

Abordagem Computacional de Mapeamento Facial e Rastreamento do Sistema Visual em População de Recém-Nascidos e Crianças

Rodrigo Ferraz Souza¹
dev.rodrigofs@gmail.com

Aicha Khalid Hassan Al-Rob¹
aicha.k2002@gmail.com

Ricardo Pires Maciel¹
ricardopimaci@gmail.com

Bruna Samantha Marchi²
brunasmarchi2@gmail.com

Giovana Pascoali Rodovanski³
gio_pascoali@usp.br

Cristiane Aparecida Moran²
cristianemor@gmail.com

Antonio Carlos Sobieranski¹
a.sobieranski@ufsc.br

¹ Laboratório de Sinais, IoT e Imagens - LSIIM,

Centro de Ciências, Tecnologias e Saúde - CTS, UFSC Araranguá,

² Laboratório Prematuridade - LAPREM,

Departamento de Ciências da Saúde, Programa de pós-graduação em Ciências da Reabilitação - PPGCR,

³ Programa de Pós-graduação em Neurociências e Comportamento.

Departamento de Psicologia Experimental. Instituto de Psicologia - USP.

Abstract. *Sistemas computacionais de mapeamento facial e rastreamento visual vêm sendo usados em diversas aplicações de interação humano-computador, como educação, entretenimento, segurança e medicina. Na medicina, esses sistemas podem ser usados para acompanhar evolução de pessoas com limitações de locomoção, micro-expressões, atipicidades da visão e acompanhar recém-nascidos prematuros. Este artigo apresenta uma abordagem computacional para mapear a face de crianças e recém-nascidos, identificar a região dos olhos e rastrear o ponto de visão para auxílio em análise clínica. A proposta se divide em: (i) detecção facial por meio de modelos convolutivos para a obtenção da malha de posicionamento facial e espacial do indivíduo; (ii) identificação da região do sistema visual do indivíduo; e (iii) rastreamento temporal do ponto de visão com o objetivo de obtenção das trajetórias visuais em relação a um protocolo clínico aplicado. Os algoritmos foram desenvolvidos sobre um dataset especificamente elaborado para o projeto em questão, sendo os vídeos coletado em ambiente laboratorial. Os resultados experimentais preliminares foram validados em função de localização do sistema visual considerando o dataset secundário de Cambridge, obtendo uma taxa média de acurácia ponderada de 74,21% (acurácia de 98,67%), e assertividade de 88,78% para uma métrica especificamente designada para computar a correspondência binária em relação ao padrão-ouro com grau de liberdade de raio $r = 9$. A abordagem proposta demonstrou preliminarmente a viabilidade de aplicação clínica como ferramenta de suporte e apoio à tomada de decisão da rotina clínica.*

1 Introdução

O sistema visual é um dos mais importantes dos cinco sentidos humanos. Possibilitando a percepção física do mundo a que o indivíduo está inserido, proporcionando a interação imersiva no ambiente. O sistema visual humano possui o seu início junto à formação do sistema nervoso central nos estágios iniciais gestacionais, e tem prosseguimento evolutivo com o nascimento até os primeiros meses de vida da criança [1]. No entanto, nascimentos que ocorrem de forma

prematura podem prejudicar a formação do sistema visual humano, resultando em má formação de funções visuais [2]. Desta forma, a intervenção humana e precoce é de fundamental importância para a correção de eventuais limitações que o indivíduo possa apresentar, assim como mitigar problemas que possam ser irreversíveis.

A análise clínica é realizada com profissionais da área de saúde, tais como fisioterapia, oftalmologia e pediatrias. No caso de crianças e prematuros, protocolos estabelecidos que objetivam a resposta à um estímulo por parte do indivíduo são empregados, tais como placas impressas de padrões de contraste em diferentes bandas de frequência e cores. Em uma das funções visuais, o campo visual é estimado a partir do movimento dos olhos para a periferia utilizando de um estímulo, com a medição da amplitude, em graus, do deslocamento da íris, assim verifica-se o estímulo do olhar para as áreas periféricas. Ainda tem-se a fixação visual, que se caracteriza por fixar o olhar a um determinado estímulo. Este é classificado estável quando dura mais de três segundos ou instável quando menos. Há também a avaliação do rastreamento visual, que é feita com um estímulo em movimento horizontal ou vertical em relação à face da criança. Com essa avaliação, classifica-se o rastreamento da criança como completo, incompleto ou breve, assim como a identificação de atipicidades como nistagmo optocinético e estrabismo[3].

