

# Análise de engajamento de alunos com TDAH usando dicas visuais de rastreador ocular

Harshit Chauhan<sup>1</sup>

hchauhan012@gmail.com HMI  
Lab, IIIT-Delhi Nova Deli,  
Índia

Anmol Prasad<sup>2</sup>

anmol16012@iiitd.ac.in HMI  
Lab, IIIT-Delhi Nova Delhi,  
Índia

Jainendra Shukla

jainendra@iiitd.ac.in HMI  
Lab, IIIT-Delhi Nova  
Delhi, Índia

## RESUMO

Neste artigo, nos concentramos em encontrar a correlação entre a atenção visual e o envolvimento de alunos com TDAH em sessões individuais com educadores especializados usando dicas visuais e dados de rastreamento ocular. **Nosso objetivo é investigar até que ponto as observações do olhar, postura, emoção e outros sinais fisiológicos podem ser usados para modelar o estado cognitivo dos sujeitos e explorar a integração de múltiplas modalidades de sensores para melhorar a confiabilidade da detecção de exibições humanas de consciência e emoção no contexto de crianças afetadas pelo TDAH.** Este é um problema novo, uma vez que nenhum estudo anterior teve como objetivo identificar marcadores de atenção no contexto de alunos afetados com TDAH. **O experimento foi projetado para coletar dados em um ambiente controlado** e, posteriormente, pode ser usado para gerar modelos de **Machine Learning** para auxiliar educadores do mundo real. Além disso, **propomos uma nova abordagem para detecção de AOI (Área de Interesse) para análise de rastreamento ocular em cenários dinâmicos usando modelos existentes de previsão de saliência e fixação baseados em aprendizado profundo.** **Nosso objetivo é usar os dados processados para extrair os recursos dos padrões de movimento dos olhos de um sujeito e usar modelos de aprendizado de máquina para classificar os níveis de atenção.**

## CONCEITOS CCS

• **Computação centrada no ser humano** e **Interação humano-computador (HCI).**

## PALAVRAS-CHAVE

Rastreamento ocular, Previsão de saliência, Detecção de olhar, Aprendizado de máquina

### Formato de Referência ACM:

Harshit Chauhan, Anmol Prasad e Jainendra Shukla. 2020. Análise de engajamento de alunos com TDAH usando pistas visuais do Eye Tracker. Na publicação complementar da Conferência Internacional de 2020 sobre Interação Multimodal (ICMI '20 Companion), 25 a 29 de outubro de 2020, evento virtual, Holanda. ACM, Nova York, NY, EUA, 5 páginas. <https://doi.org/10.1145/3395035.3425256>

<sup>1</sup>Estes autores contribuíram igualmente para esta pesquisa.

A permissão para fazer cópias digitais ou impressas de todo ou parte deste trabalho para uso pessoal ou em sala de aula é concedida sem taxa, desde que as cópias não sejam feitas ou distribuídas com fins lucrativos ou vantagens comerciais e que as cópias contenham este aviso e a citação completa na primeira página. Os direitos autorais dos componentes deste trabalho pertencentes a terceiros que não a ACM devem ser respeitados. Abstraindo com crédito é permitido. Para copiar de outra forma, ou republicar, postar em servidores ou redistribuir para listas, requer permissão específica prévia e/ou uma taxa. Solicite permissões de [permissions@acm.org](mailto:permissions@acm.org). ICMI '20 Companion, 25 a 29 de outubro de 2020, evento virtual, Holanda ©

2020 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8002-7/20/10... \$ 15,00  
<https://doi.org/10.1145/3395035.3425256>

## 1. INTRODUÇÃO

O transtorno de déficit de atenção e hiperatividade (TDAH) é um dos transtornos do neurodesenvolvimento mais comuns, afetando cerca de 3 a 7% das crianças em idade escolar em todo o mundo [34]. Caracteriza-se por baixo grau de atenção, alto grau de hiperatividade e impulsividade e incapacidade de inibir ações inadequadas. Crianças com Síndrome de Deficiência de Atenção e Hiperatividade ou TDAH enfrentam desafios complexos em suas atividades diárias e resultados escolares ruins. 1 Estatísticas recentes mostram que o número de alunos com TDAH matriculados cai de 89% na escola primária para 8,5% na secundária escola e ainda para 2,3% para o ensino secundário superior. Existe uma carência de educadores especiais em proporção à população afetada e, portanto, há necessidade de desenvolver programas para diagnosticar e monitorar o progresso das intervenções nas crianças afetadas por meio do uso de IA. Um desses métodos é o rastreamento do movimento dos olhos/olhar, que permite a objetividade e fornece dados quantitativos sobre o processo de leitura e habilidades visuais e atencionais [28]. Pesquisas recentes usaram pistas visuais para diagnosticar e monitorar o efeito de medicamentos nesses pacientes [1, 9, 11, 16, 26]. características processadas que obtemos do dispositivo de rastreamento ocular.

