

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE ENGENHARIA DE COMPUTAÇÃO

ANDREWS VIEIRA BERNI

**Uso de Inteligência Artificial
para Apoio na Identificação de Catarata**

Monografia apresentada como requisito parcial para a obtenção do grau de Bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Ronaldo Husemann

Coorientador: Prof. Marcelo Negreiros

Coorientador: Dra. Aline Lutz de Araujo

Coorientador: Dr. Rodolfo Silva

Coorientador: Dr. Roberto Umpierre

Porto Alegre

2021

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões Mendes

Vice-Reitora: Prof^ª Patricia Helena Lucas Pranke

Pró-Reitoria de Ensino (Graduação e Pós-Graduação): Prof^ª Cíntia Inês Boll

Diretora do Instituto de Informática: Prof^ª Carla Maria Dal Sasso Freitas

Diretora da Escola de Engenharia: Prof^ª Carla Schwengber Ten Caten

Coordenador do Curso de Engenharia de Computação: Prof. Walter Fetter Lages

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Bibliotecária-Chefe da Escola de Engenharia: Rosane Beatriz Allegretti Borges

AGRADECIMENTOS

Agradeço a minha mãe, Mara Matos, a minha linda esposa, Thayane Puls, e ao meu filho, Vitor Puls, pois sem eles tudo seria mais difícil de conquistar. Foram eles que me fizeram ter foco e nunca desistir.

Aos colegas de trabalho e faculdade Vicente Guedes, Thiago Martins e Guilherme Magalhães, por estarem presentes tanto nos momentos mais complicados desta jornada, quanto nos momentos de descontração, e por terem contribuído muito com a minha formação acadêmica.

A todos do Telessauders e, especialmente e principalmente, as oftalmologistas Dra Aline Lutz de Araujo e Dra Carolina Campos, que sempre se mostram dispostas na retirada de dúvidas e definições médicas para este trabalho.

Ao meu professor orientador, Ronaldo Husemann, coorientadores Marcelo Negreiros e Valter Roesler, pois graças a paciência e comprometimento deles me ajudaram a concluir essa etapa. Lembrar também de todos os professores e funcionários desta universidade, pela competência, pela dedicação, e pelos conhecimentos que continuamente transmitem a todos os alunos.

RESUMO

Catarata é a principal causadora de perda de visão (cegueira) em nível mundial. De modo geral a catarata somente é clinicamente relevante se a capacidade visual do paciente for afetada significativamente, uma vez que uma opacidade no cristalino pode existir sem qualquer sintoma. Sistemas de graduação não automática de catarata provaram ser uma maneira reprodutível e objetiva, embora possuam ainda limitações e necessitem de mais refinamento. O presente projeto propõe o uso de utilizar redes neurais para auxiliar ao diagnóstico de catarata do sujeito, a partir de imagens de exames oftalmológicos realizados no biomicroscópio ocular e no retinógrafo e um conjunto de dados clínicos obtidos a partir dos laudos das imagens.

Palavras-chave: CNN. Catarata. Inteligência Artificial.

Use of Artificial Intelligence for Support in Cataract Identification

ABSTRACT

Cataracts are the main cause of vision loss (blindness) worldwide. In general, cataracts are only clinically relevant if the patient's visual ability is significantly affected, since opacity in the lens can exist without any symptoms. Non-automatic cataract grading systems have proven to be a reproducible and objective way, although they still have limitations and need further refinement. The present project proposes the use of neural networks to aid the diagnosis of the subject's cataract, based on images of ophthalmological exams performed on the ocular biomicroscope and on the retinograph and a set of clinical data obtained from the reports of the images.

Keywords: CNN. Cataract. Artificial intelligence.

LISTA DE FIGURAS

Figura 1.1 – Principais tipos de cataratas senis	10
Figura 1.2 – Sistema LOCS III de graduação de opacidades	11
Figura 1.3 – Diagrama da solução proposta pelo projeto	12
Figura 1.4 – Imagem de fundo de olho	13
Figura 1.5 – Imagem periférica de olho	13
Figura 2.1 – Catarata cortical que não trabalha a visão.....	17
Figura 2.2 – Outra patologia que afeta a pupila	17
Figura 3.1 – Distribuição 522 das imagens do banco nos três grupos de trabalho.....	19
Figura 3.2 – Distribuição 721 das imagens do banco nos três grupos de trabalho.....	20
Figura 3.3 – Diagrama mostrando as distribuições das imagens do banco	21
Figura 3.4 – Diagrama das redes neurais convolutivas usadas na etapa experimental	23
Figura 3.5 – Algoritmo de transformação das imagens.....	28
Figura 3.6 – Exemplo de transformação para aumento artificial	28
Figura 3.7 – Estratégia de aumento artificial offline do conjunto de treinamento	29
Figura 3.8 – Estratégia de aumento artificial online do conjunto de treinamento.....	29
Figura 3.9 – Curvas de taxa de acertos para os conjuntos de treino e validação para 30.....	30
Figura 3.10 – Influência do número de épocas de treinamento na análise das curvas de treinamento: (a) 100 épocas e (b) 1000 épocas	31
Figura 4.1 – Processo de identificação da zona de interesse na imagem (pupila).....	37

LISTA DE TABELAS

Tabela 3.1: Número de imagens nas diferentes distribuições de conjuntos de treino, validação e teste	21
Tabela 3.2: Tamanho das imagens de pupila em pixels	25
Tabela 3.3: Tamanho em pixels das imagens usadas para treinamento das redes	26
Tabela 3.4 – Parâmetros de transformação das imagens para aumento artificial do conjunto de treinamento	27
Tabela 4.1 – Resultados para imagens de pupila com 70% usado para treino (CAT721)	32
Tabela 4.2 – Resultados para imagens de pupila com 50% usado para treino (CAT522)	33
Tabela 4.3 – Resultados para imagens de pupila com 50% usado para treino (CAT521)	33
Tabela 4.4 – Resultados para imagens de olho inteiro com 70% usado para treino. conjunto de teste possui 200 imagens.....	35
Tabela 4.5 – Resultados para imagens de olho inteiro com 50% usado para treino. O conjunto de teste possui 500 imagens.....	36
Tabela 4.6: Resultados para imagens de olho inteiro com 50% usado para treino. Usando o mesmo conjunto de imagens de teste do caso 721 para permitir a comparação dos resultados	36
Tabela 4.7 – Melhores resultados em cada tipo de treinamento.....	37

LISTA DE ABREVIATURAS E SIGLAS

SUS	Sistema Único de Saúde
APS	Atenção Primária à Saúde
OMS	Organização Mundial de Saúde
PSC	Posterior Subcapsular Cataract
CNN	Convolutional Neural Networks
CPU	Unidade de Central de Processamento
GPU	Unidades de Processamento Gráfico
DICOM	Digital Imaging and Communications in Medicine
PNG	Portable Network Graphics
RGB	Red Green Blue

SUMÁRIO

1	INTRODUÇÃO	9
1.1	Contextualização	9
1.2	Objetivo	11
1.3	Trabalhos Relacionados	12
1.4	Estrutura do trabalho	13
2	Banco de Dados	15
2.1	Aquisições das imagens	17
2.2	Relação imagem e laudo	17
2.3	Quantidade de amostras válidas	16
3	METODOLOGIA EXPERIMENTAL	19
3.1	Divisão das imagens para treinamento, validação e teste	20
3.2	Arquitetura das redes neurais convolucionais (CNNs)	22
3.2.1	CNN padrão	22
3.2.2	CNN com dropout	22
3.2.3	CNN pré-treinada	23
3.3	Tipo de imagem de entrada	24
3.4	Resolução das imagens	25
3.5	Aumento artificial das imagens de treinamento	27
3.5.1	Estratégia sem aumento artificial	27
3.5.2	Estratégia com aumento artificial de imagens (offline) do conjunto de treinamento	27
3.5.3	Estratégia aumento artificial de imagens (online) do conjunto de treinamento	28
3.6	Número de épocas de treinamento	29
3.7	Plataformas de desenvolvimentos	30
4	Resultados	31
4.1	Resultados para imagens de pupila	31
4.2	Resultados para imagens de olho inteiro	33
5	ANALISE DE RESULTADOS	39
6	CONCLUSÃO	40
	REFERÊNCIAS	42
	APÊNDICE A – Trabalho de Graduação I	44

1 INTRODUÇÃO

O projeto TelessaúdeRS/UFRGS foi estabelecido em 2007 e tem seu principal foco ajudar e aprimorar o SUS por meio de teleconsultorias, matérias de apoio ou cursos. Tendo esse foco na população que utiliza o SUS, em 2017 foi criado o serviço Teleoftalmo: Olhar Gaúcho, o qual efetua exames oftalmológicos em pacientes consome o serviço público de saúde. A iniciativa oferece os seguintes exames: aferição da acuidade visual, refração, medida da pressão intraocular, documentação fotográfica do segmento anterior e do fundo do olho, avaliação das pálpebras, da motilidade ocular extrínseca e dos reflexos pupilares, por meio de imagens e exames físico síncronos. Até julho de 2021 foram realizados cercar de 33 mil exames oftalmológicos usando uma ferramenta própria de vídeo colaboração e equipamentos oftalmológicos com alta tecnologia. A estruturação do serviço permite a interação em tempo real entre médicos especialistas e pacientes em pontos remotos no estado Rio Grande do Sul, cobrindo o estado com o atendimento.

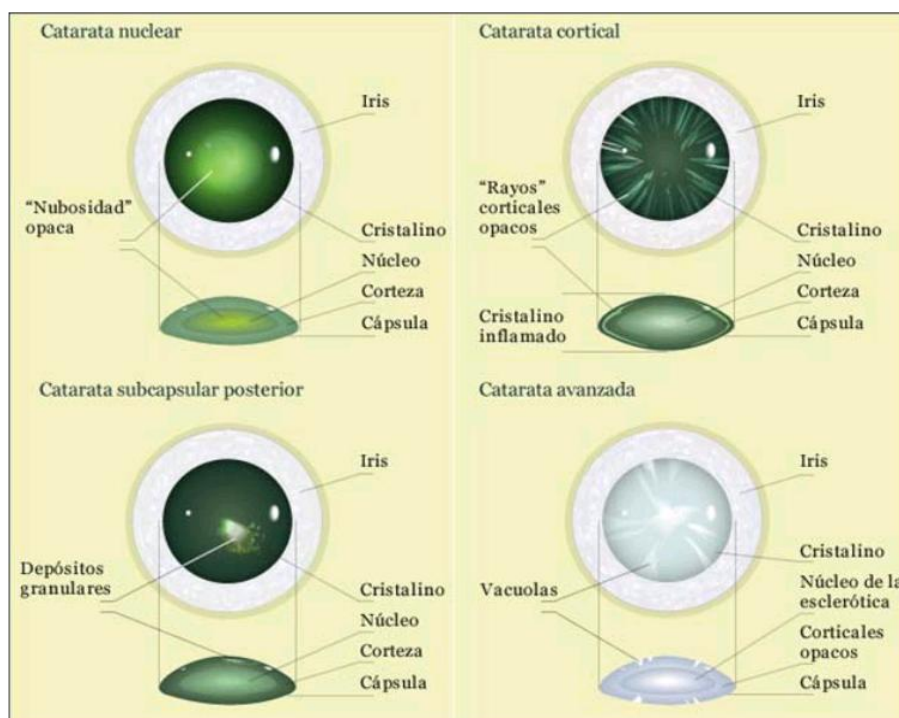
1.1 Contextualização

Catarata é a principal causadora de perda de visão (cegueira) em nível mundial e um dos principais fatores que levam à diminuição visual no mundo (PASCOLINI,2010), sendo que ainda não existe tratamento médico que possa prevenir sua formação ou progressão no cristalino do olho humano. A grande maioria de todos os tipos de cataratas são as relacionadas à idade e se constituem em um grande problema de saúde pública em nível mundial (MALHOTRA,2008).

A catarata é, com exceção dos países desenvolvidos, a maior causa de cegueira em todo o mundo e, conforme a OMS, corresponde por 48% dos casos mundiais de cegueira. A doença é definida como perda de transparência do que difrate a luz, acarretando um efeito adverso na visão. As alterações podem levar desde pequenas distorções visuais até a cegueira. O aumento gradual da expectativa de vida provocou consequente aumento da prevalência dessa enfermidade nas últimas décadas. Sua prevalência foi estimada em 2,5% entre 40 e 49 anos, 6,8% entre 50 e 59 anos, 20% entre 60 e 69 anos, 42,8% entre 70 e 79 anos e 68,3% em maiores de 80 anos (KARA-JOSÉ, 2008). Conforme a OMS, a incidência anual de catarata é estimada em 0,3% ao ano, o que no Brasil representaria cerca de 550 000 novos casos de catarata por ano (KARA-JOSÉ, 2008).

A doença pode ter etiologia senil, congênita, traumática ou secundária (KARA-JOSÉ, 2008). A principal forma da catarata é a senil, sobre a qual se concentram os estudos epidemiológicos e de prevalência. A catarata senil, com maior incidência na população acima de 50 anos (KARA-JOSÉ, 2008), é a mais comum. Dentro do grupo senil há três subtipos: nuclear, cortical e subcapsular posterior (PSC), de acordo com a localização da opacidade no cristalino. Em muitos pacientes mais de um tipo de catarata estão presentes ao mesmo tempo. A Figura 1.1 apresenta os tipos mais comuns de cataratas senis, incluindo a catarata em estágio avançado ou madura (LOPEZ; GARCIA; MENDEZ, 2016).

Figura 1.1 - Principais tipos de cataratas senis.