A análise clínica também pode ser conduzida por meio de equipamentos e recursos computacionais, que objetivam analisar amplitude do movimento e rastreamento da função do campo visual, mensurando graus de deslocamento da íris em relação à linha média de visão.

Com a evolução computacional obtida nos últimos anos, tanto em termos de hardware com elevada capacidade de processamento, como novos modelos computacionais de inteligência artificial, abordagens computacionais para rastreamento ocular tiveram a sua concepção. Os modelos em geral realizam a localização da região dos olhos para a reali-

zação do rastreamento da região de íris ou pupila a posteriori. Estes sistemas são denominados de *EyeTrackers*, e são empregados para uma vasta gama de aplicações, não necessariamente somente no contexto clínico, mas em áreas de segurança, interação humano-computador, análise mercadológica e pontos de atenção em *business*, robótica, dentre outras.

No entanto, várias limitações podem ser verificadas nas soluções em hardware e software acima mencionadas. Soluções de hardware consistem em um aparato que não é prático para o operador humano e para o indivíduo a ser analisado, além de inferir em elevados custos de aquisição. Soluções computacionais, por outro lado, são designadas para *EyeTrackers* de propósito específico, e por tal razão podem apresentar a performance reduzida em termos de identificação do sistema visual, como a pupila [4]. Modelos clássicos de visão computacional são desenvolvidos levando em consideração contextos bastante específicos, e são de complexa adaptabilidade ao problema de análise das atipicidades da visão em crianças e recém nascidos devido a tamanho de face. Desta forma, uma nova solução computacional levando em consideração o uso de protocolo clínico estabelecido e as recentes abordagens computacionais é requerida.

Este artigo apresenta uma abordagem computacional de *EyeTracking* em população de recém nascidos e crianças. A abordagem computacional considera a utilização de um *dataset* especificamente designado a partir de protocolo clínico, com a aquisição de vídeos em infra-vermelho curto, e fluxo computacional para o mapeamento facial, localização da região dos olhos, e rastreamento ocular para o fornecimento das trajetórias em domínio bi-dimensional. Os resultados experimentais preliminares foram validados em função de mapeamento facial, localização do sistema visual e rastreamento para obtenção das trajetórias para apoio à decisão clínica a posteriori. A abordagem computacional obteve uma taxa média de acurácia ponderada de 74.21% (acurácia de 98,67%), e assertividade de 88,78% para uma métrica especificamente designada para computar a correspondência binária em relação ao padrão-ouro com grau de liberdade de raio $r = 9$.

O restante deste artigo está organizado da seguinte maneira: Na Seção 2 os trabalhos correlatos à abordagem proposta são apresentados, dividindo-se em equipamentos de análise clínica, e ferramentas de software. Na Seção 3, a abordagem proposta em termos de hardware e software desenvolvidos são detalhados. Resultados preliminares experimentais são descritos na Seção 4, assim como aspectos do dataset produzido, marcação de *ground-truths* e métricas de validação. Por fim, conclusões, discussão e trabalhos futuros, são apresentados e discutidos na Seção 5.

2 Trabalhos Correlatos

Uma revisão da literatura revelou que duas classes de trabalhos correlatos puderam ser verifica-

dos: (i) equipamentos clínicos para o rastreamento e análise da visão humana, e (ii) abordagens computacionais para *EyeTracking* baseados em visão computacional ou técnicas de inteligência artificial.

