelo de reconhecimento facial para auxiliar no diagnóstico da Doença Renal Crônica (DRC), desde a fase de coleta de dados até a criação do modelo final e do aplicativo que viabiliza seu uso.

Para atingir esse objetivo, foi necessário criar um aplicativo destinado à coleta de dados dos pacientes, que incluiu informações demográficas e imagens faciais. O aplicativo se destacou como um recurso essencial no processo de coleta, sendo amplamente utilizado pelos usuários. Sua simplicidade e interface intuitiva foram elogiadas, o que contribuiu significativamente para a adesão ao projeto e para a qualidade dos dados coletados.

No que diz respeito ao modelo de aprendizado de máquina, sua principal tarefa foi explorar a capacidade de extrair informações relevantes das imagens faciais dos pacientes que pudessem contribuir para o diagnóstico de DRC. As arquiteturas desenvolvidas permitiram alcançar um modelo com desempenho sólido, superando a métrica de referência proposta.

Entretanto, uma análise mais profunda dos resultados revelou um aspecto importante: embora o modelo que utilizou imagens tenha demonstrado um desempenho ligeiramente superior ao modelo baseado apenas em dados clínicos, essa diferença não se mostrou estatisticamente significativa dentro do conjunto de dados disponível. Esse resultado sugere que, para comprovar a efetividade do uso de imagens no diagnóstico, seria necessário um conjunto de dados maior e mais diversificado, capaz de captar melhor as nuances relevantes para o problema.

Em resumo, o projeto alcançou avanços importantes, como a criação de uma ferramenta funcional e a demonstração da viabilidade técnica do modelo. Contudo, a ausência de uma diferença estatisticamente significativa nos resultados indica que há espaço para aprimoramentos, especialmente na ampliação do conjunto de dados e na exploração de outras abordagens ou características das imagens que possam enriquecer o modelo.

6.2 Contribuições

Uma das principais contribuições deste projeto é o desenvolvimento do aplicativo de coleta de dados, uma solução prática e eficiente que continua sendo utilizada para expandir o dataset disponível. Esse aplicativo não apenas facilita a coleta de dados de novos pacientes, mas também cria a base para aprimorar o modelo existente e alcançar

resultados mais robustos no futuro. Além disso, ele abre a possibilidade de ampliar os tipos de dados coletados, permitindo novas linhas de pesquisa relacionadas à Doença Renal Crônica (DRC), como a análise de outros fatores clínicos ou sociais associados ao diagnóstico e progressão da doença.

Outra contribuição significativa da equipe foi a exploração inicial de modelos de reconhecimento facial aplicados à tarefa proposta. Esse trabalho envolveu a análise de diferentes abordagens, incluindo a escolha de arquiteturas e técnicas de aprendizado de máquina para identificar características relevantes das imagens dos pacientes. A equipe avaliou cuidadosamente os resultados obtidos, destacando tanto os pontos fortes quanto as limitações de cada abordagem.

6.3 Perspectivas de Continuidade

O projeto apresenta diversas possibilidades de continuidade e expansão, que podem enriquecer os resultados obtidos e abrir novas linhas de pesquisa.

Uma das principais perspectivas é a ampliação do conjunto de dados. Com o aumento do número de registros coletados, será possível treinar os modelos de Inteligência Artificial com um dataset mais robusto e verificar se os resultados obtidos permanecem consistentes ou apresentam melhorias significativas. Um conjunto de dados maior e mais diversificado pode contribuir para superar as limitações estatísticas identificadas neste projeto, permitindo a validação mais rigorosa do impacto das imagens no diagnóstico da Doença Renal Crônica.

Outra direção promissora é a exploração de novas arquiteturas de modelos de aprendizado de máquina. Entre as possibilidades, destaca-se a utilização de modelos pré-treinados especializados em reconhecimento facial, em vez de modelos genéricos de visão computacional. Isso pode aumentar a capacidade do modelo de identificar características faciais específicas que estejam associadas à DRC.

O aprimoramento do aplicativo de coleta de dados também representa uma oportunidade importante. Com a inclusão de novos tipos de informações clínicas, o aplicativo pode se tornar uma ferramenta ainda mais abrangente. Por exemplo, dados relacionados a anemia, complicações no metabolismo ósseo e mineral, ou outros indicadores de saúde associados à DRC podem ser integrados à análise. Isso permitiria explorar como a DRC impacta diferentes sistemas do corpo, ampliando o escopo do estudo e possibilitando novas abordagens diagnósticas ou preditivas.

Essas propostas de continuidade representam um avanço significativo no uso de tecnologia e inteligência artificial para auxiliar no diagnóstico e na compreensão da Doença Renal Crônica, consolidando o legado deste projeto como um ponto de partida para

contribuições ainda mais impactantes na área da saúde.

