



CENTRO UNIVERSITÁRIO DO ESTADO DO PARÁ

PRÓ REITORIA DE GRADUAÇÃO E EXTENSÃO.

CURSO DE MEDICINA

JOÃO GABRIEL DE OLIVEIRA MENDES DA ROCHA

**UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO NEURO-
OFTALMOLÓGICO: APRENDIZADO A PARTIR DE DADOS REAIS E
DESENVOLVIMENTO DE MÉTODOS PARA ENFRENTAMENTO DE DESAFIOS
ATUAIS E FUTUROS**

BELÉM – PARÁ
2021

JOÃO GABRIEL DE OLIVEIRA MENDES DA ROCHA

**UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO NEURO-
OFTALMOLÓGICO: APRENDIZADO A PARTIR DE DADOS REAIS E
DESENVOLVIMENTO DE MÉTODOS PARA ENFRENTAMENTO DE DESAFIOS
ATUAIS E FUTUROS**

Trabalho de Conclusão de Curso apresentado
ao Centro Universitário do Estado do Pará
como requisito parcial para obtenção de Grau
em Medicina

Orientador: Professor Dr. Cláudio Eduardo
Corrêa Teixeira.

BELÉM – PARÁ
2021

Dados Internacionais de Catalogação-na-publicação (CIP)
Biblioteca do CESUPA, Belém – PA

Rocha, João Gabriel de Oliveira Mendes da.

Utilização da inteligência artificial no diagnóstico neuro-oftalmológico: aprendizado a partir de dados reais e desenvolvimento de métodos para enfrentamento de desafios atuais e futuros / João Gabriel de Oliveira Mendes da Rocha; orientador Cláudio Eduardo Corrêa Teixeira. – 2021.

Trabalho de Conclusão de Curso (Graduação) – Centro Universitário do Estado do Pará, Medicina, Belém, 2021.

1. Neurooftalmologia. 2. Fundoscopia. 3. Inteligência artificial – Aplicações médicas. I. Teixeira, Cláudio Eduardo Corrêa, orient. II. Título

CDD 23^o ed. 617.7

AGRADECIMENTO

Cada etapa vencida no curso de medicina é digna de agradecimentos, porém apresentar o Trabalho de Conclusão de Curso é um dos momentos mais gratificantes da graduação. Apesar das dificuldades para chegar até aqui, sem o apoio das pessoas que me acompanharam diariamente esta caminhada seria mais difícil e tornaram a vitória mais memorável.

Em primeiro lugar, agradeço a Deus sempre, por me acompanhar durante a minha vida, concedendo muitas bênçãos, forças, e me permitir ter familiares e amigos especiais. Agradeço também pela intercessão de nossa padroeira, Nossa Senhora de Nazaré, São João e Santo Antônio.

Aos meus pais, Edmilson e Tânia, por sempre me proporcionarem o melhor e não medirem esforços na minha formação, tanto profissional, quanto pessoal, desde o início, me ensinando e mostrando os valores do respeito, honestidade, amor e humildade, sou eternamente grato. Obrigado ao meu pai também, por fomentar em mim o meu desejo de ser médico. As minhas irmãs, Danielle e Bruna, que acompanharam toda a caminhada desde o início e sempre torceram pelo meu sucesso, agradeço por todo o companheirismo.

Ao meu padrinho, Caetano, e minha tia, Wânia (in memoriam) por terem me acolhido de forma especial, e compartilhado comigo das minhas conquistas. Agradeço a minha madrinha, Francisca, que também torce muito pelo meu sucesso pessoal e profissional. Aos meus tios, tias, primos e primas que sempre estiveram ao meu lado. Sintam-se parte dessa vitória.

Aos meus avós maternos, Emídio e Lígia, agradeço por estarem vivendo as minhas vitórias. Aos meus avós paternos, Cristino e Rosa (in memoriam), tenho certeza de que estariam felizes com minha conquista e que também estão intercedendo por mim.

Agradecimento especial ao meu orientador, Prof Dr Claudio Teixeira, pela oportunidade de fazer este trabalho junto comigo, mostrando que o conhecimento sempre será bem-vindo das diversas formas, por me estimular e proporcionar oportunidades acadêmicas únicas. A Dra Thais Mendes e ao Dr Rafael Scherer, por estarem juntos comigo no desenvolvimento desse trabalho, também fomentando ainda mais o conhecimento acerca do tema e proporcionando oportunidades únicas.

A todos os meus amigos, obrigado por serem meus companheiros e compartilharem comigo as conquistas, alegrias e tristezas durante esses anos.

Aos meus professores, que contribuíram em minha jornada acadêmica. Agradeço muito pelo conhecimento e experiência repassados, além de mostrar o verdadeiro significado de exercer uma medicina humana e competente. Obrigado a todos os funcionários do CESUPA campus João Paulo do Valle Mendes, pelo empenho e carinho com a instituição para proporcionar as melhores condições para a formação de novos profissionais da saúde.

João Gabriel de Oliveira Mendes da Rocha

RESUMO

Introdução: É importante que se tenha uma visão crítica do suporte dado por uma Inteligência Artificial (IA) no contexto médico para se confiar nesse suporte e integrar mais efetivamente essas novas ferramentas à prática diária. **Objetivo:** Medir o nível de incerteza unidimensional de um algoritmo de IA na detecção de padrões em imagens neuro-oftalmológicas, e comparar com desempenho humano. **Metodologia:** Neste trabalho (CAAE: 39292420.2.0000.5169), foi dada a um algoritmo simples de visão computacional escrito em Python (*blob detection*, OpenCV) e a um oftalmologista a tarefa de detectar um padrão bidimensional (centro do disco óptico) em mil imagens digitais de fundoscopias normais e alteradas por doenças neuro-oftalmológicas. O algoritmo realizou a tarefa 1x, o humano realizou a tarefa 2x, ambos usando marcação digital para registrar as coordenadas espaciais do padrão detectado. Enquanto que o nível de incerteza unidimensional da máquina foi medido pela respectiva comparação das coordenadas x e y registradas pela máquina e pelo humano, o nível de incerteza unidimensional do humano foi medido pela comparação das coordenadas registradas pelo próprio. A análise de dados foi feita no Programa de Computação Estatística R (<https://www.r-project.org/>). **Resultados:** Dentre as mil imagens avaliadas, apenas 2 a IA não conseguiu detectar o padrão alvo. Em média, homem e máquina mostraram maior nível de incerteza nas coordenadas y, sendo esta incerteza maior no desempenho da máquina, em ~100 unidades. O nível de incerteza na detecção foi maior em imagens de fundoscopia alteradas que em normais. **Conclusões:** Evidenciamos que a medida de incerteza de uma IA e de humanos no desempenho de uma mesma tarefa pode nos ajudar a entender suas limitações e, conseqüentemente, definir o alcance de sua utilidade enquanto ferramenta de apoio médico.

Palavras-chaves: Inteligência artificial; Detecção de padrões; Fundoscopia; Neuro-oftalmologia.

ABSTRACT

Background: It is important to have a critical view of the support provided by Artificial Intelligence (AI) in medical context, in order to trust this support. Design and setting: Cross-sectional study (CAAE: 39292420.2.0000.5169) to measure/compare unidimensional uncertainty of an AI and a human performing the same task. **Methods:** to a simple algorithm written in Python (blob detection, OpenCV) and to an ophthalmologist were given the task of detecting a two-dimensional pattern (center of the optical disc) in 1,000 digital images of normal/abnormal fundoscopies. Algorithm performed the task 1x, human performed the task 2x, both using digital register of spatial coordinates. Machine's unidimensional level of uncertainty was measured by the respective comparison of the x and y coordinates recorded by machine and human. Human's unidimensional level of uncertainty was measured by comparing the coordinates recorded by human itself. Data analysis was performed using R. **Results:** AI failed to detect the target pattern only in two images. On average, man and machine showed a higher level of uncertainty in the y coordinates, which was greater (~100 units) in machine's performance. The level of uncertainty was higher in altered fundoscopy images. **Conclusion:** the measure of uncertainty of AI and humans in the same task can help understand AI limitations and therefore define its usefulness as a medical support tool.

Keywords: Computer vision system, Ophthalmoscopies, Pattern recognition, Ophthalmology.

SUMÁRIO

1	INTRODUÇÃO	9
2	JUSTIFICATIVA	11
3	OBJETIVOS	12
3.1	Gerais	12
3.2	Específicos	12
4	METODOLOGIA	13
4.1	Aspectos éticos	13
4.2	Tipo de Estudo e Fonte de Dados.....	13
4.3	Coleta de Dados	13
4.4	Local de estudo	13
4.5	Técnica de marcação do Disco Óptico Através de Visão Computacional	13
4.6	Marcação do Disco Óptico por Oftalmologista Experiente	14
4.7	Análise dos Dados	14
5	RESULTADOS	16
6	DISCUSSÃO	35
7	CONCLUSÃO	39
	REFERÊNCIAS	40
	APÊNDICES	43
	ANEXOS.....	45

1 INTRODUÇÃO

Na década de 1990 uma grande mudança de paradigma surgiu no diagnóstico médico, intitulada de Medicina Baseada em Evidências, no qual houve uma proposta de inclusão de dados de Epidemiologia Clínica no processo de decisão profissional, antes dominado pela semiologia e propedêutica isoladas desde a organização da ciência por Hipócrates. A partir deste período revistas científicas, livros e congressos médicos difundiram termos estatísticos como Acurácia, Sensibilidade, Especificidade, Valor Preditivo Positivo, Razão de Verossimilhança, Odds Ratio, Risco Relativo, Número Necessário para Tratar e muitos outros. Tivemos todos de nos adequar a esta medicina mais precisa, tanto no âmbito da pesquisa como no ofício diário frente ao paciente¹.

Nos últimos anos, uma nova revolução está acontecendo, a inclusão das ditas Inteligências Artificiais (IA) na área da Saúde. A necessidade de aprender a lidar com estas novas ferramentas, entender seus conceitos, métricas, aplicações e limitações já existem hoje. Em diversas atividades diárias, desde o uso de um aplicativo de GPS, de tradução de texto, ou uma simples busca de termos em um navegador já possuem IA inseridas, agora, novas aplicações surgiram no âmbito da Oftalmologia, como o IDx-DR, a primeira ferramenta diagnóstica automatizada autorizada pelo Food and Drugs Administration (FDA) a executar a predição de Retinopatia Diabética por meio de Retinografias².

A IA foi conceitualizada em 1956, após um workshop no Dartmouth College. O termo Machine Learning (ML) foi posteriormente usado por Arthur Samuel em 1959 que afirmou "O computador deve ter a capacidade de aprender usando várias técnicas estatísticas, sem ser explicitamente programado." Usando o ML, o algoritmo pode aprender e fazer previsões com base nos dados que foram alimentados na fase de treinamento, usando um método de abordagem supervisionada ou não supervisionada. O ML foi amplamente adotado em aplicações como visão computacional e análise preditiva usando modelos matemáticos complexos. Com o advento das unidades de processamento gráfico (GPUs), os avanços nos modelos matemáticos, a disponibilidade de grandes conjuntos de dados e sensores de baixo custo, as técnicas de aprendizado profundo - Deep learning (DL) - posteriormente despertaram um interesse feroz e foram aplicadas em muitos setores. Assim, com o auxílio de AI os profissionais podem diminuir os custos com atendimentos

especializados além do benefício do compartilhamento na utilização dos sistemas de informação, visto que os dados gerados pelos atendimentos dos pacientes podem ser agregados e organizados produzindo um contexto que apoiará também a tomada de reconhecimento automatizado de imagens⁴.

Decisões em relação ao tratamento do paciente gerando impacto direto na redução de custos na área da saúde^{3,4}. A IA oferece um farol de esperança para melhorar algumas lacunas inerentes aos cuidados de saúde operados por seres humanos incluindo erros de diagnóstico, desperdício de recursos e tempo insuficiente entre médico e paciente, resultado do desequilíbrio entre demanda e oferta. Os avanços na IA têm o potencial de criar wavefronts por meio das vias de teleoftalmologia nos serviços nacionais de saúde. O reconhecimento automatizado de imagens permite que os computadores analisem e processem as imagens da retina e subsequentemente distingam lesões por meio de escalas de diretrizes previamente aprendidas pelo algoritmo de IA. Há uma promessa de que a conexão entre IA e teleoftalmologia fortalecerá os padrões de referência e melhorará o atendimento ao paciente⁴.

