Fall Detection System

Justin Haryanto

Nhan Nguyen

**SUBSYSTEM REPORTS**

REVISION – Draft

04 December 2022

Subsystem Reports

for

Fall Detection System

Prepared by:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Author Date

Approved by:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Leader Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

John Lusher II, P.E. Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

T/A Date

**Change Record**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Rev.** | **Date** | **Originator** | **Approvals** | **Description** |
| **-** | 04 Dec 22 | Nhan Nguyen |  | Draft Release |

Table of Contents

[Table of Contents 3](#_Toc121089158)

[1. Introduction 6](#_Toc121089159)

[2. References and Definitions 7](#_Toc121089160)

[2.1. References 7](#_Toc121089161)

[2.2. Definitions 7](#_Toc121089162)

[3. Pose Estimation Subsystem Report 8](#_Toc121089163)

[3.1. Subsystem Introduction 8](#_Toc121089164)

[3.2. Subsystem Details 8](#_Toc121089165)

[3.3. Subsystem Validation 9](#_Toc121089166)

[3.4. Subsystem Diagnostic 11](#_Toc121089167)

[3.5. Subsystem Mitigation 11](#_Toc121089168)

[4. Fall Detection Subsystem 12](#_Toc121089169)

[4.1. Subsystem Introduction 12](#_Toc121089170)

[4.2. Subsystem Details 12](#_Toc121089171)

[4.3. Subsystem Validation 15](#_Toc121089172)

[4.4. Subsystem Diagnostic 15](#_Toc121089173)

[4.5. Subsystem Mitigation 16](#_Toc121089174)

**List of Tables**

**List of Figures**

[Figure 1: Data Collected from Video at Low Quality 9](#_Toc121089108)

[Figure 2: Data Collected from Video at Low Quality 9](#_Toc121089109)

[Figure 3: Image from Fall Data Set after Processing by AlphaPose 10](#_Toc121089110)

[Figure 4: Data Collected from Fall Data Set Images 10](#_Toc121089111)

[Figure 5: Random forest code snippet and accuracy 13](#_Toc121089112)

[Figure 6: CNN code snippet 14](#_Toc121089113)

[Figure 7: CNN accuracy 14](#_Toc121089114)

# 

# Introduction

The fall detection system is designed to detect when a person has fallen in a video. There are two main subsystems: the pose estimation, and the fall detection subsystems, which have both been evaluated on their performance.

This system works by breaking down a video into its individual images and feeding it into the pose estimation subsystem. During this part, the subsystem extracts relevant key points and bounding box information, then sends this information to the fall detection subsystem to determine if a fall has occurred.

While these systems are currently not working at their best performance and more validation is needed, they could be eventually integrated into a whole system seen in the ConOps, ICD, and FSR.

# References and Definitions

## References

Fall Detection System (GitHub)

[https://github.com/Nhannguyen993/TAMU-GitHub](https://urldefense.com/v3/__https://github.com/Nhannguyen993/TAMU-GitHub__;!!KwNVnqRv!GqhHrXEEvFW1nE-B1zXPex1ZIgvYBdPPrgFjJ8sAyx6plz6v-Jtc1UqE3bzU1cLvkWOucI7YarFElLxtJ9sKnsky0VQDWKE$)

## Definitions

|  |  |
| --- | --- |
| CNN | Convolutional Neural Network |

# Pose Estimation Subsystem Report

## Subsystem Introduction

The pose estimation subsystem uses AlphaPose to extract key points and bounding box information from individual video frames for use in the fall detection subsystem. The subsystem is currently being run on Google Colab and uses the GPU resource there, with the code being located on GitHub. The system has been evaluated on its ability to work on videos and images of the training data set collected from other universities. This is done to verify that key points and bounding box information could be extracted.

## Subsystem Details

The pose estimation system can process videos of a variety of qualities with the use of AlphaPose, which is capable of whole-body pose estimation and tracking. It was found that AlphaPose performed well when ran on COCO, a renowned data set for large-scale object detection, segmentation, and captioning, achieving a mean average precision of 75. The COCO dataset has hundreds of thousands of regular scenes containing common objects placed in natural context to allow programs to detect and segment these objects.

The goal of this subsystem is to be able to take in a variety of videos and extract data from them. For training and testing, the UR Fall data set was used, which has labeled images as well as their associated video of various falling motions. The subsystem should be able to rapidly process video at various frame rates and differing quality. Since AlphaPose is already capable of running at 23 frames per second (fps), a challenge is getting it to work at 30 fps. Another challenge is to get the subsystem to work with videos of poor quality and with live video.

There are various methods used to examine the subsystem’s performance. To check its ability to provide key points and bounding box data, several hundred labeled images from a fall data set were used. To check the current limits of the subsystem, the pose estimation subsystem was given several videos of varying quality and duration.

Efforts to improve the frame rate that the subsystem runs at, operation with low-quality images and live video were put on hold to focus on the fall detection subsystem.

