

Project Proposal: Danger Detection using Pose Estimation

Team 6: Jiin Kim, Jonghyun Ko, Ji Un Hwang

October 27, 2020

Goal

The goal of this project is to enable computers to recognize human pose in real-time using a webcam to create and show alert messages in case of danger.

Motivation

The adoption of domestic CCTV systems is increasing over time. The reasons may vary, but the most common reason may be keeping one's family safe. People are still exposed to diverse types of danger even when they are at home. Usually, the most efficient way to call for help is to speak out. In some cases, however, speaking out is difficult or unavailable. For example, toddlers could slip and hit their head in the blink of an eye when they are out of the caregivers' sight. Patients with intellectual disabilities, neurological or psychiatric diseases could put themselves into danger without noticing. In such cases, constant monitoring is required, which is difficult and burdensome to caregivers.

Our project aims to develop an application that safeguards people from dangers when verbal communication is not feasible. Pose estimation using webcam would be utilized to recognize people in certain situations and to show alert messages regarding such situations.

Assumption and setting

We decided to focus on a scenario where people install domestic CCTV systems to monitor what is happening inside their houses. More specifically, we assumed two scenarios in which pose detection is more effective than verbal communication. The first scenario is that a person in danger does not have the intention or consciousness to call for help. An example could be a senior having a seizure. The second scenario is that a person in danger tries to signal for rescue but cannot talk. In case of domestic violence, victims are living in the same house with the perpetrators which makes it difficult for them to explicitly call for a rescue with their voices.

Therefore, we decided to detect a total of 4 poses: 2 unintentional and natural poses which might indicate danger, and 2 abnormal and uncommon poses that are registered by the users as a sign of an emergency. Details about the selection of poses will be discussed later.

Proposed method

We will use FrankMocap as a tool for 3D pose estimation. Using FrankMocap, we will create our own dataset of 3D skeleton images which will be used to train a neural network model for pose classification. The dataset would consist of 200 images per pose, a half of which is correct ones and the other half incorrect ones. A neural network model would be developed using tensorflow and keras. As to which model is used, we will try running a number of popular deep learning models and select the one that works the best.

Further work

We believe that danger detection using domestic CCTV systems and pose estimation would be useful especially for families with children or seniors. If pose detection from our project shows high performances, further works could aim at connecting the alert system to professional institutions such as police stations or hospitals.

Timeline

- FrankMocap installation and study (- 11/5)
- Algorithm design (- 11/15)
- Dataset creation (- 11/25)
- Model running and demo test (- 12/5)
- Paper works and submission (- 12/10)