A IA envolve DL por meio de conjuntos de dados sem programação explícita, e o DL envolve auto-treinamento adicional usando "redes neurais artificiais". O DL envolve a identificação de uma correlação entre dados em grandes redes neurais nos conjuntos de dados. Dentro de especialidades baseadas em imagens, como oftalmologia, as redes neurais convolucionais utilizam dados rotulados para reconhecer padrões, com evidências publicadas de suas capacidades para retinopatia diabética, glaucoma e retinopatia da prematuridade e degeneração macular relacionada à idade, e até mesmo analisando sua aplicação clínica na triagem de catarata⁵.

2 JUSTIFICATIVA

É importante que se tenha uma visão crítica do suporte dado por uma IA no contexto médico. Propostas no sentido de regulamentar essas ferramentas já foram lançadas, como a do FDA no seu Proposed Regulatory Framework for Modifications for Artificial Intelligence/Machine Learning Based Software as a Medical Device (SaMD), e novas regulações provavelmente surgirão. As IA são tecnologias de suporte, mas confiar em seus resultados depende de compreender aspectos muitas vezes distantes do histórico de formação profissional nas áreas biológicas e de saúde. Necessitamos, portanto, aprender a medir o desempenho das ferramentas de IA assim como sabemos medir o de um ser humano, pois reconhecendo as limitações de ambos, poderemos integrar mais efetivamente essas novas ferramentas a prática diária.

3 OBJETIVOS

2.1 Geral

Analisar a aplicabilidade de variedades de IA na avaliação de dados neuro-oftalmológicos

2.2 Específicos

- a)** Comparar estatisticamente o desempenho de diferentes ferramentas de IA na avaliação de dados neuro-oftalmológicos

4 METODOLOGIA

4.1. Aspectos éticos

Este projeto foi aprovado por Comitê de Ética em Pesquisa com Seres Humanos (CAAE: 39292420.2.0000.5169).

4.2. Tipo de Estudo e Fonte de Dados

Descritivo, Analítico, Observacional, Transversal.

4.3. Coleta de Dados

Este projeto é parte de uma linha de pesquisa, que além de comparar estatisticamente o desempenho da ferramenta de IA com o oftalmologista, visa identificar, propor correções as limitações da IA e desenvolver novas propostas a ferramenta na avaliação de dados neuro-oftalmológicos.

Todos os dados que foram utilizados neste projeto constam dos arquivos da RedCheck (ver Anexo 1), instituição colaboradora neste Projeto. A natureza dos dados é numérica e/ou categórica, sendo relativa a resultados de exames de fundoscopia normais (n = 500) e alteradas por diferentes doenças neuro-oftalmológicas (n = 500).

4.4. Local de Estudo

Belém, Pará, Brasil.

4.5. Técnica de marcação do Disco Óptico Através de Visão Computacional

Houve alimentação da máquina com as 1000 fundoscopias para detecção do centro do disco óptico. Inicialmente a imagem em formato JPEG é convertida do formato BGR para RGB, aonde as cores primárias que compõe cada pixel (Vermelho, Verde e Azul) são reordenadas na sua anotação. Em seguida a imagem é convertida para escala cinza.

A seguir é aplicado um threshold dinâmico sobre a imagem em escala cinza, basicamente, o valor de cada pixel é checado e no caso de ele estar superior a um limiar, é aumentado para o valor máximo (tornando-se totalmente branco), e caso seja inferior, é reduzido em seu valor até o mínimo (tornando-se totalmente preto).

Desta forma, sobre a imagem com threshold aplicado, é realizada a busca da região do disco óptico através de um método de Visão computacional chamado

SimpleBlobDetector, presente em uma biblioteca de livre acesso chamada OpenCV criada pela Intel®. Caso não seja encontrado, os valores de corte do threshold são ajustados novamente até que seja ou não encontrada a região.

Este método acima encontra regiões (Contours) de pixels que mantêm alguma relação entre si devido suas características de cor, segmentando a imagem em seus componentes. O Centro desses componentes é encontrado e registrado. Para que seja caracterizado como Disco Óptico, o contorno em questão tem de cumprir algumas características arbitrariamente definidas:

- 1 – Uma área mínima de 1 mil pixels e máxima de 10 mil pixels
- 2 – Uma circularidade mínima de 0,1
- 3 – Uma convexidade mínima de 0,1

A detecção do centro do disco óptico foi representada com a geração de coordenadas 'X' e coordenadas 'Y' de um plano cartesiano.

Não há necessidade de treinamento, e o custo computacional é baixo por não utilizar redes neurais artificiais. Desta forma, considerando a proposta de utilizar uma técnica para rastreamento preliminar da região do Disco Óptico, figura como uma possível opção metodológica, além disso a técnica pode ser aplicada em série ou paralelo com outros métodos para refinar a acurácia.

4.6. Marcação do Disco Óptico por Oftalmologista Experiente

Um oftalmologista experiente foi voluntário na mesma tarefa, ou seja, de detectar e marcar digitalmente com cursor o centro do disco óptico nas mesmas imagens de fundoscopia.

Utilizando as 1000 fundoscopias, sendo 500 normais e 500 com diversas alterações oftalmológicas, do banco de dados da Red Check, o oftalmologista marcou todas as retinografias em um primeiro momento, e posteriormente as marcou novamente registrando outros dados para análise comparativa com os primeiros.

4.7. Análise dos dados

Quando se pretende avaliar a concordância entre dois métodos que deveriam medir a mesma quantidade, são utilizadas análises que nem sempre estão corretas. É importante que seja evitado o uso da correlação nestas situações e que a metodologia seja utilizada de forma adequada, incluindo os limites de concordância e seus intervalos de confiança, além de comentar se os limites encontrados são

diferenças aceitáveis do ponto de vista clínico. A análise de concordância entre métodos de Bland-Altman é adequada a essa comparação, justamente por permitir a comparação entre médias de resultados obtidos com métodos de medidas diferentes e a diferença de resultados obtidos por métodos diferentes (Bland-Altman, 1983). A análise dos dados foi realizada utilizando o programa de computação estatística R (www.r-project.org).

A partir dessas coordenadas, foram feitas análises de concordância a partir do cálculo de Bland-Altman e de correlação, gerando gráficos comparativos em: 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' de fundoscopias normais, 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' de fundoscopias normais, 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' de fundoscopias alteradas, 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' de fundoscopias alteradas, 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' de fundoscopias normais conjuntamente com alteradas, 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' de fundoscopias normais conjuntamente com alteradas, 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais, 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais, 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias alteradas, 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias alteradas, 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas, 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas, 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais, 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais, 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias alteradas, 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias alteradas, 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas, 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas.

5 RESULTADOS

Em todos os gráficos, a análise de correlação foi forte, já o teste de grau de concordância teve variações nas diferenças médias em até 100 unidades aproximadamente.

Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' com fundoscopias normais, a correlação entre as observações foi forte (Figura 1) e a diferença média na análise de Bland-Altman (Figura 2) foi próxima ao 0, mostrando uma boa concordância.

Figura 1- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias normais. $R=0,96$.

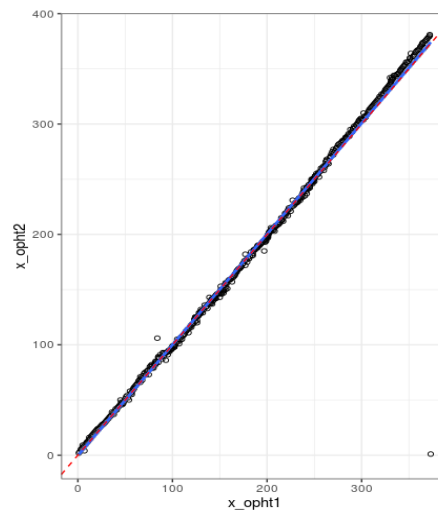
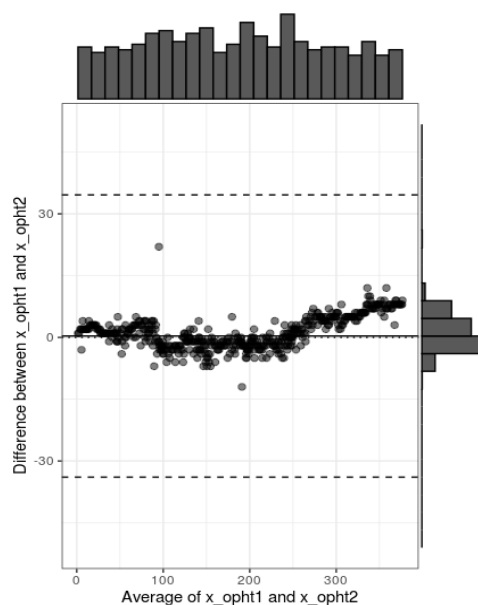


Figura 2- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias normais.



Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' com fundoscopias normais, a correlação foi forte (Figura 3) e a diferença média foi próxima ao 0 na análise de concordância (Figura 4).

Figura 3- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias normais. $R=0,96$.

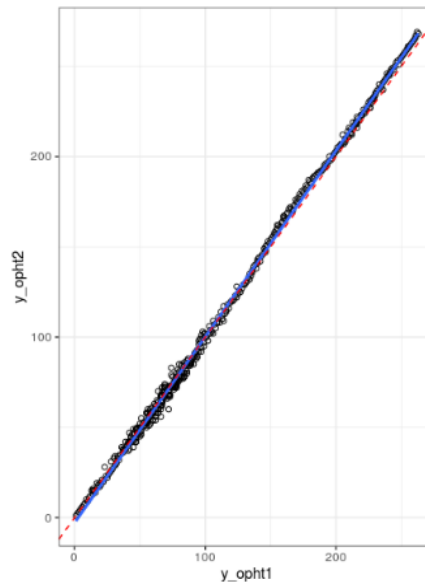
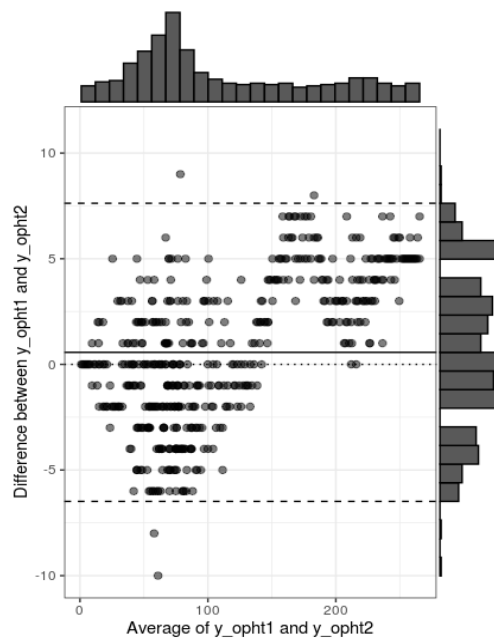


Figura 4- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias normais.



Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' com fundoscopias alteradas a correlação foi forte (Figura 5) e a diferença média foi próximo ao 0 (Figura 6).

Figura 5- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias alteradas. $R=0,96$.

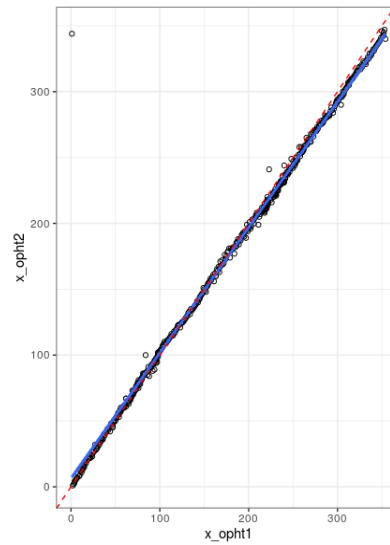
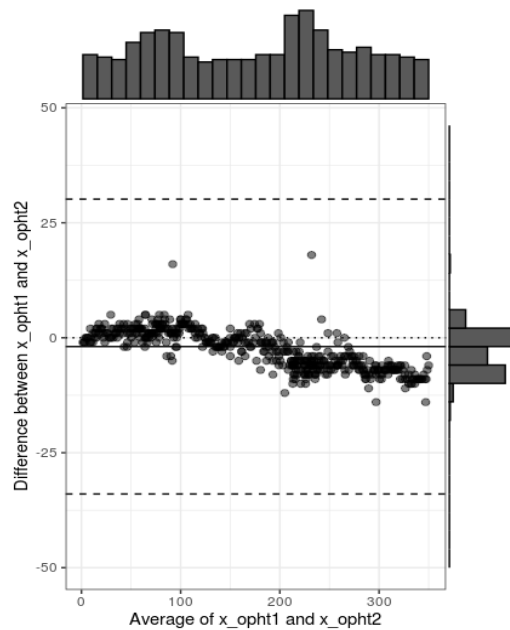


Figura 6- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias alteradas.



Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' com fundoscopias alteradas a correlação foi forte (Figura 7) e a diferença média na análise de concordância foi aproximadamente 1,25 (Figura 8).

Figura 7- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias alteradas. $R=0,96$.

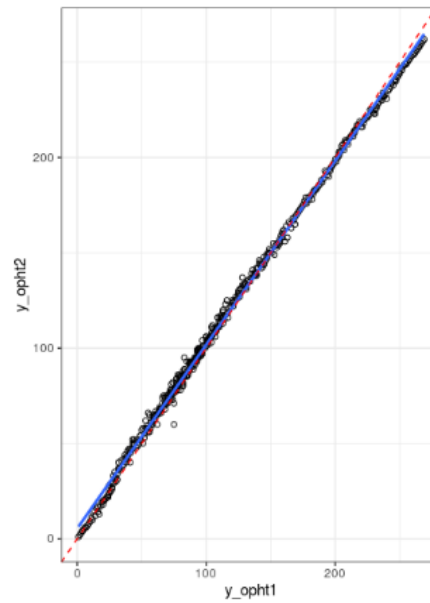
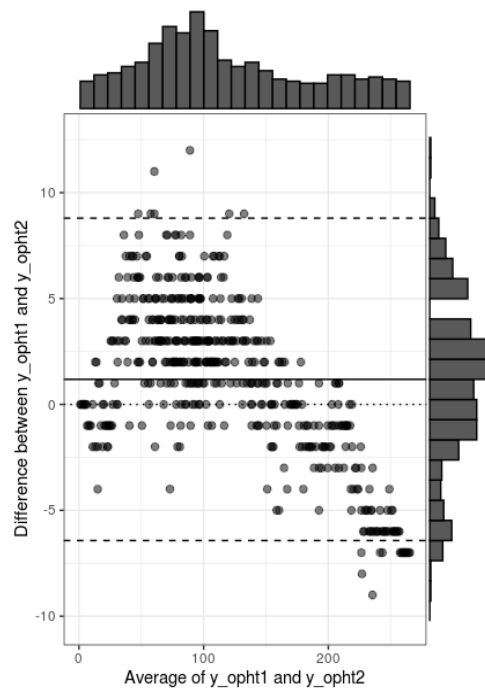


Figura 8- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias alteradas.



Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'X' com fundoscopias normais conjuntamente com alteradas a correlação foi forte (Figura 9) e a diferença média se aproximou do 0 (Figura10).

Figura 9- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias normais conjuntamente com alteradas. $R=0,96$.

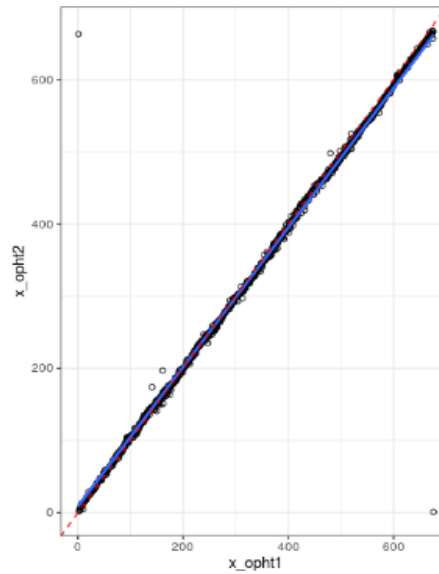
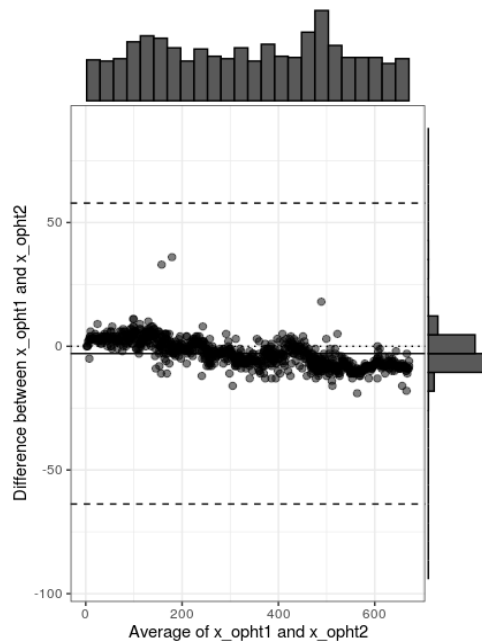


Figura 10- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada X em fundoscopias normais conjuntamente com alteradas.



Nos gráficos da 1ª observação do oftalmologista versus 2ª observação do oftalmologista coordenada 'Y' com fundoscopias normais conjuntamente com alteradas a correlação foi forte (Figura 11) e a diferença média foi em torno de 8 unidades acima do 0 (Figura 12).

Figura 11- Correlação entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias normais conjuntamente com alteradas. $R=0,96$.

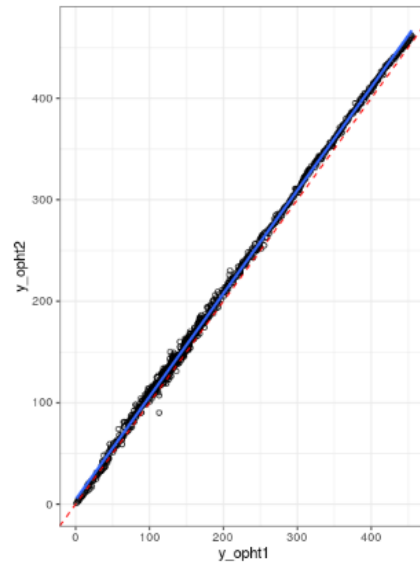
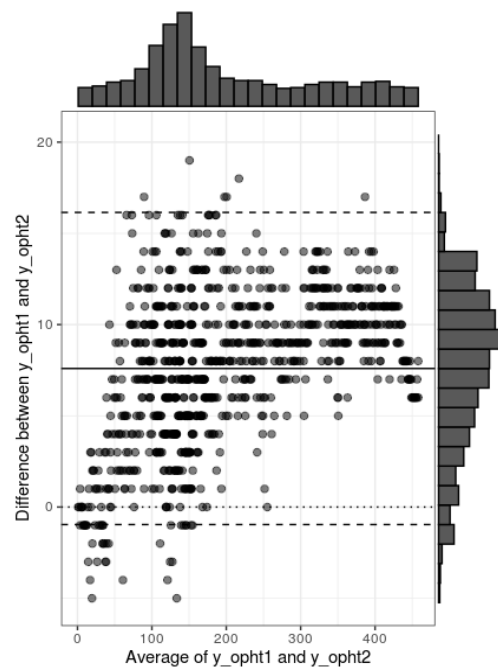


Figura 12- Grau de concordância entre a 1ª e 2ª observação do oftalmologista coordenada Y em fundoscopias normais conjuntamente com alteradas.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais a correlação foi forte (Figura 13), já a diferença média foi cerca de 35 unidades acima do 0 na análise de concordância (Figura 14).

Figura 13- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais. $R=0,96$.

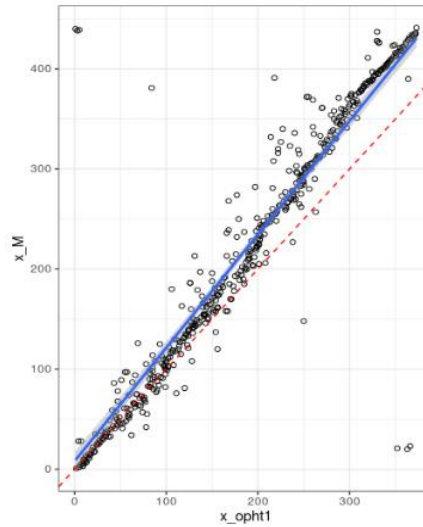
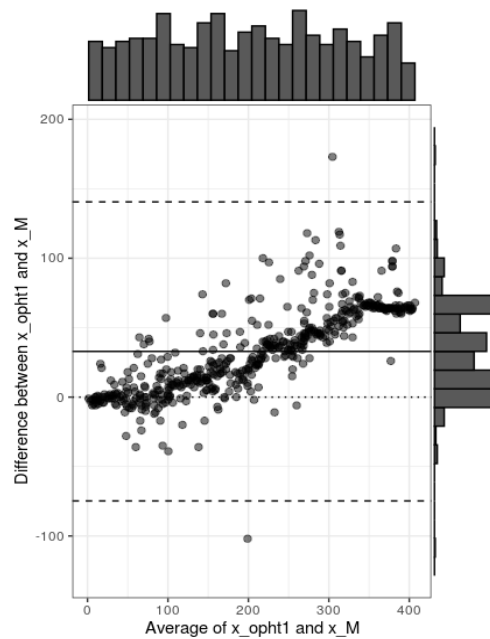


Figura 14- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais a correlação foi forte (Figura 15) e a diferença média foi cerca de 50 unidades acima do 0 (Figura 16).

Figura 15- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais. $R=0,96$.

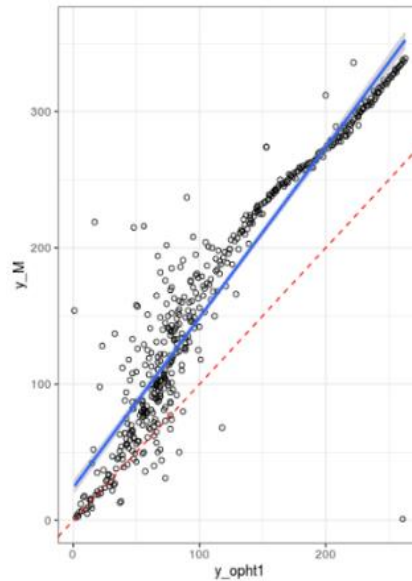
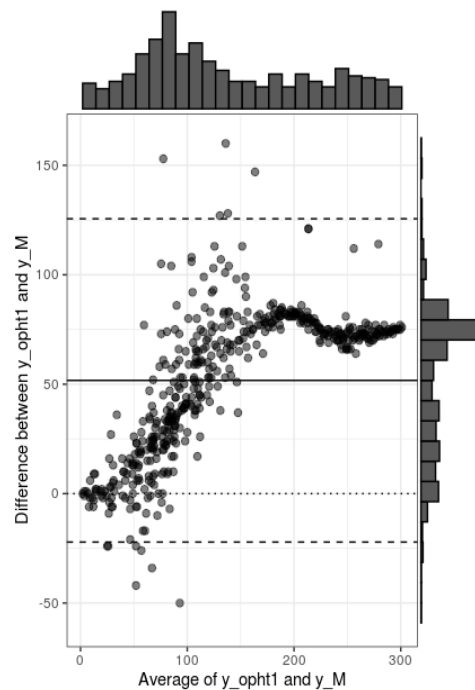


Figura 16- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias alteradas o gráfico de correlação foi forte (Figura 17). O grau de concordância teve uma diferença média de 50 unidades acima do 0 (Figura 18).

Figura 17- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias alteradas. $R=0,96$.

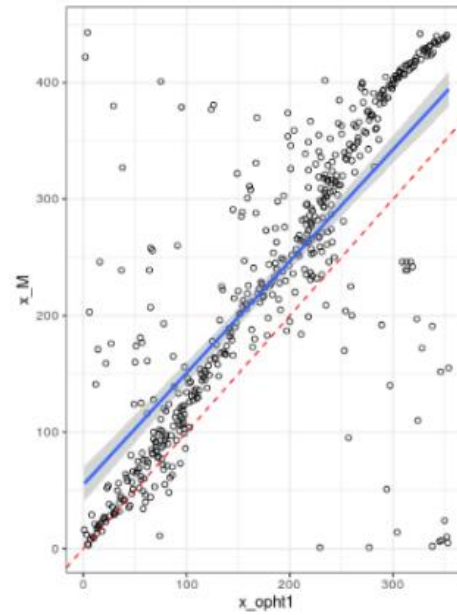
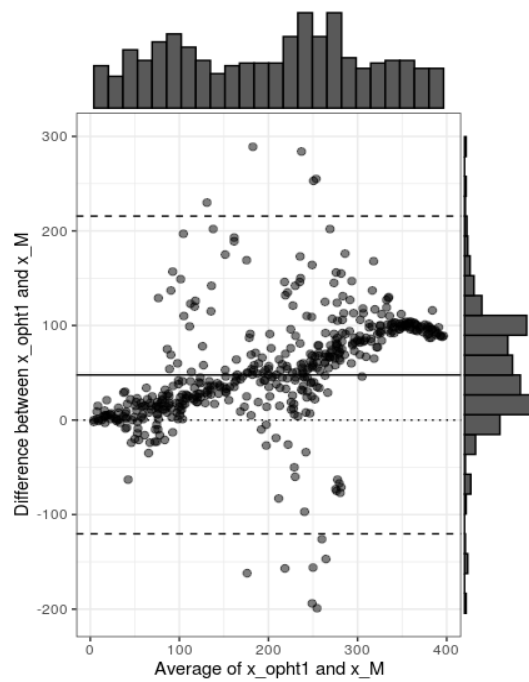


Figura 18- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias alteradas.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias alteradas a correlação foi forte (Figura 19), a diferença média teve valor de cerca de 10 unidades acima do 0 (Figura 20).

