Human Activity Detection

A project report submitted in partial fulfillment of the requirements for the award of the degree of

B.Tech. in Computer Science and Artificial Intelligence

by

Raman Kiran Nittala
(LCI2020006)
Abhinav Jalluri
(LCI2020068)
Vivek Vardhan Kova
(LCI2020078)

under the guidance of
Vishal Krishna Singh

विद्या ददाति विनयं विनयाद् याति पात्रताम्

Indian Institute of Information Technology, Lucknow
May 2023

Declaration of Authorship

we, **Rama Kiran,Abhinav and Vivek**, declare that the work presented in “**Human Activity Detection**" is our own. we confirm that:

- This work was completed entirely while in candidature for B.Tech. degree at Indian Institute of Information Technology, Lucknow.
- Where we have consulted the published work of others, it is always cited.
- Wherever we have cited the work of others, the source is always indicated. Except for the aforementioned quotations, this work is solely our work.
- I have acknowledged all major sources of information.

Signed:

(Rama Kiran) (Abhinav) (Vivek)

Date:

CERTIFICATE

This is to certify that the work entitled "**Human Activity Recognition**" submitted by **Rama Kiran, Abhinav and Vivek** who got their name registered on **Jul 2020** for the award of B.Tech. degree at Indian Institute of Information Technology, Lucknow is absolutely based upon their own work under the supervision of **Dr.Vishal Krishna Singh**, Department of Computer Science, Indian Institute of Information Technology, Lucknow - 226 002, U.P., India and that neither this work nor any part of it has been submitted for any degree/diploma or any other academic award anywhere before.

Vishal Krishna Singh
Department of Computer Science
Indian Institute of Information Technology, Lucknow
Pin - 226 002, INDIA

Acknowledgements

We would like to express our sincere gratitude to everyone who has contributed to the completion of this project. First and foremost, we would like to thank our project supervisor, **Dr. Vishal Krishna Singh** and our evaluators **Dr. Niharika Anand**,**Dr. Abhinesh Kaushik** and **Dr. Naveen Saini** for their invaluable guidance and support throughout the project. We would also be thankful to our director **Dr. Arun Mohan Sherry** for providing all the required facilities in completion of this project. We would also like to thank **Dr.Naveen saini** for special guidence and our colleagues who have provided us with their valuable feedback and insights that helped us to improve our work.

Lucknow
April 2023

Rama Kiran,Abhinav,Vivek

ABSTRACT

The first step in this project is to collect and preprocess the image dataset. The images are captured using various devices such as cameras and smartphones, and are then resized, standardized and normalized to ensure that they have the same dimensions and are in the same format. The dataset is then split into training and testing sets, with the training set used to train the YOLO model, while the testing set is used to evaluate the performance of the model.

The YOLO algorithm is a state-of-the-art object detection system that is commonly used in computer vision applications. It works by dividing the input image into a grid of cells and predicting bounding boxes and class probabilities for each cell. The algorithm then uses non-maximum suppression to filter out redundant detections and produce a final set of bounding boxes and their corresponding class probabilities. YOLO is known for its speed and accuracy, making it well-suited for real-time object detection applications.

In this project, YOLO was used to detect and classify human activities based on the visual cues present in the images. The images were preprocessed to ensure that they have the same dimensions and are in the same format, and were then fed into the YOLO algorithm for detection and classification. The YOLO model was trained on the training dataset using transfer learning techniques, where a pre-trained model was used as a starting point. The model was fine-tuned on the training dataset, and the hyperparameters were optimized using techniques such as grid search and random search.

The project focuses on detecting and classifying five different activities: smoking, football, violence, sitting, and standing. The images in the dataset are labeled accordingly, and the YOLO model is trained to recognize these activities based on the visual cues present in the images. The model is evaluated using various performance metrics, including accuracy, precision, recall, and F1 score.

The results of the project show that the YOLO model is capable of accu-

rately detecting and classifying the selected activities. The model performs well even on images with multiple activities, demonstrating its ability to handle complex situations. The results also demonstrate the importance of hyperparameter tuning, with the optimal hyperparameters resulting in significant improvements in model performance.

In conclusion, this project demonstrates the feasibility of using YOLO for human activity recognition based on image data. The YOLO model trained on the dataset is capable of accurately detecting and classifying different activities. The project highlights the importance of hyperparameter tuning in improving the accuracy of activity recognition using YOLO. The project has numerous applications in areas such as healthcare, sports analysis, and surveillance, and can be extended to recognize a broader range of human activities.

Contents

Chapter 1

Introduction

Introduction about Human activity detection

1.1 Objective

The objective of our project is to develop a human activity recognition system using YOLO [1] , a real-time object detection system. The system will be trained on a custom dataset of videos and their corresponding activity labels, and will be able to accurately detect and classify human activities in real-time.

1.2 Motivation

The motivation behind this project is to create a system that can assist in a variety of applications such as surveillance, healthcare, and sports analysis. In surveillance, the system can be used to detect and alert security personnel of suspicious activities or intruders. In healthcare, the system can be used to monitor patients and detect abnormalities in their movements. In sports analysis, the system can be used to analyze the movements of athletes and provide feedback on their performance.