Nosso progresso atual nessa área tem sido processar dados de olhar brutos do rastreador ocular usando técnicas de aprendizado profundo de última geração. Em seguida, geraremos as regiões AOI do melhor modelo para cada um dos quadros capturados durante as sessões de gravação. Os mapas de saliência gerados a partir do melhor modelo para cada quadro de vídeo, juntamente com os recursos de olhar obtidos a partir dos dados processados do rastreador ocular para medir a atenção de um sujeito durante uma tarefa de aprendizado. O objetivo final é identificar correlações entre diferentes sensores que nos permitam criar um sistema acessível e em tempo real para identificar os estados atentos do sujeito.

## 2. TRABALHO RELACIONADO

### 2.1 Atenção e olhar Os movimentos

oculares são uma fonte natural de informação para sistemas proativos que analisam o comportamento do usuário, onde o objetivo é inferir feedback implícito de relevância do olhar [21]. A observação dos movimentos oculares não é uma área nova de pesquisa dentro dos campos relacionados à psicologia, tendo sido estudada em profundidade nas últimas décadas [27-29]. A análise do olhar não se restringe apenas à atenção visual, mas também permite espiar as estratégias cognitivas de uma pessoa [13]. No entanto, não foi até recentemente que os pesquisadores começaram a analisar e introduzir sensores e técnicas de rastreamento ocular em cenários muito mais complexos. Cenários envolvendo a compreensão de padrões de jogo

<sup>1</sup><https://blog.firstcrayon.com/the-essential-guide-to-special-needs-education-india-47769fc4d234>

em jogos como xadrez, emoções induzidas por certos recursos visuais e até mesmo em setores como saúde e pesquisa de marketing [10, 14, 20, 23].

## 2.2 Acompanhamento da atenção na aprendizagem online

Com a fácil disponibilidade de recursos como a internet, o mundo está na via rápida, tudo agora está muito mais próximo e acessível do que antes. Um serviço tão grande que cresceu em uma indústria no campo do aprendizado online. Nos últimos anos, vários trabalhos investigaram a previsão de desatenção com base em vários sinais, incluindo dados de frequência cardíaca [39], dados de EEG [40], condutância da pele e temperatura [6], bem como dados de pressão de mouse de computador [38] para maximizar a saída. Mas a abordagem de destaque que pode ser aplicada em escala em um ambiente de aprendizado online em breve é o rastreamento ocular e, portanto, a maioria das pesquisas existentes sobre detecção de desatenção depende de dados de rastreamento ocular, incluindo [3-5, 8, 24, 32, 41]. E artigos recentes como [31, 33] mostram os resultados e o trabalho que tem sido feito neste campo. Nosso objetivo aqui é usar abordagens semelhantes para ajudar crianças com TDAH e educadores especiais a preencher essa lacuna com a tecnologia. Portanto, a partir de nosso estudo experimental, propomos abordar estes QRs: • QR

1: Uma modalidade não invasiva como rastreador ocular ou uma combinação de modalidades pode ser usada para prever os níveis de atenção em crianças com TDAH?

- RQ 2: Os modelos baseados em saliência de aprendizado profundo podem ser usados como proxies para marcação AOI manual? • RQ 3: Qual algoritmo de ML é o mais apropriado para o previsão de atenção?

## 3. METODOLOGIA

Esta seção descreve a configuração e o protocolo para a coleta de dados. Planejamos realizar os experimentos nas instalações educacionais na presença de educadores especializados. Tivemos várias execuções de teste do protocolo proposto em um assunto de teste pertencente ao grupo-alvo nas instalações da universidade com o consentimento das autoridades competentes.

### 3.1 Participantes

Colaboramos com uma instituição de necessidades especiais para conduzir nosso experimento. Os educadores realizarão sessões com **60 alunos com TDAH do ensino fundamental ao médio** grupo de idade.

