

Simpósio de Integração Acadêmica

“Ciências Básicas para o Desenvolvimento Sustentável”

SIA UFV 2023



IDENTIFICAÇÃO DE GESTOS PARA CONTROLE DE VEÍCULOS AUTÔNOMOS EM TEMPO REAL

Guilherme Serra Francisco Pinel - Departamento de Física - guilherme.pinel@gmail.com, Kétia Soares Moreira - Departamento de Engenharia Elétrica - ketia@ufv.br, Alexandre Santos Brandão - Departamento de Engenharia Elétrica - alexandre.brandao@ufv.br / Universidade Federal de Viçosa

Área de Tecnologia Prioritária: Habilitadoras
I - Inteligência Artificial
Palavras chave: K-NN; Classificação; AV; Robótica

Introdução

No século atual, os avanços tecnológicos têm se tornando cada vez mais presentes no cotidiano humano, abrangendo desde a medicina até a exploração espacial. Nesse cenário, destacam-se os veículos autônomos, capazes de operar sem intervenção humana direta, com aplicações que vão desde entregas até combate a incêndios. Para controlar esses veículos, a pesquisa propõe o uso de gestos humanos, capturados por sensores de movimento e classificados por uma rede neural treinada.

Objetivos

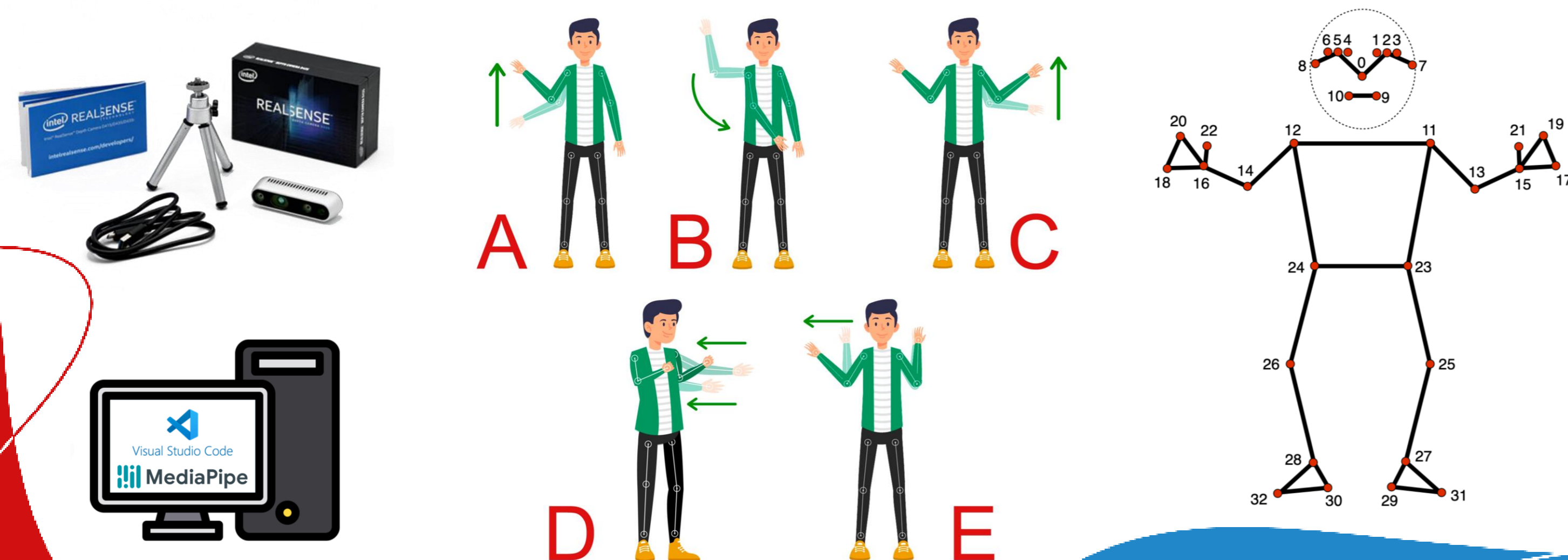
Este estudo visa criar uma abordagem prática para operar um veículo autônomo por meio de gestos, permitindo ao operador permanecer em um local remoto e controlar o veículo com gestos predefinidos. O principal objetivo é implementar um método de classificação de baixo custo e fácil utilização. A pesquisa busca desenvolver uma ferramenta que classifique ações naturais em cinco categorias usando poucos parâmetros de entrada e processamento simplificado. Para alcançar isso, os objetivos incluem a coleta de dados espaciais, a extração de atributos durante movimentos, a análise dos dados para maior robustez e a classificação das ações usando o algoritmo K-NN.

Material e Método

Este projeto visa o reconhecimento de gestos através de técnicas de aprendizado de máquina, enfocando a interação com um ambiente específico por meio de gestos simples realizados por vários usuários. Para isso, uma base de dados personalizada é criada, capturando informações usando a câmera RGB-D Intel RealSense D435i, que fornece dados de coloração (RGB) e profundidade (D). O algoritmo inicia capturando um fluxo de vídeo em tempo real da câmera e utiliza a biblioteca MediaPipe para detectar os marcos do esqueleto do usuário, como nariz, ombros, cotovelos e pulsos, fundamentais para a análise de gestos. Cinco gestos simples são pré-definidos e registrados em um documento .xlsx com coordenadas cartesianas. Dessa forma, a captura das amostras ocorre num período de 2 segundos quando o operador realiza um gesto específico. Após 3 segundos de espera, os dados são arquivados com um nome distinto, incluindo a classe do gesto (A, B, C, D ou E). Este processo é repetido para criar um banco de dados de amostras.

