AI-Powered Smart Glasses for Sensing and Recognition of Human-Robot Walking Environments

Daniel Rossos[1,4], Alex Mihailidis[2,4], and Brokoslaw Laschowski[3,4]

[1]Engineering Science, University of Toronto; [2]Biomedical Engineering, University of Toronto; [3]Mechanical Engineering, University of Toronto; [4]KITE, Toronto Rehabilitation Institute

Introduction

Environment sensing and recognition can allow robotic prosthetic legs and exoskeletons to dynamically adapt to different walking terrains. However, the development of accurate yet fast perception systems is challenging, especially using mobile and embedded devices with limited computational resources.

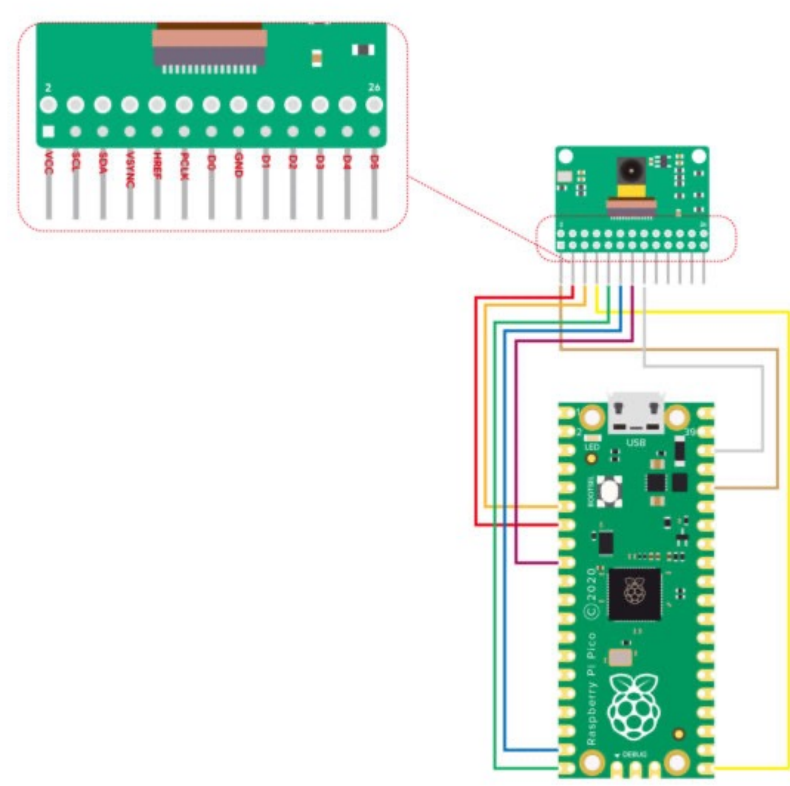
Figures 1. Examples of environment sensing and classification using our AI-powered smart glasses. Real-time predictions are shown in top left.

Objectives

Here we developed a pair of AI-powered smart glasses for onboard sensing and image classification of walking environments, including an integrated Raspberry Pi Pico microcontroller and an efficient deep convolutional neural network trained on a new large-scale image dataset. The smart glasses achieve high accuracy and low latency in the prediction of different environment states during real-world walking.

Methods

We developed our smart glasses using a Raspberry Pi Pico W microcontroller integrated with a ArduCam HM0360 VGA SPI camera sensor.



We then trained a lightweight and efficient convolutional neural network [1] for image classification using *Ego4D* developed by Meta AI - a new large-scale dataset of first-person images of real-world walking environments, which includes (among other classes) 24,189 images of indoor surfaces; 21,362 images of outdoor paved surfaces; and 16,958 images of outdoor non-paved surfaces.

Layer	Output shape	Parameters
Convolutional 2D	1x96x96x3	30
Mobilenet_V1_0.25	1x3x3x256	218,544
Global averaging	1x256	0
Full connected	1x3	771

Table 1. Lightweight and efficient convolutional neural network [1] designed for mobile and embedding computing devices.

Results

Our smart glasses were able to accurately predict real-world walking surfaces with **93.6%** image **classification accuracy** (93.7% precision) with an **inference speed** of **1.5 seconds** on the **embedded device**.

	Indoor	Outdoor non-paved	Outdoor paved
Indoor	90.04%	4.68%	5.28%
Outdoor non-paved	0.66%	96.78%	2.56%
Outdoor paved	2.80%	2.49%	94.71%

Table 2: Confusion matrix for our deep learning model on the test set, which consists of 9,408 images of real-world walking environments.

Discussion

Our AI-powered smart glasses with both onboard sensing and compute achieved high prediction accuracy on a large-scale image dataset of walking environments. This design opens new opportunities for real-world computer vision applications in human locomotion and robot control where inference on embedded devices and a low form factor (wearable integrated system) is required. Future work will focus on improving the inference speed.

References

1. Howard, et al. (2017). "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv.

Contact: brokoslaw.laschowski@uhn.ca

@KITE_UHN