Figura 19- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias alteradas. $R=0,96$.

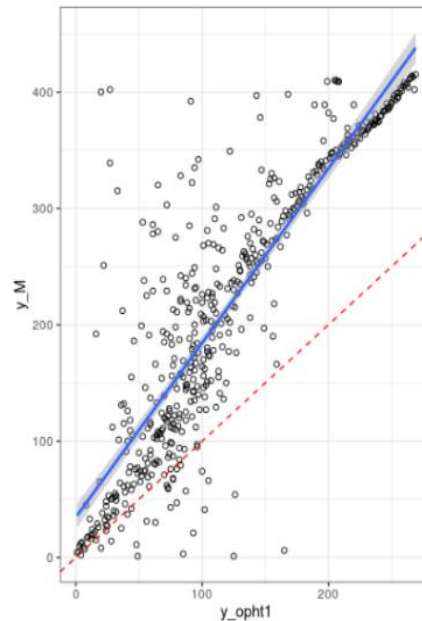
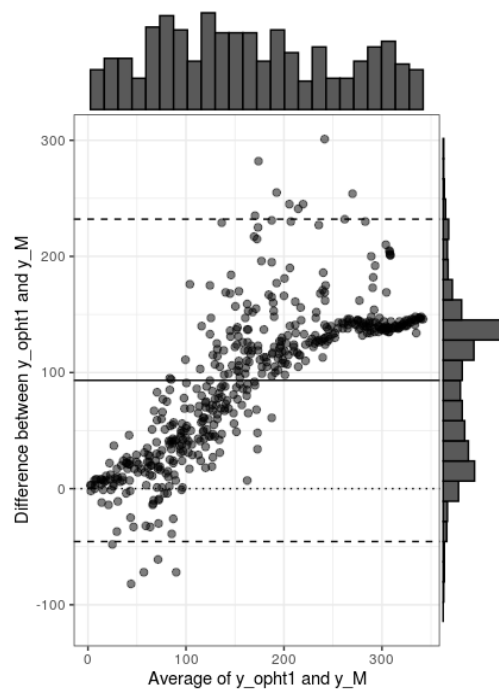


Figura 20- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias alteradas.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas teve correlação forte (Figura 21). O tamanho do erro foi de aproximadamente 75 unidades na análise de concordância (Figura 22).

Figura 21- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas. $R=0,96$.

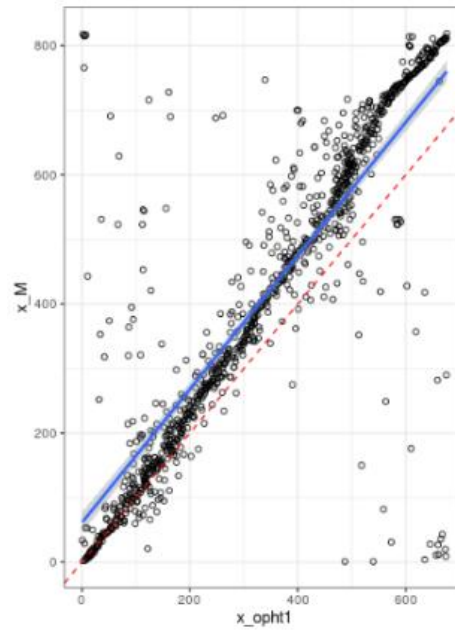
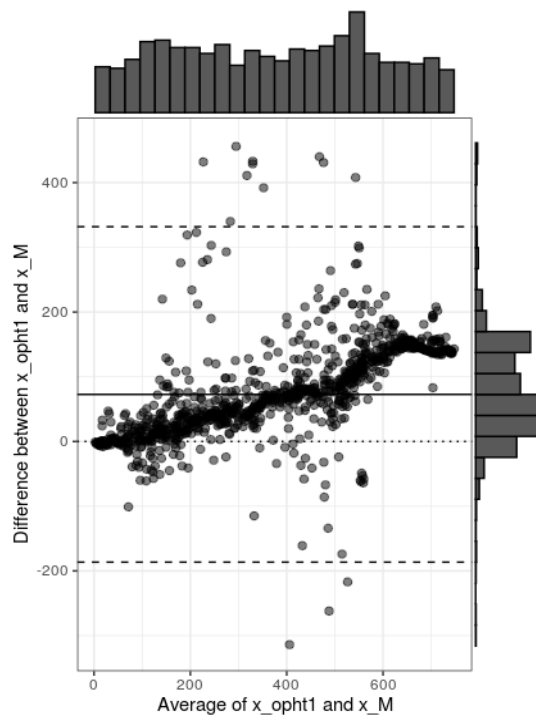


Figura 22- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas.



Nos gráficos da 1ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas teve correlação forte (Figura 23). A diferença média ultrapassou as 100 unidades distante do 0 (Figura 24).

Figura 23- Correlação entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas. $R=0,96$.

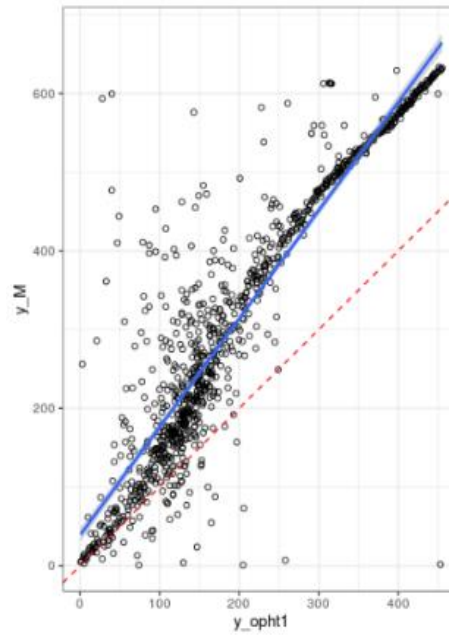
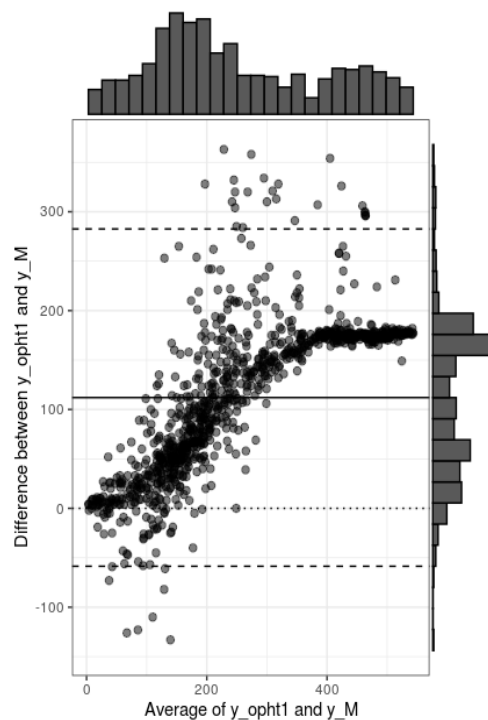


Figura 24- Grau de concordância entre a 1ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais teve forte correlação (Figura 25), e o tamanho do erro foi de cerca de 35 unidades acima do 0 (Figura 26).

Figura 25- Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais. $R=0,96$.

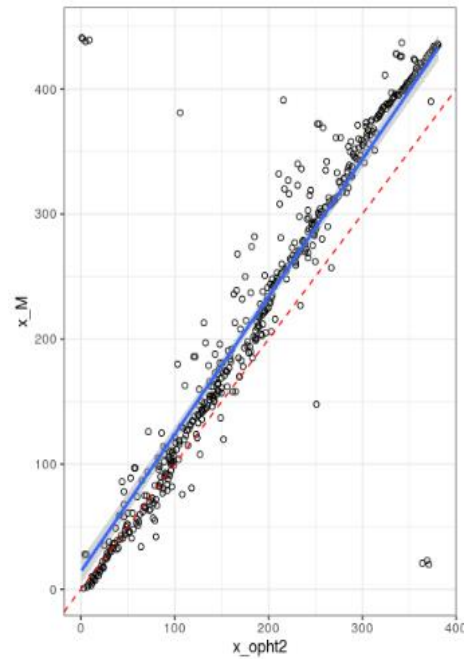
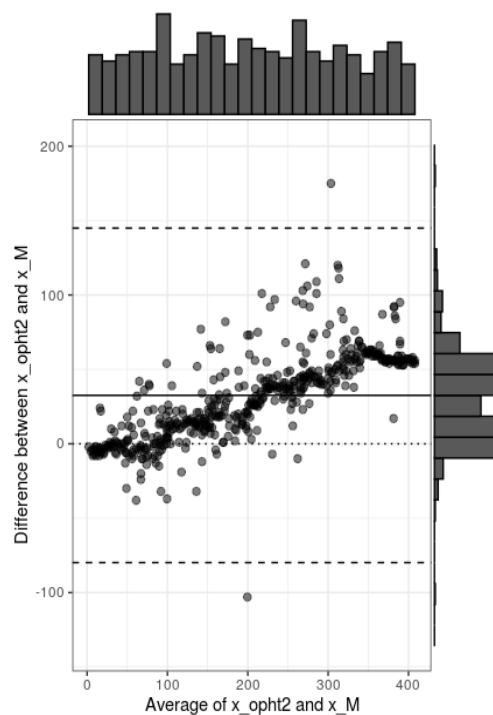


Figura 26- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias normais.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais o gráfico de correlação foi forte (Figura 27). O grau de concordância teve uma diferença média de aproximadamente 50 unidades (Figura 28).

Figura 27- Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais. $R=0,96$.

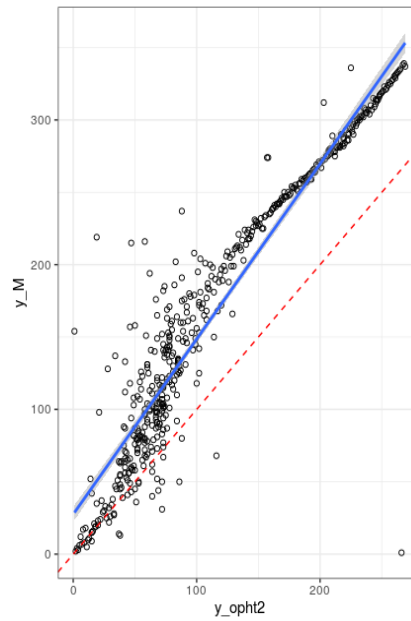
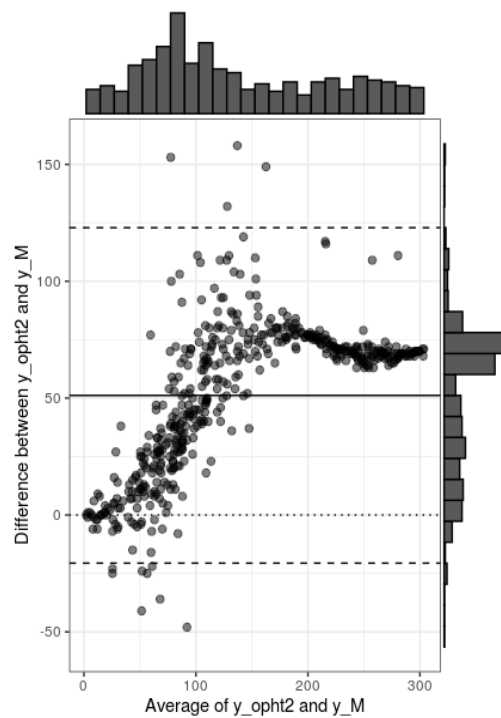


Figura 28- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias normais.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias alteradas teve forte correlação (Figura 29) e a diferença média foi aproximadamente 50 unidades acima do 0 (Figura 30).

Figura 29- Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias alteradas. $R=0,96$.