Referências

- AAD. *Kidney disease: 11 ways it can affect your skin*. American Academy of Dermatology Association, 2019. Disponível em: <<https://www.aad.org/public/diseases/a-z/kidney-disease-warning-signs>>. Acesso em: 05 mar. 2024. Citado 2 vezes nas páginas 12 e 13.
- ALCALDE, G. M. K. P. R. Expenses of the brazilian public healthcare system with chronic kidney disease. *Brazilian Journal of Nephrology*, 2018. Disponível em: <<https://www.scielo.br/j/jbn/a/7VzNY7GR8FvFHHvw3bKBRQx>>. Acesso em: 05 mar. 2024. Citado na página 11.
- BASTOS RACHEL BREGMAN, G. M. K. M. G. Doença renal crônica: frequente e grave, mas também prevenível e tratável. *Revista da Associação Médica Brasileira*, v. 56, n. 2, p. 248–253, 2010. Disponível em: <<https://www.scielo.br/j/ramb/a/3n3JvHpBFm8D97zJh6zPXbn/>>. Acesso em: 05 mar. 2024. Citado na página 12.
- DASHTBAN, A. et al. Identifying subtypes of chronic kidney disease with machine learning: development, internal validation and prognostic validation using linked electronic health records in 350,067 individuals. *eBioMedicine*, 2023. Disponível em: <<https://doi.org/10.1016/j.ebiom.2023.104489>>. Acesso em: 20 sep. 2024. Citado na página 20.
- DIVYA, C. D. et al. An efficient machine learning approach to nephrology through iris recognition. *eBioMedicine*, 2021. Disponível em: <<https://doi.org/10.1007/s44163-021-00010-4>>. Acesso em: 20 sep. 2024. Citado na página 21.
- KONG, X. et al. Automatic detection of acromegaly from facial photographs using machine learning methods. *EBioMedicine*, 2018. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5828367/>>. Acesso em: 15 mar. 2024. Citado 2 vezes nas páginas 13 e 23.
- KUMAR, K. et al. A deep learning approach for kidney disease recognition and prediction through image processing. *Applied Sciences*, 2023. Disponível em: <<https://doi.org/10.3390/app13063621>>. Acesso em: 20 sep. 2024. Citado na página 21.
- LEVINE, H. *The 5 Stages of Kidney Disease*. 2022. Health. Disponível em: <<https://www.health.com/condition/kidney-disease/kidney-disease-stages>>. Acesso em: 05 mar. 2024. Citado na página 12.
- LIN, S. et al. Feasibility of using deep learning to detect coronary artery disease based on facial photo. *European Heart Journal*, 2020. Disponível em: <<https://academic.oup.com/eurheartj/article/41/46/4400/5895010>>. Acesso em: 05 mar. 2024. Citado 2 vezes nas páginas 14 e 23.
- LUPI, O. et al. Manifestações cutâneas na doença renal terminal. *Anais Brasileiros De Dermatologia*, v. 86, n. 2, p. 319–326, 2011. Disponível em: <<https://www.scielo.br/j/abd/a/YGXjHZcNR9ZHKtrPzDtcFgH/>>. Acesso em: 05 mar. 2024. Citado na página 12.

- MALKINA, A. *Doença renal crônica*. 2023. Manual MSD. Disponível em: <<https://www.msmanuals.com/pt-br/casa/dist%C3%BArbios-renais-e-urin%C3%A1rios/insufici%C3%A2ncia-renal/doen%C3%A7a-renal-cr%C3%B4nica-drc>>. Acesso em: 05 mar. 2024. Citado na página 12.
- MINISTÉRIO DA SAÚDE. *14/3 – Dia Mundial do Rim 2019: Saúde dos Rins Para Todos*. Ceará, 2019. Disponível em: <<https://bvsmms.saude.gov.br/14-3-dia-mundial-do-rim-2019-saude-dos-rins-para-todos/>>. Acesso em: 05 mar. 2024. Citado na página 11.
- MINISTÉRIO DA SAÚDE. *Doenças Renais Crônicas*. Brasília, 2022. Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/drc>>. Acesso em: 05 mar. 2024. Citado na página 11.
- NORVIG, P.; RUSSELL, S. *Inteligência Artificial*. 3. ed. [S.l.]: Prentice Hall, 2013. Citado na página 17.
- PAN, Z. et al. Clinical application of an automatic facial recognition system based on deep learning for diagnosis of turner syndrome. *Endocrine*, 2020. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33170449/>>. Acesso em: 15 mar. 2024. Citado na página 13.
- QIANG, J. et al. Review on facial-recognition-based applications in disease diagnosis. *National Library of Medicine*, 2022. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9311612/>>. Acesso em: 05 mar. 2024. Citado 2 vezes nas páginas 13 e 22.
- QIN, B. et al. Automatic identification of down syndrome using facial images with deep convolutional neural network. *Diagnostics (Basel)*, 2020. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7400586/>>. Acesso em: 15 mar. 2024. Citado na página 13.
- SHAIKH, H. et al. Anemia of chronic renal disease. *National Library of Medicine*, 2023. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK539871/>>. Acesso em: 05 mar. 2024. Citado na página 12.
- STOREY, G. et al. 3dpalsynet: A facial palsy grading and motion recognition framework using fully 3d convolutional neural networks. *IEEE Access*, v. 7, n. 2, p. 121655–121664, 2019. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33170449/>>. Acesso em: 15 mar. 2024. Citado na página 13.
- WEI, R. et al. Deep-learning approach to automatic identification of facial anomalies in endocrine disorders. *Neuroendocrinology*, 2019. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31319415/>>. Acesso em: 15 mar. 2024. Citado 2 vezes nas páginas 13 e 24.
- WU, D. et al. Facial recognition intensity in disease diagnosis using automatic facial recognition. *National Library of Medicine*, 2021. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8621146/>>. Acesso em: 05 mar. 2024. Citado na página 22.