Além disso, foi construído um algoritmo utilizando-se do K-Nearest Neighbors (K-NN) para realizar a classificação dos gestos. Ele é um método não paramétrico de aprendizado de máquina supervisionado que se adapta continuamente a novos dados. O valor de k, o número de vizinhos a serem considerados, é definido empiricamente. O código importa bibliotecas relevantes e implementa funções para extrair características dos gestos, treinar o modelo K-NN e classificar novos gestos em formato .xlsx. A classificação é baseada em probabilidades calculadas, e gestos com probabilidades abaixo de um limiar são considerados inconclusivos.

Após classificar um conjunto de dados de validação, o código realiza uma análise dos resultados, calculando taxas de acerto, taxas de classificações inconclusivas e taxas de erro para avaliar o desempenho do modelo.



Resultados e Discussão

Um algoritmo foi desenvolvido para capturar gestos usando a biblioteca MediaPipe e a câmera Intel RealSense D435i. Duas abordagens de captura foram usadas: sequencial e aleatória, com 50 capturas por classe. As capturas foram salvas como arquivos .xlsx com coordenadas. O algoritmo permitiu a visualização em tempo real e correções durante a captura. A câmera estava a 2 metros do operador, com origem no nariz do esqueleto. Foram criados conjuntos de dados 35/15, 30/20, 25/25, 20/30 para teste/validação. Isso gerou 16 conjuntos de dados para análise. Os parâmetros de classificação usados foram $k = 5$ e $weights = \text{distância}$ (peso).

As amostras foram classificadas de diferentes maneiras para teste e validação. Matrizes de confusão foram construídas para avaliar o desempenho. O método de captura aleatória obteve uma taxa de acerto de 87,03%, um aumento de 2,22% em relação ao método sequencial. Erros e previsões inconclusivas foram observados, principalmente para gestos semelhantes. A classe D teve uma taxa de confusão de 2,9%. Dados inconclusivos foram atribuídos à escassez de amostras.



	35/15					30/20					25/25					20/30									
True Class	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
A	96,65	-	-	-	-	96,25	-	-	-	-	96,65	0,0	-	-	-	96,65	0,0	-	-	-	96,65	0,0	-	-	-
B	25,0	88,35	-	-	-	27,5	88,35	-	-	-	27,5	89,15	-	-	-	27,5	89,15	-	-	-	27,5	89,15	-	-	-
C	-	-	75,0	-	-	-	-	82,5	-	-	-	-	82,5	-	-	-	-	82,5	-	-	-	-	-	82,5	-
D	-	-	-	0,0	25,0	-	-	-	0,0	25,0	-	-	-	0,0	25,0	-	-	-	0,0	25,0	-	-	-	-	-
E	-	-	-	1,65	98,35	-	-	-	-	1,65	98,35	-	-	-	-	-	-	-	-	1,65	98,35	-	-	-	-
	-	-	-	8,35	72,3	-	-	-	-	8,35	72,3	-	-	-	-	-	-	-	-	8,35	72,3	-	-	-	-

Conclusões

Em conclusão, este estudo demonstra a eficácia do método de classificação de gestos com o algoritmo K-NN para controlar veículos autônomos em ambientes externos. O método de captura aleatória e a variedade de gestos dentro de uma classe mostraram-se vantajosos para o desempenho do classificador. Próximos passos incluem testar o método em situações de tempo real para sua implementação prática no controle de veículos autônomos.

Bibliografia

- [1]. Robotic thread is designed to slip through the brain's blood vessels. (2019, August 28). MIT News | Massachusetts Institute of Technology. <https://news.mit.edu/2019/robot-brain-blood-vessels-0828>;
- [2]. DE LIMA ANGELO, A. C. et al. Exploração Espacial com Robôs. REVISTA ACADÊMICA - ENSINO DE CIÊNCIAS E TECNOLOGIAS, v. 8, p. 127-142, [s.d.], 2021.
- [3]. DE MOURA, L. S. B. Drone de combate aos incêndios florestais. Tese de Mestrado em Engenharia Mecânica – Faculdade de engenharia da Universidade do Porto: [s.n.].
- [4]. K. B. de Carvalho, D. K. D. Villa, M. Sarcinelli-Filho, and A. S. Brandão, "Gestureteleoperation of a heterogeneous multi-robot system," The International Journal of Advanced Manufacturing Technology, vol. 118, no. 5, pp. 1999-2015, 2022.
- [5]. Pose landmarks detection task guide | MediaPipe. (n.d.). Google Developers. https://developers.google.com/mediapipe/solutions/vision/pose_landmarker/
- [6]. Shetty, R., Geetha, M., Acharya, D.U., Shyamala, G. (2022). Data Preprocessing and Finding Optimal Value of K for KNN Model. In: Reddy, V.S., Prasad, V.K., Wang, J., Reddy, K. (eds) Soft Computing and Signal Processing. ICSCSP 2021. Advances in Intelligent Systems and Computing, vol 1413. Springer, Singapore.

Agradecimentos

Agradeço a CNPq por fornecer os fundos que viabilizou a realização deste trabalho e aos meus orientadores que me instruíram e possibilitaram a capacitação para o mesmo, além do ambiente de trabalho no Núcleo de Especialização em Robótica da UFV.

Apoio financeiro