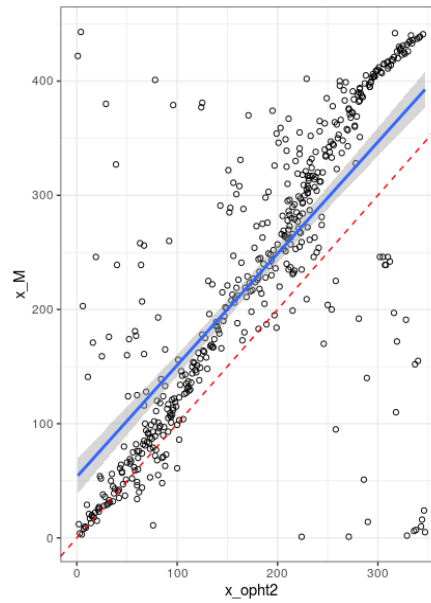
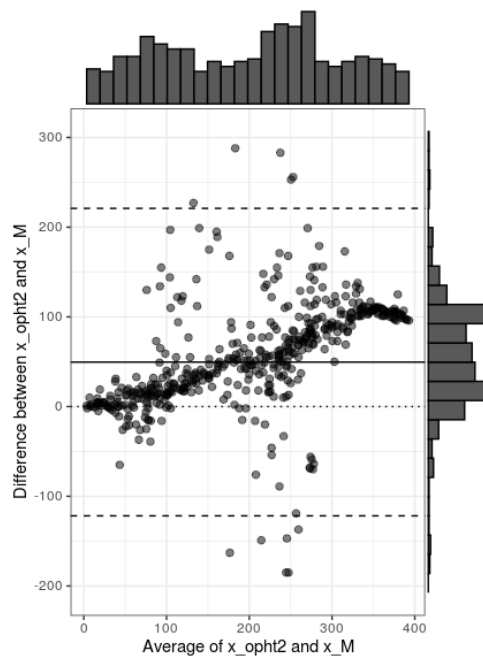


Figura 30- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias alteradas.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias alteradas a correlação foi forte (Figura 31). O tamanho da diferença entre as observações foi de aproximadamente 100 unidades (Figura 32).

Figura 31-Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias alteradas. $R=0,96$.

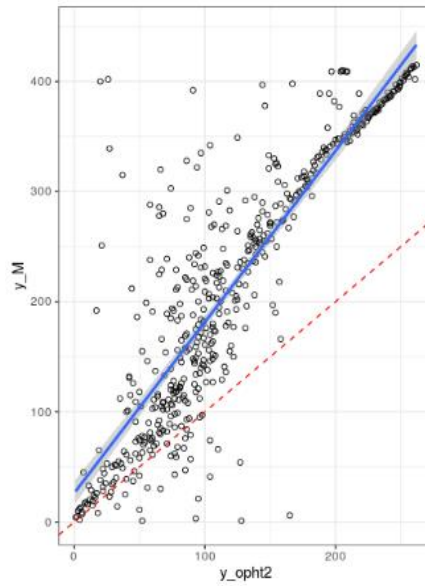
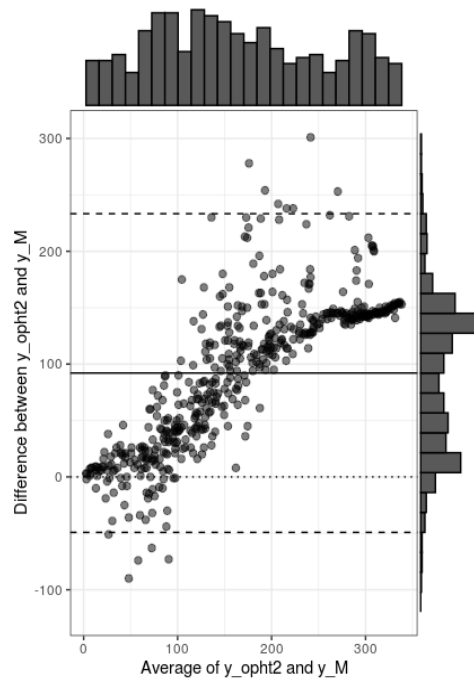


Figura 32- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias alteradas.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'X' em fundoscopias normais conjuntamente com alteradas a correlação foi forte (Figura 33) e a diferença média das observações foi de aproximadamente 75 unidades acima do 0 (Figura 34).

Figura 33- Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias conjuntamente com alteradas. $R=0,96$.

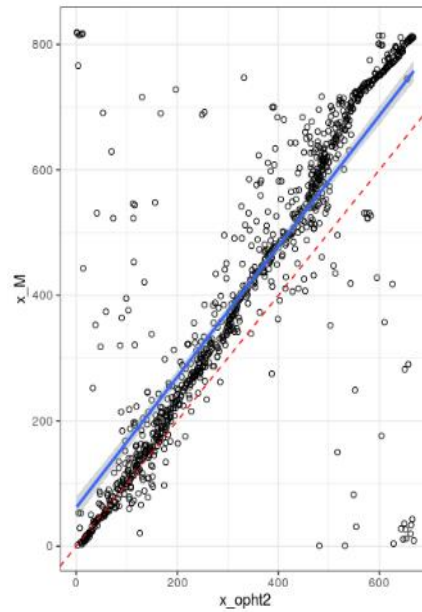
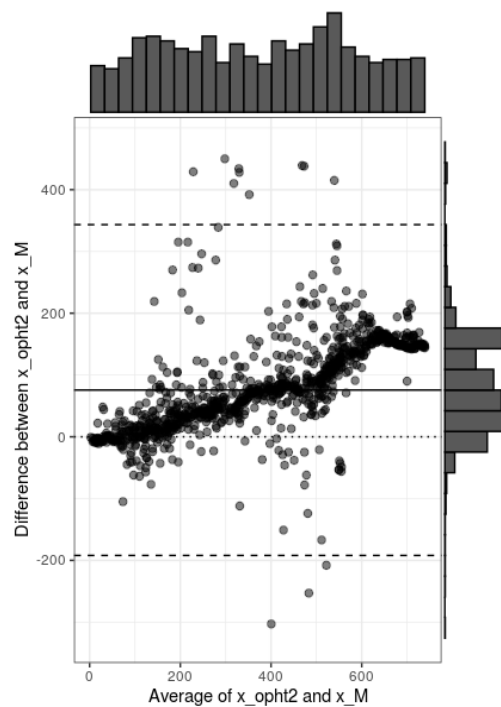


Figura 34- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'X' em fundoscopias conjuntamente com alteradas.



Nos gráficos da 2ª observação do oftalmologista versus a máquina coordenada 'Y' em fundoscopias normais conjuntamente com alteradas teve forte correlação (Figura 35), e a diferença média entre a detecção do oftalmologista e a detecção da máquina foi cerca de 100 unidades acima do 0 (Figura 36).

Figura 35- Correlação entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias conjuntamente com alteradas. $R=0,96$.

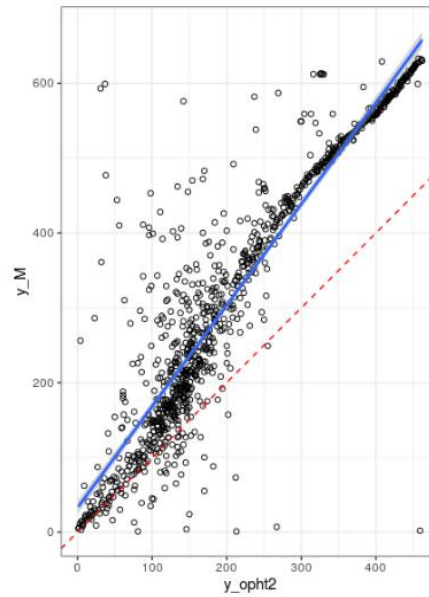
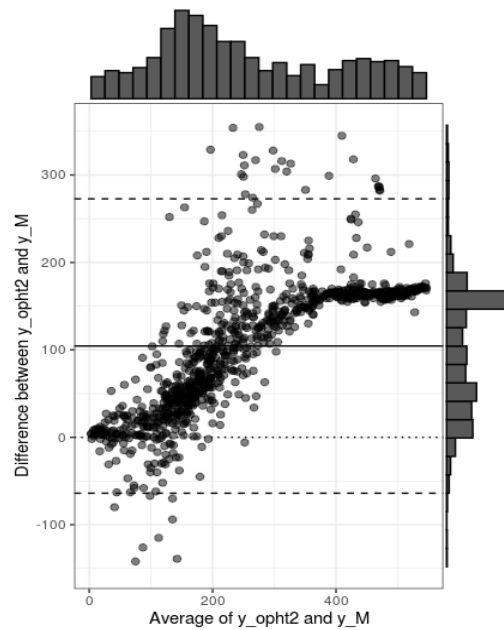
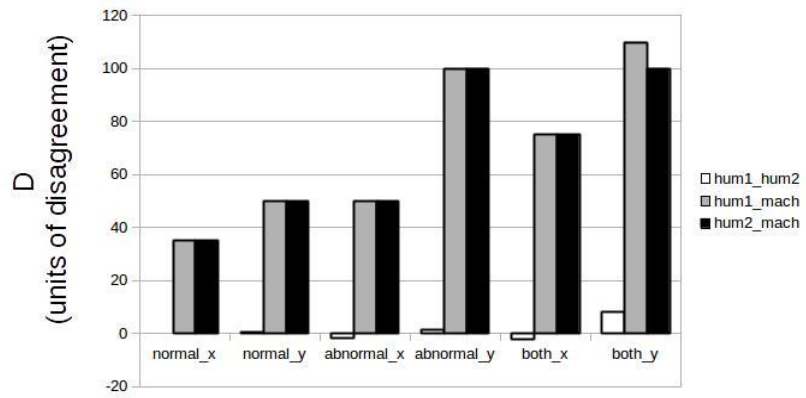


Figura 36- Grau de concordância entre a 2ª observação do oftalmologista versus máquina coordenada 'Y' em fundoscopias conjuntamente com alteradas.



A figura 37 resume os resultados das discordâncias de acordo com o grupo de fundoscopias, coordenadas e observadores.

Figura 37- Unidades de diferença média entre as coordenadas, grupos de fundoscopias e observadores.



6 DISCUSSÃO

Esse estudo teve como objetivo comparar o desempenho das observações e marcações do centro do disco óptico entre um oftalmologista e pela visão computacional através da análise de Bland-Altman. Foram gerados graus de concordância onde foi possível calcular a diferença média do tamanho do erro entre uma observação e outra.

A partir disso, foi visto que a 1ª e a 2ª observação do mesmo oftalmologista teve um grau de concordância forte em todas as coordenadas, analisando tanto fundoscopias normais, alteradas e normais conjuntamente com alteradas. A maior diferença média foi cerca de 8 unidades acima do 0, porém todos os resultados estão dentro do intervalo de confiança. Semelhante a isso, Fu *et al.*(24) propôs a segmentação do disco óptico e da escavação do disco óptico manualmente por 7 oftalmologistas em fundoscopias normais e em fundoscopias com glaucoma, sendo que na segmentação do disco óptico houve pouca diferença estatística (AUC entre 0,89 e 0,85), já na segmentação da escavação do disco óptico houve uma grande variação estatística entre os oftalmologistas (AUC entre 0,47 e 0,85).

Na comparação das duas marcações do oftalmologista em relação a marcação da máquina foi visto que as diferenças médias aumentaram significativamente todos os grupos de fundoscopias tanto da coordenada X quanto da coordenada Y. Nas fundoscopias alteradas coordenada Y e nas normais conjuntamente com alteradas nas duas coordenadas as diferenças médias chegaram a 100 unidades acima do 0. Percebeu-se que a variação do que a máquina detectava como centro do disco óptico era maior quando estavam envolvidas as fundoscopias alteradas e coordenada Y, provavelmente devido as dificuldades de detecção em imagens com baixa luminosidade, borramento, baixo contraste e das próprias patologias.

Alshayegi *et al.*²⁵ na detecção do disco óptico pela máquina também encontrou dificuldades em algumas fundoscopias e algumas imagens deixaram de ser marcadas, principalmente nas fundoscopias com contraste ruim do disco óptico, com muita variação de iluminação de fundo, similaridade de cor e contraste do disco óptico com os exsudatos da retinopatia diabética. A localização do disco óptico também foi dificultada nas fundoscopias com mais patologias.

Os estudos de identificação e segmentação de estruturas são importantes, pois são o primeiro passo para poder usar a inteligência artificial em doenças que tem biomarcadores e lesões específicas²⁶, como no caso do glaucoma. A segmentação automática pode também facilitar e otimizar o trabalho convencional de identificação estrutural do nervo óptico²⁷, pois manualmente consome muito tempo²⁸. Por essa razão esse trabalho tem finalidade de analisar o desempenho de detecção do centro do disco óptico pela máquina, com o fim de chegar à identificação de nervos ópticos glaucomatosos com essa tecnologia.