### 3.2 Configuração e protocolo

Neste estudo, os parâmetros de bio-sinal de um sujeito são registrados em uma sessão de tutorial individual on-line com um educador. Dois HP Os laptops Pavilion i5 de 8ª geração (por exemplo, L1 e L2) com todo o software necessário instalado são usados para coleta de dados dos dispositivos. Como nosso projeto está focado em uma tarefa mais ampla de análise de engajamento, integramos 3 modalidades no protocolo de coleta de dados, rastreador ocular, GSR e EEG. O participante recebe um laptop L1 com uma conexão de internet em funcionamento em uma sala supervisionada fechada composta por uma mesa e uma cadeira fixa, enquanto o educador recebe um terceiro sistema com uma conexão de internet em funcionamento configurado em uma sala diferente com uma configuração semelhante. O fone de ouvido sem fio EEG EPOC+ de 14 canais e o Gazepoint GP3 são conectados ao Laptop L1, enquanto o sensor vestível Shimmer3 GSR+ e Logitech

webcam está conectada ao laptop L2. O participante está de frente para o laptop L1 e o laptop L2 é colocado lado a lado com o laptop L1. Isso é feito para que a webcam, apesar de estar conectada ao laptop L2, possa ser montada no laptop L1 voltado para o participante. O rastreador ocular é conectado e montado no laptop L1.

Após as conexões de configuração, o fone de ouvido EEG é ajustado na cabeça do participante com faixa de cabeça, já que a maioria dos participantes são adolescentes enquanto o dispositivo GSR é preso no pulso. Open Broadcast Software (OBS) é usado para executar a webcam e a gravação da tela. O script em L1 começa com calibração de 9 pontos e, em seguida, envia um acionador usando programação de soquete em uma conexão LAN estabelecida para L2 para inicialização sincronizada dos dispositivos em diferentes laptops.

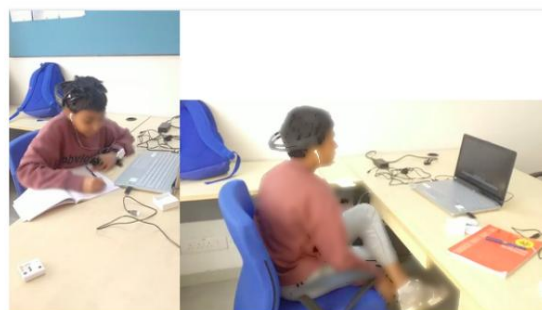


Figura 1: Protocolo experimental

O estudo é realizado nas salas de estudo de nossa instituição. O aluno tem uma videochamada individual com o educador durante a qual o educador pode compartilhar sua tela ou um vídeo para explicar conceitos. As sessões são compostas por 2 disciplinas, uma à escolha do aluno e outra à escolha do educador. Durante os intervalos de relaxamento, o aluno recebe pequenos exercícios ou quebra-cabeças ou conversa com o professor. As sessões foram divididas em partes da seguinte forma: Os intervalos das sessões foram decididos após consulta com

Tempo	Tipo de sessão
5 minutos	Sessão de relaxamento
20 minutos	Sessão educacional
5 minutos	Sessão de relaxamento
15 minutos	Sessão educacional
5 minutos	Sessão de relaxamento

Tabela 1: Detalhes da sessão

os educadores de necessidades especiais. Planejamos replicar o mesmo ambiente de aprendizado on-line que eles conduzem em suas sessões individuais normais. A razão para esta estrutura de protocolo é maximizar a coleta de dados, pois pode ser problemático manter o interesse dos alunos por períodos mais longos, uma vez que eles têm TDAH [2, 7].

Os dados coletados até agora são de 5 sessões com 1 participante, pois tivemos que interromper a coleta de dados devido ao COVID. Os resultados preliminares que mostramos são desta amostra. Nosso objetivo é coletar mais dados à medida que avançamos.

### 3.3 Anotação de Dados

A tarefa de anotação de dados é difícil no sentido de que são necessários profissionais treinados para anotar corretamente os dados. Portanto, temos cada sessão anotada manualmente por vários educadores profissionais que podem avaliar os níveis de atenção do sujeito. Para esta tarefa, usamos a ferramenta ELAN 5.9[37]<sup>2</sup>. Criamos um modelo a partir do qual o educador segmentaria e anotaria as sessões. Existem 3 grandes categorias nas quais os dados serão divididos de acordo com os níveis de atenção como Desatenção, Atenção e Moderadamente Atento. Os anotadores rotularam o vídeo de acordo com a escala de tempo variável de acordo com sua avaliação. Essas amostras são posteriormente divididas em intervalos fixos de tempo  $n$  (em segundos) e o voto da maioria é usado como o nível de atenção para um determinado intervalo de tempo.

## 4 ANÁLISE

### 4.1 Detecção de AOI

As AOIs são vitais na pesquisa de rastreamento ocular, pois o acoplamento dos dados de rastreamento ocular ao estímulo visual geralmente é feito usando as chamadas Áreas de Interesse (AOIs) [18]. Embora seja tecnicamente possível construir AOIs automaticamente, o padrão neste campo ainda é a construção manual de AOIs. Isso é provável, pois os métodos automáticos de construção de AOI são tecnicamente complexos ou não são eficazes o suficiente para pesquisas empíricas [12].