Traditional methods for human activity recognition, such as handcrafting features and training classifiers, are often time-consuming and may not generalize well to new datasets. Deep learning-based approaches, on the other hand, have shown great promise in human activity recognition. YOLO, in particular, is a popular deep learning-based approach that can efficiently detect and classify objects in real-time.

By developing a human activity recognition system using YOLO, we aim to create a system that is both efficient and accurate. The system will be able to handle a variety of activities and will be adaptable to different environments and lighting conditions. Additionally, the system will be able to operate in real-time, making it suitable for real-world applications.

Overall, the objective of our project is to develop a robust human activity recognition system using YOLO that can assist in a variety of applications. The motivation behind the project is to create a system that can improve safety, healthcare, and sports analysis, while also demonstrating the capabilities of deep learning-based approaches for human activity recognition.

Chapter 2

Literature Review

This chapter resembles the brief introduction about Human Activity Detection

2.1 Role of Human Activity Recognition

Human activity recognition (HAR) refers to the use of machine learning algorithms and sensors to identify and classify human actions or activities. The role of HAR is to provide valuable insights into people's daily lives, health and wellness, and even security.
Here are some specific roles of HAR

2.1.1 Health and wellness

HAR can be used to monitor physical activity and provide feedback on fitness goals. It can also be used to track and monitor people with certain medical conditions, such as Parkinson's disease, to better understand their movements and provide personalized care.

2.1.2 Security

HAR can be used to detect unusual or suspicious behavior, such as break-ins or intrusions, by analyzing patterns of movement and activity in a given area.

2.1.3 Human-robot interaction

HAR can be used to improve the interaction between humans and robots by allowing the robot to recognize and respond to human gestures and movements.

2.1.4 Sports and entertainment

HAR can be used to analyze and improve athletic performance by tracking movement and identifying areas for improvement. It can also be used in the entertainment industry to enhance virtual reality and gaming experiences.

Overall, HAR has the potential to improve various aspects of human life, from health and wellness to security and entertainment, by providing insights into human behavior and movement patterns.

2.2 Use of YOLO model

The YOLO (You Only Look Once) model is a popular deep learning-based object detection system that can be used for a variety of applications. Here are some of the main uses of YOLO:

2.2.1 Object detection in images and videos

YOLO can detect and localize objects in images and videos, making it useful for tasks such as surveillance, traffic monitoring, and autonomous vehicles.

2.2.2 Face detection

YOLO can be used for face detection in images and videos, which is useful for applications such as security and facial recognition.

2.2.3 Object tracking

YOLO can track objects over time, which is useful for applications such as video surveillance and sports analysis.

2.2.4 Human pose estimation

YOLO can be used for human pose estimation, which is useful for applications such as fitness tracking and gesture recognition.

2.2.5 Augmented reality

YOLO can be used for object recognition in augmented reality applications, allowing virtual objects to be placed and interacted with in the real world.

Overall, the YOLO model is a versatile and powerful deep learning tool that can be used for a wide range of applications in computer vision and artificial intelligence.

YOLOv8 uses a single shot multi-box detector (SSD) architecture, which means that the model predicts the bounding boxes and class probabilities for an object in a single forward pass. This makes YOLOv8 fast and efficient, allowing it to run in real-time on relatively low-powered devices.

The YOLOv8 architecture is based on the ResNet-50 backbone, which is a deep convolutional neural network (CNN) that has been pre-trained on the ImageNet dataset. The ResNet-50 backbone is then fine-tuned on the COCO dataset to learn to detect objects in 80 different categories.

Chapter 3

Methodology

The proposed method for human activity detection using YOLO consists of the following steps:

3.1 Dataset collection:

A custom dataset consisting of images of human activities such as smoking, football, violence, sitting was collected. The images were collected from various sources such as publicly available datasets and online sources.

3.2 Dataset annotation:

The images were annotated using the LabelImg tool to create bounding boxes around the regions of interest (ROIs) for each activity. The annotations were saved in the YOLO format, which includes the class label, bounding box coordinates, and image dimensions.

3.3 Data pre-processing:

The annotated dataset was split into training and validation/Testing. The images were resized images and converted to the YOLO format.

3.4 Model selection:

We selected YOLOv8 model as our basse model for Human Activity Recognition and detection,we choose this model for its high accuracy and efficiency in object detection tasks.

3.5 Model Configuration:

We set the no of classes to the number of human activities we want to recognize and set the anchor sizes based on the size of the object we want to detect

3.6 Model training:

The YOLOv8 model was used to train the dataset on a GPU-enabled system(Google colab). The model was trained for 25 epochs with a batch size of 4 . The model weights were saved after epochs.

3.7 Model evaluation:

The trained model was evaluated on the validation set to measure the model's accuracy in detecting and classifying the activities.