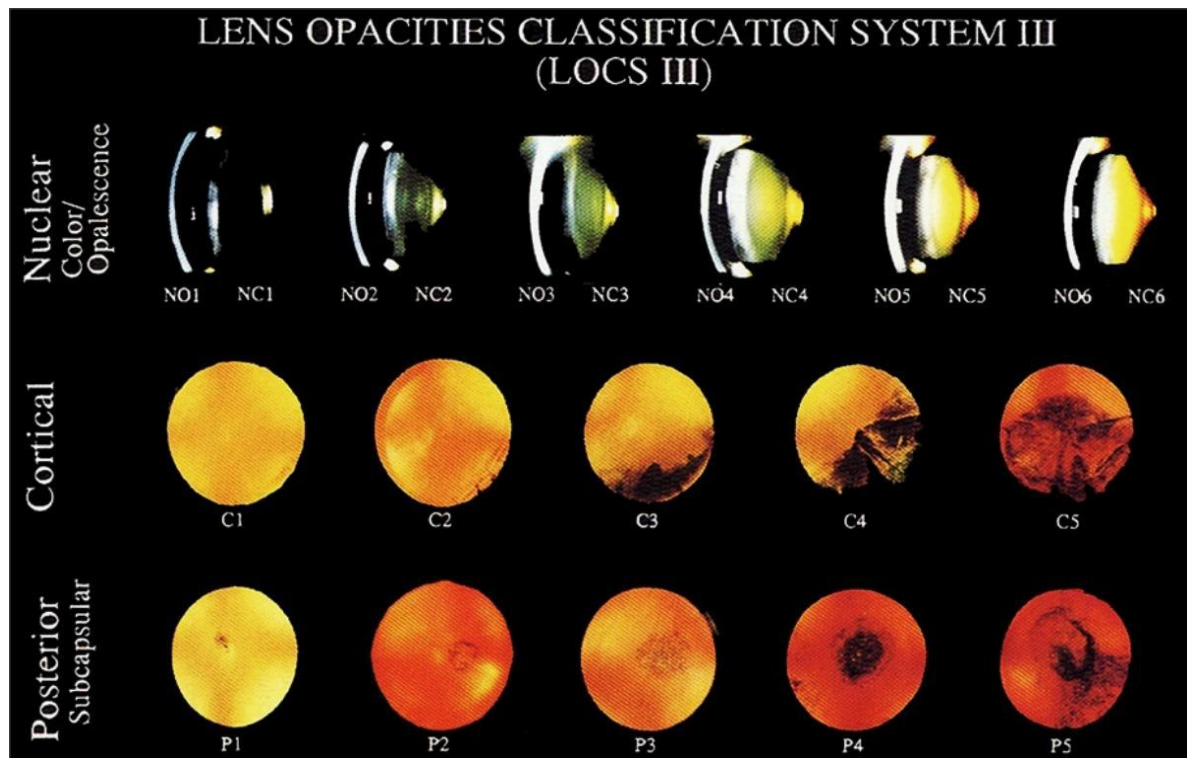


Fonte: LOPEZ; GARCIA; MENDEZ (2016).

Para ser possível detectar a presença da doença e estudar o seu progresso e permitir comparação entre casos existe a necessidade de se determinar se a doença está em um dos estágios iniciais ou num estágio mais avançado. Existem diversos sistemas para classificação de opacidades na literatura e a Figura 1.2 apresenta um deles, o sistema de graduação LOCS III (LOPEZ; GARCIA; MENDEZ, 2016). O sistema utiliza uma série de imagens de referência para permitir a classificação do grau de desenvolvimento da doença, baseados em imagens obtidas nos exames oftalmológicos de lâmpada de fenda e de retroiluminação. No sistema LOCSIII o exame de lâmpada de fenda é usado para a graduação de cataratas nucleares (primeira linha na Figura 1.2) e as imagens do exame de retroiluminação são usadas para

gradação de cataratas corticais e posterior subcapsulares (linhas 2 e 3 na Figura 1.2). Todos exames são feitos com a pupila dilatada.

Figura 1.2 - Sistema LOCS III de gradação de opacidade

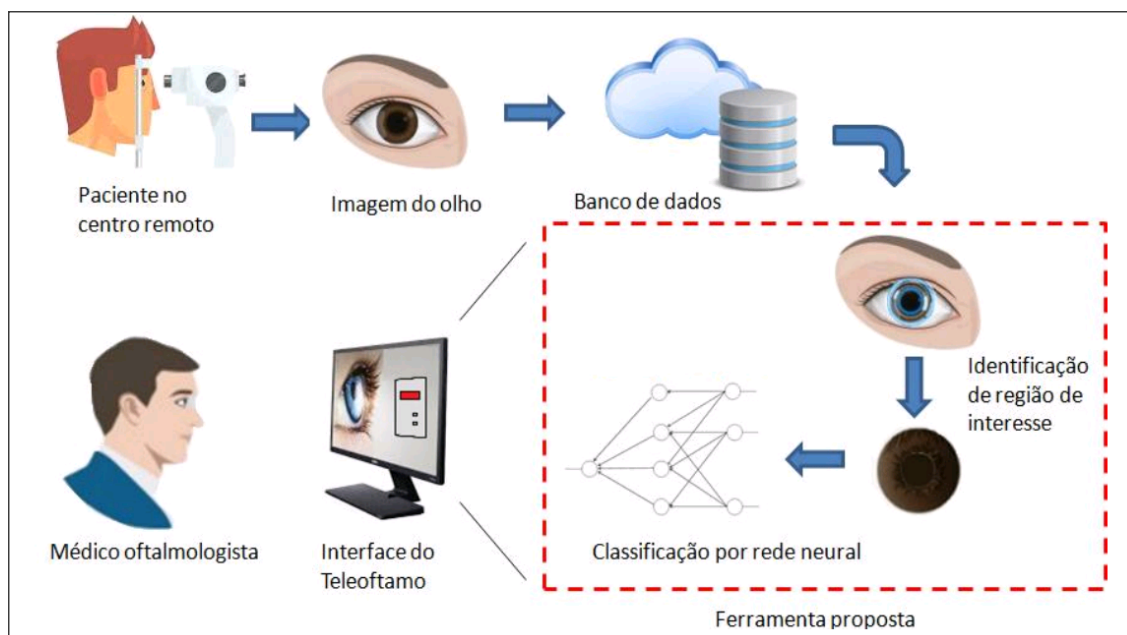


Fonte: Cataract Detection Techniques: A Review (2016).

1.2 Objetivo

Sistemas que utilizam inteligência artificial (IA) têm se mostrado capazes de detectar diferentes doenças a partir de imagens e dados médicos incluindo a doença catarata (GAO; LIN; WONG,2015). Logo, neste trabalho foi proposto o desenvolvimento de um projeto piloto que integre uma ferramenta de IA sobre a solução de vídeo colaboração do Teleoftalmo para auxiliar no diagnóstico da doença catarata, de acordo com o diagrama apresentado na Figura 1.3.

Figure 1.3 - Diagrama da solução proposta pelo projeto.



Fonte: HUSEMANN (2019)

O projeto buscou utilizar técnicas de inteligência artificial para processar imagens oftalmológicas de exames de retroiluminação com o objetivo final de criar uma ferramenta que auxilie no diagnóstico da doença catarata, obtendo uma classificação automática entre paciente normal ou com catarata. A fim de obter um elevado índice de acerto as redes neurais profundas exigem um treinamento com muitas imagens, na ordem de dezenas a centenas de milhares de amostras, ocasionando uma alta demanda de tempo e de recursos computacionais. Para reduzir o tempo de treinamento pode ser usada uma rede neural convolucional (CNN) pré-treinada que seja aberta e disponível para a comunidade. Este tipo de rede já foi extensamente treinado em alguns casos por alguns milhões de imagens. Ao utilizar essa estratégia apenas os estágios finais da rede são refinados para a classificação desejada.

1.3 Trabalhos Relacionados

Existem alguns trabalhos (CARNEIRO; SANTOS,2020), relacionados a detecção de cataratas com redes neurais. Porém, as imagens utilizadas para treinamentos são imagens de fundo de olho (Figura 1.4) na qual os equipamentos que tiram esse tipo de imagem não são de fácil acesso para quem utiliza o SUS. As imagens usadas no projeto são imagem de simples acesso da população em geral, Figura 1.5 por exemplo.

Figura 1.4 - Imagem de fundo de olho.



Fonte: Elaborado pelo autor

Figura 1.5 - Imagem periférica de olho.



Fonte: HUSEMANN (2019)

Além disso, o assunto do desse trabalho foi inicializada com a pesquisa TeleDIAC (HUSEMANN,2019) que foi uma pareceria entre o TelessaudeRS e Microsoft para o uso da plataforma Azure. No TeleDIAC o tipo de imagem usada foi a periférica, porém foi feito um filtro para usar somente a pupila (parte mais central do olho).

1.4 Estrutura do trabalho

O Capítulo 2 traz uma explicação de como foi a montagem do banco de dados para o projeto. Como o projeto tinha participação de outros pesquisadores, esta etapa foi a qual tive mais participação e tempo alocado, visto que por ser um colaborador do projeto TelessaudeRS entendia melhor da estrutura do serviço Teleoftalmo, uma vez que para o treinamento da IA

seriam necessárias as imagens e seus laudos, o que não estava estruturado de forma fácil e era fundamental para o sucesso do treinamento.

A forma como os dados serão usados será explicado no Capítulo 3. Esse capítulo descreve as divisões utilizadas para treinamento, teste e validação, além do tipo de arquitetura da rede neural, o agrupamento das imagens, o tamanho das imagens, as épocas de processamentos, qual algoritmo usado e plataformas usadas para codificação.

Já no Capítulo 4 são apresentados os resultados dos processos realizados nos capítulos anteriores (com a estruturação do banco de dados e a análise de algumas redes). Por conseguinte, no capítulo 5 é apresentada uma análise dos resultados do tópico anterior. Após toda a explanação do projeto, o último capítulo traz uma conclusão de todas as etapas anteriores. Além do mais, descreve possíveis melhorias e trabalhos futuros.

2 BANCO DE DADOS

Nesta seção é apresentado o desenvolvimento do banco de imagens usado para o treinamento das redes. Essa foi uma das etapas mais longas do projeto devido ao elevado número de imagens a serem analisadas e da necessidade de desenvolver ferramentas específicas de busca e seleção de imagens devido à problemas de integração entre as bases de dados de imagens e de laudos presentes no Teleoftalmo.

2.1 Aquisições das imagens

Na área médica, o formato para guardar as imagens e metadados é o DICOM. Esse formato é padrão internacional e tem um conjunto de normas que torna a comunicação e armazenamento das informações médicas em um formato eletrônico único, estruturado num protocolo. Com esses protocolos as informações associadas às imagens geradas por equipamentos de tomografias, ressonâncias magnéticas, radiografias podem ser trocadas entre si. No mais, é mantida a qualidade das imagens para uma melhor avaliação.

As imagens capturadas pelo programa Teleoftalmo incluem as imagens de exames oftalmológicos de 2 tipos: retinografia e biomicroscopia. Na rotina do exame, o paciente faz as imagens nos equipamentos e o técnico do ponto de atendimento envia para o oftalmologista gerar o laudo. Para montar o nosso banco de imagens (detecção de catarata) são mais relevantes as imagens de pupila dilatada, pois permitem a visualização de uma área maior do cristalino do olho. As imagens de pupila dilatada são adquiridas no retinógrafo (Zeiss Visucam 224) usando um conjunto padrão de configurações (ângulo 30°, flash 6 e foco manual na íris). Todas as imagens foram adquiridas por técnicos especializados do programa em diferentes postos de atendimento e enviados, no protocolo DICOM, para um servidor de imagens.

As imagens não ficam junto com o laudo gerado e não tem uma indexação direta, toda a análise médica é confirmando nome e data de nascimento do paciente. Nesse ponto criamos um código, em Python, que, a partir de uma lista de pacientes marcado com catarata, pudéssemos extrair as imagens e qual o olho pertence. Como há muitas imagens, tanto de todos pacientes quanto do próprio paciente alvo do estudo, a busca das imagens se deu por uma comparação do nome do paciente, data de nascimento, ângulo da foto, flash e foco. Em seguida, extraímos as imagens em formato PNG, na qual a maior vantagem é a compactação sem muitas perdas de resolução. As extraídas tinham o tamanho mínimo de 2448x2448 pixels

e retirada de alguma forma de identificação da imagem, anonimizando as imagens para poder conectar com o laudo médico.

2.2 Relação imagem e laudo

Durante um exame o oftalmologista deve observar o olho do paciente e anotar as alterações morfológicas de cada olho, direito e esquerdo, e assim definir a conduta a ser seguida. Na hipótese da catarata, a presença da doença em um paciente pode ou não impactar a visão do mesmo. No caso de impactar a visão o especialista encaminha o paciente para avaliação de cirurgia em serviço de oftalmologia da rede SUS. Já no cenário do impacto ser pequeno existe um critério subjetivo do especialista em registrar a descrição da patologia que pode variar entre os especialistas, já que não é feita uma graduação da doença catarata para cada paciente.

Os laudos são gerados por paciente e não por olho. Como foi usado um banco verdadeiro, ou seja, não foi projetado especificamente para este trabalho, uma das dificuldades de gerar o banco foi analisar dentro do laudo qual olho, direito ou esquerdo, tem a indicação de catarata. Para decidir isso, foi necessário analisar outros dados, tais como acuidade visual ou uma análise manual. Nessa etapa também foi criado um pequeno código para linkar as imagens e laudo. Isso foi necessário porque a medida que a pesquisa ia avançando, novas imagens e laudos eram realizados.