Para estímulos em movimento (telas com elementos em movimento), os AOIs devem ser construídos para cada quadro de vídeo. Houve tentativas feitas para tornar o processo de marcação de AOIs automático. [17] validaram um método de construção de AOI totalmente automático para vídeos de faces com base na detecção de marcos faciais OpenFace e Voronoi tessellation. Apesar da melhoria, essas abordagens são subjetivas e orientadas para tarefas específicas, portanto, não podem ser generalizadas para todos.

Fuhl et al [15] sugerem que as anotações manuais estão sujeitas ao julgamento e viés do analista de dados e o nível de detalhe não é claro (por exemplo, anotar um rosto pode ser feito como todo o rosto ou apenas sub-regiões de olhos e boca) e, portanto, esse método pode produzir resultados incorretos. Técnicas automatizadas comumente usadas usam os dados do olhar gravados dos participantes que receberam o estímulo para identificar regiões salientes criando mapas de saliência. Os autores aprimoram ainda mais essas informações usando as frequências de fixação das regiões, bem como gradientes nos mapas de saliência para identificar a AOI.

Nosso objetivo é usar métodos de aprendizado profundo pré-existentes para criar as regiões de interesse. As abordagens existentes mostraram que os mapas de saliência podem fornecer um proxy das informações sobre as possíveis regiões de interesse, uma vez que são construídos a partir dos dados do olhar do sujeito. Isso não é totalmente preciso, conforme mencionado por Fuhl et al [15], e requer algum processamento adicional para obter resultados muito precisos. Portanto, primeiro procuramos comparar os resultados de vários modelos SOTA em um pequeno conjunto de dados com mapas de verdade anotados manualmente para obter previsões de saliência, que comparamos com a verdade usando várias métricas estabelecidas. O conjunto de dados consiste em 250 imagens de 50 vídeos divididos em 2 categorias:

- Pessoa: Vídeos contendo regiões de texto, escrita ou imagem, bem como o educador. (Fig. 2.a)

<sup>2</sup><https://archive.mpi.nl/tla/elan>

- Não Presencial: Vídeos contendo apenas texto e imagens com sobreposição de voz do educador. (Fig. 2.b)

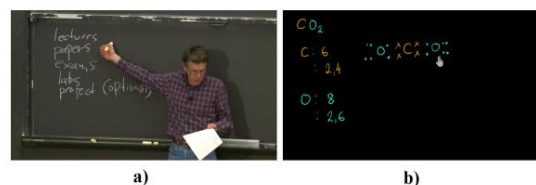


Figura 2: Exemplo de imagem da categoria a) Pessoa b) Não Pessoa

Os vídeos no conjunto de dados têm de 2 a 8 minutos de duração e a sobreposição entre as duas categorias é reduzida ao mínimo. Em seguida, 5 quadros aleatórios distintos são retirados de cada um dos vídeos para anotação. Cada um dos vídeos é anotado pelos 5 membros da equipe usando a FERRAMENTA DE ANOTAÇÃO VGG4 e a interseção da sobreposição é considerada como a máscara de verdade de campo final. A saliência da nossa tarefa de anotação é em termos do contexto semântico do vídeo. A Previsão de Fixação (FP) visa prever onde os humanos se fixam durante a visualização livre da cena. A detecção de objetos salientes (SOD) visa destacar regiões de objetos salientes em imagens. Diferente do FP que é originário de comunidades de pesquisa cognitiva e psicológica, o SOD é uma tarefa de visão computacional conduzida por aplicativos em nível de objeto [36]. Apenas alguns estudos [36] exploraram como os humanos escolhem explicitamente objetos salientes. Eles confirmaram quantitativamente que os julgamentos de saliência do objeto concordam com os movimentos do olho humano. De acordo com as análises em [36], existe uma forte correlação entre julgamentos de saliência explícita e fixações de visão livre, que podem ser vistos como dois proxies de atenção visual.

Comparamos os resultados dos seguintes modelos de previsão de fixação SOTA baseados em aprendizagem profunda do MIT Saliency Benchmark e modelos de previsão de saliência:

- (1) Deepgaze-II [22]: DeepGaze II é um modelo de saliência para prever fixações em imagens. Ele usa aprendizado de transferência da rede neural profunda VGG-19 para obter desempenho de ponta no benchmark MIT300.
- (2) ASNet [36]: Attentive Saliency Network (ASNet) aprende a detectar objetos salientes a partir de fixações. A ASNet é baseada em uma hierarquia de convLSTMs que oferece um mecanismo recorrente eficiente para refinar sequencialmente os recursos de saliência em várias etapas.
- (3) EML-Net [19]: O modelo proposto visa melhorar os resultados atuais do SOTA usando uma configuração de codificador-decodificador na qual o codificador usa vários modelos de aprendizado profundo existentes para extração de recursos para gerar os mapas de saliência.