Apêndices

APÊNDICE A – Brochura da Pesquisa submetido para a Plataforma Brasil



UNIVERSIDADE DE SÃO PAULO

Escola Politécnica

Graduação em Engenharia de Computação

Gabriel Zambelli Scalabrini

João Victor Texeira Degelo

Johan Su Kwok

TRABALHO DE CONCLUSÃO DE CURSO

Uso de Reconhecimento Facial para Detecção de Doença Renal Crônica

São Paulo - São Paulo

2024

DESENHO

O projeto proposto visa desenvolver um modelo de inteligência artificial para auxiliar médicos nefrologistas no diagnóstico de doença renal crônica (DRC). Este modelo receberá imagens dos pacientes como entrada, além de informações sobre seu estado de saúde e sua meta é prever se o paciente possui DRC.

Para aprimorar o treinamento do modelo, é essencial coletar dados das faces de pacientes com DRC, assim como daqueles sem essa condição, a fim de estabelecer um grupo de controle. Essa coleta será realizada em colaboração com a Clinefro - Clínica de Nefrologia de Juazeiro, os quais serão responsáveis pela captura das fotos e pelo preenchimento do formulário proposto. Para auxiliar na coleta dessas informações, o grupo desenvolverá um aplicativo que será utilizado para o preenchimento do formulário de pesquisa e na captura de fotos. A pesquisa pretende coletar dados de cerca de 500 pacientes para alimentar o modelo.

O modelo de visão computacional adotado utilizará a técnica de aprendizado supervisionado. Para seu treinamento, será empregado um dataset composto pelas imagens dos pacientes rotuladas com seus respectivos diagnósticos. Esses dados serão divididos aleatoriamente em conjuntos de treino e teste, permitindo a seleção do melhor algoritmo e análise detalhada sobre o desempenho do modelo. Os detalhes sobre a captura das fotos dos pacientes, aplicação do questionário e o cronograma de execução das etapas do projeto estão descritos em seções posteriores deste documento. A coleta e o armazenamento dos dados dos pacientes será feita com os devidos cuidados de modo a garantir a privacidade dos pacientes e cumprir com as diretrizes estabelecidas pela legislação brasileira de proteção de dados pessoais.

INTRODUÇÃO

A doença renal crônica (DRC) é uma condição de longa duração que afeta os rins, órgãos que são responsáveis pela filtragem do sangue, remoção de resíduos metabólicos, impurezas, e regulação do equilíbrio de líquidos e pH do corpo. A DRC é caracterizada pela perda gradual e irreversível da função renal ao longo do tempo. Estima-se que aproximadamente dez milhões de brasileiros tenham a doença, dos quais 90 mil estão em diálise (estágio terminal da doença)⁽⁹⁾, e os gastos com o tratamento da DRC, e de outras doenças relacionadas, correspondem a 12,97% das despesas em saúde⁽¹⁰⁾.

Os rins são os órgãos responsáveis pela remoção de resíduos metabólicos, hormônios e substâncias tóxicas do corpo. Em particular, os rins das pessoas acometidas pela DRC não conseguem realizar adequadamente essa função. Os estágios da doença

renal crônica são classificados de acordo com a taxa de filtração glomerular (TFG), que indica a capacidade dos rins de filtrar o sangue. Sintomas iniciais podem passar despercebidos, mas conforme a DRC progride, podem surgir sintomas como fadiga, edema, pressão arterial elevada e alterações na frequência urinária. Complicações graves podem ocorrer, incluindo insuficiência renal, que exige tratamentos como diálise ou transplante renal⁽¹⁾.