2.3 Quantidade de amostras válidas

O projeto teve acesso a um conjunto de mais de 6500 imagens que foram analisadas pela equipe do projeto, procurando analisar sua adequação como base para um ambiente de aprendizagem artificial. Foi necessário fazer uma pré-seleção devido ao número elevado de imagens com problemas que poderiam interferir no treinamento das redes, como imagens fora de foco, imagens com oclusão da pupila pelas pálpebras (olho fechado), com iluminação incorreta, com lentes artificiais e com patologias diversas que se manifestem na região da pupila. Na pesquisa o foco foi apenas nos olhos que não sofreram cirurgia de catarata (sem lentes artificiais) e outras patologias que influenciam a área da pupila não foram considerados também.

Ainda na etapa de seleção das imagens, algumas laudadas como normal (sem catarata) foram descartadas pois demonstram inconsistências visuais claras, ou seja, teve que passar

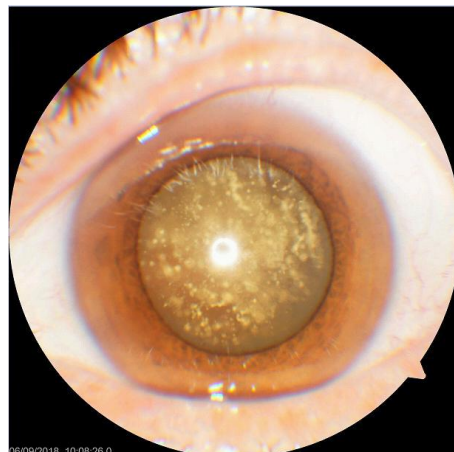
novamente por uma avaliação oftalmológica. Por exemplo, a Figura 2.1 foi laudada como normal dado que embora exista catarata cortical ela não atrapalha a visão. Na Figura 2.2 existe a presença de outra patologia, atrás do cristalino, que afeta a área da pupila durante o exame de retroiluminação.

Figura 2.1: Catarata cortical que não atrapalha a visão



Fonte: HUSEMANN (2019)

Figura 2.2: Outra patologia que afeta a pupila



Fonte: HUSEMANN (2019)

Como resultado da pré-seleção foram selecionadas 1020 imagens de olhos com catarata e 1020 imagens sem catarata de acordo com os laudos disponíveis, num total de 2040 imagens de um conjunto total de mais de 6500 imagens. Precisa ser enfatizado que esse conjunto de imagens selecionado não teve o laudo médico revisado especificamente para a

pesquisa em questão, apenas foram usados os dados disponíveis durante a operação normal do Teleoftalmo até meados de junho de 2019.

3 METODOLOGIA EXPERIMENTAL

Estruturado o banco de dados é possível iniciar o treinamento das redes neurais para a identificação da doença catarata. Na ferramenta o foco da etapa experimental consiste em verificar qual a melhor taxa de acertos que pode ser esperada usando uma rede neural de estrutura CNN, contando unicamente com a imagem do exame de retroiluminação. Futuros trabalhos usando outros tipos de imagens ou simplesmente utilizando mais informações poderão ser comparados ao desempenho obtido neste trabalho.

Mesmo restringindo as informações de entrada para rede neural, as imagens dos exames necessitam ser avaliadas de formas diferente, para poder retirar o máximo de resultado positivo da rede. O que tange as imagens, foi escolhido 2 tipos: pupila e olho inteiro, pois essa divisão nos permite identificar se as áreas externas a pupila (pálpebra, cílios, pele e esclera) impactam na análise da rede. O tamanho da imagem, devido a sua quantidade de informação, foi um ponto de teste também. Além disso, tínhamos poucas imagens validadas, ou seja, laudo e exame corretos, precisou-se realizar um aumento, ou não, do *dataset* disponível. Outra estratégia analisada, foi o tipo de arquitetura, porque muitas delas podem levar muito tempo ou até não convergir para resultados satisfatório. Nesse trabalho usou-se 3 abordagens que serão explicadas na sessão 3.2. Outro parâmetro avaliado foi a divisão do conjunto dados em treino, teste e validação. Esses parâmetros são necessários para evitar *overfitting* e *underfitting*, ou seja, os dados se adaptarem de forma ruim ao conjunta de dados que possuímos. Por último, o número de épocas (distribuição e apresentação dos dados para a rede) que impactam da assertividade dos dados de validação.

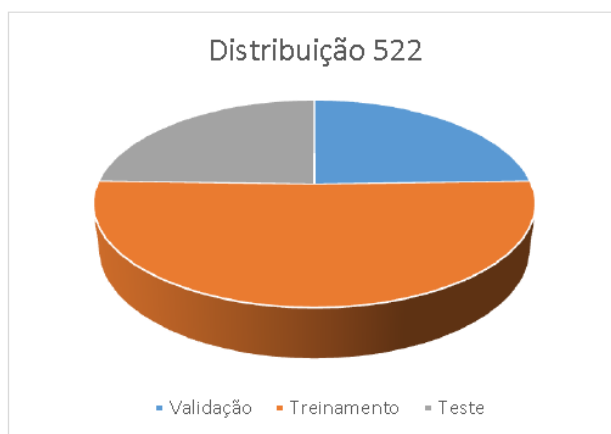
Sendo assim, a variação causada por cada mudança de parâmetro precisa ser avaliada gera em torno de 216 treinamentos diferentes. Como alguns treinamentos podem levar mais de 10 horas fica difícil a abordagem de força bruta e precisa ser feita uma seleção dos casos a serem simulados. Note-se que devido à aleatoriedade nos treinamentos teria de ser feita também uma média de mais de 30 treinamentos para cada caso, para que o valor médio e a variância associada a casa caso pudesse ser feita, mas devido a curta duração do projeto isso não foi possível. Deste modo os resultados apresentados a seguir são o resultado de um único treinamento.

3.1 Divisão das imagens para treinamento, validação e teste

Quando se implementa um projeto de inteligência artificial baseada em rede neural o banco de dados precisou ser dividido em três grupos: treinamento, validação e teste. O grupo de treinamento é aquele que irá ser apresentado repetidamente à rede, ajustando seus coeficientes internos da forma mais adequada para que consiga reproduzir os diagnósticos informados. O grupo de validação serve para verificar, a cada rodada de treinamento, se a rede está de fato convergindo de forma eficiente verificando os diagnósticos das imagens de validação. Vale destacar que as imagens do grupo de validação não são usadas para ajuste de coeficientes da rede. Por fim, após executar várias rodadas de treinamento e validação, denominadas de épocas, a rede é considerada como um modelo pronto para uso. Neste momento a rede é submetida às imagens do grupo de teste que serve para verificar a taxa de acertos da rede. Como os grupos tem funções distintas no processo, deve-se garantir que cada imagem só possa pertencer exclusivamente a um grupo. Embora outros métodos de divisão dos conjuntos de imagens existam eles não foram usados nesse trabalho.

Não existe uma regra clara de como dividir as imagens nos três grupos. Desta forma, foram analisadas duas distribuições distintas para a divisão do conjunto de imagens. A primeira, nomeada internamente no projeto pelo sufixo 522, usa aproximadamente 50% das imagens para treinamento, 25% para validação e 25% para teste. Uma representação gráfica desta distribuição encontra-se na Figura 3.1.

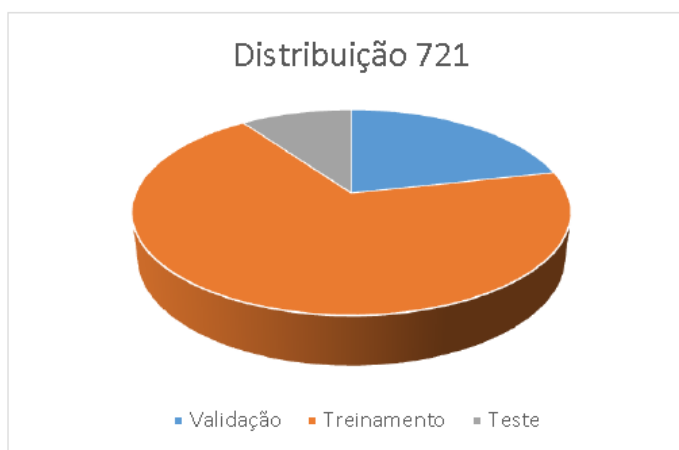
Figura 3.1: Representação da distribuição 522 das imagens do banco nos três grupos de trabalho.



Fonte: HUSEMANN (2019)

A segunda estratégia, nomeada internamente no projeto pelo sufixo 721, usa cerca de 70% das imagens para treinamento, 20% para validação e 10% para teste. Uma representação gráfica desta distribuição encontra-se na Figura 3.2.

Figura 3.2: Representação da distribuição 721 das imagens do banco nos três grupos de trabalho.



Fonte: HUSEMANN (2019)

Cada distribuição, tanto a 721 quanto a 522, possui um número de imagens de teste diferentes. Deste modo a comparação direta dos resultados de avaliação obtidos não é adequada, uma vez que os resultados são avaliados sob diferentes conjuntos de imagens durante o teste. Para permitir uma comparação justa dos modelos de rede treinados foi criado uma terceira distribuição, denominada 521, que difere da 522 apenas pelo fato de usar o mesmo conjunto de imagens de teste que a estratégia 721. Note que a distribuição 521 apresenta o mesmo conjunto de imagens de treino e validação que a distribuição 522, apenas o conjunto de teste é alterado, removendo as imagens de teste extras. A Figura 3.3 apresenta a divisão do banco de imagens nas três distribuições conforme utilizados nesse trabalho e a Tabela 3.1 apresenta o número de imagens para cada distribuição usada.

Tabela 3.1: Número de imagens nas diferentes distribuições de conjuntos de treino, validação e teste

Distribuição	Tipo	Treino	Validação	Teste	subtotal	Total
522	com catarata	520	250	250	1020	2040
	sem catarata	520	250	250	1020	
721	com catarata	700	220	100	1020	2040
	sem catarata	700	220	100	1020	
521	com catarata	520	250	100	870	1740
	sem catarata	520	250	100	870	

Fonte: HUSEMANN (2019)

Figura 3.3: Diagrama mostrando as distribuições das imagens do banco.

banco	com catarata (1020)		
	sem catarata (1020)		
721	treinamento	validação	teste
	(1400)	(440)	(200)
522	treinamento	validação	teste
	(1040)	(500)	(500)
521	treinamento	validação	não-usado
	(1040)	(500)	(300)
			teste
			(200)

Fonte: HUSEMANN (2019)

3.2 Arquitetura das redes neurais convolucionais (CNNs)

Foram usadas redes neurais convolucionais por serem os tipos de redes neurais mais indicadas para trabalho com imagens. Três arquiteturas de redes neurais convolutivas foram avaliadas: uma CNN padrão, uma CNN padrão com uma camada adicional de *dropout* e uma CNN pré-treinada com as últimas camadas ajustadas para a aplicação atual. A escolha das redes foi inspirada em uma aplicação semelhante utilizada na referência (CHOLLET,2017).

3.2.1 CNN padrão

A estrutura de rede CNN padrão é composta por uma sequência de 4 camadas de rede convolucionais (*conv2d + max_pooling*), uma etapa de planificação (*flatten*) e duas camadas densas (*dense*) no topo, com uma saída binária para indicar a presença da doença catarata na imagem de entrada. Um diagrama dessa rede é apresentado na primeira coluna da Figura 3.4. Essa rede CNN possui internamente 3.453.121 parâmetros para serem treinados quando usada com uma imagem de entrada colorida de 150x150 pixels e 24 bits de cor por pixel.

3.2.2 CNN com *dropout*

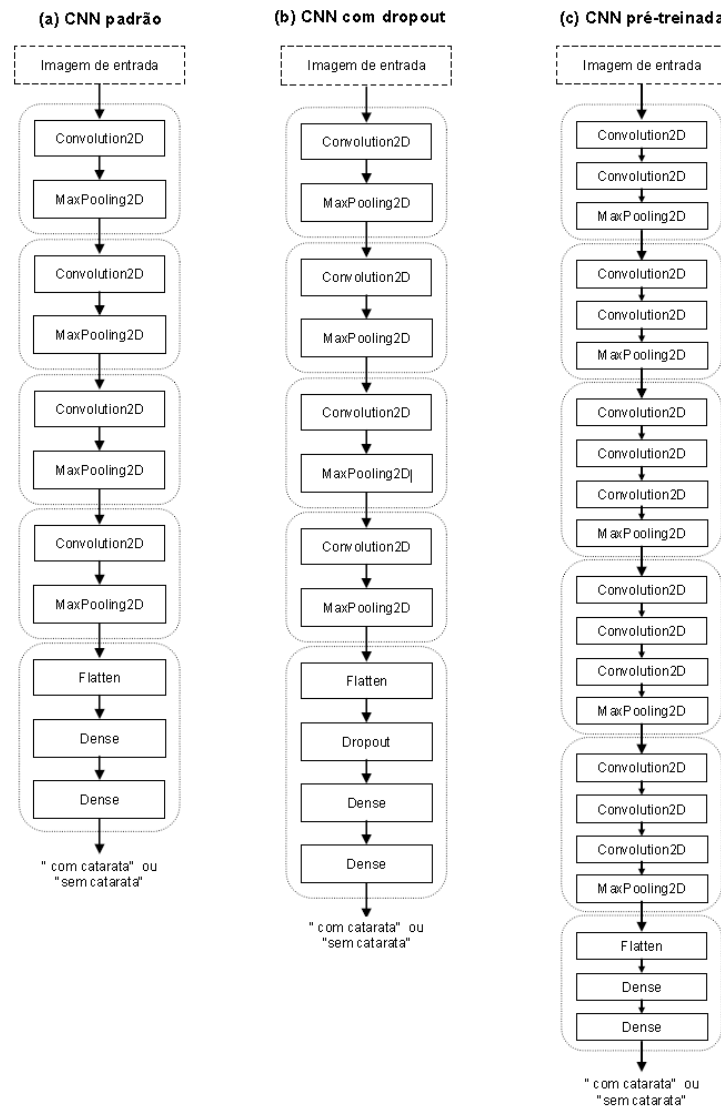
A estrutura de rede CNN com *dropout* possui a mesma composição da rede CNN padrão exceto pela camada adicional de *dropout* após a etapa de planificação (*flatten*). Essa técnica, desativa alguns dos neurônios da camada associada com alguma probabilidade p . Desativar um neurônio significa mudar o valor de saída para 0. No final, os neurônios que

sofreram *dropout* têm os parâmetros reajustados para que seja possível encontrar os melhores pesos (NIELSEN, 2015). Um diagrama dessa rede é apresentado na segunda coluna da Figura 3.4. Essa rede também possui internamente 3.453.121 parâmetros para serem treinados quando usada com uma imagem de entrada colorida de 150x150 pixels e 24 bits de cor por pixel.