Após inspeção visual (Fig. 3), descobrimos que esses resultados são satisfatórios para nossa tarefa. No entanto, planejamos realizar alguns testes estatísticos em um conjunto de dados maior para verificar nossa hipótese de que os mapas de saliência podem funcionar tão bem quanto a anotação manual. Como as anotações são divididas em intervalos conforme mencionado na seção 3.3, haverá um conjunto de quadros em vez de um único quadro para seleção AOI.

<sup>3</sup>Os quadros foram extraídos de © 2020 MIT 6.824: Distributed Systems, Lecture 1 e © 2020 Khan Academy India - inglês, Covalent bonds under CCPL. <sup>4</sup>[https://www.robots.ox.ac.uk/vgg/software/via/via\\_demo.html](https://www.robots.ox.ac.uk/vgg/software/via/via_demo.html)

Para resolver isso, planejamos estimar uma média móvel com um tamanho de janela fixo  $w$  sobre o conjunto de quadros.

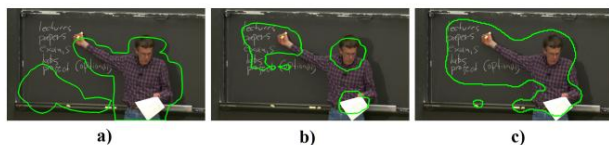


Figura 3: Previsões AOI a)ASNet b)EML-Net c)DeepGaze II

## 4.2 Métricas

É claro que uma única métrica não é suficiente para obter uma comparação justa de diferentes modelos[30]. Assim, usamos uma variedade de métricas diferentes, como Kullback-Leibler Divergence (KLD), que é mais sensível a FNs, Normalized Scanpath Saliency (NSS) e Linear Correlation Coeficiente (CC), que penalizam igualmente FPs e FNs, AUC- Judd (AUC-J), que coloca mais atenção em FPs com valores altos, ignorando FPs de baixo valor, Similarity Metric (SIM) e Information Gain (IG), que são comumente usados em tarefas de avaliação de saliência.

A curva PR e a medida F são referidas no artigo original da ASNet [36]. A curva PR é gerada da seguinte forma:

**Definição 4.1.** PR-Curve Dado um mapa de saliência com valores contínuos normalizados no intervalo de 0 e 255, primeiro calculamos os mapas binários correspondentes usando todos os limites de inteiros fixos possíveis. Deixe  $B$  denotar a máscara binária correspondente a um mapa de saliência contínuo  $S$  usando um limite e  $G$  indicar a máscara de verdade do solo. A precisão e o recall são calculados como precisão =  $|B \cap G|/|B|$ , e recall =  $|B \cap G|/|G|$ , respectivamente, onde  $| \cdot |$  acumula as entradas diferentes de zero em uma máscara. Em seguida, calculamos os pares de precisão/recall de todos os mapas binários para plotar a curva PR por um valor médio sobre todos os mapas de saliência em um determinado conjunto de dados.

## 4.3 Extração de recursos

A análise de fixações e movimentos sacádicos durante a execução de certas tarefas está relacionada à atenção de várias maneiras, concluindo que os mecanismos oculomotores dependem da atenção para alguns aspectos do controle do movimento ocular [14]. A AOI gerada acima é usada para mascarar as coordenadas do olhar atípico. Usamos o algoritmo [35] para gerar os rótulos de movimento ocular para o olhar bruto do rastreador ocular. Todos os limites usados para fixações de merge e drop ping são referidos no artigo Tobii IVT [25]. Segue a tabela de características:

Parâmetros de movimento dos olhos	Recursos extraídos
Duração da fixação (ms)	Média, desvio padrão
Sacada	Média, desvio padrão da duração da sacada (ms) e amplitude da sacada (°)
Estatísticas do evento	frequência de fixação, duração média da fixação, dispersão total da fixação, tempo até a primeira fixação, frequência sacada, duração sacada média, amplitude sacada média e latência sacada média

Tabela 2: Características extraídas do Gaze

## 4.4 Modelo de Classificação

Esta parte ainda está em fase experimental, pois iremos proceder a ela após a verificação de nossa hipótese. A dependência de parâmetros como o tamanho do intervalo ( $n$ ) e recursos adiciona mais dimensionalidade a esse problema de classificação supervisionada. Planejamos explorar e testar esses dados para diferentes técnicas de seleção de recursos e aprendizado de máquina, como SVM multiclasse e outros. Este trabalho pode fornecer uma linha de base para nossos empreendimentos futuros.