Estágios da DRC

O primeiro estágio da DRC ocorre quando a TFG é superior a 90, o que indica que os rins estão funcionando adequadamente, porém apresentam sinais de danos leves. Alguns sintomas possíveis são pressão alta, presença de proteína na urina e inchaço nas pernas. O segundo estágio ocorre quando o nível da TFG está entre 60 e 89, o que representa que os rins ainda estão funcionando, porém com sinais claros de danos. Os sintomas ainda são leves, e alguns pacientes podem não apresentá-los, se caracterizando pelo acúmulo de fluidos, principalmente nas mãos e pés. No terceiro estágio, quando o TFG está entre 30 e 59, os rins já apresentam danos sérios. Além dos sintomas dos primeiros 2 estágios, pode ser observado a anemia, e distúrbio mineral ósseo, caracterizado pela não-filtragem de minerais responsáveis pela manutenção óssea, como cálcio e fósforo, o que pode afetar vasos sanguíneos e o coração. O quarto estágio é o último antes da falência renal, e ocorre quando os níveis de TFG estão entre 15 e 29, e diversos sintomas podem ocorrer, como icterícia, fadiga, inchaços, náusea, câimbra ou problemas cardíacos. O tratamento pode consistir na hemodiálise, em que o sangue circula por uma máquina que filtra e remove resíduos, ou diálise peritoneal, em que a diálise é feita através do abdômen, e pode ser realizada em casa. Por fim, o quinto estágio é caracterizado pela falência renal, quando a TFG é inferior a 15, indicando que os rins não funcionam mais. Neste caso, os resíduos metabólicos se acumulam no sangue e no corpo, gerando mudanças de coloração, ausência de urina, dificuldades respiratórias, e podendo levar ao óbito. Os tratamentos são a hemodiálise e o transplante renal.

O diagnóstico precoce é crucial para retardar a progressão da doença e gerenciar os sintomas. Medidas preventivas, como controle da pressão arterial, monitoramento dos níveis de glicose e adoção de um estilo de vida saudável, são fundamentais na gestão da doença renal crônica⁽²⁾.

Sintomas

A presença prolongada dos resíduos metabólicos no sangue gera alterações em diversos órgãos do corpo, e também na pele e face^(3,4). Dentre estas alterações, pode-se citar:

- *Palidez*⁽⁵⁾: Os rins dos pacientes, por não funcionarem corretamente, não conseguem produzir eritropoietina, hormônio responsável pelo estímulo à produção de glóbulos vermelhos, de maneira suficiente, o que leva à anemia. A maioria dos pacientes desenvolve anemia, que pode ocorrer no início da doença e piorar com o tempo.
- *Mudanças de coloração da pele*⁽⁶⁾: Além da palidez, o acúmulo de toxinas no corpo, devido à sua não-filtração pelos rins, pode levar à diversas mudanças na coloração da pele, como manchas escuras, tonalidade cinza, cistos e manchas que se assemelha à espinhas, e pele amarelada, com inchaços e linhas profundas.
- *Icterícia*⁽⁶⁾: A partir do estágio 4 da DRC, o paciente pode apresentar icterícia, incluindo o amarelamento do branco dos olhos.
- *Inchaço do rosto*⁽⁶⁾: Os rins removem o excesso de fluido e sal do corpo. Quando eles não conseguem mais fazer isso, os fluidos e o sal se acumulam no corpo, que é responsável pelo inchaço.
- *Xerose*⁽⁶⁾: Em casos nos quais a doença renal crônica está em estado terminal, necessitando de diálise e transplante de rim, é comum que a pele do paciente se torne muito seca, desenvolvendo rachaduras e escamas
- *Coceira aguda*⁽⁶⁾: Esse é um sintoma comum da doença renal crônica avançada. A coceira pode variar de irritante a incapacitante, podendo afetar toda ou a maior parte da pele o tempo todo. É comum os pacientes apresentarem marcas de arranhões, feridas, ou inchaços.

Além desses sintomas na face, esta doença também pode afetar a aparência das mãos e dos pés⁽⁶⁾. Além do inchaço, e da presença de bolhas, pessoas com doença renal avançada podem desenvolver:

- Coloração branca na parte superior das unhas, e coloração normal ou marrom avermelhada na parte inferior (*unhas de Lindsay* ou meio-e-meio).
- *Unhas pálidas*.
- Faixas brancas percorrendo uma ou mais unhas (*unhas de Muehrcke*).