3.2.3 CNN pré-treinada

A estrutura de rede CNN pré-treinada usa redes do tipo ResNet, AlexNet, DenseNet e VGG (SIMONYAN; ZISSERMAN, 2014) sem as camadas finais do topo, e adiciona uma camada de planificação e duas camadas densas no final, como feito nas demais redes. Diferentemente das arquiteturas anteriores, apenas essas camadas do topo são treinadas. Um diagrama dessa rede é apresentado na terceira coluna da Figura 3.4. Essa rede possui 14.714.688 parâmetros que já foram treinados anteriormente e ficarão congelados, mais 2.097.665 parâmetros das camadas de topo para serem treinados, num total de 16.812.353 parâmetros quando usada com uma imagem de entrada colorida de 150x150 pixels e 24 bits de cor por pixel.

Figura 3.4: Diagrama das redes neurais convolutivas usadas na etapa experimental.



Fonte: HUSEMANN (2019)

3.3 Tipo de imagem de entrada

Conforme apresentado na introdução, a doença catarata consiste de uma opacificação do cristalino. Para que a imagem de entrada para o sistema contenha a maior área do cristalino possível as imagens do olho do paciente devem ser adquiridas com a pupila dilatada. As imagens do banco de dados são fotografias completas do olho do paciente, contendo regiões de pele, pálpebra, cílios, esclera, íris e pupila correspondendo ao exame de retroiluminação conforme descrito na seção 2.

No presente trabalho foi adotado dois modos de entrada de imagem: só a pupila e a imagem inteira (com cílios, pálpebra, pele e esclera). De modo a isolar apenas a região de interesse de cada imagem, o que deveria evitar interferências e erros adicionais na etapa de classificação, foi usado um algoritmo de detecção automática de pupila (HUSEMANN; NEGREIROS; TOMAGGI; ARAUJO; ROESLER), que analisa cada imagem e identifica a região da pupila informando o centro e raio desta região, assumindo uma zona circular. Além das imagens de pupila foram empregadas na etapa de classificação também as imagens originais de olho inteiro para permitir a comparação do desempenho das redes. Na seção seguinte são analisadas as resoluções das imagens usadas na etapa de classificação.

3.4 Resolução das imagens

As imagens originais do banco de dados são de resolução mínima de 2448x2448 pixels, tendo sido detectado o centro e o diâmetro das pupilas para cada uma das 2040 imagens do banco. Após o processo de refinamento manual as imagens das pupilas foram recortadas, mascaradas com uma máscara circular e salvas em arquivo, sem alterar a resolução da área da pupila, que foi mantida no mesmo tamanho de cada imagem original. Cada imagem recortada está centralizada na pupila e o diâmetro da pupila é igual à largura ou à altura da imagem, que sempre apresenta largura e altura iguais, já que a área detectada da pupila é sempre circular. Os tamanhos das imagens de pupila são apresentados na Tabela 3.2.

Tabela 3.2: Tamanho das imagens de pupila em pixels.

	Com_Catarata	Sem_Catarata	Ambos
Mínimo	486	602	486
Máximo	1740	1694	1740
Médio	934	1046	990

Fonte: HUSEMANN (2019)

Em relação ao banco de imagens de olhos inteiros, as imagens originais de tamanho mínimo de 2448x2448 foram reduzidas para o tamanho intermediário de 800x800 pixels e mascaradas com uma máscara circular. O retinógrafo usado para aquisição pode ou não colocar informações extras escritas na área mascarada com caracteres brancos, como data e hora do exame e o nome do paciente. Usando o mascaramento essas informações extras são removidas, garantindo que apenas características das imagens sejam levadas em conta durante os treinamentos das redes.

Para o processamento com as redes o tamanho das imagens de pupila e de olho inteiro precisa ser uniforme, já que o tamanho da camada de entrada da rede neural é fixo. As imagens de pupila foram reduzidas para a dimensão de 150x150, uma escolha baseada em (CHOLLET,2017) e que fica abaixo do diâmetro médio das pupilas na base de dados. Para avaliar se o tamanho da imagem é uma influência significativa ou não também foram avaliadas imagens de resolução 250x250.

As imagens de olho inteiro foram dimensionadas para manter a mesma proporção do tamanho da pupila que as imagens contendo apenas a área da pupila. Conforme mostrado na Tabela 3.3, o diâmetro médio das imagens de pupila no banco é de 990 pixels. As pupilas sofrem uma diminuição média de tamanho de $990/150=6,6$ vezes devido ao redimensionamento das imagens para 150x150 pixels. Usando o mesmo número para redimensionar a imagem de olho inteiro, de tamanho mínimo 2448x2448 pixels, temos $2448/6,6=370,9$, valor que foi arredondado para 370. Deste modo as imagens de olho inteiro apresentam uma resolução de 370x370. As imagens de olho inteiro também foram avaliadas para efeito de aumento de tamanho. O aumento das imagens contendo apenas pupilas foi de 150 para 250 pixels, um aumento de 1667 vezes. No caso das imagens de olho inteiro esse valor ficaria $370*1,66=614$ pixels, mas foi arredondado para 600x600, um aumento de 1,6 vezes.

O tamanho final das imagens de pupila e de olho inteiro utilizadas na etapa de classificação é apresentado na Tabela 3.3. Perceba que o redimensionamento das imagens pode ser feito usando os módulos de processamento de imagens da biblioteca Keras (TensorFlow), dispensando a criação de bancos de imagens de tamanho específico. No entanto é necessário que o mascaramento das imagens seja feito antes, já que o Keras não disponibiliza funções específicas para mascaramento.

Tabela 3.3: Tamanho em pixels das imagens usadas para treinamento das redes.

Tamanho	Pupila	Olho Inteiro
padrão	150x150	370x370
grande	250x250	600x600

Fonte: HUSEMANN (2019)

3.5 Aumento artificial das imagens de treinamento

Um dos maiores problemas no treinamento de redes neurais convolucionais usando conjuntos de dados pequenos é o chamado *overfitting* (CHOLLET,2017), tornando o modelo treinado de rede neural incapaz de obter bom desempenho com imagens novas, que não participaram do treinamento. A melhor solução para esse problema é o aumento do conjunto de dados de entrada. Quando isso não é possível pode-se usar uma técnica de aumento de dados artificial aplicando transformações aleatórias às imagens de entrada, fazendo com que a rede receba um número maior de imagens durante o treinamento.

Nesse trabalho a etapa de treinamento das redes verificou o desempenho de estratégias de aumento artificial do conjunto de treinamento. Em primeiro lugar foi usado o banco de dados sem aumento artificial de imagens no treino. Em segundo lugar foi usada uma estratégia de aumento artificial, denominada de aumento *offline*, com aumento de 10 vezes. Finalmente uma terceira estratégia de aumento artificial de imagens de treino foi usada, denominada de aumento *online*. A seguir são explicadas as técnicas usadas.

Para aumentar artificialmente o número de imagens apresentadas para as redes durante o treinamento uma série de transformações aleatórias são aplicadas às imagens originais. A partir das imagens originais (pupila ou olho inteiro) são criadas outras imagens com pequenas alterações aleatórias que deslocam as coordenadas do centro, variam a ampliação da imagem, aplicam uma pequena rotação e também podem aplicar ou não um espelhamento horizontal à cada imagem. O conjunto de transformações e sua variação estão documentados na Tabela 3.4.

Tabela 3.4: Parâmetros de transformação das imagens para aumento artificial do conjunto de treinamento.

Parâmetro	Varição
coordenadas do centro da imagem	-10% a +10%
ampliação (<i>zoom</i>)	0% a +14.2%
rotação da imagem	-15 graus a +15 graus
espelhamento horizontal da imagem	aplicado a 50% das imagens

Fonte: HUSEMANN (2019)

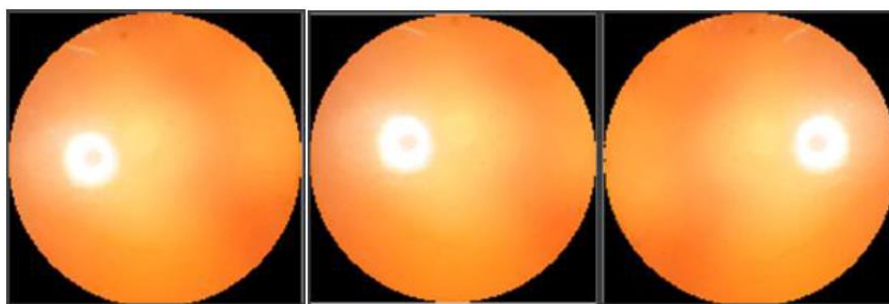
Uma vez determinado o conjunto de transformações aleatórias a serem aplicadas à imagem, o algoritmo na Figura 3.5 descreve o conjunto de operações que devem ser executadas. O aspecto das imagens de pupila transformadas é apresentado na Figura 3.6.

Figura 3.5: Algoritmo de transformação das imagens.

1. Avalia transformações aleatórias;
2. Recorta parte da imagem original;
3. Aplica rotação;
4. Aplica inversão horizontal se selecionado;
5. Redimensiona;
6. Aplica máscara.

Fonte: HUSEMANN (2019)

Figura 3.6: Exemplo de transformação para aumento artificial.



Fonte: HUSEMANN (2019)

3.5.1 Estratégia sem aumento artificial

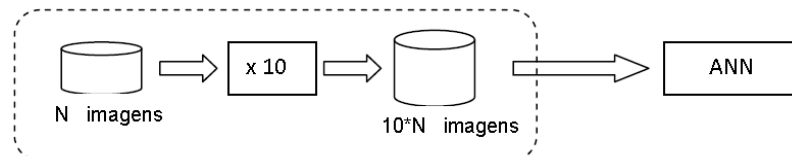
Na estratégia sem aumento artificial as redes são treinadas apresentando-se as imagens originais de treino diretamente para as redes. As únicas operações feitas são o redimensionamento para a resolução utilizada pela rede (150x150, 250x250, ...) e a normalização dos valores RGB, originalmente de 0 a 255 (números em formato 8 bits sem sinal), para números em ponto flutuante de 0 a 1 (valores originais multiplicados por 1/255). Nessa estratégia, a cada época de treinamento o conjunto total de imagens de treino é apresentado para as redes (um total de 1040 ou 1400 dependendo da distribuição usada).

3.5.2 Estratégia com aumento artificial de imagens (*offline*) do conjunto de treinamento

Na estratégia de aumento *offline* do conjunto de imagens de treinamento as transformações aplicadas às imagens originais são salvas em novas imagens, produzindo um novo conjunto de imagens formado por pequenas transformações de cada imagem original. Essa estratégia foi usada para produzir um aumento artificial de 10x no conjunto de imagens

de treino. Essa estratégia requer mais espaço em disco e a mesma imagem pode ser apresentada à rede mais de uma vez, já que o conjunto de treinamento é pequeno. Como vantagem temos menor uso de CPU ou GPU no pré-processamento das imagens durante o treino. Essa estratégia está ilustrada na Figura 3.7.

Figura 3.7: Estratégia de aumento artificial *offline* do conjunto de treinamento

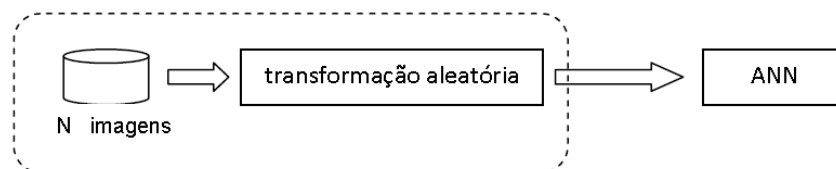


Fonte: HUSEMANN (2019)

3.5.3 Estratégia aumento artificial de imagens (online) do conjunto de treinamento

No aumento online do banco de imagens de treinamento, antes de apresentar uma imagem à rede, as operações de aumento descritas no algoritmo da Figura 3.7 são feitas em Python, sem salvar o resultado em disco, apenas apresentando-o diretamente à rede. Neste trabalho estas operações foram feitas usando as bibliotecas PIL e numpy. Diferentemente do aumento offline, a rede neural nunca observa diretamente as imagens de treino, apenas as versões transformadas aleatoriamente dentro dos parâmetros especificados na Tabela 3.5. Devido ao caráter aleatório usado para as transformações dificilmente a mesma imagem vai ser aplicada à rede neural mais de uma vez. A estratégia de aumento artificial online está ilustrada na Figura 3.8.

