SEGMENTATION TO DETECT THE PRESENCE OF HUMAN DURING NATURAL DISASTERS

By

CHETHANA H. T * POOJA C. A ** P. VAMSHIKA ***

*-*** Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Date Received: 21/12/19 Date Revised: 26/03/20 Date Accepted: 28/04/20

ABSTRACT

Nowadays, object detection is widely used for various purposes. There is a need for detecting living objects so that various problems like humans getting stuck in an undesirable place, and to find the suspected objects out of a group of other objects. Image segmentation is a process of dividing a digital picture into many segments, which helps in getting meaningful and well analyzed data that can be used in obtaining a well-defined output. Detection of living beings during the natural disasters is the most challenging task. Our proposed system makes use of You Only Look Once (YOLOv3), the quickest and most accurate among object detection system. During natural disasters, objects/living beings are detected using YOLOv3 model and the same information will be sent to the rescue team so that lives of human beings will be saved quickly and effectively.

Keywords: Convolutional Neural Network, Mask-RCNN, Image Segmentation, Feature MAP, Object Detection.

INTRODUCTION

Human posture recognition is now widely used as it has its applications in various fields. It is a system that is based on 3D motion that uses various means of media such as video inputs from CCTV. A2D video-based system captures movements of human beings using the available video camera and loads that into a computer. The system is affordable to all types of users. However, it cannot perform detailed movement analysis automatically (Rosenberg et al., 2010). Human posture detection is now widely used and implemented in many applications, like content-based retrieval, indoor, surveillance and outdoor monitoring.

Nowadays, we have seen thousands of humans and animals suffering in the disaster since they were not identified in that chaotic place, due to which many of them could not get timely help when required. To avoid this, it is important to detect humans during natural disasters using segmentation of images.

The main aim of this paper is to develop an algorithm that can detect the living beings in the scene quickly and accurately.

1. Related Work

In this section, we briefly discuss existing literature survey on object detection and also discuss various methods applied for object detection.

Pulse Coupled Neural Network (PCCN) is applied to detect humans moving in a disturbed background. It was proposed by (Gulli & Pal, 2017) for recognizing human, where analyzing the body parts is vital. Another working system was proposed by Baumgart and Mies (2007) for tracking and detecting the movements of humans by using a combination of different fields of Artificial intelligence. A Hidden Markov Model (Rosenberg et al., 2010) was proposed using infrared method for monitoring smart homes for recognizing human postures. Based on the characteristics of human posture, a training data model was produced, which classifies the human posture effectively. Figure 1 shows the Human Posture Recognition with key points.

Many techniques were proposed for detecting animals entering the roads giving a warning to the drivers via audio and visual signals. It was useful for real time applications. Some method like edge based matching, animal

Figure 1. Human Posture Recognition Key Points (Tensor Flow, n.d.)

detection, object matching, template matching, frame differencing, skeleton extraction (Keras Documentation, n.d.) are utilized.

Automatic covert camera traps are being used to efficiently monitor wild animals in various natural habitats. This method was helpful in wildlife conservation to take appropriate management decisions (Patel et al., 2017). Due to agricultural mowing operations, many animals are harmed every year. To promote wildlife-friendly farming and to reduce wildlife deaths, there is much importance for the detection and recognition of wildlife within the agricultural fields. Objects are detected based on threshold and it is adjusted to each frame. Thermal feature extraction algorithm is used for classifying animals about the heat characteristics of objects and are described using thermal signature, which is parameterized using Discrete Cosine Transform (DCT). To discriminate animals from non-animals, k-Nearest-Neighbor (kNN) classifier is applied (Arasanipalai et al., 2018).

Many research based techniques were proposed which helps in detecting humans via audio and visual signals. Methods like Phantom and Support Vector Machine, DCNN, Template Matching Algorithm, Pulse Coupled Neural Network, Decision tree, HOG, CNN, and MLBP were applied. Thus there is a scope for the development of new approaches in the field of object detection during natural disasters. The development of new approaches in this field of object detection during the natural disasters has a wide range of scope.

2. Proposed System

In our proposed system, we use You Only Look Once for object detection that alerts the user about the possible intrusion. You Only Look Once (YOLOv3) is the quickest and most accurate among object detection systems. But there is a tradeoff between accuracy and speed. Our proposed system involves h-level design (HLD), low level design, and system architecture. Flow diagram of our proposed system is shown in Figure 2.

