



**INSTITUTO FEDERAL
DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA**
São Paulo

BRENO FIGUEROA ARAKAKI

BRUNO SANCHES

JOÃO VICTOR DOS SANTOS PEREIRA

VICTOR ALEXANDRE SCURA

SANDROVITOR FERREIRA DA SILVA

CTII 448

**DASHBOARD PARA RECONHECIMENTO DO ESTADO AFETIVO DOS
ESTUDANTES**

CUBATÃO

2024

INTRODUÇÃO:

A investigação sobre o aprendizado tem destacado a importância de compreender as reações dos estudantes aos estímulos educacionais, incluindo aspectos emocionais. A emoção, uma característica humana essencial, é manifestada por expressões faciais e comportamentos que refletem a resposta aos conteúdos educativos.

Este trabalho propõe uma ferramenta para analisar as reações faciais e corporais dos estudantes durante o aprendizado, composta por: (1) um módulo que estima o engajamento, as emoções básicas, a posição da cabeça, o fluxo óptico e as Unidades de Ação (AUs) da face e (2) um painel de controle para visualização desses dados. Essa ferramenta permite o upload de vídeos gravados durante as interações dos estudantes com materiais pedagógicos, extrai as informações relevantes e exibe as estimativas no dashboard.

A estrutura deste artigo é a seguinte: a Seção 1 apresenta os conceitos e terminologias utilizados; a Seção 2 descreve as etapas e procedimentos da pesquisa; a Seção 3 detalha o desenvolvimento da ferramenta proposta; e, por fim, a Seção 4 traz as considerações finais.

O engajamento é amplamente entendido como a participação ativa dos indivíduos em atividades específicas. Segundo Fredricks (2014), o engajamento abrange três dimensões principais: comportamental, cognitiva e afetiva/emocional. O engajamento comportamental se refere às manifestações visíveis de interesse, ou seja, às respostas físicas e corporais a determinados estímulos.

Já os engajamentos cognitivo e emocional, também conhecidos como engajamentos psicológicos, têm uma abordagem mais interna. O engajamento cognitivo está relacionado ao investimento mental no processo de aprendizado, incluindo a adoção de estratégias eficazes e a definição de objetivos claros para o que se pretende alcançar. O engajamento emocional, por sua vez, reflete a conexão emocional com a atividade; um aluno engajado em uma aula, por exemplo, tende a demonstrar emoções como entusiasmo ou alegria (Toni, 2023).

Para estimar o engajamento, implementamos um modelo preditivo usando redes neurais convolucionais (CNN) com base no conjunto de dados Daisee. A fim de otimizar o modelo, foram realizadas várias tentativas de balanceamento, visto que o dataset naturalmente possui uma quantidade maior de vídeos rotulados como “Muito engajado” e “Engajado”, distribuídos nas pastas de treinamento, teste e validação. Para equilibrar as classes, adotamos a estratégia de remover os vídeos rotulados como

“Muito engajado” e “Engajado” apenas da pasta de treinamento, enquanto os vídeos rotulados como “Desengajado” e “Muito desengajado” foram removidos de todas as pastas.

Utilizamos uma CNN baseada na biblioteca Keras do Python, uma ferramenta de código aberto desenvolvida para fornecer uma interface de alto nível para a construção, treinamento e avaliação de modelos de aprendizado profundo.

A partir dos estudos das emoções de estudantes em ambiente escolar, observou-se que as emoções básicas não são as prin

1. Materiais e Métodos

Nesta seção, abordamos as etapas realizadas para desenvolver o dashboard. Primeiramente, utilizamos a biblioteca PyFeat, que oferece modelos pré-treinados para extração de emoções básicas (raiva, nojo, medo, alegria, tristeza, surpresa e neutro) e posições de cabeça (pitch, roll e yaw). O PyFeat, descrito por Cheong et al. (2023), é uma ferramenta amplamente empregada em estudos de comportamento humano e visão computacional. Em nossa análise, ele emprega Máquinas de Vetores de Suporte (SVM) para classificar a orientação da cabeça, a partir de uma linha de base de dados previamente rotulados.

1.1 Testes com o PyFeat

A Figura 1 mostra um exemplo da utilização do Pyfeat na estimativa da pose da cabeça. Podemos observar nas Tabelas 1, 2 e 3, respectivamente, os valores referentes a pose da cabeça, Aus e emoções básicas obtidos a partir da imagem exibida na Figura 1.