Existem diversos equipamentos comuns de exame da visão. Alguns deles são:

- O oftalmoscópio direto, e o binocular indireto, por exemplo, tem a função de emitir uma luz que passa por lentes convergentes, e assim permitir o estudo e análise da retina. Além desse, o retinógrafo captura imagens que permitem a observação direta da retina, também auxiliam em um exame ocular [5];
- Existe também a lâmpada de fenda, um tipo de microscópio que emite uma luz de alta densidade para foca-la no globo ocular e assim examina-lo com mais precisão [5].
- Ferramenta de van Hof-van Duin para avaliação do campo visual da população infantil, que utiliza do perímetro de dois arcos cinéticos posicionados de forma perpendicular e de duas esferas iguais, uma na linha média e outra que movimenta-se da extremidade para o centro, até o olhar encontrá-la. [6]

Trabalhos com enfoque em abordagens computacionais para *EyeTracking*:

- Em ExCuSe[7] é apresentado um algoritmo para detecção de pupilas em cenários reais, baseado na filtragem de borda e análise de histogramas;
- Em Starburst[8] é apresentado um algoritmo para detecção de pupilas em vídeos em infra-vermelho utilizando um sistema de rastreamento ocular montado na cabeça. Ele utiliza um método iterativo para estimar o contorno da pupila, enviando raios em todas as direções e marcando pontos conforme suas diretrizes. Então o Starburst utiliza o RANSAC¹ para o ajuste da elipse;
- Em Swirski[9] é apresentado um método para identificação da pupila em 3 estágios. Primeiro, utiliza o modelo HAAR para obter a região da pupila na imagem, a fim de reduzir o espaço de busca. Em seguida, executa uma segmentação da região da pupila utilizando clusters de K-Means do histograma e faz uma detecção de borda com o Detector de Bordas de Canny. Por último, aplica o RANSAC para estimar o contorno da pupila;
- Em George and Routray[10] é apresentado um modelo de detecção de pupilas que inicia fazendo uma estimativa aproximada com base no filtro de convolução anelar[11], então utilizam o Kernel Schaar para calcular uma imagem gradiente. Em seguida, fazem uma análise de desvio padrão para, por fim, aplicarem o RANSAC.

¹RANdom SAMple Consensus

3 Materiais e Métodos

3.1 Protocolo de Captura para Geração do Dataset

O protocolo de captura dos vídeos para a geração do *dataset* consiste em utilizar protótipo de equipamento específico de operação na faixa de infravermelho curto e câmera auxiliar, descritos na próxima subseção. A captura consiste em posicionar a criança ou recém-nascido em frente ao equipamento de captura (25 centímetros de distância), e um conjunto de testes de funções visuais para diagnóstico cinético funcional são realizados com o objetivo de averiguar a fixação² por parte do indivíduo. A criança deve virar os olhos ou a cabeça e os olhos, em direção ao estímulo. São feitos 4 movimentos de teste:

- **1: Deslocamento horizontal lento do estímulo - rastreamento horizontal:** utiliza-se um alvo em movimento horizontal lento, com fixação nas periferias, intercalando os lados para a captura em movimentos radiais em relação à face do indivíduo, que deve estar fixada (com auxílio da mãe ou equipe de trabalho);
- **2: Deslocamento abrupto ao estímulo - visão periférica:** intercala-se o alvo na periferia da visão do indivíduo, com o objetivo de mensurar amplitude em graus, com a cabeça fixada;
- **3: Deslocamento vertical lento ao estímulo - rastreamento vertical:** utiliza-se um alvo em movimento vertical lento, com fixação em cima e em baixo, intercalando no eixo da ordenada para a captura em movimentos radiais em relação à face do indivíduo, que deve estar fixada;
- **4: Deslocamento lento da cabeça - reflexo vestibulo-ocular:** Consiste na estabilização do olhar de forma fixa em um alvo, posicionado central, enquanto a cabeça é movimentada de forma súbita e em pequena amplitude para o lado esquerdo e direito.

O objetivo dos movimentos supramencionados é averiguar os dados para possibilitar a classificação quanto à fixação visual, que é a capacidade de fixar em um determinado estímulo por mais de 3 segundos. O Rastreamento horizontal e vertical pode ser classificado em completo, quando realiza o acompanhamento com os olhos no estímulo até a posição extrema dos olhos, e incompleto quando não atinge a posição extrema dos olhos, ou ainda breve, quando o acompanhamento acontece por poucos graus. Contudo, a análise manual implica muitas limitações no diagnóstico, pois desta forma não é possível avaliar, em graus, quanto a iris deslocou pra encontrar o estímulo nas extremidades durante o rastreamento, tornando muito difícil a classificação da amplitude da visão periférica.