2.1 High Level Design

It describes the architecture which is used for developing a software product. An architecture diagram describes the overview of an entire system, which identifies the main components developed for the product and their interfaces.

2.2 System Architecture

It is a conceptual model that defines the structure, behavior, and more views of a system. An architecture description gives a brief formal description and representation of a system, which is organized in a way that supports the reasoning about behaviors and structures of the system. The architecture of the proposed system is shown in Figure 3. The system architecture can be divided in to two parts namely Object Detection and Storage.

2.3 Low Level Design

It is a step by step process of refinement and also termed as component-level design. This process can be used for designing software architecture, data structures, required source code and ultimately, performance algorithms

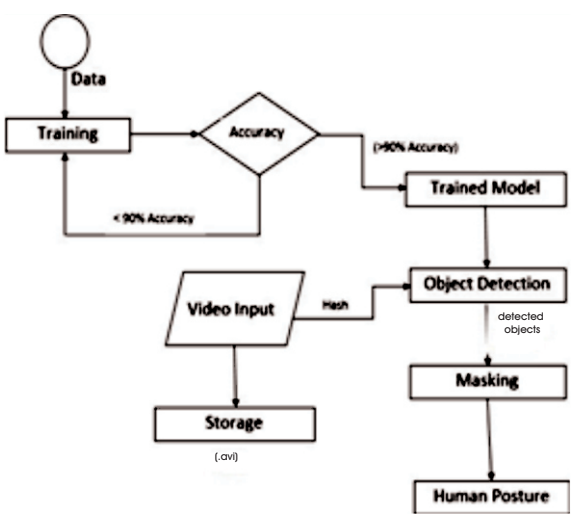

Figure 2. Flow Diagram of the Proposed System

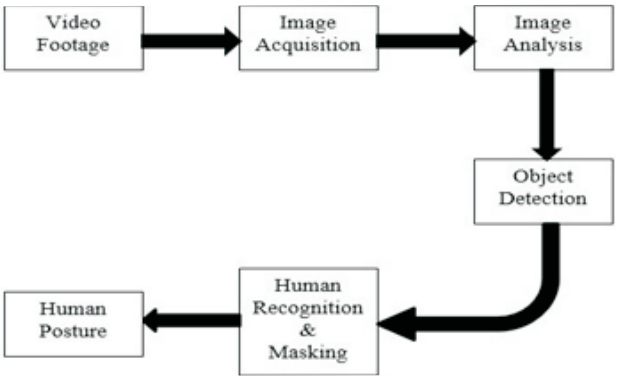

Figure 3. System Architecture (Tensor Flow, n.d.)

(Tutorials Point, n. d.).

Pseudocode is an informal low-level description of the operating principle of a computer program or other algorithm as shown in Figure 4. It uses the structural conventions of a normal programming language, but is intended for human reading rather than machine reading.

3. Training Process

The training process is done by making use of Mask R-CNN method as shown in Figure 4. Here outputs are obtained as a binary mask which predicts whether a given pixel is a part of an object or not. The branch in white in the above image, as before, is just a Fully Convolutional Network on top of a CNN based feature map. Here are some of the inputs and outputs

3.1 Inputs

Feature Map of CNN.

3.2 Outputs

The pixels that are belonging to the object are represented by 1s in the Matrix and 0s elsewhere. This is called as a binary mask (Valchanov et al., 2018).

