MedAssist: An Automated Solution for The Assessment of Medication Intake

Adem Kaya, George Lalidis, Giedrius Mirklys, Tam Van

21-06-2025

1 Introduction

For the course AI in the Professional Workfield (SOW-MKI76), we were challenged to develop a project for a company in a selection of companies provided on the Masters Challenge platform. Given our shared background and passion for societal impact and healthcare, we ended up with a company called MedAssist.

1.1 MedAssist

MedAssist is a company that is concerned with building medication dispensary devices. Not only do their devices feature automatic release of medication, such that it capable of helping its patients to take their medication on time, but the devices also feature a build in camera that is capable of recoding videos of the patients whenever they are exactly in front of the device.

1.2 The Problem

As of now the video material that is collected by the deviced is manually reviewed by humans to check whether the patient in question has sucessfully taken their medication. The problem lays in the time consuming nature of this process. In order to provide a solution to this time consuming approach, MedAssist has reached out to us to build an AI which is capable of automatically assessing whether the patient has taken their medication or not.

1.3 The Goal

In order to provide a solution to the problem, the goal is to build an AI which is capable of automatically assessing whether the patient has taken their medication or not to its best extent. Not only should we aim to maximize the accuracy of the AI, but we should also carefully aim to minimize the amount of false positives as we do not want the AI to make it seem like the patient has taken their medication even though they have not.

2 Methods

2.1 Data Labelling and Processing

MedAssist had a collection of videos filmed using their medication tracking device. There were 86 videos in total where 71 videos displayed a person taking a pill or similar. It must be noted that 20 of them were difficult to decipher whether the medication was taken in or not, yet the ones that resembled an eating action or putting an item near the mouth closer were still labelled as positive instances of medication intake.

Moreover, the videos were approximately 30 s long, and their frame rate was around 60 frames per second. This leads to approximately 1800 frames per video. Due to limited computational resources, we lowered the frame count by extracting 10 frames at uniformly spaced intervals across the video's duration. These frames were saved as a NumPy zipped array, and each frame was preprocessed during training and testing using HAR model from Hugging Face[1].

2.2 Human Activity Detection (HAD)

2.3 Finetuning

Evidently, there were 15 negative cases, making the dataset highly imbalanced

3 Results

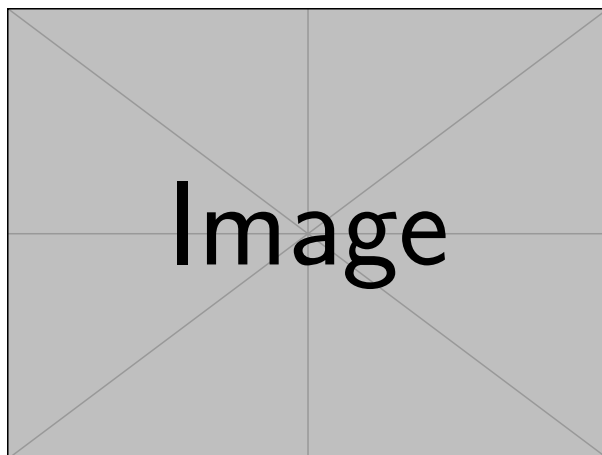Here is how you include an image (make sure the image is in the same directory or provide the path):

Figure 1: Sample Image

4 Tables

You can include tables like this:

Column 1	Column 2	Column 3
A	B	C
1	2	3

Table 1: Sample Table

5 Conclusion

Summarize the results and discuss the implications of your findings.

6 Discussion and Recommendations

References