Figura 3.8: Estratégia de aumento artificial *online* do conjunto de treinamento.



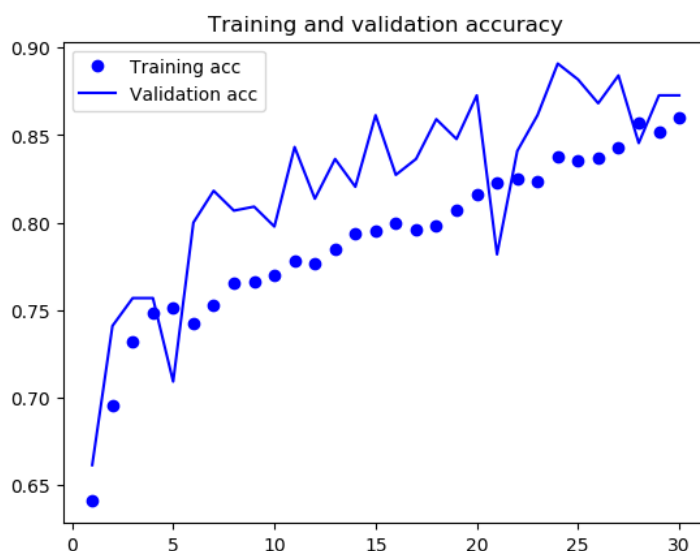
Fonte: HUSEMANN (2019)

3.6 Número de épocas de treinamento

A cada época de treinamento uma determinada quantidade de imagens e suas classificações são apresentadas à rede neural, em geral o conjunto completo de imagens

disponíveis para o treino. A taxa de acertos é avaliada para o conjunto de imagens de treino e para o conjunto de imagens de validação, que não participam do treinamento. A Figura 3.9 mostra o comportamento durante o treinamento de uma rede por 30 épocas. Pode-se notar pela figura que tanto as taxas de acertos do conjunto de treinamento quanto do conjunto de teste parecem não ter atingido um patamar, indicando que um aumento no número de épocas possa melhorar os resultados.

Figura 3.9: Curvas de taxa de acertos para os conjuntos de treino e validação para 30 épocas.

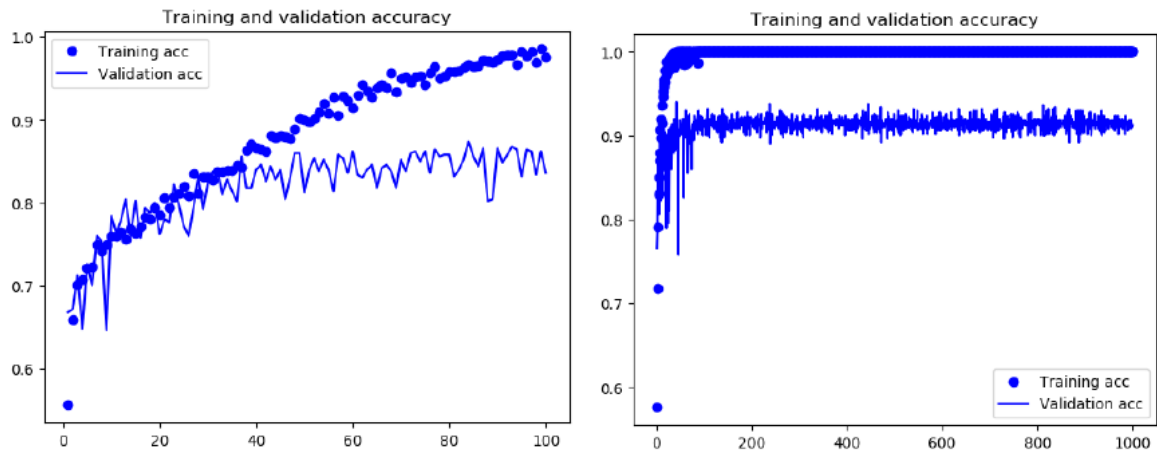


Fonte: HUSEMANN (2019)

O primeiro gráfico da Figura 3.10 apresenta o resultado de um treinamento com 100 épocas. Observa-se a tendência da curva de taxa de acertos de treinamento (*Training acc*) em chegar a valores próximos à unidade, mas a taxa de acertos do conjunto de validação parece atingir um patamar em cerca de 85%. De modo a verificar se as tendências das curvas realmente atingem determinados patamares é necessário um número de épocas ainda maior. O segundo gráfico da Figura 3.10 é um treinamento com 1000 épocas é apresentado. Atente-se que pelo gráfico que essa rede atinge um patamar de taxa de acertos no conjunto de validação acima de 90%, mas são necessárias mais de 100 épocas de treinamento para verificar a existência desse patamar.

Figura 3.10: Influência do número de épocas de treinamento na análise das curvas de treinamento:

(a) 100 épocas e (b) 1000 épocas.



Fonte: HUSEMANN (2019)

Neste trabalho foram estabelecidos os valores de 100 épocas e 1000 épocas para avaliação dos treinamentos. A escolha de 1000 épocas deve garantir tempo suficiente para as redes atingirem o seu melhor desempenho e evidenciar quaisquer limiares nas taxas de acertos para os conjuntos de treinamento e de validação.

3.7 Plataformas de desenvolvimentos

Para o desenvolvimento dos treinamentos é necessária uma máquina já configurada para o trabalho com inteligência artificial e com recursos de hardware que permitam acelerar os treinamentos das redes. Desse modo foram escolhidas como plataformas de desenvolvimento as máquinas virtuais Azure com GPU e o Google Colab com GPU. Cabe ressaltar que a pesquisa começou em parceria Microsoft, então tinha-se tempo e processamento para os testes. Depois do encerramento da parceria, optou por fazer os testes no modelo gratuito do Google Colab, o que restringe a quantidade de processamento diário. No trabalho o foco é nos resultados, então configurações das máquinas são omitidos.

O ambiente de desenvolvimento utilizou Keras (2.0.8), Python (3.6.8), Tensorflow (1.13.1) otimizado para acesso à GPU e Cuda (10.0). O ambiente Python foi gerenciado através do Anaconda.

4 RESULTADOS

De acordo com a proposta do projeto a etapa de classificação da doença catarata avalia o desempenho usando redes neurais convolutivas CNN e também usando redes neurais pré-treinadas. Foram avaliados 28 configurações de redes com imagens de pupila e 16 configurações de redes com imagens de olho inteiro num total de 44 avaliações (das 144 combinações possíveis discutidas no capítulo anterior). A seguir são apresentados os resultados obtidos sendo feita a apresentação separando-se os resultados por tipo de imagem (pupila ou olho inteiro) e pela distribuição dos conjuntos de treino, validação e teste.

4.1 Resultados para imagens de pupila

Foram avaliadas 28 configurações de redes com imagens de pupila. Os resultados obtidos são apresentados na Tabela 4.1 para a distribuição de 70% dos dados usados para treinamento, conforme explicado na seção 3.1. O melhor resultado usou uma rede pré-treinada e aumento *offline* dos dados de treinamento, com uma taxa de acertos de 88,5% (23 erros em 200 imagens).

Tabela 4.1: Resultados para imagens de pupila com 70% usado para treino (CAT721).

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros	Taxa Acerto
padrão	150x150	100	nenhum	1400	44	76,5
padrão	150x150	100	online	1400	67	66,5
padrão	150x150	100	offline	14000	39	80,5
padrão	150x150	200	online	1400	38	81,0
padrão	250x250	100	nenhum	1400	41	79,5
padrão	250x250	100	online	1400	58	71,0
com dropout	150x150	100	online	1400	50	75,0
com dropout	150x150	100	offline	14000	43	78,5
pretrained	150x150	100	online	14000	26	87,0
pretrained	150x150	100	online	1400	29	85,5
pretrained	150x150	100	nenhum	1400	32	84,0
pretrained	150x150	100	offline	14000	23	88,5
pretrained	150x150	1000	online	1400	32	84,0
pretrained	150x150	1000	nenhum	1400	28	86,0
pretrained	250x250	100	online	1400	32	84,0

Fonte: HUSEMANN (2019)

Na Tabela 4.2 são apresentados os resultados para a distribuição de 50% das imagens usadas para treinamento. Nessa distribuição o melhor resultado foi também de uma rede pré-treinada, sem aumento artificial, mas com um treinamento de 1000 épocas, atingindo 91% com 45 erros em 500 imagens de teste. Para podermos comparar as redes neurais treinadas com diferentes distribuições de imagens (50% e 70%) foi feita a avaliação da taxa de acertos da rede treinada com a distribuição 50% usando as mesmas imagens do teste da distribuição 70% e os resultados são apresentados na Tabela 4.3. Nesse conjunto de teste o melhor resultado foi de uma taxa de acertos de 88% correspondendo a 24 erros em 200.

Tabela 4.2: Resultados para imagens de pupila com 50% usado para treino (CAT522).

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros	Taxa Acerto
padrão	150x150	100	nenhum	1040	87	82,6
padrão	150x150	100	offline	10400	92	81,6
com dropout	150x150	100	offline	10400	81	83,8
pretrained	150x150	100	online	10400	46	90,8
pretrained	150x150	100	online	1040	55	89
pretrained	150x150	100	nenhum	1040	61	87,8
pretrained	150x150	100	offline	10400	53	89,4
pretrained	150x150	1000	online	1040	58	88,4
pretrained	150x150	1000	nenhum	1040	62	87,6
pretrained	150x150	1000	nenhum	1040	45	91
pretrained	250x250	100	online	10400	76	84,8
pretrained	250x250	100	online	1040	66	86,8
pretrained	250x250	1000	online	1040	61	87,8

Fonte: HUSEMANN (2019)

Tabela 4.3: Resultados para imagens de pupila com 50% usado para treino (CAT521).

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros	Taxa Acerto
padrão	150x150	100	nenhum	1040	44	78
padrão	150x150	100	offline	10400	44	78
com dropout	150x150	100	offline	10400	41	79,5
pretrained	150x150	100	online	10400	27	86,5
pretrained	150x150	100	online	1040	36	86,5
pretrained	150x150	100	nenhum	1040	30	82
pretrained	150x150	100	offline	10400	33	85
pretrained	150x150	1000	online	1040	35	83,5
pretrained	150x150	1000	nenhum	1040	24	82,5
pretrained	150x150	1000	nenhum	1040	46	88
pretrained	250x250	100	online	10400	46	77
pretrained	250x250	100	online	1040	32	84
pretrained	250x250	1000	online	1040	30	85

Fonte: HUSEMANN (2019)

É notório que em todos os resultados uma rede pré-treinada tem muito mais vantagem que outro tipo de arquitetura. E isso acontece, pois, as imagens tem uma forma geométrica muito bem definida, o que ajuda a definir alguns pesos das redes profundas e deixa a maior análise para os neurônios que analisam a diferença interna da pupila. Mas tarde, vamos poder averiguar que acontece igualmente quando a imagem é o olho completo, ou seja, contém mais informação.

Outro ponto importante é o banco de dados não ter passado por uma revisão oftalmológica. Obviamente, a maioria dos casos sem catarata e com catarata foram classificados corretamente, porém se tivéssemos uma maior revisão dos casos, como por exemplo, mesmo não atrapalhando a visão indicar que há uma deformidade no cristalino, o erro seria menor para todos conjuntos de treinamento.

4.2 Resultados para imagens de olho inteiro

As imagens de pupila precisam ser segmentadas previamente. Essa operação garante que outros aspectos da imagem não serão levados em conta pelas redes neurais, mas é necessário um pré-processamento complexo das imagens. Um dos objetivos da pesquisa foi também avaliar o desempenho das redes com imagens de olho inteiro, sem a detecção da pupila.

Foram avaliadas 16 configurações de redes com imagens de olho inteiro. Os resultados obtidos são apresentados na Tabela 4.4 para a distribuição de 70% dos dados usados para treinamento. O melhor resultado usou uma rede pré-treinada e aumento *online* dos dados de treinamento, com uma taxa de acertos de 86,5% correspondendo a 27 erros em 200 imagens.

Tabela 4.4: Resultados para imagens de olho inteiro com 70% usado para treino. Conjunto de teste possui 200 imagens

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros(200)	Taxa de acerto
padrão	370x370	100	nenhum	1400	54	73
padrão	370x370	100	online	1400	45	77,5
com dropout	370x370	100	nenhum	1400	41	79,5
com dropout	370x370	100	online	1400	45	77,5
com_dropout	370x370	100	offline	14000	36	82
pretrained	370x370	100	online	14000	33	83,5
pretrained	370x370	100	online	1400	34	83
pretrained	370x370	100	online	1400	27	86,5
pretrained	370x370	100	offline	14000	30	85
pretrained	370x370	1000	online	1400	49	75,5
pretrained	600x600	100	online	1400	32	84

Fonte: HUSEMANN (2019)

Na Tabela 4.5 são apresentados os resultados para a distribuição de 50% das imagens usados para treinamento. Nessa distribuição o melhor resultado foi também de uma rede pré-treinada, com aumento artificial *offline* e taxa de acerto de 87,8% correspondendo a 61 erros em 500. Para podermos comparar as redes neurais treinadas com diferentes distribuições de imagens (50% e 70%) foi feita a avaliação da taxa de acertos da rede treinada com a distribuição 50% usando as mesmas imagens do teste da distribuição 70% e os resultados são apresentados na Tabela 4.6, atingindo 85,5% de taxa de acerto (29 erros em 200).