## Subsystem Validation

The pose estimation subsystem should be able to extract bounding box and key point data regardless of the quality of the video so that the fall detection subsystem could operate properly. The capability of the AlphaPose system to work with lower-quality video was demonstrated, but not thoroughly tested to work on the fall detection subsystem. Figure 1 shows a video clip, adjusted for quality, that was processed.



Figure : Data Collected from Video at Low Quality

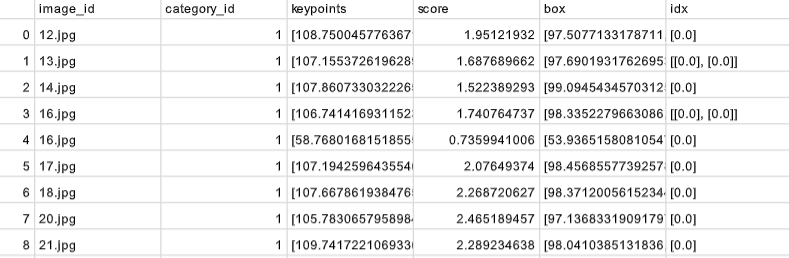


Figure : Data Collected from Video at Low Quality

In addition to this, the pose estimation subsystem was tested with images from a labeled fall data set to see if bounding boxes and key points information could be obtained. Figure 4 confirms the subsystem’s ability to supply the fall detection subsystem with bounding boxes and key point data.



Figure : Image from Fall Data Set after Processing by AlphaPose

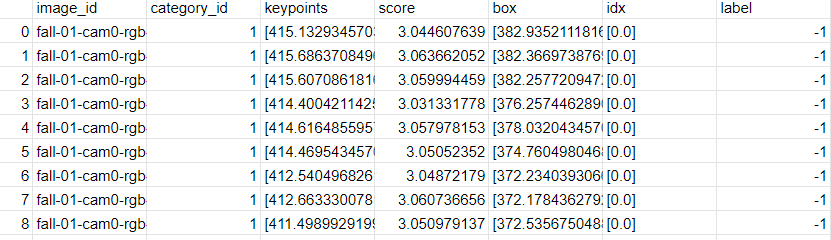


Figure : Data Collected from Fall Data Set Images

## Subsystem Diagnostic

The pose estimation subsystem has not faced rigorous testing with videos of different frame rates or quality. It also does not possess the interface for live video feed to be tested in its current configuration.

## Subsystem Mitigation

An expanded data set with videos at various frame rates and quality could be applied to the subsystem to examine its performance. Any improvements from there would be made by adjusting the underlying model of AlphaPose. For live video feed, an interface could be created to operate with the code either on Google Colab or directly on a machine.

* 1. ***Subsystem Conclusion***

The subsystem's ability to provide bounding boxes and key point data is working correctly. Its ability to handle videos at differing quality, frame rate, or as live feed has not been thoroughly tested or improved upon. When interfaced with the rest of the subsystems, it will allow a steady flow of data collected from inputted videos for the fall detection subsystem to work with.

# Fall Detection Subsystem

## Subsystem Introduction

The fall detection subsystem takes the data processed by the pose estimation subsystem and runs it through two machine learning models to determine if a fall has occurred. The results of the fall detection programs are then combined to give the final verdict on whether a fall has occurred or not. The result is then sent out to the external system owned by a client for them to use.

## Subsystem Details

For determining falls, the random forest and convolutional neural network (CNN) machine learning models were used, where they were chosen based on the data to be given. Random forest is a simple model that is sufficient for the analysis of key point information. In contrast, the CNN model’s image processing capability made it suitable for the analysis of video image frames alongside the bounding box information.

Training and testing were performed using the UR Fall Detection dataset, which contains a series of falling videos with their respective labeled image frames. After the pose estimation subsystem generates the bounding box and key points information, the data is processed by the two machine learning models for analysis. Then the results of the models are compared, where a fall has occurred only if both models agree, and the verdict is outputted to an external system. For our demonstration, the external system is represented by a video overlaid with text indicating if a person has fallen or not.