Li et al.²⁹ estudou a eficácia de detecção da neuropatia do disco óptico em glaucoma usando fundoscopias normais e com glaucoma através da tecnologia de deep learning com um $n=48116$, e mostrou um resultado robusto, com $AUC=0,986$, sensibilidade de 95,6% e especificidade de 92% na detecção de fundoscopias referidas com neuropatia do disco óptico glaucomatosa pelos 21 especialistas. Durante a escolha das fundoscopias as que tinham qualidade ruim foram excluídas do estudo. O maior motivo pelo surgimento de falsos negativos ($n=87$) foi pela coexistência de outros fatores nas imagens, incluindo os patológicos ou alta miopia.

Da mesma forma, Christopher et al.³⁰ teve uma alta acurácia ($AUC=0,91$) na identificação de nervo óptico glaucomatoso (previamente classificado por um especialista.), e chegou a ter uma performance melhor em fundoscopias com glaucoma consideradas moderada a severa ($AUC=0,97$). Porém, a qualidade inadequada da retinografia com borramento ou baixo contraste pode fazer o modelo gerar erros e ter resultados de baixa confiança ($AUC<0,5$).

Nesse estudo, somente em 2 imagens o algoritmo utilizado não conseguiu identificar e gerar marcações, o que gerou boa perspectiva para o modelo. As retinografias estão representadas na Figura 38 e Figura 39. Em ambas as imagens podemos perceber a presença de borramento e um menor contraste, dando uma intensidade luminosa homogênea para as fundoscopias.

Figura 38- Retinografia de olho direito que não se obteve marcação pelo algoritmo.

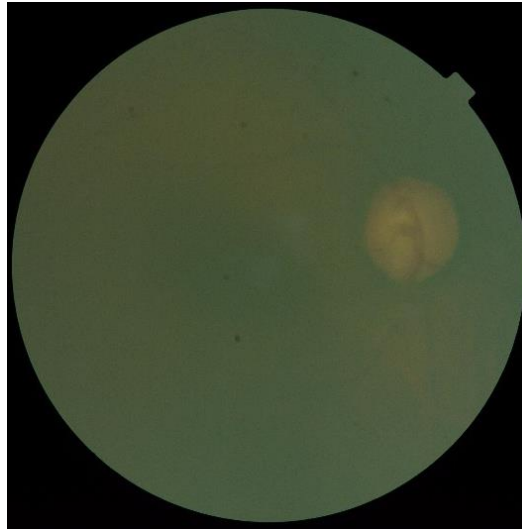
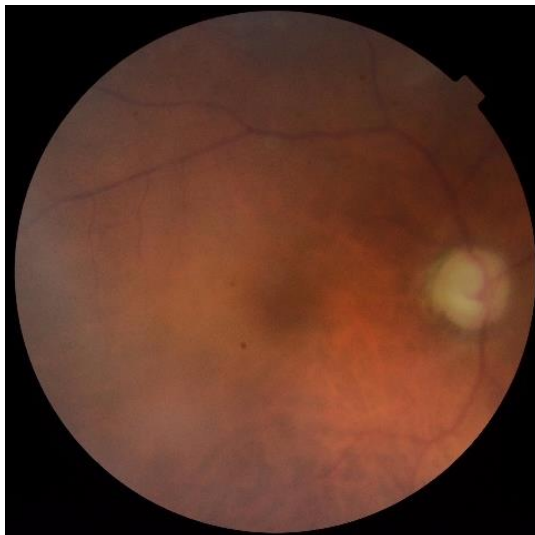


Figura 39- Retinografia de olho direito que não se obteve marcação pelo algoritmo.



Apesar de somente 2 imagens não terem sido identificadas por esse modelo, algumas coordenadas identificadas pela máquina não conseguiram ser dentro do disco óptico, como mostra a Figura 40 que o disco óptico não é representado como a área mais luminosa da imagem e na Figura 41 em que a marcação das coordenadas foi feita no segundo ponto mais luminoso da retinografia.

Figura 40- Retinografia de olho esquerdo em que a marcação pelo algoritmo foi fora dos limites do disco óptico.

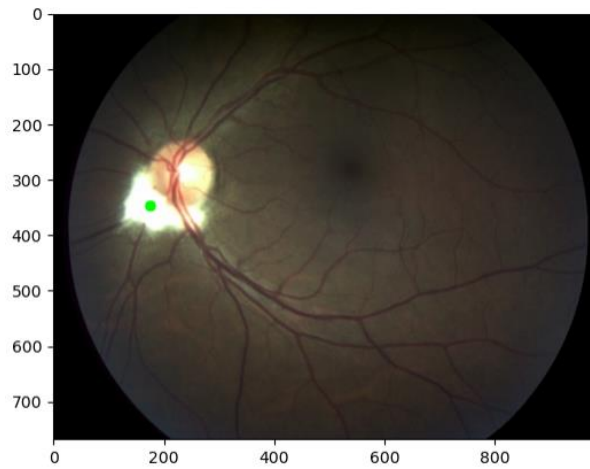
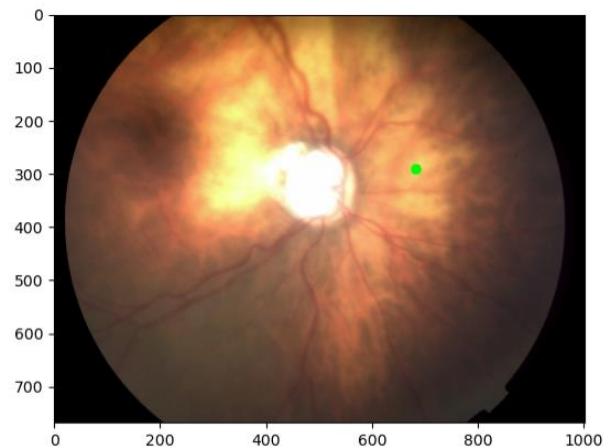


Figura 41- Retinografia de olho esquerdo em que a marcação pelo algoritmo foi fora dos limites do disco óptico.



Liu et al.³¹ também propôs a utilização de deep learning na identificação de nervos ópticos glaucomatosos em diferentes populações. As fundoscopias utilizadas passaram por controle de qualidade, a qual tentou diminuir as imagens com significativos artefatos, baixa resolução e iluminação inadequada. Na população local (Chinese Glaucoma Study Alliance) a AUC chegou a alcançar 0,996, com especificidade de 97,7% e sensibilidade de 96,2%. Nas demais populações houve uma AUC maior que 0,964, especificidade e sensibilidade maior que 80,8% e 87,7%, respectivamente. Quando o modelo foi utilizado em imagens de qualidade variadas obtidas na internet o resultado foi: AUC=0,823, Especificidade de 70,4% e Sensibilidade de 82,2%.

7 CONCLUSÃO

O estudo da comparação de detecção do disco óptico em fundoscopias entre um especialista e um algoritmo de IA permite nos concluir que o modelo utilizado é promissor.

Podem ser feitas algumas análises com os graus de concordâncias gerados pelos gráficos de Bland-Altman. Quando é realizada a comparação intra-especialista a discordância média em todos os grupos de fundoscopias e coordenadas foram semelhantes e próximas ao 0.

Na comparação das duas observações do oftalmologista versus máquina, a discordância entre os dois ficou evidenciada com os aumentos das diferenças médias. Porém em alguns grupos de fundoscopias as detecções do centro do disco óptico pela máquina não tiveram discrepâncias unitárias tão fortes quanto outras. Nas fundoscopias sem patologias a detecção da máquina tendeu a se aproximar mais das marcações feitas pelo especialista nas duas observações.

O grupo de fundoscopias alteradas e normais conjuntamente com alteradas foram onde as discordâncias foram mais acentuadas. Somado a isso, quando a coordenada Y estava envolvida, as discordâncias alcançaram os maiores valores unitários, chegando a ultrapassar as 100 unidade acima do 0. Portanto, há uma relação quanto a coexistência de diversos fatores presentes na imagem, incluindo os patológicos, no desempenho do modelo utilizado.

Um dos fatores que contribuíram para o aumento da diferença média foi a luminosidade das retinografias, que é bastante descrito na literatura. O algoritmo usado pela máquina se baseia também na iluminação que o disco óptico transmite, o que pode ser confundido com outros focos de iluminação produzidos por certas alterações patológicas ou prejudicado pelo contraste inadequado da imagem.

Apesar de um certo grau de discordância entre o especialista e a máquina, o algoritmo usado nesse estudo mostrou que tem grande potencial de detecção do centro do disco óptico. A discussão em torno das várias possibilidades de investigação é necessária para aumentar a capacidade do modelo e tornar a diferença a visão computacional cada vez mais próxima de um especialista.

REFERÊNCIAS

1. Lopes A.A.. Medicina Baseada em Evidências: a arte de aplicar o conhecimento científico na prática clínica. Rev. Assoc. Med. Bras. [Internet]. 2000 Sep [cited 2020 Mar 03] ; 46(3): 285-288. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-42302000000300015&lng=en. <https://doi.org/10.1590/S0104-42302000000300015>.
2. Ting, D.S.W., Peng, L., Varadarajan, A.V., Keane, P.A., Burlina, P., Chiang, M.F., Schmetterer, L., Pasquale, L.R., Bressler, N.M., Webster, D.R., Abramoff, M., Wong, T.Y., Deep learning in ophthalmology: The technical and clinical considerations, Progress in Retinal and Eye Research (2019), doi: <https://doi.org/10.1016/j.preteyeres.2019.04.003>.
3. Grewal, Parampal & Oloumi, Faraz & Rubin, Uriel & Tennant, Matthew. (2018). Deep learning in ophthalmology: a review. Canadian Journal of Ophthalmology. 53. 10.1016/j.cjco.2018.04.019.
4. An Ophthalmologist's Guide to Deciphering Studies in Artificial Intelligence Ting, Daniel S.W.Lee, Aaron Y.Wong, Tien Y. et al. Ophthalmology, Volume 126, Issue 11, 1475 – 1479
5. Rao, Chetan & Raman, Rajiv. (2019). Artificial intelligence applications for Ophthalmology: Current status. Nepalese Journal of Ophthalmology. 11. 1-4. 10.3126/nepjoph.v11i1.25409.
6. De Fauw, J., Ledsam, J.R., Romera-Paredes, B. et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 24, 1342–1350 (2018). <https://doi.org/10.1038/s41591-018-0107-6>
7. Siegfried K. Wagner, Dun Jack Fu, Livia Faes, Xiaoxuan Liu, Josef Huemer, Hagar Khalid, Daniel Ferraz, Edward Korot, Christopher Kelly, Konstantinos Balaskas, Alastair K. Denniston, Pearse A. Keane; Insights into Systemic Disease through Retinal Imaging-Based Oculomics. Trans. Vis. Sci. Tech. 2020;9(2):6. doi: <https://doi.org/10.1167/tvst.9.2.6>.
8. Edward Korot, Siegfried K. Wagner, Livia Faes, Xiaoxuan Liu, Josef Huemer, Daniel Ferraz, Pearse A. Keane, Konstantinos Balaskas; Will AI Replace Ophthalmologists?. Trans. Vis. Sci. Tech. 2020;9(2):2. doi: <https://doi.org/10.1167/tvst.9.2.2>.
9. Mohammadi, Mahdi & Al-Azab, Fadwa & Raahemi, Bijan & Richards, Gregory & Jaworska, Natalia & Smith, Dylan & de la Salle, Sara & Blier, Pierre & Knott, Verner. (2015). Data mining EEG signals in depression for their diagnostic value. BMC medical informatics and decision making. 15. 108. 10.1186/s12911-015-0227-6.