Figure 5 shows an image provided as an input to CNN. Selective search algorithm is applied to get feature map. To identify the region proposal, a separate network is used and these predicted region proposals are reshaped using a Region of Interest (RoI) pooling layer method. The images inside a proposed region are classified using RoI layer and then predicts the offset values for all the

```
import os

import model as modellib
from model import log
ROOT_DIR = os.getcwd()
MODEL_DIR = os.path.join(ROOT_DIR, "mylogs")
# Local path to trained weights file
COCO_MODEL_PATH = os.path.join(ROOT_DIR, "mask_rcnn_coco.h5")
# Run one of the code blocks

# Shapes toy dataset
# import shapes
# config = shapes.ShapesConfig()

# MS COCO Dataset
import coco
config = coco.CocoConfig()
COCO_DIR = "D:/Github/FastMaskRCNN/data/coco"  # TODO: enter value here
# Load dataset
assert config.NAME == "coco"
# Training dataset
# load person keypoints dataset
train_dataset_keypoints = coco.CocoDataset(task_type="person_keypoints")
train_dataset_keypoints.load_coco(COCO_DIR, "train")
train_dataset_keypoints.prepare()

val_dataset_keypoints = coco.CocoDataset(task_type="person_keypoints")
val_dataset_keypoints.load_coco(COCO_DIR, "val")
val_dataset_keypoints.prepare()

print("Train Keypoints Image Count: {}".format(len(train_dataset_keypoints.image
print("Train Keypoints Class Count: {}".format(train_dataset_keypoints.num_class
for i, info in enumerate(train_dataset_keypoints.class_info):
    print("{:3}. {:50}".format(i, info['name']))

print("Val Keypoints Image Count: {}".format(len(val_dataset_keypoints.image_ids
print("Val Keypoints Class Count: {}".format(val_dataset_keypoints.num_classes))
for i, info in enumerate(val_dataset_keypoints.class_info):
    print("{:3}. {:50}".format(i, info['name']))
```

Figure 4. Sample Pseudocode of the Proposed System

bounding boxes obtained (Arasanipalai et al., 2018; Common Objects in Context, n.d.).

4. Results and Discussions

Most of the human presence detected are true positives and we have 85% accuracy. We have-trained our model using COCO dataset. Figures 6-9 represent the snapshots of detecting humans during a disaster.

In our model, the human posture is calculated by taking the key features like, distance between the arm shoulders.

When the video is uploaded, the output is in the form of a VI file that is instantly created and saved in the folder of the program.

5. Advantages of Proposed System

Our proposed system has the following advantages.

Cost effective solution for storage.

Rendering and analyzing of video footage in real time.

Advanced intrusion detection system with option to classify over 80 most common items (COCO dataset) (Common Objects in Context, n.d.).

Saves precious time of manually checking every video footage in case of any mishap.

Model is versatile and can be trained for specific cases like face detection.

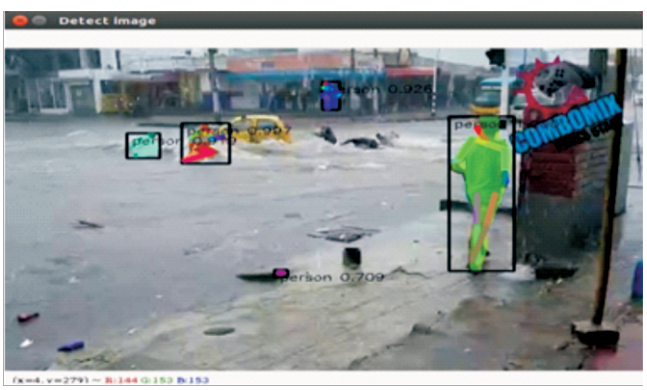

Figure 5. Mask R-CNN training model Arasanipalai et al. (2018)

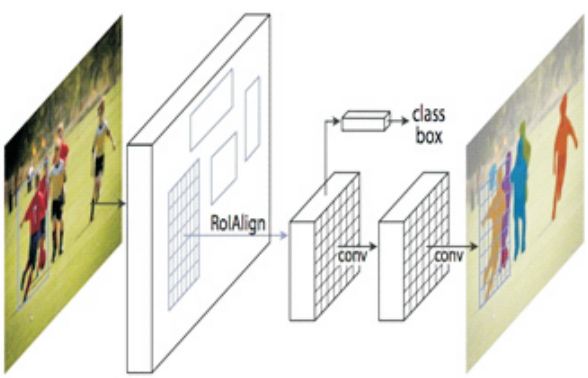

Figure 6. Snapshot from the Input Video

Figure 7. Result displaying the Detected Humans

Figure 8. Snapshot Displaying False Positive

Figure 9. Snapshot of Detected Humans in Another Scenario

Decentralized storage comes with added advantage of less management issues and cost effectiveness and attacks like DDoS will be ineffective against IPFS by considering more number of peers participating in the network.

Conclusion

Many research techniques were analysed which helps in detecting humans via visual images. In our proposed system, we automated the process of detect living beings during natural disasters using image segmentation. The collected inputs are segmented using the concepts of Mask-RCNN and the living beings are identified. This helps us in reducing the time, physical effort, and the accuracy of identification of living beings during natural disasters is also increased. Our proposed system has developed an algorithm which recognize it in an efficient way.