## 5 AVALIAÇÃO

Métrica/Modelo	SIM	AUC-J	IG	NSS	KL-D	CC	AUC (PR)	F-medida
DG-II	0,52	0,90			-1,10	0,85	0,76	0,67
EML-NET	0,52	0,86			-7,01	0,66	4,87	0,48
ASNet	0,55	0,84			-5,54	0,71	3,85	0,54

Tabela 3: Resultados de diferentes modelos no conjunto de dados

Os resultados de todas as 3 redes são mostrados na Tabela 3. Deve-se notar que exceto KL-Divergência, em todas as métricas um valor maior implica em um melhor resultado. Conforme explicado na seção 4.2, as diferentes métricas penalizam falsos positivos e falsos negativos igualmente (no caso de NSS e CC) ou penalizam falsos positivos (no caso de AUC) ou penalizam falsos negativos (no caso de KL, IG e SIM). Em cinco das seis métricas baseadas em fixação, o Deepgaze-II superou o ASNet e o EML Net. ASNet supera as outras redes na métrica de similaridade (SIM). O Deepgaze-II também supera as outras redes em ambas as métricas baseadas em saliência. Portanto, continuaremos com o Deepgaze-II para experimentos futuros assim que verificarmos estatisticamente os resultados em um conjunto de dados maior.

## 6 DISCUSSÃO E TRABALHOS FUTUROS

Nosso estudo descreve o protocolo para coleta de dados multimodais para alunos com TDAH em um ambiente de aprendizagem online. Também discutimos as técnicas de extração de recursos que estamos usando para processar os dados baseados no rastreador de olhar. Entre os 3 modelos SOTA avaliados para a tarefa de seleção automatizada de AOI, encontramos o DeepGaze II superior ao ASNet e EML-Net. As descobertas iniciais sugerem que o processo pode ser automatizado e pode fornecer bons resultados em nosso caso de uso. Nosso objetivo é usar diferentes técnicas de extração de recursos manuais e baseadas em aprendizado profundo nos dados coletados para finalmente criar um modelo capaz de prever o nível de atenção de um aluno em tempo real. Nossa abordagem inicial é usar todas as modalidades para esta tarefa. Para trabalhos futuros, pretendemos usar apenas a webcam juntamente com sensores vestíveis para aumentar a viabilidade e escalabilidade do projeto.

## 7 RECONHECIMENTO

Esta pesquisa é parcialmente apoiada pelo Infosys Center of Artificial Intelligence e TCS – Center for Design and New Media.