10. Kaushik S. (2019). Commentary: Linear discriminant score for differentiating early primary open angle glaucoma from suspects. *Indian journal of ophthalmology*, 67 (1), 81–82. https://doi.org/10.4103/ijo.IJO_1506_18
11. García-Gonzalo, Esperanza & Fernández-Muñiz, Zulima & Garcia Nieto, Paulino Jose & Sánchez, Antonio & Menéndez, Marta. (2016). Hard-Rock Stability Analysis for Span Design in Entry-Type Excavations with Learning Classifiers. *Materials*. 9. 531. 10.3390/ma9070531.
12. Tucker, Allan & Vinciotti, Veronica & Liu, Xiaohui & Garway-Heath, David. (2005). A spatio-temporal Bayesian network classifier for understanding visual field deterioration. *Artificial intelligence in medicine*. 34. 163-77. 10.1016/j.artmed.2004.07.004.
13. Jake VanderPlas. 2016. *Python Data Science Handbook: Essential Tools for Working with Data* (1st. ed.). O'Reilly Media, Inc.
14. Matthäus Pilch, Knut Stieger, Yaroslava Wenner, Markus N. Preising, Christoph Friedburg, Erdmuthe Meyer zu Bexten, Birgit Lorenz; Automated Segmentation of Pathological Cavities in Optical Coherence Tomography Scans. *Invest. Ophthalmol. Vis. Sci.* 2013;54(6):4385-4393. doi: <https://doi.org/10.1167/iovs.12-11396>.
15. Scikit-learn: Machine Learning in Python, Pedregosa et al. , *JMLR* 12, pp. 2825-2830, 2011.
16. Smadja, David & Touboul, David & Cohen, Ayala & Doveh, Etti & Santhiago, Marcony & Mello, Glauco & Krueger, Ronald & Colin, Jesus. (2013). Detection of Subclinical Keratoconus Using an Automated Decision Tree Classification. *American journal of ophthalmology*. 156. 10.1016/j.ajo.2013.03.034.
17. 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(22):2402–2410. doi:10.1001/jama.2016.17216
18. Tabuchi, Hitoshi & Kamiura, Notake & Saitoh, Aumu & Isokawa, Teijiro & Matsui, Nobuyuki. (2014). Ophthalmological Examination Determination Using Data Classification Based on Support Vector Machines and Self-Organizing Maps. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*. 26. 559-572. 10.3156/jsoft.26.559.
19. Tran, Bach & Ha, Giang & Vu, Giang & Vuong, Quan-Hoang & Ho, Tung & Vuong, Thu-Trang & La, Viet-Phuong & Ho, Toan & Nghiem, Kien-Cuong & Nguyen, Huong & Latkin, Carl & Tam, Wilson & Cheung, Ngai & Nguyen, Hong Kong & Ho, Cyrus & Ho, Roger. (2019). Global Evolution of Research in Artificial Intelligence in Health and Medicine: A Bibliometric Study. *Journal of Clinical Medicine*. 8. 10.3390/jcm8030360.

20. Livia Faes, Xiaoxuan Liu, Siegfried K. Wagner, Dun Jack Fu, Konstantinos Balaskas, Dawn A. Sim, Lucas M. Bachmann, Pearse A. Keane, Alastair K. Denniston; A Clinician's Guide to Artificial Intelligence: How to Critically Appraise Machine Learning Studies. *Trans. Vis. Sci. Tech.* 2020;9(2):7. doi: <https://doi.org/10.1167/tvst.9.2.7>.
21. Saito T, Rehmsmeier M (2015) The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* 10(3): e0118432. <https://doi.org/10.1371/journal.pone.0118432>
22. Jun. (Mar, 3 2020). Deep Learning Algorithms <https://medium.com/jun-devpblog/dl-1-learning-algorithms-and-basic-terms-of-dl-65d46ceb1b0a>
23. Altman DG, Bland JM (1983). "Measurement in medicine: the analysis of method comparison studies". *The Statistician*. 32(3): 307–317.
24. FU, Huazhu et al. A retrospective comparison of deep learning to manual annotations for optic disc and optic cup segmentation in fundus photographs. **Translational vision science & technology**, v. 9, n. 2, p. 33-33, 2020.
25. Alshayeji, Mohammad et al. Optic disc detection in retinal fundus images using gravitational law-based edge detection. **Medical & biological engineering & computing**, v. 55, n. 6, p. 935-948, 2017.
26. Schmidt-Erfurth, Ursula et al. Artificial intelligence in retina. **Progress in retinal and eye research**, v. 67, p. 1-29, 2018.
27. Zheng, Yuanjie et al. Optic disc and cup segmentation from color fundus photograph using graph cut with priors. In: **International Conference on Medical Image Computing and Computer-Assisted Intervention**. Springer, Berlin, Heidelberg, 2013. p. 75-82.
28. Almazroa, Ahmed et al. Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey. **Journal of ophthalmology**, v. 2015, 2015.
29. Li, Zhixi et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. **Ophthalmology**, v. 125, n. 8, p. 1199-1206, 2018.
30. Christopher, Mark et al. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. **Scientific reports**, v. 8, n. 1, p. 1-13, 2018.
31. Liu, Hanruo et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. **JAMA ophthalmology**, v. 137, n. 12, p. 1353-1360, 2019.

**APÊNDICE A – PARECER DO ORIENTADOR SOBRE A VERSÃO PRELIMINAR
DO TC PARA ANÁLISE PRÉVIA**

NOME DOS AUTORES

JOÃO GABRIEL DE OLIVEIRA MENDES DA ROCHA

TÍTULO DO TRABALHO

**“UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO NEURO-
OFTALMOLÓGICO: APRENDIZADO A PARTIR DE DADOS REAIS E
DESENVOLVIMENTO DE MÉTODOS PARA ENFRENTAMENTO DE DESAFIOS
ATUAIS E FUTUROS”**

Declaro junto a Coordenação do Trabalho de Curso do Cesupa que li a versão preliminar do TC que tem como título: “Utilização de Inteligência Artificial no Diagnóstico Neuro-Oftalmológico: Aprendizado a partir de dados reais e desenvolvimento de métodos para enfrentamentos de desafios atuais e futuros” e considero que a mesma se encontra em condições de submissão à análise prévia por parte de um dos membros da banca examinadora, assim como, aguardo sugestões para correção e/ou aprimoramento da qualidade do trabalho.

Belém, 24 / 05 / 2021


Prof.(a) Dr(a) NOME: _____
Orientador(a)

**APÊNDICE B - PARECER DO ORIENTADOR SOBRE A VERSÃO DO TC PARA
DEFESA PÚBLICA**

JOÃO GABRIEL DE OLIVEIRA MENDES DA ROCHA

**UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO NEURO-
OFTALMOLÓGICO: APRENDIZADO A PARTIR DE DADOS REAIS E
DESENVOLVIMENTO DE MÉTODOS PARA ENFRENTAMENTO DE DESAFIOS
ATUAIS E FUTUROS**

Declaro junto a Coordenação do Trabalho de Curso do CESUPA que li a versão final do TC que tem como título: “UTILIZAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO DIAGNÓSTICO NEURO-OFTALMOLÓGICO: APRENDIZADO A PARTIR DE DADOS REAIS E DESENVOLVIMENTO DE MÉTODOS PARA ENFRENTAMENTO DE DESAFIOS ATUAIS E FUTUROS” e considero que a mesma se encontra em condições de submissão à banca examinadora durante a Jornada de Defesa de Trabalho de Curso do Curso de Bacharelado em Medicina do Cesupa.

Belém, ___/___/___



Prof. Dr. Cláudio Eduardo Corrêa Teixeira.

ANEXO A – APROVAÇÃO DO CEP DO PROJETO DE PESQUISA



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Utilização de Inteligência Artificial no diagnóstico neuro-oftalmológico: aprendizado a partir de dados reais e desenvolvimento de métodos para enfrentamento de desafios atuais e futuros

Pesquisador: Claudio Eduardo Corrêa Teixeira

Área Temática:

Versão: 2

CAAE: 39292420.2.0000.5169

Instituição Proponente: Centro Universitário do Pará - CESUPA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.418.779

Apresentação do Projeto:

As informações contidas nos campos Apresentação do Projeto, Objetivo da Pesquisa e Avaliação dos Riscos e Benefícios foram retiradas do documento

"PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1603053.pdf", datado de 10/11/2020, em que lê-se, no item:

INTRODUÇÃO:

Na década de 1990 uma grande mudança de paradigma surgiu no diagnóstico médico, intitulada de Medicina Baseada em Evidências, no qual houve uma proposta de inclusão de dados de Epidemiologia Clínica no processo de decisão profissional, antes dominado pela semiologia e propedêutica isoladas desde a organização da ciência por Hipócrates. A partir deste período revistas científicas, livros e congressos médicos

difundiram termos estatísticos como Acurácia, Sensibilidade, Especificidade, Valor Preditivo Positivo, Razão de Verossimilhança, Odds Ratio, Risco Relativo, Número Necessário para Tratar e muitos outros. Tivemos todos de nos adequar a esta medicina mais precisa, tanto no âmbito da

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

UF: PA

Município: BELEM

CEP: 66.060-232

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

pesquisa como no ofício diário frente ao paciente¹. Nos últimos anos, uma nova revolução está acontecendo, a inclusão das ditas Inteligências Artificiais (IA) na área da Saúde. A necessidade de aprender a lidar com estas novas ferramentas, entender seus conceitos, métricas, aplicações e limitações já existem hoje. Em diversas atividades diárias, desde o uso de um aplicativos de GPS, de tradução de texto, ou uma simples busca de termos em um navegador já possuem IA inseridas, agora, novas aplicações surgiram no âmbito da Oftalmologia, como o IDx-DR, a primeira ferramenta diagnóstica automatizada autorizada pelo Food and Drugs Administration (FDA) a executar a predição de Retinopatia Diabética por meio de Retinografias². A IA foi conceitualizada em 1956, após um workshop no Dartmouth College (Figura 1). O termo Machine Learning (ML) foi posteriormente usado por Arthur Samuel em 1959 que afirmou "O computador deve ter a capacidade de aprender usando várias técnicas estatísticas, sem ser explicitamente programado." Usando o ML, o algoritmo pode aprender e fazer previsões com base nos dados que foram alimentados na fase de treinamento, usando um método de abordagem supervisionada ou não supervisionada. O ML foi amplamente adotado em aplicações como visão computacional e análise preditiva usando modelos matemáticos complexos. Com o advento das unidades de processamento gráfico (GPUs), os avanços nos modelos matemáticos, a disponibilidade de grandes conjuntos de dados e sensores de baixo custo, as técnicas de aprendizado profundo - Deep learning (DL) - posteriormente despertaram um interesse feroz e foram aplicadas em muitos setores. Assim, com o auxílio de AI os profissionais podem diminuir os custos com atendimentos especializados além do benefício do compartilhamento na utilização dos sistemas de informação, visto que os dados gerados pelos atendimentos dos pacientes podem ser agregados e organizados produzindo um contexto que apoiará também a tomada de reconhecimento automatizado de imagens permite que os computadores analisem e processem as imagens da retina e subsequentemente distingam lesões por meio de escalas de diretrizes previamente aprendidas pelo algoritmo de IA. Há uma promessa de que a conexão entre IA e teleoftalmologia fortalecerá os padrões de referência e melhorará o atendimento ao paciente⁴. A IA envolve DL por meio de conjuntos de dados sem programação explícita, e o DL envolve auto-treinamento adicional usando "redes neurais artificiais". O DL envolve a identificação de uma

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@oesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

correlação entre dados em grandes redes neurais nos conjuntos de dados. Dentro de especialidades baseadas em imagens, como oftalmologia, as redes neurais convolucionais utilizam dados rotulados para reconhecer padrões, com evidências publicadas de suas capacidades para retinopatia diabética, glaucoma e retinopatia da prematuridade e degeneração macular relacionada à idade, e até mesmo analisando sua aplicação clínica na triagem de catarata⁵. Decisões em relação ao tratamento do paciente gerando impacto direto na redução de custos na área da saúde^{3,4}. A IA oferece um farol de esperança para melhorar algumas lacunas inerentes aos cuidados de saúde operados por seres humanos incluindo erros de diagnóstico, desperdício de recursos e tempo insuficiente entre médico e paciente, resultado do desequilíbrio entre demanda e oferta.

Os avanços na IA têm o potencial de criar wavefronts por meio das vias de teleoftalmologia nos serviços nacionais de saúde. O reconhecimento automatizado de imagens permite que os computadores analisem e processem as imagens da retina e subsequentemente distingam lesões por meio de escalas de diretrizes previamente aprendidas pelo algoritmo de IA. Há uma promessa de que a conexão entre IA e teleoftalmologia fortalecerá os padrões de referência e melhorará o atendimento ao paciente⁴. A IA envolve DL por meio de conjuntos de dados sem programação explícita, e o DL envolve auto-treinamento adicional usando "redes neurais artificiais". O DL envolve a identificação de uma correlação entre dados em grandes redes neurais nos conjuntos de dados.

Dentro de especialidades baseadas em imagens, como oftalmologia, as redes neurais convolucionais utilizam dados rotulados para reconhecer padrões, com evidências publicadas de suas capacidades para retinopatia diabética, glaucoma e retinopatia da prematuridade e degeneração macular relacionada à idade, e até mesmo analisando sua aplicação clínica na triagem de catarata...