Referências de reconhecimento facial para diagnóstico de doenças:

Para enriquecer o entendimento do contexto do projeto e explorar diferentes abordagens, conduzimos uma revisão da literatura sobre o emprego de algoritmos de reconhecimento facial no diagnóstico de doenças. No entanto, observamos uma lacuna de estudos específicos sobre a

aplicação desses algoritmos no diagnóstico da Doença Renal Crônica (DRC), tornando este trabalho pioneiro nesse campo. Durante a revisão, destacaram-se três artigos relevantes que foram consultados.

O primeiro deles é o “Review on Facial-Recognition-Based Applications in Disease Diagnosis”⁽¹¹⁾, que realizou uma análise da literatura para identificar os principais estudos na área de reconhecimento facial aplicados ao diagnóstico de doenças. Este estudo destacou práticas recomendadas para a captura de fotos, incluindo padronização da expressão facial dos pacientes, posicionamento da câmera e condições ambientais. Além disso, foram apresentados os principais algoritmos utilizados no diagnóstico e as circunstâncias em que cada um deles é mais apropriado.

O segundo artigo é o “Facial Recognition Intensity in Disease Diagnosis Using Automatic Facial Recognition”⁽¹²⁾, que realizou um compilado de 141 estudos que utilizaram o reconhecimento facial para o diagnóstico de doenças, e uma análise estatística sobre os fatores que influenciam a eficiência dos modelos de inteligência artificial. Os autores deste estudo definiram o conceito de Facial Recognition Intensity (FRI), um indicador que descreve a diferença entre as feições faciais da doença estudada e o grupo de controle. O FRI é calculado pela fórmula $N_f \times P_{max}$, onde N_f representa o número de fenótipos faciais relevantes para a doença, ou seja, o número de características faciais que são afetadas pela doença, e P_{max} , que representa a penetrância máxima destas características, isto é, o máximo percentual de indivíduos acometidos pela doença que apresentam esta característica. Por exemplo, uma doença como Down apresenta 9 fenótipos, e uma penetrância máxima de 100%, o que resulta em um FRI de 9.

O estudo em questão identificou que há uma forte correlação entre o FRI e a eficiência do modelo, enquanto outros fatores, como resolução da imagem utilizada no treino, tamanho do dataset utilizado para treino e algoritmo de inteligência artificial utilizado, não apresentam uma correlação tão significativa. Além disso, para alcançar uma eficiência maior no diagnóstico de doenças que apresentam um FRI baixo, o tamanho do dataset utilizado para treino deve ser bem maior do que para doenças com FRI elevado. Estas informações são de extrema relevância, visto que impactam em como a abordagem para este projeto deve ser realizada.

O terceiro artigo que se destacou, pela metodologia empregada no estudo, foi o “Feasibility of using deep learning to detect coronary artery disease based on facial photo”⁽¹³⁾, que analisou a utilização de um algoritmo de reconhecimento facial baseado em deep learning para a detecção da doença arterial coronária (CAD).

Além do treinamento do modelo de inteligência artificial para realizar o diagnóstico, os autores também realizaram diversas análises sobre quais eram os fatores que mais impactavam a eficiência deste. A primeira análise comparava o diagnóstico quando eram fornecidas ao modelo,

além das fotos faciais, informações sobre os pacientes tais como idade, sexo, sedentarismo, tempo de trabalho, utilização de álcool ou cigarro, entre outros, e quando somente eram fornecidas ao modelo as fotos faciais. O resultado observado era que a presença de informações adicionais sobre os pacientes não gerava uma diferença significativa na eficiência do modelo.

A segunda avaliação realizada foi sobre quais eram as feições faciais mais determinantes para o diagnóstico do modelo. Para isso, foi comparada a eficiência do diagnóstico utilizando as fotos sem alteração, e utilizando as fotos com determinadas partes do rosto, como olhos, nariz, ou região bucal, sendo cobertas. Com base nos resultados, foi construído um mapa de calor do rosto, indicando quais as características mais importantes no diagnóstico.

HIPÓTESE

É possível obter um diagnóstico preliminar de doença renal crônica de um paciente a partir de uma análise visual realizada por um algoritmo treinado por aprendizado de máquina.

PERGUNTAS DE PESQUISA

- 1) Em que medida o reconhecimento facial obtido através de algoritmos de IA aplicados a dataset com imagens rotuladas pode identificar o paciente com doença renal crônica?