Future Enhancements

We believe that we have trained our model to its best for the specific use with the available dataset. But for the special cases like facial recognition and car model recognition, we need specific datasets. Out model is also specific, and we can train the RCNN for any purpose if the dataset is available for that purpose. We can further use the model for more generic use and medical studies so that they can be used in various object detections.

Block chain

Further enhancement would be in the storage services and the distribution of medical images and eventually created market place for the medical studies. And also using block chain like Ethereum to store the hashes securely. Block chain is an underlying boom after advent of bit coin and it stores our data securely. Since hash is a very important asset of our system, we would love to store it in more secure way using block chain.

Voice

Recommendation: Conversion of text to voice for physically challenged people.

Robust back end

Here, we have used light weight back end with least and basic uses of it. Further enhancement is by using robust services like Django for back end so that our APIs are more secure and is observed for securing of transmission of the data across the REST network.

References

[1]. Arasanipalai, A. U. (2018). *State of the art deep learning: An introduction to Mask R-CNN.* University of Illinois, Urbana-Champaign, Urbana, IL. https://www.freecodecamp.org/news/mask-r-cnn-explained-7f82bec890e3/

[2]. Baumgart, I., & Mies, S. (2007, December). S/kademlia: A practicable approach towards secure key-based routing. In 2007 *International Conference on Parallel and Distributed Systems* (pp. 1-8). IEEE. https://doi.org/10.1109/ICPADS.2007.4447808

[3]. Common Objects in Context. (n.d.). *SE(3) Computer Vision Group,* Cornell Tech. https://vision.cornell.edu/se3/projects/microsoft-coco/

[4]. Gulli, A., & Pal, S. (2017). *Deep Learning with Keras.* Packt Publishing Ltd.

[5]. Keras Documentation. (n.d.). *The Sequential Model*. Retrived from https://keras.io/getting-started/sequential-model.

[6]. Patel, P., Bhatt, B., & Patel, B. (2017, February). Human body posture recognition—A survey. In 2017 *International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 473-477). IEEE. https://doi.org/10.1109/ICIMIA.2017.7975660

[7]. Pooja, C. A., Vamshika, P., Jain, R. B., Jain, V. K., & Chethana, H. T. (2018, December). Comprehensive survey on detection of living or dead humans and Animals using different approaches. In *International Conference on Computer Networks, Big Data and IoT* (pp. 94-100). Springer, Cham. https://doi.org/10.1007/978-3-030-24643-3_10

[8]. Rosenberg, J. (2010). {i}Interactive connectivity establishment (ICE): A protocol for network address translator (NAT) traversal for offer/answer protocols {i}, *Internet Engineering Task Force (IETF),* RCF 5245. Retrieved from https://www.hjp.at/doc/rfc/rfc5245.html

[9]. Tensor Flow. (n.d.). *Object detection.* Retrived fromhttps://www.tensorflow.org/lite/models/object_detection/overview

[10]. **Tutorials Point. (n.d.).** *System Analysis and Design - Overview.* Retrived from https://www.tutorialspoint.com/system_analysis_and_design/system_analysis_and_design_overview.htm

[11]. **Valchanov, I. (2018).** False positive and false negative. *Towards Data Science.* Retrived from https://towardsdata science.com/false-positive-and-false-negative-b29df2c60aca

ABOUT THE AUTHORS

Chethana H. T is currently working as an Assistant Professor in the Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India and pursuing Ph.D. in JSS Science and Technological University, Mysuru, Karnataka, India. She received her Bachelor's degree in Computer Science and Engineering from Coorg Institute of Technology, Ponnampet, in 2011 and M.Tech degree in Software Engineering from Visvesvaraya Technological University, in 2015. She has 5 years of experience in teaching. She is a life member of Indian Society for Technical Education. Her current research interests include Image Processing, Machine Learning and Forensics.

Vamshika P received her Bachelor's degree in Computer Science and Engineering from Vidyavardhaka College of Engineering, Mysuru, Karnataka, India, in 2019. Her current research interests include Image Processing, Machine Learning and Deep Learning.

Pooja C A received her Bachelor's degree in Computer Science and Engineering from Vidyavardhaka College of Engineering, Mysuru, Karnataka, India, in 2019. Her current research interests include Image Processing, Machine Learning and Deep Learning.