Tabela 4.5: Resultados para imagens de olho inteiro com 50% usado para treino. O conjunto de teste possui 500 imagens.

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros(500)	Taxa de acerto
com_dropout	370x370	100	nenhum	1040	107	78,6
com_dropout	370x370	100	offline	10400	107	78,6
pretrained	370x370	100	online	1040	79	84,2
pretrained	370x370	100	nenhum	1040	98	80,4
pretrained	370x370	100	offline	10400	61	87,8
pretrained	600x600	100	online	1040	83	83,4

Fonte: HUSEMANN (2019)

Tabela 4.6: Resultados para imagens de olho inteiro com 50% usado para treino. Usa o mesmo conjunto de imagens de teste do caso 721 para permitir a comparação dos resultados.

Arquitetura	Resolução	Épocas	Aumento	Imagens/ Época	Erros(200)	Taxa de acerto
com_dropout	370x370	100	nenhum	1040	49	75,5
com_dropout	370x370	100	offline	10400	50	75
pretrained	370x370	100	online	1040	36	82
pretrained	370x370	100	nenhum	1040	44	78
pretrained	370x370	100	offline	10400	29	85,5
pretrained	600x600	100	online	1040	46	77

Fonte: HUSEMANN (2019)

4.4 Melhores resultados por imagem e distribuição

A Tabela 4.7 apresenta os melhores resultados de cada tipo de imagem e distribuição de dados, com erros avaliados sob o mesmo conjunto de 200 imagens de teste.

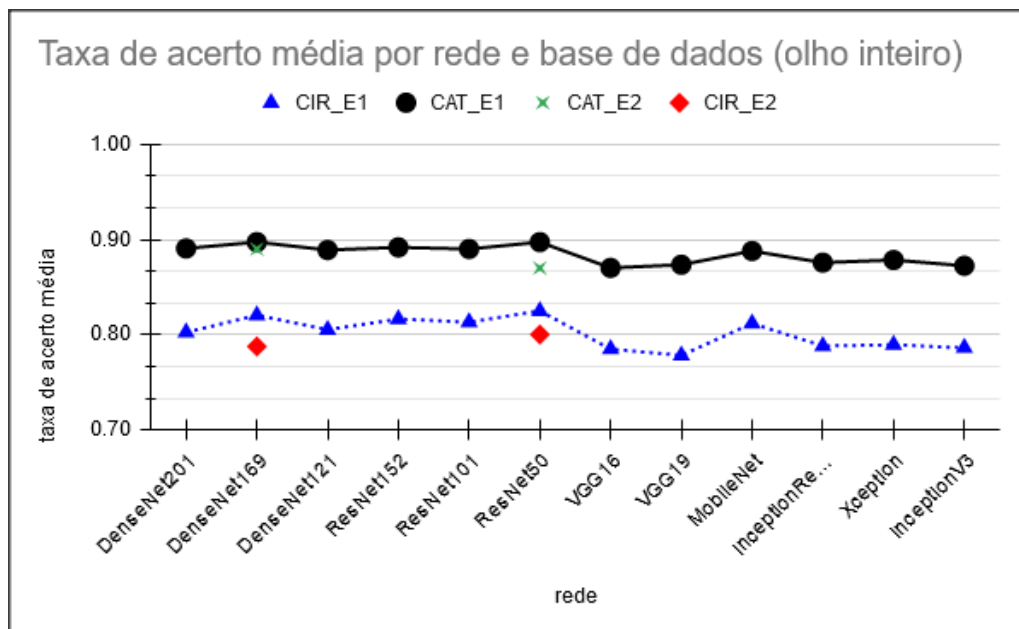
Tabela 4.7: Melhores resultados em cada tipo de treinamento.

Pupila	70%	pretrained	150x150	100	offline	14000	23	88,5
	50%	pretrained	150x150	1000	nenhum	1040	24	88
Olho inteiro	70%	pretrained	370x370	100	online	1400	27	86,5
	50%	pretrained	370x370	100	offline	10400	29	85,5

Fonte: HUSEMANN (2019)

Houve outras conjecturas de testes, como por exemplo, 711 (70 para treinamento, 15 para teste e 15 para validação), mas todas obtiveram resultado piores que a formatação da tabela acima. Além disso, para todos os modelos foram testados com um tipo de rede CNN diferente, Figura 4.1, e as que melhor obtivemos resultado foram DenseNet169, ResNet50 e MLP.

Figura 4.1: Gráfico cada conjunto de treinamento.



Fonte: Próprio autor

Basicamente, cada uma dessas redes tem sua peculiaridade de profundidade e complexidade, sendo interessante dependendo o tipo de dado. Outro cuida foi em repetir os mesmos modelos de teste com a mesma CNN porque embora fossem as mesmas

configurações variava. Essa variação ocorre por causa do peso que é calculado em cada camada da rede.

Foi usado, nos treinamentos, o método de *kfolds*. Esse método consiste em dividir a base de dados de forma aleatória em K subconjuntos (em que K é definido previamente) com aproximadamente a mesma quantidade de amostras em cada um deles. A cada iteração, treino e teste, um conjunto formado por K-1 subconjuntos são utilizados para treinamento e o subconjunto restante será utilizado para teste gerando um resultado de métrica para avaliação (ex: acurácia). Esse processo garante que cada subconjunto será utilizado para teste em algum momento da avaliação do modelo. No próximo capítulo iremos concluirmos o que obtivemos de resultado final de todos os treinamentos e técnicas utilizadas.

5 ANÁLISE DE RESULTADOS

Na proposta de projeto uma das premissas era de que a imagem da pupila previamente segmentada facilitaria o desempenho da rede, já que a doença catarata se manifesta na região da pupila. As redes treinadas com imagens de pupila realmente obtiveram melhor desempenho que as redes treinadas com imagens de olho inteiro, conforme apresentado na Tabela 4.7. Observa-se que o desempenho das redes que usaram imagens inteiras está bastante próximo do desempenho das redes que usam imagens de pupila, o que não era esperado. Uma vez que a segmentação da zona de pupila não está completamente automatizada, o uso de imagens de olho inteiro pode ser uma opção viável numa futura implementação prática. O único problema de usar uma imagem tão grande é o desempenho, visto que para cada pixel da imagem é atribuído um neurônio. Como as imagens do projeto são 2448x2448 pixels, levaria um alto tempo para treinar ou no caso retreinar o modelo com inputs novos.

Seria de se esperar que as redes treinadas com mais imagens apresentassem melhor desempenho na etapa de teste. Os conjuntos com mais imagens de treinamento (721) obtiveram melhor desempenho que os conjuntos que usaram menos imagens no treinamento (522 / 521). Novamente o desempenho ainda está próximo, com cerca de 2 a 3% de diferença. Caso se alcance uma taxa de acertos melhor no futuro esse valor pode se tornar significativo. Com relação ao aumento artificial das imagens de treinamento, os treinamentos de redes com imagens aumentadas obtiveram 3 dos 4 melhores resultados. Precisa ser salientado que tivemos problemas em algumas configurações de aumentos artificiais do tipo online que causaram instabilidades no treinamento das redes. Os gráficos de convergência durante o treinamento precisam ser inspecionados para verificação de tais problemas antes do uso dos modelos treinados.

A influência da resolução da imagem foi analisada em diversos treinamentos. Os melhores resultados foram obtidos com menores resoluções de imagens. No caso de imagens de olhos inteiros, as imagens maiores, de resolução 600x600, causaram problemas nas instâncias NV6, que não conseguiram treinar as redes por falta de memória. Essas redes foram treinadas com sucesso nas instâncias NC6, que possuem GPUs com mais memória interna (11GB VS 8GB).

Os resultados finais apresentaram um limiar abaixo de 90% para os índices de acerto das redes treinadas considerando o conjunto de teste de 200 imagens. Esse valor é bem inferior ao obtido em problema similar de classificação binária apresentado na referência

usada nesse trabalho (CHOLLET,2017), da ordem de 97%. O problema parece estar relacionado ao próprio banco de imagens. É conhecido na literatura o fato de que as imagens de retroiluminação, usadas neste projeto, não podem ser usadas para a graduação de cataratas nucleares (HARRIS,1993). No entanto o objetivo da pesquisa foi avaliar a "presença" e não efetuar uma graduação da catarata nuclear, então o baixo resultado poderia indicar uma limitação do tipo de imagem. Outro fator limitante para o diagnóstico pela rede neural pode ser devido à falta de informações de entrada. Durante o exame oftalmológico o especialista tem acesso a outras imagens e informações do paciente para efetuar o laudo médico, mas neste trabalho a rede neural não recebe outras informações além da imagem de retroiluminação de entrada.

Mais um fator limitante é a versão do banco de dados usado neste trabalho, que não foi desenvolvida exclusivamente para fins acadêmicos e contou com um número relativamente pequeno de imagens (1020 com catarata e 1020 sem catarata, num total de 2040 imagens). Deste modo ainda não é possível descartar a hipótese de que o limiar de cerca de 90% na taxa de acertos das melhores redes esteja relacionado à problemas do próprio banco de imagens.

Foi verificado no capítulo anterior que algumas imagens do banco de dados tiveram taxas de erro de classificação consistentemente altas durante os testes, tanto nas redes que usaram olho inteiro como pupila. Por exemplo, sabe-se que as imagens que mais apresentaram problemas correspondem aos dois olhos de um mesmo paciente e o diagnóstico original das imagens já foi confirmado. Essas imagens devem ser analisadas no futuro para verificar se é possível identificar a razão dos erros de classificação.

Os resultados apresentados no capítulo 4 para cada caso são de apenas um único treinamento, a Figura 4.1 demonstra como todos os testes deveriam ser realizados. O caráter aleatório dos treinamentos e as variações observadas na taxa de acertos do conjunto de validação sugere que exista uma variação percentual significativa de resultados entre múltiplos treinamentos da mesma rede. Para avaliar essa variação seria necessário efetuar um conjunto de simulações de cada caso e depois verificar o valor médio e o desvio padrão de cada conjunto para confirmar que as diferenças entre os treinamentos são significativas no sentido estatístico.

6 CONCLUSÃO

O uso de técnicas de inteligência artificial tem aumentado muito nos últimos anos, com diversas aplicações envolvendo a área biomédica, incluindo a área oftalmológica e trabalhos relacionados à detecção e graduação da doença catarata. A demanda computacional dessas técnicas pode ser grande no caso de trabalhar diretamente com imagens e redes neurais convolucionais.

Neste trabalho o uso de máquinas virtuais sob a plataforma Azure da Microsoft e google Colab, se mostraram adequados à proposta do projeto e possibilitou o treinamento de diversas arquiteturas de redes neurais convolucionais de maneira flexível permitindo um aumento da capacidade computacional sob demanda e permitindo o uso de recursos computacionais adequados ao problema, dispensando a necessidade de compra, configuração e manutenção de servidores e máquinas especiais no laboratório. O uso de máquinas virtuais permitiu a implementação rápida dos exemplos encontrados na literatura de inteligência artificial (CHOLLET,2017) usando ferramentas como Keras e Python e sua aplicação ao banco de dados. Obviamente, para uma continuidade do projeto e a criação da ferramenta, necessitaria de uma máquina robusta para o alto processamento que exige o algoritmo.

Os resultados obtidos usando imagens de retroiluminação foram de 88,5% para a melhor rede usando imagens de pupila, e de 86,5% usando imagens de olho inteiro, utilizando o banco de dados com 2040 imagens, sendo 200 imagens utilizadas para o teste. Todos os melhores resultados fizeram uso de redes pré-treinadas, onde as camadas de topo foram substituídas e treinadas, mantendo congelados os pesos das demais camadas da rede.

A partir do momento que o projeto começou a ser realizado e encontramos problemas no banco, a equipe de oftalmologia padronizou as os laudos, ou seja, mesmo que a pessoa não tenha dificuldade visual, mas a imagem mostra alguma alteração isso é marcado no laudo. Além disso, em um campo aberto do laudo é indicado qual olho a doença atinge. Por conta da pandemia e processo de trabalho não poderá ser feita uma revisão exame por exame. Logo, os dados desse trabalhando dão forte indícios que é uma solução válida ao serviço Teleoftalmo.

Uma outra possibilidade que se levantou para melhorar a taxa de acerto da rede foi o uso de informações extras, como idade do paciente, ou mesmo de outros tipos de imagens. Já se iniciou essa hipótese usando de início as informações de idade e raça do paciente. Resultados preliminares mostraram uma melhor acurácia, atingindo 91% de acerto. Por fazer parte da equipe do Telessauders e trabalhos futuro criarei um aplicativo que auxilie os oftalmologistas a diagnosticar a doença catarata a partir das imagens de olho inteiro.

Além da ferramenta, outra análise necessária é os cálculos dos falsos positivos e falsos negativos que, em aplicações médicas, necessitam de um ponto de alerta. Por se tratar de aplicações médica, o tempo de resposta da rede é uma etapa de grande impacto na criação da ferramenta. Seria necessário, igualmente aos falsos positivos e falso negativos, ter um estudo para otimizar esse tempo de resposta de uma imagem ao vivo. Nos 2 casos foi analisado de forma superficial, necessitando mais trabalhos e testes.