OBJETIVO PRIMÁRIO

Criar um modelo de inteligência artificial que determinará, a partir da imagem da face, a probabilidade de uma pessoa possuir doença renal crônica

OBJETIVO SECUNDÁRIO

- 1) Analisar as imagens coletadas e buscar um melhor entendimento sobre quais são os indicativos de doença renal crônica que aparecem na face de pacientes;

METODOLOGIA PROPOSTA

A metodologia proposta para a coleta de dados dar-se-á através das seguintes etapas:

- 1) Apresentação do projeto ao Comitê de Ética da Escola;
- 2) Tendo a aprovação do CONEP, realizar coleta de dados dos pacientes com Doença Renal Crônica na Clinefro (Clínica de Nefrologia de Juazeiro). Tais dados serão obtidos por meio de fotografias e respostas a um questionário. Serão tiradas ao todo quatro fotos de cada participante: frontal, perfil direito, vista superior sobre a cabeça

e, por fim, das mãos dos participantes. É importante que o ambiente esteja bem iluminado com luz branca, com fundo branco e que o indivíduo mantenha uma expressão neutra. Além desses dados, também será preenchido um questionário contendo informações da saúde de cada pessoa. Ao final, serão coletados os mesmos dados de outros pacientes que sabidamente não possuam doença renal crônica para compor o grupo de controle. Para diminuir o trabalho manual nesta etapa, será desenvolvido um aplicativo com o objetivo de facilitar a obtenção dos dados;

- 3) Realizar uma análise dos dados coletados buscando um melhor entendimento sobre o público que foi atingido. Analisar eventuais desbalanceamentos dos dados e como resolvê-los.
- 4) Testar diversos algoritmos de inteligência artificial, buscando a abordagem que apresenta o melhor desempenho. Para avaliar cada algoritmo testado serão utilizadas técnicas de validação cruzada;
- 5) Análise do melhor algoritmo no grupo de dados separado para teste do modelo. Nessa análise, busca-se entender o que afetou seu desempenho e quais foram os dados mais relevantes para o diagnóstico da doença renal crônica.

TIPO DE PESQUISA

Pesquisa descritiva, uma vez que a pesquisa tem um objetivo claro de auxiliar a detecção de doença crônica renal por meio de análise das fotos.

TÉCNICAS DA PESQUISA

Pesquisa de levantamento com questionários e obtenção de fotos.

CORPUS

As fotos dos pacientes com DRC e o questionário das informações dos pacientes associados a um identificador numérico único, o software que analisará a foto e retornará o diagnóstico.

CRITÉRIO DE INCLUSÃO:

Os participantes deverão ser pacientes que contenham doença renal crônica, e pacientes que contenham ou não outras doenças para compor o grupo de controle. Os participantes deverão ser maiores de 18 anos e menores de 70 anos.

CRITÉRIO DE EXCLUSÃO:

Pessoas com defeitos faciais, doenças genéticas que levam a defeitos faciais, pacientes que tenham sofrido algum tipo de remodelação facial estética ou por trauma. Pacientes submetidos a paratireoidectomia (remoção cirúrgica da paratireoide). Pacientes menores de 18 anos ou maiores de 70 anos.

RISCOS: Existe a possibilidade de baixa eficiência do modelo no diagnóstico da doença.

BENEFÍCIOS: Entre os benefícios estão a oportunidade de desenvolvimento de uma ferramenta que facilitará o diagnóstico precoce de médicos nefrologistas, potencialmente melhorando a saúde e qualidade de vida da população.

CUSTOS ESTIMADOS DA PESQUISA

Para garantir que os pacientes que integram o grupo de controle não apresentem doença renal crônica, é necessário a realização de um exame de creatinina. Esse exame tem um custo médio de R\$10,00, dessa forma, considerando que 50% dos pacientes entrevistados sejam parte do grupo de controle, serão necessários 250 exames de creatinina. Com isso, ao final da pesquisa, o custo estimado total desse procedimento será de R\$2500,00.

METODOLOGIA PARA ANÁLISE DE DADOS

Nesta fase, serão examinadas várias informações coletadas dos pacientes, incluindo idade, peso, histórico familiar de diabetes e hipertensão, padrões alimentares e atividades físicas, e uso de medicamentos. A pesquisa em questão é de natureza quantitativa, centrada na avaliação de múltiplos modelos de inteligência artificial para a detecção de doença renal crônica. Além disso, serão conduzidas análises para determinar a relevância de cada informação (tais como características faciais, respostas a questionários, etc.) para o desempenho do modelo de inteligência artificial.

DESFECHO PRIMÁRIO

O desfecho primário deste projeto é o desempenho do modelo de inteligência artificial no reconhecimento e detecção de doença renal crônica (DRC) por meio da análise de imagens faciais, e será avaliada com base na capacidade do modelo em identificar corretamente os padrões faciais associados à DRC na população de estudo.

