International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Detection System Using Fall and Violence Detection

Debrup Sengupta[1], Twinkle Mohanty[2]

[1,2]Kalinga Institute Of Industrial Technology

ABSTRACT

Emergency detection in places and areas that are not entirely private is really important. We need to be able to respond in an emergency to save lives. This article is about a system that can detect emergencies using videos. It knows what is happening around it. The system can detect when someone falls down and when someone is being violent. It uses a method called YOLOv8 to detect falls in real time. It also uses a way to look at videos and figure out when someone is being violent. This system can do things at the same time and it uses many different signs to automatically send out emergency warnings. Emergency detection systems like this one are important, for saving lives in emergency situations like falls and violent incidents. People looked at results from public fall detection challenges. They also checked the violence dataset. This shows that the system we are talking about is really good at finding things that happen in videos like when someone falls or gets hurt. The system is effective, in detecting these kinds of things in videos.

KEYWORDS: Fall Detection, Violence Detection, Computer Vision, Deep Learning, Emergency Alert System, YOLOv8, CNN

1. INTRODUCTION

The public and the health of people need to be protected. This is becoming more important because surveillance cameras and intelligent monitoring systems are being used more and more. In places senior care homes and residential areas people often get hurt from falls or violent events. Watching the images manually is not a way to do this. It takes a lot of time and mistakes can happen.

Surveillance cameras and intelligent monitoring systems are being used to help with this. The health of individuals and the public need to be safeguarded. Falls and violent events are the kinds of emergencies that happen a lot in zones and residential areas. Manual monitoring of images from surveillance cameras is not very good. It is slow. Mistakes can be made.

People are starting to use surveillance cameras and intelligent monitoring systems often to safeguard the public and the health of individuals. This is a thing because falls and violent events are common, in public places and senior care institutions. Manual monitoring is not the way to keep people safe. Surveillance cameras and intelligent monitoring systems can help to protect the public and the health of people.

There have been some advances in deep learning techniques that are used for convolutional neural networks and real-time object detection algorithms. This means we can now look at visual patterns in a really effective way.However most of the designs and developments in this area are limited to one situation like detecting when someone falls or when there is violence. This makes them not very useful, in life situations.

In this paper we are talking about a computer vision system that can automatically start emergency response measures when it detects someone falling or violence. The system is designed to handle both falling and violence which are two important things that require emergency response measures.

2. Related Work

Fall detection is something that people have been looking into a lot. They have been using sensors that you wear and systems that use cameras to detect falls. Using cameras to detect falls with something called CNNs and pose estimation is also a way to do it because it does not bother the person.
Fall detection using cameras is an option.
People have also been trying to detect violence in videos. They have been using features that look at space and time 3D CNNs and something called RNNs to detect violence in videos. Fall detection and violence detection are things to work on.
People have been trying to use YOLO-based architectures to detect when someone falls because YOLO-based architectures can process things in time. YOLO-based architectures are good for this.
Light CNNs are also used to figure out when someone is being violent because they make things easier to process.
This study is trying to do two things: detect when someone falls and detect violence. The study wants to do both of these things in one system that can respond quickly. The system is, for detecting falls and violence.

3. Proposed System Architecture

The system we are talking about has four parts: the video input part, the fall detection part, the violence detection part and the emergency alert part.
The system takes video frames from a camera or some other video source. It looks at them right away.
It is always checking the video input part. The other parts like the fall detection part and the violence detection part to see what is happening.
The video frames are, like pictures that the system uses to figure out what is going on so it can send out an emergency alert if it needs to.
The fall detection module uses a YOLOv8 model to find fall patterns and boxes around people who have fallen. This is really important.
The violence detection module has a kind of computer program called a CNN. It looks at video frames to see if a scene is violent or not.
When either of these modules sees something bad happening and it is pretty sure, about it the emergency alert module sends out a warning to the ambulance or the police. This happens with the fall detection module and the violence detection module.

4. Dataset and Preprocessing

There are two public datasets used for training and testing purposes:

- Fall detection image dataset marked with images indicating fall and non-fall occurrences.
- Dataset of real-life violence situations along with video clips labeled into violent or non-violent situations.

The datasets which we used for research work are from Kaggle

1. For Falling : https://www.kaggle.com/datasets/shahliza27/ur-fall-detection-dataset

2. For Violence:https://www.kaggle.com/datasets/mohamedmustafa/real-life-violence-situations-dataset

To figure out if there is violence, in a video we look at parts of the video at times. We make sure the pictures are all the same before we use them to train the system. We do not want the system to be unfair so we make sure the data is balanced. We take samples from all parts. When we are trying to detect if someone has fallen we label the tags in a way that works with YOLO.

5. Model Training

The fall detection model was trained using the YOLOv8 architecture.The violence detection model uses a kind of neural network that is called a CNN. This CNN is based on MobileNet. It helps the model work faster. The fall detection model and violence model which are made are trained with 50 epochs each. When we trained these models we used something called binary cross-entropy and the Adam optimizer to help them learn better.We used these things to make sure the fall detection model and the violence detection model work well. The training process took a rounds and it worked out really well in a some amount of time. This is because we used -trained weights and did the preprocessing correctly. We saved the models so we can use them for real-time processing and to make predictions with the models. Then we integrated both the models for making a well emergency service system for the safety of the people in roads if any person fall in road due to the some accident or any medical issue camera will detect the fall and direct contact to the ambulance, as ambulance contact varies from country to country so ambulance contact is not preloaded and if any violence captures on camera it also contact police helpline number for quick help.

6. Experimental Results and Discussion

The system was checked to see how well it works. This was done by using videos that were recorded ahead of time to make it seem like the system was getting video for keeping an eye on things. The fall detection model was really good at finding out when someone fell. The violence classification model was also very good, at figuring out what was happening in the test videos. The fall detection model and the violence classification model were both tested with videos. The fall detection model did a job of detecting falls and the violence classification model did a great job of classifying what was happening.

We use the pre recorded video : https://www.shutterstock.com/video/search/falling-man

Resultsis

The console shows us that it found and responded to an emergency situation. This tells us that our system is working the way it should. We added this system to reduce warnings. Now the system only sends alerts when it is really sure something is wrong. What we learned is that using models that do things at once to find unusual events is better than using models that only do one thing when we want to make good guesses. The results show that these multi-task models for finding events are better than single-task models, at making predictions.

7. Applications

The system can be used in surveillance systems, elderly care systems, safety systems, hospitals and home security systems. This system is really good, for these things. The system can do a lot of things too like find problems that are not normal which is what the system is made for to detect anomalies in smart surveillance systems and other safety systems and home security systems.

8. Limitations and Future Work

The current way of doing things is pretty good. It has some problems. It really depends on how good the video's where the cameras are. In the future researchers will try to make it work better when things are blocked from view. They will also try to make a model that can understand what happens over time to detect violence. The researchers will also try to use video and audio to make the violence detection more accurate. They want to use sensors, like video and audio to make the violence detection better.

9. Conclusion

This research work is, about creating a system that can detect emergencies using cameras. It focuses on finding out when someone has fallen down or when there is violence happening. The system uses computer programs called deep learning algorithms to do this. We used a program called YOLO to detect falls and another program called CNN to figure out when violence is happening. When we combined these two programs we were able to make a system that can automatically respond to emergencies. The results of our experiment show that this system can really work in life situations where we need to watch what is happening.

References

1. Redmon, J., et al., "You Only Look Once: Unified, Real-Time Object Detection," CVPR.
2. Simonyan, K., Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," ICLR.