## REFERÊNCIAS

- [1] Jac N Airdrie, Kate Langley, Anita Thapar e Stephanie HM van Goozen. 2018. Reconhecimento de emoções faciais e olhar fixo no transtorno de déficit de atenção/hiperatividade com e sem transtorno de conduta comórbido. *Jornal da Academia Americana de Psiquiatria Infantil e Adolescente* 57, 8 (2018), 561–570.
- [2] Othman Asiry, Haifeng Shen e Paul Calder. 2015. Estendendo o período de atenção de crianças com TDAH por meio de uma interface de usuário adaptável dirigida por rastreador ocular. Nos *Anais da 24ª Conferência de Engenharia de Software da ASWEC 2015 (Adelaide, SA, Austrália) (ASWEC '15 Vol. II)*. Association for Computing Machinery, Nov 2015, 2824–2827.
- [3] Stylianos Asteriadis, Kostas Karpouzis e Stefanos Kollias. 2014. Foco visual de atenção em ambientes não calibrados usando estimativa do olhar. *Jornal Internacional de Visão Computacional* 107, 3 (2014), 293–316.
- [4] Robert Bixler e Sidney D'Mello. 2014. Rumo à detecção totalmente automatizada e independente de divagações mentais. Em *Conferência Internacional sobre Modelagem, Adaptação e Personalização do Usuário*. Springer, 37–48.
- [5] Robert Bixler e Sidney D'Mello. 2016. Detecção automática independente do usuário, baseada no olhar, de divagações mentais durante a leitura computadorizada. *Modelagem do usuário e interação adaptada ao usuário* 26, 1 (2016), 33–68.
- [6] Nathaniel Blanchard, Robert Bixler, Tera Joyce e Sidney D'Mello. 2014. Detecção automática baseada em fisiologia da divagação da mente durante o aprendizado. In *Conferência Internacional sobre Sistemas Tutores Inteligentes*. Springer, 55–60.
- [7] BH Cho, JM Lee, JH Ku, DP Jang, JS Kim, IY Kim, JH Lee e SI Kim. 2002. Sistema de Melhoria da Atenção usando realidade virtual e biofeedback EEG. Em *Proceedings IEEE Virtual Reality* 2002. 156–163.
- [8] Matthieu Pereira Da Silva, Vincent Courboulay, Armelle Prigent e Pascal Estrailier. 2008. Rastreamento facial em tempo real para jogos adaptativos com atenção. Em *Conferência Internacional sobre Sistemas de Visão Computacional*. Springer, 99–108.
- [9] Yarden Dankner, Lilach Shalev, Marisa Carrasco e Shlomit Yuval-Greenberg. 2017. Inibição pré-estímulo de sacadas em adultos com e sem transtorno de déficit de atenção/hiperatividade como um índice de expectativas temporais. *Psychological Science* 28, 7 (2017), 835–850.
- [10] Jakob De Lemos. 2007. Atenção visual e detecção de resposta emocional e sistema de exibição. Pedido de Patente dos EUA. 11/685.552.
- [11] Pamela Deans, Liz O'Laughlin, Brad Brubaker, Nathan Gay, Damon Krug, e outros. 2010. Uso do rastreamento do movimento ocular no diagnóstico diferencial de transtorno de déficit de atenção e hiperatividade (TDAH) e dificuldade de leitura. *Psicologia* 1, 04 (2010), 238.
- [12] Abhishek Dutta e Andrew Zisserman. 2019. O software de anotação VIA para imagens, áudio e vídeo. In *Proceedings of the 27th ACM International Conference on Multimedia*. 2276–2279.
- [13] Maria K Eckstein, Belén Guerra-Carrillo, Alison T Miller Singley e Silvia A Bunge. 2017. Além do olhar: o que mais o rastreamento ocular pode revelar sobre cognição e desenvolvimento cognitivo? *Neurociência cognitiva do desenvolvimento* 25 (2017), 69–91.
- [14] Maite Frutos-Pascual e Begonya Garcia-Zapirain. 2015. Avaliando a atenção visual usando sensores de rastreamento ocular em terapias cognitivas inteligentes baseadas em jogos sérios. *Sensores* 15, 5 (2015), 11092–11117.
- [15] Wolfgang Fuhl, Thomas C Kübler, Thiago Santini e Enkelejda Kasneci. 2018. Geração Automática de Áreas de Interesse Baseadas em Saliência para Visualização e Análise de Dados de Rastreamento Ocular. In *VMV*. 47–54.
- [16] Begonya Garcia-Zapirain, Isabel de la Torre Díez e Miguel López-Coronado. 2017. Sistema duplo para aprimorar as habilidades cognitivas de crianças com TDAH usando tecnologias de movimento de salto e rastreamento ocular. *Journal of medical systems* 41, 7 (2017), 1–8.
- [17] Roy S Hessels, Jeroen S Benjamins, Tim HW Cornelissen e Ignace TC Hooge. 2018. Uma validação de áreas de interesse geradas automaticamente em vídeos de um rosto para pesquisa de rastreamento ocular. *Fronteiras em psicologia* 9 (2018), 1367.
- [18] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka e Joost Van de Weijer. 2011. *Eye tracking: Um guia completo para métodos e medidas*. OUP Oxford.
- [19] Sen Jia e Neil DB Bruce. 2020. Eml-net: Uma rede multicamada expansível para previsão de saliência. *Computação de Imagem e Visão* (2020), 103887.
- [20] Jennifer L Kellough, Christopher G Beevers, Alissa J Ellis e Tony T Wells. 2008. Curso de tempo de atenção seletiva em adultos jovens clinicamente deprimidos: um estudo de rastreamento ocular. *Pesquisa comportamental e terapia* 46, 11 (2008), 1238–1243.