Figura 1. Imagem de teste submetida ao pyfeat



Fonte: Próprio Autor

Tabela 1. Dados retornados sobre a posição da cabeça.

Tipo de movimento	Resultado
Pitch	0.14709
Roll	-0.655515
Yaw	1.007429

Na Tabela 1 podemos observar os valores dos parâmetros Pitch, Roll e Yaw, respectivamente, relacionados a inclinação frontal, inclinação lateral e rotação da cabeça.

Os termos "pitch", "roll" e "yaw" são relativos à orientação espacial de objetos, descrevendo a movimentação dos mesmos em relação aos seus eixos (ver Figura 1). Como podemos observar na Figura 1, *pitch* refere-se ao movimento de inclinação da face para cima ou para baixo em relação a um eixo horizontal na altura das orelhas; *roll* descreve o movimento de rotação da face em relação a um eixo que passa verticalmente pelo nariz e *yaw* se refere ao movimento de rotação da face em relação a um eixo vertical que atravessa de cima para baixo pela cabeça. Com esses dados, é

possível estimar se uma pessoa está com o rosto em direção a tela ou não, isto é, se está engajada comportamentalmente (Gualberto, 2018).

Da mesma forma, podemos notar na Tabela 2 um exemplo de intensidades de ações faciais retornadas pelo Pyfeat.

Tabela 2. Níveis de intensidade das Au's, retornados pelo Pyfeat.

Action Unit	Intensidade
AU01	0.283622
AU02	0.443742
AU04	0.585005
AU05	0.327326
AU06	0.131561
AU07	1
AU09	0.182703
AU10	0.02788
AU11	0.464094
AU12	0.128843

Action Unit	Intensidade
AU14	0.290371
AU15	0.402487
AU17	0.600018
AU20	0
AU23	0.458389
AU24	0.467561
AU25	0.007398
AU26	0.102922
AU28	0.337822
AU43	0.032426

Fonte: Próprio Autor.

Na tabela 2 é possível observar o nível de intensidade relativo às Au's AU01, AU02, AU04, AU05, AU06, AU07, AU09, AU10, AU11, AU12, AU14, AU15, AU17, AU20, AU23, AU24, AU25, AU26, AU28 e AU43.

Outra função do Pyfeat é reconhecer e retornar a porcentagem de chance de determinado indivíduo estar representando alguma das emoções básicas, conforme demonstrado na Tabela 3.

Tabela 3. Porcentagem das emoções básicas estimadas pelo PyFeat

Emoção	Porcentagem
Raiva	0.000140
Nojo	0.000084
Medo	0.000105
Alegria	0.000040
Tristeza	0.006921
Surpresa	0.000102
Neutro	0.992606

A Tabela 3 demonstra os resultados de todas as emoções consideradas pelo Pyfeat, a saber, raiva, nojo, medo, alegria, tristeza, surpresa e neutro. Os números relativos a cada emoção remetem a uma porcentagem de chance, logo, nesse caso, a pessoa tem 99,2% de chance de estar neutra. Para o Pyfeat, a rotulagem de determinada emoção é baseada nos resultados das AUs.

Seguimos os conceitos apresentados por Paul Ekman (1992), na qual emoções básicas são emoções universais, expressas de maneira semelhante por pessoas de diferentes culturas. Para todo estímulo há uma resposta do corpo humano representando essas emoções. As emoções básicas podem tanto ser positivas quanto negativas, sendo elas representadas por raiva, nojo, medo, felicidade, tristeza e estado neutro.

1.2 Implementação das Funcionalidades do Dashboard em relação às métricas retornadas pelo PyFeat

Além dos parâmetros retornados pela biblioteca PyFeat, analisamos o fluxo óptico dos vídeos o que nos permitiu verificar mudanças nos pixels de um frame para outro indicando se houve ou não movimento nas posturas corporais.

Pose da cabeça:

Com os dados de pose da cabeça do Pyfeat, foi possível elaborar um modelo de aprendizado de máquina que possibilitou estimar se uma determinada pessoa está ou não com a cabeça direcionada à tela. O modelo foi treinado com o dataset 'ICCVW Frame Engagement Annotations' e baseou-se em um algoritmo de Support Vector Machine (SVM).