3.8 Model testing:

The trained model was tested on a separate test set of images to evaluate its performance in detecting and classifying activities in real-world scenarios.

3.9 Alerting:

When any odd activity or harmful activity is discovered, an alert message will pop out.

3.10 Results

The results were analyzed and compared with the state-of-the-art methods in the field of human activity recognition.

Chapter 4

Simulation and Results

Activity Detection can be used to see the unusual activities in college

4.1 Dataset:

We used a custom dataset consisting of images and videos of various human activities, such as sitting, smoking, fighting, and walking collected from a variety of sources. The dataset was split into a training set and a validation set to train and evaluate the YOLO model.

4.2 Preprocessing:

4.2.1 Resizing

All the images and videos(are extracted in frames) were preprocessed. We used data augmentation techniques such as random scaling, flipping, and rotation to increase the size of our dataset and to make our model more robust to different scenarios.

4.2.2 Labellimg and yaml file:

To label the preprocessed photos and create boundary boxes around the areas of the image where we wish to look for an action or object, we utilised the LabelImg [2] application.Accordingly, those images and labels are divided into train and testing, and we also made a yaml file to specify the locations of train and validation and the classes involved.

4.3 Model architecture:

We used the YOLOv8 architecture as our base model for human activity recognition and detection. We customized the last layer of the model to output the probability of different human activities and the location of the person in the image.

4.4 Training:

We trained our custom YOLO model on Google colab using GPU. We used a batch size of 4 and trained the model for 25 epochs. We used the Adam optimizer with a learning rate of 0.001.

4.5 Real-world testing:

We tested our model on real-world videos of people performing different activities, such sitting, smoking, fighting, and walking and observed that our model was able to detect and recognize these activities accurately.

4.6 Alerting:

We used an Telegram bot to send an alert when an unusual activity is detected.

4.7 Conclusion:

In conclusion, we have successfully trained a custom YOLO model for human activity recognition and detection. Our model achieved good accuracy on both the validation set and real-world testing, indicating that it can be used for various applications such as surveillance and sports analysis. Future work could involve fine-tuning the model on larger datasets or using more advanced architectures to improve the accuracy of the

Chapter 5

Conclusion and Future Work

In this project, we proposed a method for human activity detection using the YOLOv8 algorithm. The proposed method was trained and tested on a custom dataset consisting of images of human activities such as smoking, football, violence, sitting, and standing. The model achieved high accuracy rates in detecting and classifying these activities in real-world scenarios. The YOLOv8 algorithm proved to be effective in detecting and classifying human activities with high accuracy rates, and the proposed method could be used in various applications such as video surveillance and sports analysis.

5.1 Future Work:

There are several directions for future work that can improve the proposed method for human activity detection using YOLOv8. These include:

5.1.1 Augmenting the dataset:

The accuracy of the model can be improved by increasing the size of the dataset and adding more diverse images of human activities.

5.1.2 Fine-tuning the model:

Fine-tuning the model with additional data or adjusting the hyperparameters can improve the model's performance.

5.1.3 Incorporating temporal information:

Currently, the model processes each frame independently. Incorporating temporal information from multiple frames could improve the accuracy of the model in detecting human activities.

5.1.4 Real-time implementation:

The proposed method can be implemented in real-time using efficient hardware such as GPUs or FPGAs. Real-time implementation would be useful in applications such as video surveillance and sports analysis.

5.1.5 Extending to other domains:

The proposed method can be extended to detect human activities in other domains such as healthcare, retail, and transportation.

In conclusion, the proposed method for human activity detection using YOLOv8 proved to be effective in detecting and classifying human activities with high accuracy rates. The future work outlined above can improve the accuracy and efficiency of the proposed method and extend it to other domains.

Appendix Title Here

5.2 Dataset Description:

We collected images from different sources and created a custom data set and labelled each of the image accordingly to train our model. we have...no of images....images.

5.3 Model Architecture:

This appendix can provide a more detailed description of the architecture of the YOLOv8 model used in the project, including the number of layers and the size of each layer.

5.4 Training and Validation Curves:

This appendix can include graphs and plots of the loss and accuracy curves during training and validation of the YOLO model.

5.5 Confusion Matrix:

This appendix can show a confusion matrix that provides information on the misclassifications made by the YOLO model during testing

5.6 Sample Outputs:

This appendix can include sample outputs from the YOLO model, such as images with bounding boxes around detected activities and the corresponding class labels.

5.7 Hardware and Software Configurations:

This appendix can provide details about the hardware and software used in the project, including the computer specifications, operating system, and software libraries used.

5.8 Code:

This appendix can include a copy of the code used to implement the YOLO model for human activity detection.

5.9 Further Analysis:

This appendix can provide additional analysis of the results obtained, including insights into the distribution of activity classes in the dataset, the relative performance of the YOLO model on different classes, and the potential sources of error in the model.

These appendix titles can provide additional details and insights into the project and help readers to better understand the methods and results.

.

Bibliography

[1] "Yolo v8 model," 2023.

[2] "labelimage," 2015.