²É a capacidade de manter o olhar fixo em um estímulo.

3.2 Protótipo de Hardware para Aquisição dos Vídeos

O protótipo de hardware desenvolvido e utilizado para a aquisição possui as seguintes partes componentes: (a) um emissor de infra-vermelho é utilizado, posicionado para o indivíduo a ser amostrado. Em (b) é utilizada uma câmera USB do Modelo HardLine Cutie 6809, Marca HardLine, resolução espacial de 640 x 480 pixels, taxa de amostragem 30 *fps*, tendo seu filtro infra-vermelho removido fisicamente para possibilitar aquisição e melhor resposta na faixa infra-vermelho curto do espectro de onda. Esta câmera é posicionada frontalmente em relação ao indivíduo. Em (c), uma segunda câmera USB é utilizada, posicionada atrás do indivíduo a ser coletado, com o propósito de visualização dos movimentos do operador para eventual correlação dos movimentos da primeira câmera. Esta correlação está fora do escopo da presente abordagem descrita neste artigo. Em (d) é demonstrado um estímulo, que pode corresponder a uma placa com padrões de contraste definidos, **Figura 1**, que tem por objetivo chamar a atenção da criança, gerando o estímulo visual para os quatro movimentos anteriormente mencionados. Com as câmeras

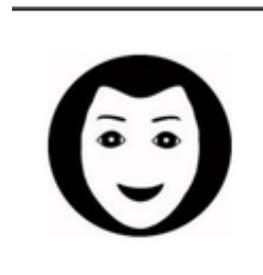


Figura 1: Placa para estímulo visual impressa fisicamente[12].

conectadas a um computador, um software especificamente desenvolvido para adquirir vídeos de ambas as câmeras de forma síncrona é utilizado. Este software em sua atual versão está desenvolvido em linguagem C++, com utilização da biblioteca OpenCV, gerando dois vídeos para cada coleta, e em cada tipo de movimento do protocolo estabelecido. Os vídeos são inspecionados e armazenados no *dataset* com identificação numérica, *timestamp*, dentre outras informações pertinentes à sua identificação.

3.3 Abordagem Computacional Proposta para Rastreamento do Sistema Visual

A abordagem computacional de *EyeTracking* para população de indivíduos em idade infantil e pré-maturos proposta pode ser definida a partir de: (a) modelo de reconhecimento facial para obtenção da malha facial dos pontos; (b) localização e reconhecimento do sistema visual do indivíduo; (c) rastreamento do movimento ocular para obtenção das coordenadas bidimensionais. Salienta-se que o escopo da abordagem proposta é fornecer as trajetórias correspondentemente aos

da malha 3D mesmo na presença de pontos de oclusão (mão da mãe fixando o rosto da criança, e estimativa dos pontos faltantes faciais).

3.3.2 Reconhecimento do Sistema Visual do Indivíduo

Uma vez aplicada a malha 3D facial e utilizado os 10 *landmarks* mencionados na subseção anterior, é possível obter com precisão a região dos olhos da criança. Do conjunto de pontos obtidos, 5 correspondem a cada um dos olhos, sendo que 4 realizam uma aproximação com as bordas da íris, e 1 ponto para a posição central. Com essas referências é possível realizar um recorte na imagem original e enviar para um modelo de reconhecimento secundário, responsável pela detecção do círculo que melhor descreve a aproximação com a pupila. Esta informação é relevante, uma vez que existe a possibilidade da mudança de tamanho da pupila ao longo da coleta com a aproximação da criança em relação à câmera e movimentação da cabeça resultando em rotações rígidas da malha, correlacionando com alguma eventual atipicidade da visão humana a ser verificada a posteriori.

O modelo de reconhecimento secundário é baseado em análise de histograma na cena, partindo-se do pressuposto que a região com menor nível de intensidade corresponde à pupila. Como a análise é realizada levando-se em consideração a informação da íris dentro de um círculo, a premissa é satisfeita. Exemplo de detecção pelo modelo de reconhecimento secundário para um histograma de pupila é apresentado na **Figura 4**, para linha e colunas (intensidade média da cor) da subimagem.

