

Trabalho 2 - Introdução à Visão Computacional

Leonardo Miranda e Guilherme Scherer

11 de Junho de 2025

Objetivo

O segundo trabalho busca integrar os conceitos aprendidos ao longo do semestre relacionados à visão computacional com redes neurais e machine learning, buscando um avanço do trabalho anterior, ou até mesmo, um tema diferente, porém de complexidade maior.

Enunciado

Cada grupo deverá realizar uma apresentação de 10 minutos para a turma sobre a continuidade do trabalho I (ou uma mudança de tema, desde que envolva visão computacional). O objetivo da apresentação é que o grupo ensine sobre o pipeline aos colegas, tanto de maneira teórica quanto prática. Para esse trabalho será possível incluir técnicas de redes neurais.

Introdução

Analisando nosso tema e trabalho anteriores, tentamos montar uma estratégia de melhoria para o desenvolvimento do T2. Ao final de contas, exploramos um novo contexto de aplicação para as técnicas de V.C, mas ainda pensando no contexto automobilístico.

Com isso em mente, elaboramos um novo tema, pensando em ajudar os motoristas que, por muitas vezes, acabam caindo no sono durante uma de suas viagens, colocando em risco sua própria vida e dos motoristas. Para isso, utilizamos as bibliotecas OpenCV e MediaPipe (além do NumPy para auxílio adicional na transformação de arrays).

Desenvolvimento

Neste projeto, utilizamos técnicas de visão computacional para detectar sinais de sonolência em motoristas. A aplicação foi construída com as bibliotecas OpenCV, MediaPipe e NumPy, e realiza o mapeamento facial em tempo real, a partir de imagens capturadas por webcam. A estratégia adotada baseia-se em métricas fisiológicas observáveis visualmente, como o EAR (Eye Aspect Ratio), o MAR (Mouth Aspect Ratio) e a frequência de piscadas. Essas medidas foram em literaturas científicas para indicar cansaço ou indicações de fadiga visual.

EAR (Eye Aspect Ratio)

O EAR é uma métrica geométrica que quantifica a abertura dos olhos a partir da razão entre distâncias euclidianas verticais e horizontais da pálpebra (SOUKUPOVA et al., 2016). O cálculo parte de seis pontos anotados pela biblioteca MediaPipe ao redor de cada olho. A fórmula usada é a seguinte:

$$EAR_{dir} = \frac{\|p_{160} - p_{144}\| + \|p_{158} - p_{153}\|}{2 \cdot \|p_{33} - p_{133}\|}$$
$$EAR_{esq} = \frac{\|p_{385} - p_{380}\| + \|p_{387} - p_{373}\|}{2 \cdot \|p_{362} - p_{263}\|}$$

Quando os olhos estão abertos, o valor de EAR aumenta, e quando estão fechados, o valor diminui. A literatura costuma sugerir um valor de limiar de $EAR = 0.3$ como indicativo de fechamento ocular. Em nosso sistema, utilizamos a média dos dois EARs, para termos apenas um

valor. Para a aplicação desenvolvida, quando o EAR fica abaixo deste valor por mais de 1.5 segundos, interpretamos como sinal de sonolência.

Esse tempo é controlado usando timestamps associados ao momento em que os olhos fecham. Se o tempo entre o fechamento e a reabertura ultrapassar o limite configurado, uma mensagem de alerta é exibida ao motorista.

MAR (Mouth Aspect Ratio)

O MAR mede a proporção de abertura da boca utilizando a razão entre distâncias euclidianas entre o lábio superior e inferior e a largura da boca. A fórmula usada foi:

$$MAR = \frac{\|p_{82} - p_{87}\| + \|p_{13} - p_{14}\| + \|p_{312} - p_{317}\|}{2 \cdot \|p_{78} - p_{308}\|}$$

Esse índice é útil para detectar se o motorista está rindo, e não considerar o fechamento dos olhos durante o riso como um sinal de sonolência. Quando o MAR ultrapassa o limiar definido (0.12) consideramos que a boca está aberta, assim, evitando emitir alerta de sonolência – pois este padrão é característico de riso e não de cansaço, onde os olhos se fecham levemente e os pontos da boca se afastam, gerando uma distância maior.

Essa verificação ajuda a tornar o sistema muito mais confiável e adequado ao uso real, evitando alarmes indevidos quando o motorista estiver apenas rindo, conversando ou sorrindo.

Contagem de Piscadas

Outro elemento fundamental no sistema é a contagem de piscadas. A piscada é identificada quando o EAR cai rapidamente abaixo do limiar de 0.3 e volta a subir logo em seguida. Cada piscada é registrada com um timestamp, e as últimas ocorrências são analisadas para calcular a frequência de piscadas por minuto.

Normalmente, uma pessoa pisca de 15 a 20 vezes por minuto. Uma diminuição na frequência de piscadas pode ser um indicativo de cansaço, de acordo com estudos (PASARIBU et al., 2018).

Utilizamos essa informação em conjunto com o EAR: se o número de piscadas por minuto é menor que 10, o sistema também alerta um sinal de detecção de sonolência.

Resultados e Conclusão

De acordo com nossa avaliação e testes intensos, chegamos à conclusão que o projeto foi satisfatório, e atingiu nosso objetivo, que era de avisar o motorista que ele estava sonolento.

O sistema desenvolvido mostra-se funcional e promissor para a detecção de sinais de sonolência a partir de expressões faciais. Ao utilizar métricas simples como EAR e MAR, conseguimos criar um modelo que, mesmo sem utilizar aprendizado de máquina, consegue interpretar sinais básicos de fadiga.

Além disso, o uso de bibliotecas como o MediaPipe, que fornece pontos de referência faciais de forma precisa e eficiente, torna a implementação prática e viável mesmo em sistemas com hardware modesto.

Como possíveis melhorias futuras, podemos citar a inclusão de um alarme sonoro para reforçar o alerta de sonolência, e a implementação de um histórico de comportamento do usuário ao longo do tempo.

Em suma, o projeto cumpre seu objetivo de explorar novas aplicações de visão computacional no contexto da segurança veicular e abre portas para soluções mais robustas e inteligentes nesse campo.

Referências

- [1] Link do GitHub do Grupo com os códigos utilizados:
<https://github.com/leomiranda16/T2-IVC>
- [2] SOUKUPOVA, Tereza et al. Real-Time Eye Blink Detection using Facial Landmarks: 21st Computer Vision Winter Workshop. Czech Technical University in Prague, 2016. Disponível em:
<https://vision.fe.uni-lj.si/cvww2016/proceedings/papers/05.pdf>. Acesso em: 09 jun. 2025.
- [3] PASARIBU, Novie Theresia Br. et al. Proceedings of the 1st International Conference on Life, Innovation, Change and Knowledge: Drowsiness Detection According to the Number of Blinking Eyes Specified From Eye Aspect Ratio Value Modification. Atlantic Press, 2018. Disponível em:
<https://www.atlantis-press.com/proceedings/icontrol-18/125913292>. Acesso em: 06 jun. 2025.