REFERÊNCIAS

- Projeto TelessaúdeRS - Teleoftalmo. Disponível em: <<https://www.ufrgs.br/telessauders/telediagnostico/teleoftalmo/>>. Acessado em 30 de junho de 2021.
- HARZHEIM, E. et al., **Telehealth in Rio Grande do Sul, Brazil: Bridging the Gaps**. *Telemedicine Journal and e-Health*, v. 22, p. 1-7, 2016. Doi:10.1089/tmj.2015.0210.
- PASCOLINI, D.; MARIOTTI, S. P., **Global estimates of visual impairment: 2010**, *British Journal of Ophthalmology* 2012; 96: 614-618. Doi: 10.1136/bjophthalmol-2011-300539.
- MALHOTRA, R., Chapter 1: The ageing lens and classification of cataracts, **Cataract**, 1. ed. Elsevier, 2007.
- LOPEZ, H. I. M.; Garcia, J. C. S.; MENDEZ, J. A. D., **Cataract Detection Techniques: A Review**. *IEEE Latin America Transactions*, vol. 14, 2016. Doi: 10.1109/TLA.2016.7587604
- Chylack LT, Wolfe JK, Singer DM, et al. **The lens opacities classification system III**. *Arch Ophthalmol* 1993; 111:831–836.
- GALI, Helena E.; SELLA, Ruti; AFSHARI, Natalie A., **Cataract grading systems: a review of past and present**, *Current Opinion in Ophthalmol* 2019. p. 13–18. Doi:10.1097/ICU.0000000000000542.
- GAO, X., et al., (2013) **Automatic Grading of Cortical and PSC Cataracts Using Retroillumination Lens Images**. In: Lee K.M., Matsushita Y., Rehg J.M., Hu Z. (eds) *Computer Vision – ACCV 2012*. ACCV 2012. Lecture Notes in Computer Science, vol 7725. Springer, Berlin, Heidelberg. Doi:10.1007/978-3-642-37444-9_20.
- GAO, X.; LIN, S.; WONG, T. Y. **Automatic feature learning to grade nuclear cataracts based on deep learning**. *IEEE Trans Biomed Eng* 2015. Doi: 10.1109/TBME.2015.2444389.
- CHOLLET, F. **Deep Learning With Python**, Manning, 2017.
- CHOLLET, F. et al. **Keras**. 2015. Disponível em: <<https://keras.io>>. Acesso em: 22 jun. 2018.
- HUSEMANN, R.; **"TeleDIAC-Sistema de IA para Diagnóstico de Catarata do TeleOftalmo-RS"** Workshop RNP 2019. Disponível em:<<https://eventos.rnp.br/sites/default/files/activity/activitypresentation/wrnp-2019-telediac-v2-husemann.pdf>>. Acessado em 15 de junho de 2019.
- HUSEMANN, R.; **Poster do Projeto Telediac no Workshop RNP 2019**. Disponível em:<<https://wrnp.rnp.br/sites/wrnp2019/files/Telediac.pdf>>. Acessado em 16 de junho de 2019.
- HUSEMANN, R.; NEGREIROS, M.; TOMAGGI, H.; ARAUJO, A. L.; ROESLER, V.; et all. **"Desenvolvimento de uma Ferramenta para Auxílio ao Diagnóstico de Catarata em Telemedicina"**, aceito para publicação em : VI Workshop “O Futuro da Videocolaboração” (WCT-Video 2019), Rio de Janeiro, Brasil. Anais Estendidos do Simpósio Brasileiro de

Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2019. Doi.org/10.5753/webmedia_estendido.2019.8154.

"GPU optimized virtual machine sizes". Disponível em: < <https://docs.microsoft.com/pt-br/azure/virtual-machines/windows/sizes-gpu>>. Acessado em 15 de agosto de 2019.

HARRIS, M. L, et al. **Analysis of Retro-Illumination Photographs for Use in Longitudinal Studies of Cataract**. Eye, 1993. Doi: 10.1038/eye.1993.124.

KARA-JOSÉ, N. ; BICAS, HEA; CARVALHO, RS. **Cirurgia de catarata: necessidade social**.2.ed., São Paulo, 2008.

TALEB, A. ; ÁVILA, M; MOREIRA, H. **As condições de saúde ocular no Brasil**.1ed., 2019. Disponível em: < https://www.cbo.com.br/novo/publicacoes/condicoes_saude_ocular_brasil2019.pdf >. Acessado em 17 de agosto de 2019.

CARNEIRO, HUGO V. ; SANTOS, LUCAS P. **Detecção de catarata por meio de imagens utilizando redes neurais convolucionais**. 2020. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Software) – Universidade Católica do Salvador, Salvador, 2020.

ALMANÇA, A. C. D. ; JARDIM, S. P. ; DUARTE, S. R. M. P.. **Perfil epidemiológico do paciente submetido ao mutirão de catarata**. Scileo, 2018. Doi.org/10.5935/0034-7280.20180055.

SIMONYAN, K.; ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. arXiv preprint arXiv:1409.1556, 2014.

TensorFlow. Disponível em: <<https://www.tensorflow.org/about?hl=pt-br>>. Acessado em 23 de agosto de 2019.

Microsoft. **"Introduction to Azure Data Science Virtual Machine for Linux and Windows"**. Online: Disponível em:< <https://docs.microsoft.com/en-us/azure/machine-learning/data-science-virtualmachine/overview>>. Acessado em 12 de junho de 2019.

NIELSEN, M. A. **Neural networks and deep learning**. [S.l.]: Determination press USA, 2015.

APÊNDICE A – TRABALHO DE GRADUAÇÃO I

Uso de Inteligência Artificial para Apoio na Identificação de Catarata

Andrews V. Berni¹, Ronaldo Husemann², Marcelo Negreiros², Aline Lutz de Araujo³

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brasil

²Departamento de Engenharia Elétrica – Universidade Federal do Rio Grande do Sul(UFRGS)

³Médica Oftalmologista
TelessaúdeRS-UFRGS – Porto Alegre, RS – Brasil

andrews.vieira@inf.ufrgs.br, {rhusemann,marcelo.negreiros,alinelutz.a}@gmail.com

Abstract. *Cataracts are the main cause of vision loss (blindness) worldwide, with about 51 % of cases. In general, cataracts are only clinically relevant if the patient's visual ability is significantly affected, since opacity in the lens can exist without any symptoms. Automatic cataract grading systems have proven to be a reproducible and objective way, although they still have limitations and need further refinement. The present project proposes the use of deep learning techniques to aid the diagnosis of the subject's cataract, based on images of ophthalmological exams performed in the ocular biomicroscope and retinograph and a set of clinical data obtained from the reports of the images.*

Resumo. *Catarata é a principal causadora de perda de visão (cegueira) em nível mundial. De modo geral a catarata somente é clinicamente relevante se a capacidade visual do paciente for afetada significativamente, uma vez que uma opacidade no cristalino pode existir sem qualquer sintoma. Sistemas de graduação automática de catarata provaram ser uma maneira reprodutível e objetiva, embora possuam ainda limitações e necessitem de mais refinamento. O presente projeto propõe o uso de utilizar redes neurais para auxiliar ao diagnóstico de catarata do sujeito, a partir de imagens de exames oftalmológicos realizados no biomicroscópio ocular e no retinógrafo e um conjunto de dados clínicos obtidos a partir dos laudos das imagens.*

1. Introdução

Catarata é a principal causadora de perda de visão (cegueira) em nível mundial, com cerca de 51% de casos [PASCOLINI,2010]. Catarata também é um dos principais fatores mundiais que levam à diminuição visual, sendo responsável por 33% dos casos, ficando em segundo lugar apenas para os erros refrativos não corrigidos (43 %) [PASCOLINI,2010]. Dados semelhantes foram divulgados em estudo anterior [RESNIKOFF, 2004] e também no site da Organização Mundial da Saúde (World Health Organization - WHO)[WHO, 2019]. Catarata e erro refrativo não corrigido combinados correspondem a 55% das causas de cegueira e a 77% das causas de diminuição visual em adultos de 50 anos ou mais segundo dados de 2015 [FLAXMAN,2017]. Existe variação nessa tendência dependendo do país, com a proporção de problemas de visão atribuídos à catarata sendo menor em países mais desenvolvidos [WHO, 2019] [BCSC 2018-2019].

Atualmente áreas da fisiologia e bioquímica das lentes do olho humano (cristalino) são ainda áreas ativas de pesquisa. Não há tratamento médico que possa prevenir a formação ou a progressão da catarata no cristalino de um adulto e teorias sobre a formação de cataratas e formas inovativas de gerenciamento continuam controversas. Embora vários fatores de risco para o desenvolvimento de catarata tenham sido identificados, tais como idade, composição genética, exposição à luz ultravioleta e diabetes, dados para prevenir o desenvolvimento de cataratas ainda não são conclusivos [BCSC 2018-2019].

A catarata ocorre pela opacificação do cristalino de modo a causar uma diminuição da visão em graus variados, inclusive levando à cegueira. Sendo assim, o cristalino tem a função primária de transmitir a luz incidente e efetuar o seu foco sobre a retina. Para que isso ocorra o cristalino deve ser transparente. No ser humano recém-nascido o cristalino é incolor e transmite praticamente 100% da luz incidente. Com a idade, o espalhamento e a absorção da luz pelo cristalino aumenta, com a diminuição da transparência e resultando num aspecto amarelado para o cristalino [MALHOTRA,2008].

O fator de risco mais importante para catarata é a idade. Cataratas relacionadas à idade constituem a grande maioria de todos os tipos de cataratas e se constituem em um grande problema de saúde pública em nível mundial [MALHOTRA,2008]. De modo geral a catarata somente é clinicamente relevante se há dificuldade para enxergar com nitidez, uma vez que uma opacidade no cristalino pode existir sem qualquer sintoma [ASBELL,2005]. Existem três tipos principais de catarata relacionada à idade: nuclear, cortical e subcapsular posterior. Uma catarata nuclear se forma na zona central (núcleo) da lente. A catarata cortical é caracterizada por opacidades brancas em forma de cunha que começam na periferia da lente e chegam ao centro de maneira semelhante a raios. Esse tipo de catarata ocorre no córtex da lente, que é a parte da lente que circunda o núcleo central. Uma catarata subcapsular ocorre na parte posterior do cristalino. Pessoas com diabetes (doença crônica na qual o corpo não produz insulina) ou que tomam altas doses de esteroides têm maior risco de desenvolver uma catarata subcapsular. Em muitos pacientes mais de um tipo de catarata estão presentes [BCSC 2018-2019], [MALHOTRA,2008].

Esquemas para a classificação objetiva de catarata usam padrões fotográficos para subdividir os principais tipos em diferentes graus. Esses graus podem ser baseados na densidade e na cor (no caso do núcleo) ou de acordo com a área anatômica da catarata (no caso de áreas cortical ou subcapsular posterior). Pode ser feita a comparação do cristalino do paciente conforme visto pelo exame lâmpada de fenda, exame que emite uma luz muito brilhante para analisar estruturas internas e externas, (Slit-lamp Examination [NEMA,2009]) com uma fotografia de vários padrões (graduação clínica). Também pode ser feita uma fotografia do cristalino para depois classificá-lo (graduação fotográfica).

Avanços na remoção cirúrgica de cataratas, incluindo cirurgia de pequena incisão, uso de viscoelásticos e o desenvolvimento de lentes intraoculares tornaram o tratamento muito efetivo e a recuperação visual muito rápida na maioria dos casos. Praticamente uma visão normal pode ser obtida depois da cirurgia de catarata com o uso de uma lente intraocular [ASBELL 2005]. Aproximadamente 3 milhões de cirurgias de catarata são feitas anualmente nos Estados Unidos, fazendo com que esse seja o procedimento cirúrgico atual mais comum realizado em nível ambulatorial [BCSC 2018-2019]. Já no Brasil foram realizados 601 mil cirurgias de cataratas pelo Sistema Único de Saúde (SUS) [CBO-2020].

A medida que a população mundial aumenta, ela também envelhece [WORLD-METERS 2021]. A população Brasileira também acompanha esse crescimento [IBGE 2019] e com isso o número de pessoas afetadas por problemas de visão aumenta substancialmente, necessitando de um grande aumento no atendimento nos serviços médicos para minimizar os problemas de visão da população [FLAXMAN,2017].

A Academia Americana de Oftalmologia (AAO) posiciona-se favoravelmente ao uso da tecnologia da informação na especialidade, incluindo métodos para rastreamento, refração e outros testes diagnósticos baseados na internet, desde que obedecidos os mesmos critérios usados para os demais exames médicos [AAO 2014].

Sistemas que utilizam inteligência artificial têm se mostrado capazes de detectar diferentes doenças a partir de imagens e dados médicos [LECUN,2015], [RAJKOMAR,2018], tais como melanoma [ESTEVA,2017] e retinopatia diabética [GULSHAN,2016], [GARGEYA,2017]. Nestes exemplos, a performance do sistema é comparável à interpretação diagnóstica feita por médico especialista na respectiva área. Sistemas de gradação automática de catarata provaram ser uma maneira reprodutível e objetiva, embora possuam ainda limitações e necessitem de mais refinamento [GALI,2019]. Técnicas de aprendizado profundo (Deep learning) e redes neurais prometem ser um caminho para avançar o campo de diagnóstico e gradação de catarata [GALI 2019].

O projeto “Teleoftalmologia como Estratégia de Atenção Integral à Saúde Ocular junto aos Médicos e Pacientes da Rede de Atenção Primária à Saúde do Rio Grande do Sul: Projeto Olhar Gaúcho” (Teleoftalmo) faz atendimentos oftalmológicos para pacientes do SUS de todo estado do Rio Grande do Sul. Como a fila para um exame presencial no estado pode demorar em média 12 meses [SAUDE.RS,2017] e baseado numa maior utilização de inteligência artificial na área da saúde, o presente projeto propõe o uso de técnicas de aprendizado profundo no auxílio ao diagnóstico de catarata do sujeito, a partir de imagens de exames oftalmológicos realizados no biomicroscópio ocular e no retinógrafo, além de um conjunto de dados clínicos obtidos a partir dos laudos das imagens do TeleOftalmo.