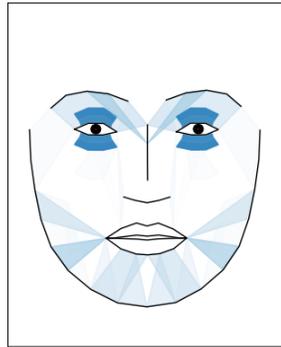
O 'ICCVW Frame Engagement Annotations' (Delgado *et al.*, 2021), uma base de dados criada com frames retirados de vídeos de alunos da University of Massachusetts Amherst, enquanto estavam resolvendo problemas de matemática. O dataset contém 18.721 frames de 19 alunos diferentes, sendo eles divididos em três classes: 'looking at screen', que seriam as pessoas olhando diretamente para a tela; 'looking at paper', essa classe é relativa às pessoas que estão lendo o papel que se encontra na mesa em que estão sentados, logo estão olhando para baixo; além de 'wandering', que seria quando uma determinada pessoa está olhando para os lados ou para cima.

A fim de facilitar o entendimento das pessoas que fossem utilizar futuramente o modelo e também para alcançar uma acurácia maior, o dataset 'ICCVW Frame Engagement Annotations', foi binarizado. Sendo assim, sua representação foi reduzida as seguintes classes: 'olhando para tela' (0) ('looking at screen') e 'olhando para fora da tela' (classes 'looking at paper' e 'wandering'). Essas classes foram balanceadas ficando cada uma com cerca de cinco mil frames.

Aus:

A biblioteca PyFeat permite que as “action units” extraídas sejam visualizadas através de imagens que representam a face de um indivíduo, facilitando a compreensão e a análise das micro expressões faciais de forma isolada (Ver Figura 4).

Figura 4. Representação dos dados da tabela 1 graficamente.



Fonte: Próprio Autor

Na imagem é possível estabelecer relações adicionais. Usando como exemplo a AU 07, relacionada ao levantamento dos músculos corrugadores superciliares (levantamento da sobrancelha), ao observarmos a Tabela 1, notamos a intensidade 1 (intensidade máxima) para esta AU. Ao analisarmos a Figura 4 percebemos uma maior intensidade de azul ao redor dos olhos.

Além de poderem ser visualizadas na tela, todas essas informações podem também serem impressas. Implementamos o relatório utilizando a biblioteca ReportLab. Nesse relatório, exibimos as informações de cada frame analisado. O tamanho do PDF depende de quantos segundos tem o vídeo e de quantos frames por segundo ele está sendo analisado.

Emoções básicas:

Com a utilização do PyFeat, ficou claro a importância da criação de novas ferramentas para trabalharem em conjunto ao modelo de predição do engajamento, logo, foi idealizado a construção de um gráfico (Ver Seção XX) para analisar as movimentações do estudante em meio ao ambiente de aprendizado. Para construção dos gráficos utilizamos a biblioteca matplotlib (Ver Seção XX).

Optical Flow

Implementamos o Fluxo Óptico a fim de identificar se o estudante está constantemente se movimentando ou não, e se sim, o quão bruscos são esses movimentos. Com os resultados do mesmo foi possível criar um gráfico a fim de mostrar o quanto os rastreadores do fluxo óptico se moveram a cada segundo, representando assim, o teor abrupto, ou não, do movimento.

O fluxo óptico, do termo em inglês *optical flow*, é um algoritmo que consegue rastrear a trajetória de objetos em movimento, além de detectar movimentos anormais ou mudanças repentinas. De acordo com a documentação da OpenCV, biblioteca utilizada para a implementação dessa métrica, o conceito por trás do *optical flow* se baseia em dois principais parâmetros. Primeiramente, a intensidade dos *pixels* de um objeto não muda entre *frames*. Segundo é que, ao analisarmos um determinado pixel, seus vizinhos terão um movimento similar. Seguindo esses dois conceitos e algumas fórmulas matemáticas, por exemplo, é possível estabelecer se a cabeça de um indivíduo se mexeu durante cinquenta frames, ou não.