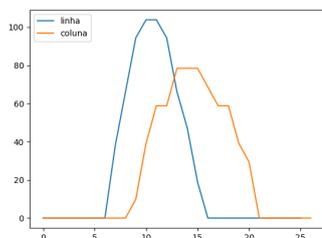


Figura 4: Exemplo de detecção por meio de Histograma da Pupila Direita (olho esquerdo da criança) pelo modelo de reconhecimento secundário.



Figura 5: Exemplo de detecção da Pupila

Através da análise dos padrões estatísticos como média e desvio padrão das curvas normais, o

centro e raio que melhor descreve a pupila são então obtidos. Exemplo da detecção da região da pupila é ilustrado através da Figura 5.

3.3.3 Rastreamento Temporal do Movimento do Sistema Visual

O rastreamento temporal das coordenadas bidimensionais na cena que correspondem ao sistema de *EyeTracking* é realizado de duas maneiras distintas: (a) rastreamento relativo ao frame original na cena, de forma que as coordenadas correspondem de maneira fidedigna ao seu posicionamento na imagem original, e (b) rastreamento normalizado com transformada de perspectiva, pertinente à malha facial independente do seu posicionamento na cena.

Ambas as abordagens utilizam inicialmente dos centros de íris fornecidos pelo método secundário anteriormente descrito, sendo os mesmos armazenados como padrão de referência do frame em arquivos do tipo *csv*. Para a primeira abordagem em (a), as coordenadas de cada olho são rastreadas temporalmente e armazenadas. Já a segunda abordagem, descrita em (b), o rosto é centralizado de forma a manter o seu posicionamento fixo, para que os dados salvos fossem apenas referentes à movimentação global, e eliminando o ruído da movimentação da cabeça do indivíduo. Para realizar esta operação, apenas foi necessário com que fosse definido um tamanho padrão para a distância entre um ponto na localizador testa e outro no queixo da criança, para que todos os rostos apresentem o mesmo tamanho. Após isto, é realizada a estabilização da imagem através do alinhamento dos dois olhos. Desta forma sempre se estará analisando o movimento em rostos do mesmo tamanho e com o olhar fixo e alinhado horizontalmente.



Figura 6: Rastreamento dos olhos nos frames 25, 48 e 72

4 Resultados Experimentais Preliminares

4.1 Critérios de Validação Aplicados

Para a etapa de validação, foram considerados selecionados alguns vídeos que tiveram sucesso na execução integral do protocolo. Por execução integral, subentende-se os probandos que conseguiram realizar as quatro categorias de movimentos, já que em alguns casos a reprodução do protocolo pode não ter uma resposta da criança, por distração ou por situações de desconforto.

O seguinte protocolo de validação preliminar foi aplicado:

- O conjunto de coordenadas Padrão-ouro serão utilizadas para confrontar a identificação automatizada do método proposto, considerando a simples correspondência das coordenadas sugeridas pelo método computacional X padrão-ouro;
- Os pontos de correspondência, por considerarem unidades métricas em pixel, devem possibilitar um valor de desvio padrão a ser considerado. Para tal, pontos de correspondência dentro de um raio de 9 pixels em relação ao padrão-ouro serão considerados verdadeiros positivos (true-positive - TP) no cálculo do centro, e 5 pixels de margem para o comprimento do raio do método proposto;

4.2 Validação do Método Computacional

4.2.1 Reconhecimento Facial e Mesh de Pontos

Com os testes executado foi verificado a robustez do modelo, como é possível de verificar na **Figura 7**, onde ele é submetido à uma situação de estresse e ainda consegue identificar satisfatoriamente os pontos de interesse.



Figura 7: Exemplo de robustez do FaceMesh

O modelo de Facemesh[13] utiliza o modelo de Rede Neural BlazeFace[14], o qual possui uma precisão de 98,61%.