- [21] Arto Klami. 2010. Inferir regiões de imagem relevantes para a tarefa a partir de dados do olhar. Em 2010, IEEE International Workshop on Machine Learning for Signal Processing. IEEE, 101–106.
- [22] Matthias Kümmerer, Thomas SA Wallis e Matthias Bethge. 2016. DeepGaze II: Fixações de leitura de recursos profundos treinados em reconhecimento de objetos. pré-impressão arXiv arXiv:1610.01563 (2016).
- [23] Jian Li, Li Su, Bo Wu, Junbiao Pang, Chunfeng Wang, Zhe Wu e Qingming Huang. 2016. Previsão de saliência de página da Web com fusão de vários recursos. Em 2016, IEEE International Conference on Image Processing (ICIP). IEEE, 674–678.
- [24] Caitlin Mills, Robert Bixler, Xinyi Wang e Sidney K D'Mello. 2016. Detecção automática de divagação baseada no olhar durante a compreensão de filme narrativo. Sociedade Internacional de Mineração de Dados Educacionais (2016).
- [25] Anneli Olsen. 2012. O filtro de fixação Tobii I-VT. *Tecnologia Tobii* (2012), 1–21.
- [26] Ebrahim Pishyareh, Mehdi Tehrani-Doost, Javad Mahmoodi-Gharaie, Anahita Khorrami e Saeid Reza Rahmdar. 2015. Um estudo comparativo do viés atencional sustentado no processamento emocional em crianças com TDAH com imagens com rastreamento ocular. *Jornal iraniano de neurologia infantil* 9, 1 (2015), 64.
- [27] Keith Rayner. 1978. Movimentos oculares na leitura e processamento de informações. *Boletim psicológico* 85, 3 (1978), 618.
- [28] Keith Rayner. 1998. Movimentos oculares na leitura e processamento de informações: 20 anos de pesquisa. *Boletim psicológico* 124, 3 (1998), 372.
- [29] Keith Rayner. 2009. A 35ª Palestra de Sir Frederick Bartlett: Movimentos oculares e atenção na leitura, percepção de cena e busca visual. *Jornal trimestral de psicologia experimental* 62, 8 (2009), 1457–1506.
- [30] Nicolas Riche, Matthieu Duvinage, Matei Mancas, Bernard Gosselin e Thierry Dutoit. 2013. Saliência e fixações humanas: estado da arte e estudo de métricas de comparação. In *Proceedings of the IEEE International Conference on Computer Vision*. 1153–1160.
- [31] Tarmo Robal, Yue Zhao, Christoph Lofi e Claudia Hauff. 2018. Rastreamento de atenção baseado em webcam no aprendizado online: um estudo de viabilidade. In 23ª Conferência Internacional sobre Interfaces de Usuário Inteligentes. 189–197.
- [32] Kshitij Sharma, Patrick Jermain e Pierre Dillenbourg. 2014. Como os alunos aprendem usando MOOCs: uma visão de rastreamento ocular. Relatório técnico.
- [33] Kshitij Sharma, Patrick Jermain e Pierre Dillenbourg. 2014. "With-me-ness": Uma medida de olhar para a atenção dos alunos em MOOCs. In *Proceedings of the international conference of the learning sciences 2014*. ISLS, 1017–1022.
- [34] María Solé Puig, Laura Pérez Zapata, Laura Puigcerver, Neus Esperalba Iglesias, Carmen Sanchez Garcia, August Romeo, Josep Cañete Crespillo e Hans Supér. 2015. Vergência ocular relacionada à atenção medida em crianças com transtorno de déficit de atenção e hiperatividade. *PLoS One* 10, 12 (2015), e0145281.
- [35] Mikhail Startsev, Ioannis Agtzidis e Michael Dorr. 2019. 1D CNN com BLSTM para classificação automatizada de fixações, sacadas e perseguições suaves. *Behavior Research Methods* 51, 2 (2019), 556–572.
- [36] Wenguan Wang, Jianbing Shen, Xingping Dong, Ali Borji e Ruigang Yang. 2019. Inferir objetos salientes a partir de fixações humanas. *Transações IEEE sobre análise de padrões e inteligência de máquina* (2019).
- [37] Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann e Han Sloetjes. 2006. ELAN: uma estrutura profissional para pesquisa multimodal. Na 5ª Conferência Internacional sobre Recursos Linguísticos e Avaliação (LREC 2006). 1556–1559.
- [38] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper e Rosalind Picard. 2009. Tutores conscientes do afeto: reconhecendo e respondendo ao afeto do aluno. *Jornal Internacional de Tecnologia de Aprendizagem* 4, 3-4 (2009), 129–164.
- [39] Xiang Xiao e Jingtao Wang. 2017. Compreender e detectar a atenção dividida na aprendizagem mooc móvel. Nos *Anais da Conferência CHI 2017 sobre Fatores Humanos em Sistemas de Computação*. 2411–2415.
- [40] Thorsten O Zander, Christian Kothe, Sabine Jatzhev e Matti Gaertner. 2010. Aprimorando a interação humano-computador com informações de interfaces cérebro-computador ativas e passivas. Em *Interfaces cérebro-computador*. Springer, 181–199.
- [41] Yue Zhao, Christoph Lofi e Claudia Hauff. 2017. Detecção de divagação escalável para MOOCs: uma abordagem baseada em webcam. Na Conferência Europeia sobre Aprendizagem Reforçada por Tecnologia. Springer, 330–344.