Conclusões

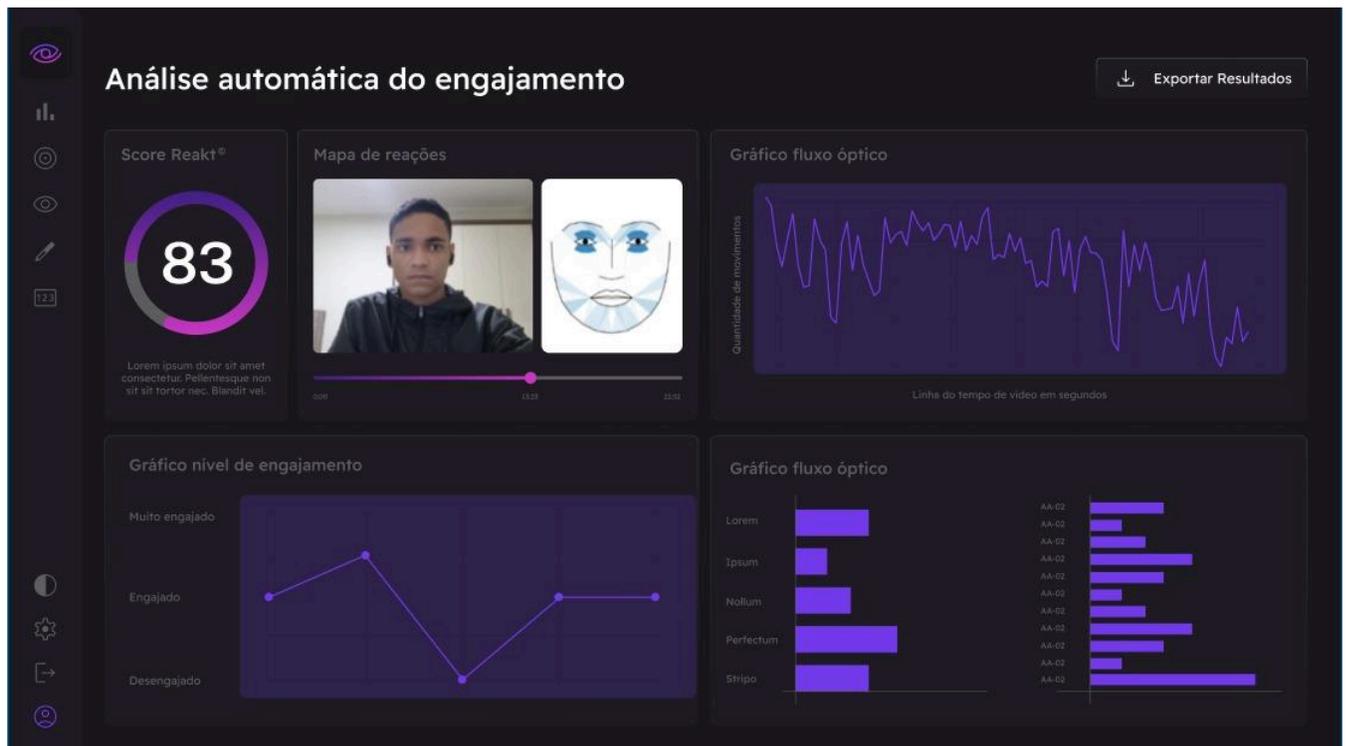
O dashboard desenvolvido contribui para a análise do comportamento e engajamento dos estudantes em ambientes de aprendizagem virtual. A capacidade de estimar o nível de engajamento e identificar as emoções básicas proporciona insights valiosos para educadores e pesquisadores, que podem adaptar os métodos de ensino conforme as necessidades emocionais e cognitivas dos alunos.

O modelo de engajamento permite aos professores observar como os estudantes interagem com o conteúdo e identificar momentos de maior ou menor interesse. A análise das emoções básicas revela nuances emocionais que, somadas às AUs, oferecem uma visão completa das respostas dos estudantes. Adicionalmente, a posição da cabeça e o fluxo óptico ajudam a identificar períodos de distração ou movimentos bruscos, que podem indicar reações de surpresa, confusão ou ansiedade.

Esses resultados fortalecem o entendimento das respostas emocionais dos estudantes e podem aprimorar o ambiente de aprendizado, tornando-o mais adaptável e eficaz para diferentes perfis de alunos.

Analisar engajamento online com IA

Enviar Vídeo



Fonte: Próprio Autor