4.2.2 Reconhecimento do Sistema Visual

Para avaliar a precisão do método proposto para a identificação da pupila foi utilizado o dataset da Universidade de Cambridge[15] para analisar 11382 imagens de olhos em RGB com resolução de 120x80px. Apesar da validação ser conduzida em função de um Dataset externo ao projeto, trata-se de um situação mais complexa, uma vez que imagens em padrão RGB apresentam maior complexidade em informação de cores em relação à uma Near Infrared. Desta forma o cálculo de erro foi realizado com as seguintes métricas:

- Erro Centro: Distância entre os centros, com margem de erro de 9 pixels;
- Erro do Raio: Diferença entre o comprimento calculado e o real, com margem de erro de 5 pixels;
- Áreas de Intersecção da pupila real com a calculada, vide **Figura 8**, onde a linha vermelha

representa o *Ground-Truth* e o azul a estimativa do modelo. As métricas calculadas neste método são:

- True Positive (TP): Cor rosa;
- True Negative (TN): Cor Branca;
- False Positive (FP): Cor Vermelha;
- False Negative (FN): Cor Azul;
- True Positive Rate (TPR);
- False Positive Rate (FPR);
- True Negative Rate (TNR);
- False Negative Rate (FNR);
- Accuracy (ACC);
- Sensitivity;
- Specificity;
- Acurácia Ponderada (weighted_acc);



Figura 8: Exemplo de análise de sobreposição da estimativa com o Ground Truth

Um vídeo com uma sequência de estimativas do modelo com a sobreposição com o Ground Truth pode ser encontrado neste Link

Utilizando as métricas acima citadas foram gerados os resultados dispostos na **Tabela 2** e representados graficamente na **Figura 9**. O dataset da Cambridge foi escolhido por conta das imagens terem sido geradas artificialmente, desta forma o padrão-ouro está livre de vieses de especialistas humanos.

Com a evolução computacional obtida nos últimos anos, tanto em termos de hardware com elevada capacidade de processamento, como novos modelos computacionais de inteligência artificial, abordagens computacionais para rastreamento ocular tiveram a sua concepção. Os modelos em geral realizam a localização da região dos olhos para a realização do rastreamento da região de íris ou pupila a posteriori.

5 Conclusão e Discussão

O presente artigo apresentou uma abordagem computacional para o mapeamento facial em crianças e recém-nascidos, assim como o rastreamento do sistema visual destes indivíduos em um sistema de *Eye-Tracking*. Para tal, foi elaborado um protótipo de hardware e software para aquisição de vídeos em ambiente clínico levando-se em consideração *frames* capturados em espectro de infra-vermelho curto, com aparato modificado. Vem sendo elaborado *dataset* próprio em ambiente clínico e acadêmico de modo a coletar vídeos de crianças e recém-nascidos, como o objetivo de aprimorar o processo de tomada de decisão clínica e o desenvolvimento de ferramentas de pré-triagem e apoio à decisão na busca por atipicidades da visão humana.

O presente artigo considerou para o momento a descrição do ambiente experimental e protótipo de

	TP	FP	TN	FN	TPR	FPR	TNR	FNR	ACC	sensitivity	specificity	weighted_acc
mean	58.2531	65.4732	9416.7099	59.5636	0.4912	0.0069	0.9930	0.5087	0.9869	0.4912	0.9930	0.7421
std	35.5701	71.9689	76.9639	35.1838	0.2708	0.0075	0.0075	0.2708	0.0077	0.2708	0.0075	0.1348
min	0.0	0.0	8886.0	0.0	0.0	0.0	0.9366	0.0	0.936875	0.0	0.9366	0.4867
25%	33.0	14.0	9384.0	35.0	0.2962	0.0014	0.9900	0.3185	0.9827	0.2962	0.9900	0.6451
50%	56.0	40.0	9438.0	59.0	0.4955	0.0042	0.9957	0.5044	0.9889	0.4955	0.9957	0.7447
75%	81.0	94.0	9470.0	81.0	0.6814	0.0099	0.9985	0.7037	0.9929	0.6814	0.9985	0.8404
max	217.0	601.0	9551.0	226.0	1.0	0.0633	1.0	1.0	1.0	1.0	1.0	1.0
total de imagens	11382											

	sensitivity	Specificity	weighted_acc
mean	0.4912	0.9930	0.7421
std	0.2708	0.0075	0.1348
min	0.0	0.9366	0.4867
25%	0.2962	0.9900	0.6451
50%	0.4955	0.9957	0.7447
75%	0.6814	0.9985	0.8404
max	1.0	1.0	1.0
total de imagens	11382		