A eficiência do modelo será medida usando métricas padrão de avaliação de modelos de inteligência artificial, como a sensibilidade, especificidade, precisão e área sob a curva ROC (Receiver Operating Characteristic). A sensibilidade indicará a capacidade do modelo em identificar corretamente os casos positivos de DRC entre os indivíduos afetados pela doença. A especificidade avaliará a habilidade do modelo em descartar corretamente os casos negativos de DRC entre os indivíduos saudáveis. A precisão mostrará a proporção de predições corretas em relação ao total de predições feitas pelo modelo. E a área sob a curva ROC fornecerá uma medida global da capacidade discriminativa do modelo.

Avaliações adicionais serão conduzidas para investigar o impacto de fatores como idade, sexo e raça na precisão do modelo. Além disso, serão realizadas análises para determinar se existem diferenças na precisão do modelo entre diferentes estágios de progressão da DRC.

CRONOGRAMA DE EXECUÇÃO

	MAI/ 2024	JUN/ 2024	JUL/ 2024	AGO/ 2024	SET/ 2024	OUT/ 2024	NOV/ 2024	DEZ/ 2024
Submissão ao comitê de ética								
Desenvolvimento de aplicativo para coleta de dados								
Coleta de dados dos pacientes*								
Análise dos dados coletados e desenvolvimento do modelo de inteligência artificial								
Desenvolvimento de aplicativo para disponibilização do modelo aos médicos								
Apresentação do projeto de formatura								

*A pesquisa com os participantes somente terá início após a aprovação do Comitê de Ética.

BIBLIOGRAFIA

1. MINISTÉRIO DA SAÚDE. *Doenças Renais Crônicas*. Brasília, 2022. Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/d/drc>>. Acesso em: 05 mar. 2024.
2. BASTOS RACHEL BREGMAN, G. M. K. M. G. Doença renal crônica: frequente e grave, mas também prevenível e tratável. *Revista da Associação Médica Brasileira*, v. 56, n. 2, p. 248–253, 2010. Disponível em: <<https://www.scielo.br/j/ramb/a/3n3JvHpBFm8D97zJh6zPXbn/>>. Acesso em: 05 mar. 2024.
3. MALKINA, A. *Doença renal crônica*. 2023. Manual MSD. Disponível em: <<https://www.msdmanuals.com/pt-br/casa/dist%C3%BArbios-renais-e-urin%C3%A1rios/insufici%C3%A2ncia-renal/doen%C3%A7a-renal-cr%C3%B4nica-drc>>. Acesso em: 05 mar. 2024.
4. LUPI, O. et al. Manifestações cutâneas na doença renal terminal. *Anais Brasileiros De Dermatologia*, v. 86, n. 2, p. 319–326, 2011. Disponível em: <<https://www.scielo.br/j/abd/a/YGXjHZcNR9ZHKtrPzDtcFgH/>>. Acesso em: 05 mar. 2024.
5. SHAIKH, H. et al. Anemia of chronic renal disease. *National Library of Medicine*, 2023. Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK539871/>>. Acesso em: 05 mar. 2024.
6. AAD. Kidney disease: 11 ways it can affect your skin. *American Academy of Dermatology Association*, 2019. Disponível em: <<https://www.aad.org/public/diseases/a-z/kidney-disease-warning-signs>>. Acesso em: 05 mar. 2024.
7. WEI, R. et al. Deep-learning approach to automatic identification of facial anomalies in endocrine disorders. *Karger*, 2019. Disponível em: <<https://karger.com/nen/article-abstract/110/5/328/220661/Deep-Learning-Approach-to-Automatic-Identification>>. Acesso em: 05 mar. 2024.

8. LEVINE, H. The 5 Stages of Kidney Disease. 2022. *Health*. Disponível em: <<https://www.health.com/condition/kidney-disease/kidney-disease-stages>>. Acesso em: 05 mar. 2024.

9. MINISTÉRIO DA SAÚDE. 14/3 – *Dia Mundial do Rim 2019: Saúde dos Rins Para Todos*. Ceará, 2019. Disponível em: <<https://bvsmis.saude.gov.br/14-3-dia-mundial-do-rim-2019-saude-dos-rins-para-todos/>>. Acesso em: 05 mar. 2024.

10. ALCALDE, G. M. K. P. R. Expenses of the Brazilian public healthcare system with chronic kidney disease. *Brazilian Journal of Nephrology*, 2018. Disponível em: <<https://www.scielo.br/j/jbn/a/7VzNY7GR8FvFHHvw3bKBRQx>>. Acesso em: 05 mar. 2024.

11. QIANG, J. et al. Review on facial-recognition-based applications in disease diagnosis. *National Library of Medicine*, 2022. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9311612/>>. Acesso em: 05 mar. 2024.

12. WU, D. et al. Facial recognition intensity in disease diagnosis using automatic facial recognition. *National Library of Medicine*, 2021. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8621146/>>. Acesso em: 05 mar. 2024.