METODOLOGIA PROPOSTA:

Os principais pontos para a análise seguem abaixo:
3.1 - Os métodos do estudo foram pré-definidos?
A definição dos métodos aplicados ao desenvolvimento da IA previamente a sua criação reduz dois vieses:
(1) Estudos Negativos - Aquelas IA que falham em sua tarefa de rejeitar a hipótese de nulidade frequentemente

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

não são publicadas em periódicos, desta forma, as evidências podem ser utopicamente direcionadas a favor de tecnologias com alto desempenho; (2) De outra forma, as publicações que não tem métodos prédefinidos tendem a publicar somente os aspectos positivos dos desfechos encontrados. Destarte, deve-se atentar para o planejamento estatístico do artigo, seus desfechos primários e secundários esperados, abordagem para validação, critérios de inclusão/exclusão, aplicação do teste isoladamente ou em exames multimodais, bem como em qual etapa da linha de cuidados do paciente, além da definição do seu uso como exame de triagem ou diagnóstico. 4,143.2 - A base de dados do estudo foi constituído dentro de uma linha de cuidados? Em geral, exames como a Angiofluoresceinografia somente são indicadas em pacientes com achados suspeitos ao exame clínico e OCT, portanto, a probabilidade pré-teste (de possuir doença) em pacientes submetidos a

Angiofluoresceinografia tende a ser maior do que aqueles submetidos ao OCT. Isto demonstra como a linha de cuidado de um paciente apresenta diversos valores de probabilidade pré- e pós-teste a cada etapa do diagnóstico e assistência, sendo assim, nenhum exame, nem mesmo aqueles que envolvem IA devem ser considerados fora da linha de cuidado. Definir em qual etapa da linha de cuidado foi aplicada a IA que está agora sendo utilizada em consultório é fundamental para reproduzir os indicadores de acurácia determinados em sua criação. Atentar também para base de dados públicos utilizados na criação de uma IA, como, por exemplo, MESSIDOR para fundus photographs dataset, pois, via de regra, estes bancos não dispõem de detalhes sobre suas imagens, como por exemplo, em quem e onde foram coletadas. 4,143.3 - De que forma os dados foram

divididos? Em geral, a prática comum no desenvolvimento de uma IA é a divisão da base de dados

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.772

em

grupos de Treinamento (Ensino da IA), Teste (Afinar seus parâmetros), e Validação (confirmar seu poder de generalização), em proporções que frequentemente giram em torno de 60/20/20%, respectivamente. Os grupos de Treinamento e Teste são de certa forma analisados pela IA na sua criação, enquanto o grupo de Validação deve se constituir de imagens nunca vistas durante este processo, de forma que se prove com o desempenho em meio a estes dados que a IA aprendeu de fato algo (consegue generalizar), e não simplesmente gravou os dados dos grupos de treino/teste. Problemas comuns a atentar nesta etapa são os seguintes: (1) Frequentemente as bases de dados para criação de IA se constituem de mais de uma fotografia do mesmo paciente, portanto, não se pode permitir que no grupo de validação estejam fotografias do mesmo paciente dos demais grupos, mesmo que sejam imagens de momentos diferentes, isso reduz o valor da validação dos dados. (2) Embora a validação interna dos dados tenha seu valor, o mais importante é que os dados de acurácia da IA em questão tenham validação externa, o que significa que foi feita análise de seu desempenho em base de dados totalmente diversa daquela usada em sua criação, preferencialmente distinta geograficamente, temporalmente e realizada por outro grupo de pesquisa, talvez até mesmo em outra publicação científica.4,143.4 - Os dados foram divididos conforme o estado Real da doença? Como foram divididas as imagens da base de dados entre normais e com doença? Conforme registro eletrônico ou manualmente? Prospectivamente ou Retrospectivamente? (em geral a análise retrospectiva tem maior valor devido a presença da linha completa de cuidados do paciente no histórico). Outro fato importante é o de que, embora em áreas como a Oncologia o padrão-ouro seja o...

METODOLOGIA DE ANÁLISE DE DADOS:

Métricas de Desempenho vs. Indicadores Bioestatísticos A criação de IAs frequentemente envolve profissionais de áreas fora das ciências biomédicas, como cientistas da computação e de dados, pouco habituados a termos de bioestatística, frequentemente gerando confusão ao lermos artigos sobre o tema, além da complexidade já inerente aos modelos matemáticos das ferramentas.4.1 Matriz de

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

UF: PA

Telefone: (91)4009-9100

Município: BELEM

CEP: 66.060-232

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

Confusão Chamada de tabela de

contingência em estatística, referem-se às conhecidas figuras abaixo, utilizadas para representar medidas de desempenho de um teste diagnóstico.

4.2 Recall Nada mais é do que o que conhecemos como Sensibilidade, o percentual de indivíduos doentes detectados pela IA, a proporção que ela atinge no screening da população doente.

4.3 Precision Conhecida na medicina como Valor Preditivo Positivo, a proporção de indivíduos doentes dentre aqueles com resultado positivo pela IA, o quanto um resultado positivo realmente significa estar doente.

4.4 Loss A grosso modo é uma medida relacionada à diferença do valor predito pela IA do valor real durante o treino.

4.5 Medidas de Acurácia A acurácia nada mais é do que a taxa de acertos global, o quanto a IA alcança o diagnóstico sendo ele normal ou alterado em proporção. F1-Score, outro indicador muito utilizado é a média harmônica do Precision e Recall.

4.6 Como avaliar a curva de desempenho de uma IA Conhecida como curva ROC (receiveroperating curve), ela traça a variação da especificidade de um teste diagnóstico conforme se aumenta a proporção de doentes rastreados (sensibilidade). É excelente para definir o melhor ponto de corte de um exame, bem como para medir seu desempenho global (área sob a curva ROC). No entanto, quando em IAs se utilizam grupos muito desbalanceados na validação, o que é frequente, devido a falta de casuística em subgrupos mais graves de doenças, as curvas ROC são de pouco valor, oferecendo muitas vezes resultados otimistas falsos, como demonstrado na figura abaixo, sendo melhor aplicáveis as curvas de Precision-Recall, que cruzam a sensibilidade com o Valor Preditivo Positivo, menos susceptíveis a dados de grupos desbalanceados. Importante avaliar todos estes parâmetros em cada grupo da base de dados, desde o treinamento até o grupo de validação. Sendo a diferença entre esse grupos indicador de Overtraining/Overfitting como explicado abaixo. Os parâmetros melhor

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

correlacionados

a prática diária são os do grupo de validação externa. Também vale ficar atento ao intervalo de confiança de 95%, frequentemente utilizado em estudos e ensaios clínicos.^{4,144.7} Mapas de Calor Quando disponíveis, esses mapas demonstram quais regiões da imagem foram vitais para que a IA definisse o seu diagnóstico, sendo

fundamental para reduzir um dos dilemas do DL, chamado Black Box Dilemma, que consiste em sabermos o resultados da IA sem saber os meios pelos quais os produziu, muito frequente em Redes Neurais Artificiais. Com estes mapas, identificamos, pelo menos, com base em que região da imagem a IA classificou aquele exame.^{4,145} - Overtraining e Overfitting Overfitting se refere a uma situação muito importante no desenvolvimento de uma IA, muitas vezes, apesar de resultados otimistas no artigo de criação da rede, o desempenho real da mesma, seja na validação externa, quando existe, ou na prática diária, ficam abaixo do esperado. Isso acontece pois um treino excessivo (overtraining) ou com uma base de dados insuficiente tende a tornar o poder de generalização da IA muito restrito, ela praticamente grava as características das imagens com as quais foi treinada e não consegue transferir este aprendizado para imagens diversas. Nem sempre mais

tempo de treino é melhor, e acompanhar gráficos de desempenho como os abaixo nos ajuda a identificar qual seria o ponto ótimo de treinamento.^{4,14}.

Objetivo da Pesquisa:

Objetivo Primário:

Analisar a aplicabilidade de variedades de IA na avaliação de dados neuro-oftalmológicos.

Objetivo Secundário:

- Comparar estatisticamente o desempenho de diferentes ferramentas de IA na avaliação de dados neurooftalmológicos

- Identificar e propor

correções as limitações das ferramentas de IA na avaliação de dados neuro-oftalmológicos

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

- Desenvolver novas propostas de ferramenta de IA para avaliação de dados neuro-oftalmológicos.

Avaliação dos Riscos e Benefícios:

Riscos:

A metodologia de avaliação a ser empregada apresenta riscos mínimos, principalmente por ser tratar de avaliação de desempenho de ferramentas de IA na análise dados de pacientes constantes dos arquivos de instituições colaboradoras neste projeto, cujos procedimentos analíticos desidentificam os dados por agrupamento quantitativo dos mesmos. Entretanto, para prevenir o risco de eventuais perdas ou divulgações desses dados fora do escopo do presente projeto, todos os arquivos serão codificados e permanecerão sob a guarda do pesquisador responsável, proponente deste projeto.

Benefícios:

Os resultados desta pesquisa poderão ajudar e esclarecer melhor sobre o desempenho de ferramentas de IA na avaliação de dados neurooftalmológicos, o que, além de contribuir para a literatura correlata, também contribuirá para a inserção dessas ferramentas na prática médica neurooftalmológica, o que potencialmente facilitará e agilizará o processo de diagnóstico.

Comentários e Considerações sobre a Pesquisa:

É importante que se tenha uma visão crítica do suporte dado por uma IA no contexto médico. Propostas no sentido de regulamentar essas ferramentas já foram lançadas, como a do FDA no seu Proposed Regulatory Framework for Modifications for Artificial Intelligence/Machine Learning Based Software as a Medical Device (SaMD), e novas regulações provavelmente surgirão. As IA são tecnologias de suporte, mas confiar em seus resultados depende de compreender aspectos muitas vezes distantes do histórico de formação profissional nas áreas biológicas e de saúde. Precisamos, portanto, aprender a medir o desempenho das ferramentas de IA assim como sabemos medir o de um ser humano, pois reconhecendo as limitações de ambos, poderemos integrar mais efetivamente essas novas ferramentas a prática

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA

Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



CENTRO UNIVERSITÁRIO DO
PARÁ - CESUPA



Continuação do Parecer: 4.418.779

diária.

Considerações sobre os Termos de apresentação obrigatória:

Documentos anexados e preenchidos adequadamente:

1. Projeto no formato Plataforma Brasil (PB).
2. Declaração de Aceite do orientador
3. TCUD
4. Declaração de Autorização da Instituição coparticipante
5. Projeto Original
6. Folha de Rosto

Recomendações:

Anexar na Plataforma Brasil o formulário Checklist.

Conclusões ou Pendências e Lista de Inadequações:

O Protocolo submetido em 10/11/2020 atende aos pressupostos das Resoluções CNS 466/2012 e 510/2016, e Norma Operacional 001/2013.No entanto, acatar as seguintes recomendações:

1. Anexar na Plataforma Brasil o formulário Checklist.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1603053.pdf	10/11/2020 17:32:53		Aceito
Outros	aceitorientador.pdf	10/11/2020 17:26:22	Claudio Eduardo Corrêa Teixeira	Aceito
Outros	tcudados.pdf	10/11/2020 17:25:40	Claudio Eduardo Corrêa Teixeira	Aceito
Declaração de Instituição e Infraestrutura	declaracaodecolaboracaoredcheck.pdf	07/10/2020 12:54:19	Claudio Eduardo Corrêa Teixeira	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO.pdf	14/09/2020 18:36:40	Claudio Eduardo Corrêa Teixeira	Aceito
Folha de Rosto	folhaderosto.pdf	14/09/2020 18:33:46	Claudio Eduardo Corrêa Teixeira	Aceito

Endereço: Av. Governador José Malcher, 1963

Bairro: São Brás

CEP: 66.060-232

UF: PA Município: BELEM

Telefone: (91)4009-9100

E-mail: cep@cesupa.br



Continuação do Parecer: 4.418.779

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

BELEM, 24 de Novembro de 2020

Assinado por:
PATRICK ABDALA FONSECA GOMES
(Coordenador(a))

Endereço: Av. Governador José Malcher, 1903
Bairro: São Brás **CEP:** 66.060-230
UF: PA **Município:** BELEM
Telefone: (91)4002-9100 **E-mail:** cep@cesupa.br

ANEXO B – TERMO DE COLABORAÇÃO



Rua Dr. Ovídio Pires de Campos, 75. 05401-000
Cerqueira César - São Paulo/SP
CNPJ: 27.961.641/0001-69

São Paulo, 14 de agosto de 2020

DECLARAÇÃO

Declaro para os devidos fins, que a RedCheck Intermediação e Tecnologia LTDA, aceita colaborar na execução do projeto intitulado "Utilização de Inteligência Artificial no diagnóstico neuro-oftalmológico: aprendizado a partir de dados reais e desenvolvimento de métodos para enfrentamento de desafios atuais e futuros", realizados sob a coordenação do Prof. Dr. Cláudio Eduardo Corrêa Teixeira do Centro Universitário do Estado do Pará - CESUPA.

Atenciosamente,

Rafael Scherer
Diretor Executivo - RedCheck

Tel.: (48) 98838-9256
redcheck@redcheck.com.br
www.redcheck.com.br