Tabela 1: Resultados do Teste do modelo da Pupila com o *Cambridge SynthesEyes Dataset* - Método de sobreposição geométrica

	center_error	radius_error
mean	4.78817688	1.49947129
std	3.05080547	1.18803249
min	0.07159801	0.00058018
25%	2.42593429	0.63205647
50%	4.12061737	1.26092259
75%	6.64960240	2.03862444
max	17.16407710	9.37054717
% dentro da margem de erro	88.78%	98.08%
Total de imagens	11382	

Tabela 2: Resultados do Teste do modelo da Pupila com o *Cambridge SynthesEyes Dataset* - Método de Margem de erro

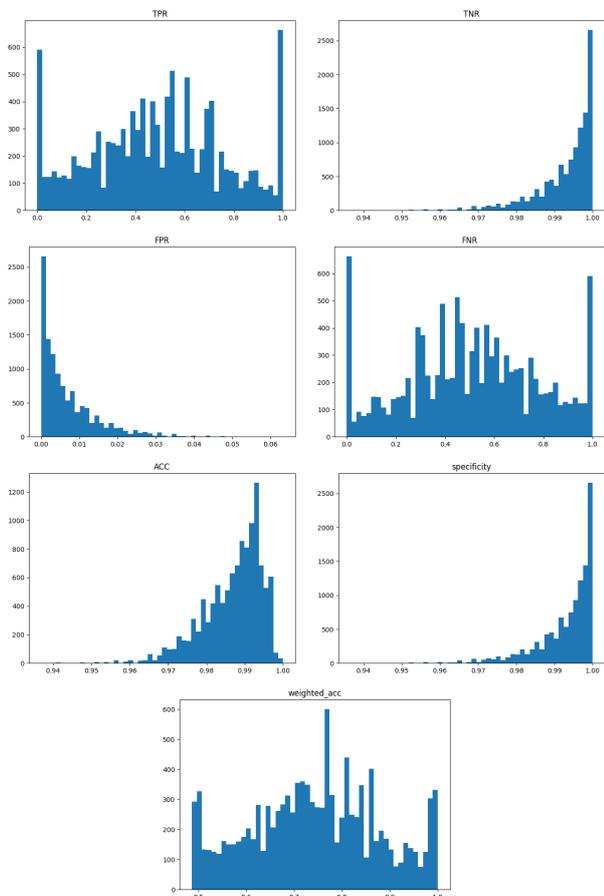


Figura 9: Histogramas de quantidade de resultados por métrica

captura, assim como o desenvolvimento de métodos para o mapeamento facial, identificação da região dos olhos e rastreamento ocular, em vídeo. Os resultados experimentais preliminares utilizaram as trajetórias adquiridas pela abordagem computacional proposta, e validadas em relação ao *dataset* de Cambridge³, que já possui coordenadas exatas da posição dos olhos dos indivíduos. O processo de validação para o protocolo acima definido considerou a simples correlação das coordenadas providas pelo método computacional proposto, em razão das coordenadas padrão-ouro. A taxa de precisão para a correta localização das pupilas foi de 88.78%, indicando que a abordagem proposta é promissora em termos de mapeamento facial 3D a partir de imagens e vídeos, assim como a detecção dos olhos para estimativa de coordenadas de rastreamento de forma temporal.

Trabalhos futuros pertinentes à abordagem proposta neste trabalho direcionarão a apresentar um modelo computacional que seja capaz de estimar as etapas decisórias clínicas dos especialistas médicos em relação às atipicidades encontradas durante as coletas. Através da análise de determinados padrões de trajetórias, assim como a sua regularidade e tempo de permanência em determinada posição, se buscará por correlações e evidências que possibilitem classificar de forma automatizada a informação temporal e converter em eventuais atipicidades da visão humana para a população de crianças e recém nascidos.