2. Objetivos

2.1. Geral

Determinar se o diagnóstico de presença ou não de catarata pode ser obtido com o uso de técnicas de inteligência artificial baseado em redes neurais a partir de imagens de biomicroscopia ocular em lâmpada de fenda e no retinógrafo, assim auxiliando num menor tempo de diagnóstico, diminuindo o tempo de espera por tratamento (cirúrgico ou preventivo).

2.2. Específicos

1. Segmentar base de dados, separando pacientes em duas categorias principais (com e sem catarata), a fim de treinar a rede de aprendizagem;
2. Subdividir cada grupo em treinamento, teste e validação
3. Treinar rede com base de dados segmentada e gerar modelo de rede treinada;
4. Validar modelo gerado subgrupo de validação a fim de determinar a taxa de acertos da rede gerada.

3. Métodos

Trata-se de um estudo de desenvolvimento e validação de modelo preditivo para detecção de catarata. Serão utilizadas técnicas computacionais de aprendizado profundo em parceria com os dados do projeto Teleoftalomo, conforme detalhado a seguir.

População: Dados clínicos e imagens oftalmológica de pacientes avaliados no Projeto Teleoftalmologia - Olhar Gaúcho, o qual já foi aprovado pelo Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre, CAAE 64499316.1.0000.5327, parecer consubstanciado número 2.757.077. O termo de consentimento do mesmo prevê a utilização dos dados dos pacientes para outros estudos. Os pacientes não necessitam ser reavaliados para o presente estudo, já que apenas os seus dados coletados para o TeleOfalmo serão utilizados.

Fator em estudo: Algoritmo de processamento de dados para predição do diagnóstico de catarata. Esse algoritmo irá utilizar dados clínicos e demográficos (idade, sexo, cor de pele, descrição do cristalino e acuidade visual), pois esses dados são determinantes para um diagnóstico final. Além disso, necessidade de encaminhamento, conduta do oftalmologista e imagens obtidas na avaliação por biomicroscópio e retinografia.

Desfecho de interesse: Catarata definido pela avaliação do oftalmologista conduzindo as avaliações no estudo Olhar Gaúcho.

Fluxo de desenvolvimento: O presente projeto envolve o desenvolvimento de um algoritmo de reconhecimento de padrões a partir das imagens médicas que identifiquem pacientes com catarata. O fluxo de desenvolvimento prevê três atividades, não sequenciais, principais: (i) identificação e separação de parte de interesse da imagem, (ii) identificação ou não da presença de catarata e (iii) refinamento da solução.

A seguir se descrevem estas atividades.

Atividade (i): identificação de parte de interesse da imagem: tem por finalidade buscar e isolar regiões relevantes da imagem. Esta estratégia de selecionar regiões de interesse e servirá para testar a tempo de convergência do algoritmo, pois caso as imagens possam ser puras, sem uma pré manipulação, poderá agilizar o tempo da resposta do diagnóstico da ferramenta proposta. Na prática, o foco principal será identificar as regiões de íris e pupila. Claramente trata-se de um algoritmo de reconhecimento de padrões [LU 2011]. Sua implementação poderá ser feita por algoritmos tradicionais de segmentação ou uma rede de inteligência artificial profunda (deep learning) [GULSHAN,2016]. Optando-se pela segunda alternativa, deve-se buscar operar com redes neurais convolutivas (CNN – Convolutional Neural Network), que tem demonstrado bons resultados para aplicações de processamento de imagens.

Atividade (ii): identificação da doença: A segunda etapa do projeto será responsável pela identificação, ou seja, interpretar as imagens focando nas regiões de interesse identificadas na etapa anterior. Neste caso se prevê o uso de uma estrutura de IA baseada em redes neurais, na qual se mostrou a estratégia mais eficiente [Rumelhart,1986] para a identificação de padrões. A fim de obter um elevado índice de acerto, redes profundas exigem um treinamento complexo, ou seja, devem ser alimentadas por bancos de dados com milhares de amostras. Isto impacta em um tempo normalmente grande de treinamento [VARDARAJAN,2018]. Para o caso do projeto em questão, pretende-se

reduzir este tempo com o uso de uma rede CNN pré-treinada, ou seja, uma rede com reconhecimento em alguns padrões geométricos (círculos por exemplo) e que seja aberta e disponível para a comunidade. Ao utilizar essa estratégia, apenas os estágios finais da rede deverão ser refinados para a identificação desejada.

Atividade (iii): refinamento da solução: os resultados até então obtidos serão analisados para se definir a arquitetura mais eficiente – por exemplo, definir se será utilizada uma única rede para implementar todas as funções citadas ou se serão implementadas duas redes distintas, uma para segmentação de regiões de interesse e outra segunda para identificação. É importante citar que a solução final gerada deverá servir como informação de apoio ao profissional oftalmologista, em via de regra apenas fornecendo subsídios adicionais que possam vir a facilitar sua tomada de decisão no diagnóstico.

Análise estatística: Os desempenho do algoritmo será avaliado em termos da especificidade, acurácia global, bem como valores preditivos negativo e positivo.

Tamanho amostral: Sabe-se que algoritmos de predição baseados em imagens têm melhor desempenho quanto maior a base de treinamento disponível. Para esse estudo, será utilizada todas as imagens interpretáveis, ou seja, possíveis de serem avaliadas e disponíveis no estudo do Teleoftalmo - Olha Gaúcho. Estima-se que serão cerca de 2000 pacientes, com cada pacientes tendo de 4 a 6 imagens distintas.

4. Aspectos Éticos

Como trata-se de uma análise de dados retrospectivos e de pacientes que já anuíram com a participação do estudo Olhar Gaúcho, não será aplicado novo termo de consentimento. Os pesquisadores se comprometem em lidar com os dados de forma sigilosa e não expor os participantes. Além disso, os pesquisadores preencherão o termo de consentimento de uso de dados (TCUD), conforme recomendado pelo comitê de ética em pesquisa do HCPA.

O banco de dados deste projeto é mantido na Central de Processamento de Dados da UFRGS e acessado pelos pesquisadores do Núcleo de Telessaúde Técnico-Científico da Universidade Federal do Rio Grande do Sul (TelessaúdeRS-UFRGS). A identidade e demais informações pessoais dos participantes não interessam ao presente projeto e portanto serão mantidas em confidencialidade. Os dados clínicos acessados serão utilizados somente para os fins do estudo aqui proposto.

5. Resultados Esperados

Os resultados esperados para este projeto são:

1. Geração de base de dados segmentada, a partir de análise técnica, separada entre pacientes com e sem catarata.
2. Treinamento de rede neural (técnicas de inteligência artificial) realizada a partir da base previamente segmentada.
3. Análise dos resultados obtidos avaliando se a aplicação poderia gerar uma ferramenta no auxílio do diagnóstico da presença de catarata em pacientes a partir de imagens médicas.
4. A possibilidade de aplicar o mesmo algoritmos para outras doenças (retinopatia por exemplo).

6. CRONOGRAMA

A seguir se apresenta cronograma resumido do projeto de pesquisa.(Tabela 1)

Tabela 1. Cronograma de desenvolvimento

Atividade	1	2	3	4	5
	Elaboração Projeto e Aprovação CEP	X			
Atividade (i) - segmentação de imagens	X	X	X		
Atividade (ii) - classificação de imagens		X	X	X	
Atividade (iii) - refinamento da solução e testes			x	X	X
Redação de artigo científico					X

7. Referências

American Academy of Ophthalmology. Information statement: Use of innovative technologies in diagnosing eye diseases and conditions. September 2014.< <https://www.aao.org/clinical-statement/innovative-technologies-in-diagnosing-eye-diseases>>.Acessado em 20 de março 2021.

Age-Related Eye Disease Study Research Group.(AREDS) The age-related eye disease study (AREDS) system for classifying cataracts from photographs: AREDS report no. 4. American Journal of Ophthalmology, Volume 131, Issue 2, February 2001, Pages 167-175

Asbell PA, Dualan I, Mindel J, et al. Age-related cataract. Lancet, 2005, vol. 365, p.599-609.< [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(05\)17911-2/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)17911-2/fulltext)>. Acessado em 25 de março 2021.

Basic and Clinical Science Course(BCSC), Section 11: Lens and Cataract, Volume 11 de 2018-2019 BCSC basic and clinical science course, American Academy of Ophthalmology.

Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115–118.

Flaxman SR, et al. Vision Loss Expert Group of the Global Burden of Disease Study. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta analysis. Lancet Glob Health. 2017;5(12):e1221-e1234.

Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology. 2017;124: 962–969.

Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402–2410.

Helena Gali;Ruti Sella;Natalie Afshari; Cataract grading systems: a review of past and present *Current Opinion in Ophthalmology*. 30(1):13–18, JAN 2019 DOI: 10.1097/ICU.0000000000000542 Issn Print: 1040-8738.

LeCun Y, Yoshua B, Geoffrey H. Deep learning. *Nature*. 2015; 521:436–444.

Lu S. Accurate and efficient optic disc detection and segmentation by a circular transformation. *IEEE Trans Med Imaging* 2011;30(12):2126–33.

Malhotra, 2008, Cataract,Raman Elsevier Chapter 1: The ageing lens and classification of cataracts.

Nema,HV; Nema,Nitin Jaypee Brothers Diagnostic procedures in Ophthalmology Medical Publishers, 2nd edition, 2009.

Pascolini D, Mariotti SP, Global estimates of visual impairment: 2010,*British Journal of Ophthalmology* 2012; 96: 614-618.<<http://dx.doi.org/10.1136/bjophthalmol-2011-300539>>.Acessado em 24 de março 2021.

Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *Npj Digital Medicine*. 2018; 18. doi:10.1038/s41746-018-0029-1 .

Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., Mariotti, S. P. (2004). Global data on visual impairment in the year 2002. *Bulletin of the World Health Organization*, 82(11), 844-51.

Varadarajan AV, Poplin R, Blumer K, et al. Deep Learning for Predicting Refractive Error From Retinal Fundus Images. *Invest Ophthalmol Vis Sci*. 2018;59(7):2861-2868.

World Health Organization Web Site, Blindness and vision impairment.<<https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>>. Acessado em 26 de março de 2019.

Número de cirurgias de catarata realizadas pelo SUS dobra desde 2009 e oftalmologistas alertam sobre importância da prevenção.**CBO(Conselho Brasileiro de Oftalmologia)**,2020.Disponível em: < <https://cbo.net.br/2020/numero-de-cirurgias-de-catarata-realizadas-pelo-sus-dobra-desde-2009-e-oftalmologistas-alertam-sobre-importancia-da-prevencao> >. Acesso em: 01 de fevereiro de 2021.

WORLDMETERS.World Demographics,2021.Disponível em:<<https://www.worldometers.info/demographics/world-demographics/median-age>>.Acessado em: 20 de março de 2021.

IBGE.PIRÂMIDE ETÁRIA,2021.Disponível em:<<https://educa.ibge.gov.br/jovens/conhecendo-brasil/populacao/18318-piramide-etaria.html>>.Acessado em: 20 de março de 2021.

SAUDE.RS.Consultórios remotos vão reduzir fila de espera por consulta oftálmica,2017.Disponível em:<<https://saude.rs.gov.br/consultorios-remotos-vaoreduzir-fila-de-espera-por-consulta-oftalmica> >.Acessado em: 20 de março de 2021.

Rumelhart DE; Hinton GE; Williams RJ. Learning internal representation by error propagation. In Rumelhart, D.E., and McClelland, J.L. (eds) *Parallel Distributed Processing*. MIT Press, Cambridge, 1986.

HOSPITAL DE CLÍNICAS DE PORTO ALEGRE
Grupo de Pesquisa e Pós Graduação

Carta de Aprovação

Projeto

2020/0013

Pesquisadores:

ROBERTO NUNES UMPIERRE

Andrews Vieira Berni

PIETRO BAPTISTA DE AZEVEDO

ALINE LUTZ DE ARAUJO

CARLOS ANDRÉ AITA SCHMITZ

RONALDO HUSEMANN

DIMITRIS RUCKS VARVAKI
RADOS

CYNTHIA GOULART MOLINA

MARCELO RODRIGUES
GONCALVES

CAROLINA WILTGEN CAMPOS

RODOLFO SOUZA DA SILVA

Marcelo Negreiros

ERNO HARZHEIM

Número de Participantes: 2000

Título: Uso de Inteligência Artificial para Apoio na Identificação de Catarata

Este projeto foi APROVADO em seus aspectos éticos, metodológicos, logísticos e financeiros para ser realizado no Hospital de Clínicas de Porto Alegre.

Esta aprovação está baseada nos pareceres dos respectivos Comitês de Ética e do Serviço de Gestão em Pesquisa.

- Os pesquisadores vinculados ao projeto não participaram de qualquer etapa do processo de avaliação de seus projetos.

- O pesquisador deverá apresentar relatórios semestrais de acompanhamento e relatório final ao Grupo de Pesquisa e Pós-Graduação (GPPG).

25/09/2020



Assinado digitalmente por:

PATRICIA ASHTON PROLLA

Grupo de Pesquisa e Pós-graduação

26/09/2020 10:23:23

<https://registro-externo.hcpa.edu.br/pesquisa/publico/cadastroapoio/conferenciaArquivo.xhtml?coDarq=1104497>