13. LIN, S. et al. Feasibility of using deep learning to detect coronary artery disease based on facial photo. *European Heart Journal*, 2020. Disponível em: <<https://academic.oup.com/eurheartj/article/41/46/4400/5895010>>. Acesso em: 05 mar. 2024.

APÊNDICE B – Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) Senhor(a),

O(a) Sr(a). está sendo convidado(a) a participar da pesquisa: **“Uso de Reconhecimento Facial para Detecção de Doença Renal Crônica”**, que tem por objetivo principal o desenvolvimento de um modelo de inteligência artificial capaz de auxiliar no diagnóstico de doença renal crônica a partir de imagens da face e outras informações da saúde do participante. Para possibilitar o treino desse modelo de inteligência artificial, é necessário a coleta dessas informações de pessoas, de modo a compor o banco de dados que será utilizado no projeto.

Nesta pesquisa serão realizadas capturas fotográficas do seu rosto e mãos. Além disso, solicitaremos o preenchimento de um questionário de perfil, com perguntas as quais englobam históricos de doenças e hábitos físicos e alimentares.

Os participantes ficarão resguardados das seguintes garantias:

- a Será garantida a plena liberdade do participante da pesquisa de recusar-se a participar ou retirar seu consentimento, em qualquer fase da pesquisa, sem penalização alguma.**
- b As imagens fotografadas e os dados de saúde dos participantes são confidenciais e não serão divulgados, sendo somente utilizados para o treinamento e teste do modelo a ser desenvolvido.**
- c O participante terá direito a ressarcimento, caso exista despesas decorrentes da sua participação.**
- d O participante terá direito de buscar indenização conforme as leis brasileiras, caso tenha algum dano/prejuízo decorrente de sua participação na pesquisa.**

Entre os possíveis riscos da pesquisa estão a possibilidade do participante sentir algum desconforto ao fornecer as imagens da sua face e mãos, e informações sobre sua saúde física para a utilização durante o projeto.

Assinatura do participante

Assinatura do pesquisador responsável

Data: ____/____/____

Declaramos que a guarda e o armazenamento de seus dados pessoais ficarão sob responsabilidade dos pesquisadores e que você terá o direito a uma cópia deste Termo. Sempre que necessário você poderá entrar em contato com os pesquisadores para esclarecimento de eventuais dúvidas. A pesquisadora responsável para contato é: Anarosa Alves Franco Brandão, telefone 11-3091-0626 e-mail: anarosa.brandao@usp.br integrante do **Laboratório de Técnicas Inteligentes - LTI**, Escola Politécnica da Universidade de São Paulo - EPUSP, situado na Av. Prof. Luciano Gualberto, trav.3, n.158, Prédio de Engenharia Elétrica, sala C2-50, Cidade Universitária, CEP: 05508-970 - São Paulo - SP/Brasil, Telefone: +55-11-3091-5397 Fax: +55 11 3091 5294.

Declaramos que este Termo de Consentimento Livre e Esclarecido está em conformidade com as disposições estabelecidas na Resolução CNS N° 510/2016, que regulamenta as pesquisas envolvendo seres humanos. Todos os participantes foram devidamente informados sobre os objetivos, métodos, riscos e benefícios da pesquisa, garantindo assim o seu pleno entendimento e consentimento voluntário.

Sua participação é importante e voluntária e gerará informações importantes para o desenvolvimento de métodos mais eficazes para apoiar o diagnóstico da doença renal crônica, beneficiando assim a comunidade médica e a prevenção de tal doença. Este termo será assinado em duas vias, pelo(a) senhor a) e pela responsável pela pesquisa, ficando uma via em seu poder.

Para esclarecimentos adicionais ou em caso de dúvidas sobre este estudo ou sobre seus direitos como participante, entre em contato com o Comitê de Ética em Pesquisa (CEP) da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH-USP) nos canais descritos embaixo. O CEP é um órgão colegiado interdisciplinar e independente, constituído nos termos da Resolução no. 466, do Conselho Nacional de Saúde, criado para defender os interesses dos participantes de pesquisas em sua integridade e dignidade.

E-mail: cep-each@usp.br

Telefone: (11) 3091-1046

Endereço: Arlindo Bettio, 1000. CEP: 03828-000, Vila Guaraciaba, São Paulo/SP, Prédio 1 | Sala T1

Horários de Atendimento: Segundas às sextas-feiras: das 09:00 às 11:00 e das 8414:00 às 16:00

De posse das informações sobre a pesquisa intitulada **“Uso de Reconhecimento Facial para Detecção de Doença Renal Crônica”** concordo voluntariamente em participar dela, de forma livre e esclarecida.

Assinatura do participante

Assinatura do pesquisador responsável

Data: ____/____/____