6 Agradecimentos

FAPESC, Fundação de Amparo a Pesquisa de Santa Catarina, pelo auxílio financeiro de apoio ao projeto vinculado ao edital 26/2020.

Referências

- [1] Vann Chau, Margot J Taylor, and Steven P Miller. Visual function in preterm infants: visualizing the brain to improve prognosis. *Documenta Ophthalmologica*, 127(1):41–55, 2013.

³Cambridge SynthesEyes Dataset[15]

- [2] Nora Tusor, Tomoki Arichi, Serena J Counsell, and A David Edwards. Brain development in preterm infants assessed using advanced mri techniques. *Clinics in perinatology*, 41(1):25–45, 2014.
- [3] Daniela Ricci, Domenico M Romeo, Francesca Gallini, Michela Groppo, Laura Cesarini, Silvia Pisoni, Francesca Serrao, Patrizia Papacci, Ilaria Contaldo, Francesca Perrino, et al. Early visual assessment in preterm infants with and without brain lesions: correlation with visual and neurodevelopmental outcome at 12 months. *Early human development*, 87(3):177–182, 2011.
- [4] Wolfgang Fuhl, David Geisler, Thiago Santini, Wolfgang Rosenstiel, and Enkelejda Kasneci. Evaluation of state-of-the-art pupil detection algorithms on remote eye images. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, UbiComp '16, page 1716–1725, New York, NY, USA, 2016. Association for Computing Machinery.
- [5] Moises Arthur et al. Técnicas de análise da retina humana: Revisão breve sobre os recursos e as renovações propedêuticas disponíveis à prevenção e o diagnóstico precoce das complicações vasculares da retina associadas às doenças sistêmicas. *Ensaios e Ciência: Ciências Biológicas, Agrárias e da Saúde*, 2011.
- [6] J Van Hof-van Duin, DJ Heersema, F Groenendaal, W Baerts, and WPF Fetter. Visual field and grating acuity development in low-risk preterm infants during the first 2 1/2 years after term. *Behavioural brain research*, 49(1):115–122, 1992.
- [7] Wolfgang Fuhl, Thomas Kübler, Katrin Sippel, Wolfgang Rosenstiel, and Enkelejda Kasneci. Excuse: Robust pupil detection in real-world scenarios. In George Azzopardi and Nicolai Petkov, editors, *Computer Analysis of Images and Patterns*, pages 39–51, Cham, 2015. Springer International Publishing.
- [8] Dongheng Li, D. Winfield, and D.J. Parkhurst. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, pages 79–79, 2005.
- [9] Lech undefinedwirski, Andreas Bulling, and Neil Dodgson. Robust real-time pupil tracking in highly off-axis images. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, page 173–176, New York, NY, USA, 2012. Association for Computing Machinery.
- [10] Anjith George and Aurobinda Routray. Fast and accurate algorithm for eye localisation for gaze tracking in low-resolution images. *IET Computer Vision*, 10(7):660–669, 2016.
- [11] TJ Atherton and DJ Kerbyson. Size invariant circle detection image and vision computing. *I*, 999(1):7.
- [12] Daniela Ricci, Laura Cesarini, Domenico MM Romeo, Francesca Gallini, Francesca Serrao, Michela Groppo, Agnese De Carli, Francesco Cota, Domenico Lepore, Fernando Molle, et al. Visual function at 35 and 40 weeks' postmenstrual age in low-risk preterm infants. *Pediatrics*, 122(6):e1193–e1198, 2008.
- [13] Google. Mediapipe face mesh. https://google.github.io/mediapipe/solutions/face_mesh.html, 2019.
- [14] Valentin Bazarevsky, Yury Kartynnik, Andrey Vakunov, Karthik Raveendran, and Matthias Grundmann. Blazeface: Sub-millisecond neural face detection on mobile gpus. *CoRR*, abs/1907.05047, 2019.
- [15] Erroll Wood, Tadas Baltrusaitis, Zucong Zhang, Yusuke Sugano, Andreas Bulling, and Peter Robinson. Syntheseyes dataset. 2020.