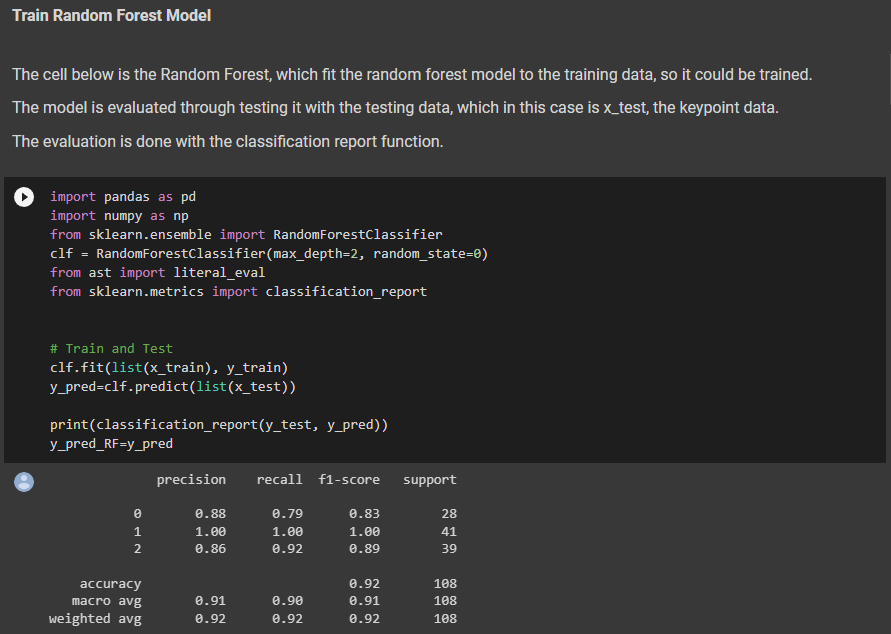


Figure : Random forest code snippet and accuracy

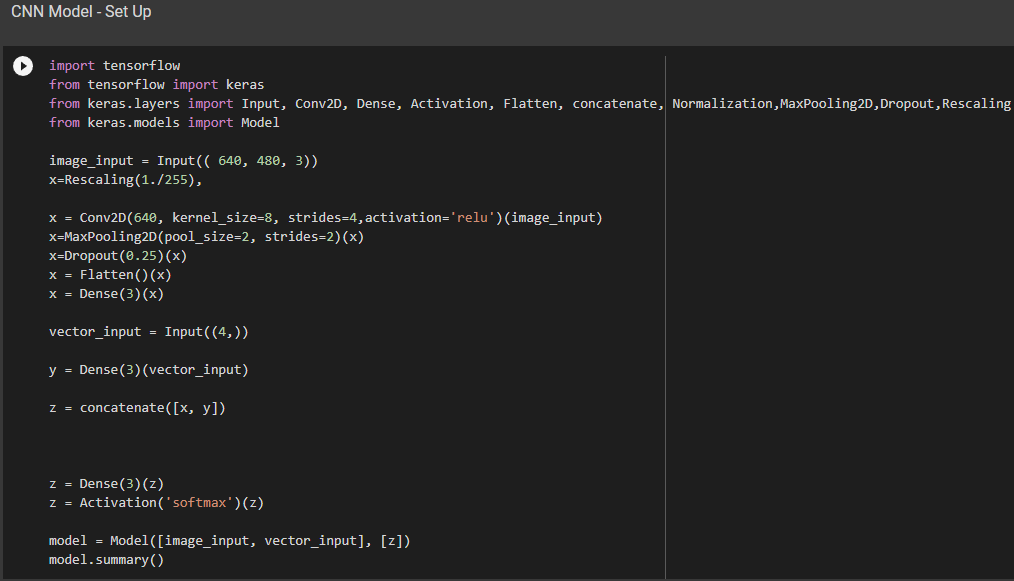


Figure : CNN code snippet

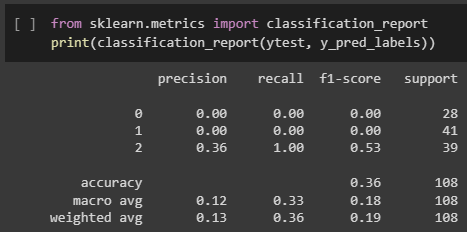


Figure : CNN accuracy

## Subsystem Validation

The following are the validation criteria for the fall detection subsystem:

|  |  |  |
| --- | --- | --- |
| **Test** | **Success Criteria** | **Methodology** |
| Fall Detection Accuracy | >90% Correct | Compare results to dataset |
| Random Forest Accuracy | >90% Correct | Analyze key points from the dataset, compare results to dataset |
| CNN Accuracy | >90% Correct | Analyze bounding boxes from the dataset, compare results to dataset |

The focus of the subsystem is to increase its categorization accuracy when determining if a fall has occurred or not. It should determine if a person has fallen regardless of any fall-like movements. This was validated through training on a hundred images from the UR Fall data set and measuring the accuracy of the models on a test set of another couple hundred images. The models work correctly if they are both able to correctly identify over 90% of the images associated with falls. The random forest was able to meet this performance for this test. The CNN performance was poor though this will be corrected with future adjustments.

## Subsystem Diagnostic

False positives can occur when the subsystem determines a fall has occurred when there wasn’t. On the other hand, false negatives occur when a fall has occurred, but the subsystem has not detected it. Based on the output, it was found that the subsystem could not decide between falling or not falling despite the person falling, resulting in a blinking of the text “fall” on the output video. However, the way the result is decided makes false negatives occur more often.

Another problem that was found was the processing time for the fall detection subsystem was slow, meaning that it will not keep up with a video.

## Subsystem Mitigation

False results will be reduced when both models are given more training and diverse datasets with a variety of poses, especially ones with fall-like movements. The parameters of the machine learning models can also be modified to tune the subsystem.

There are two ways where the speed of the fall detection subsystem could be improved. The first is to use a GPU to run the code, which allows for parallel processing. The second is to either make optimized random forests and CNN models that suit our purposes or decrease the size of the program.

* 1. ***Subsystem Conclusion***

The random forest subsystem appears to be working correctly, though the CNN model has poor performance. With refinement, the subsystem will be ready for integration in the future.