InternationalJournalofScientificResearchandEngineeringDevelopment-– Volume5Issue4,July-August2022

RESEARCHARTICLE OPENACCESS

PERSON DETECTION AND AERIAL IMAGE CLASSIFICATION FOR RESCUE OPERATION

Greeshma P[1], Amrutha Vishnupriya S[2]

1. P G scholar, Department of Computer Applications, Musaliar College of Engineering and Technology, Pathanamthitta
2. Assistant Professor, Department of Computer Science and Engineering, Musaliar College of Engineering and Technology, Pathanamthitta

Abstract:

Due to a growing number of people who carry out various adrenaline activities or adventure tourism and stay in the mountains and other inaccessible places, there is an increasing need to organize a search and rescue operation (SAR) to provide assistance and health care to the injured person. The goal of SAR operation is to search the largest area of the territory in the shortest time possible and find a lost or injured person. Because of the high speed and accuracy and the small number of false detections, the YOLOv4 detector was chosen for further examination. YOLOv4 model results related to different network sizes, different detection accuracies, and transfer learning settings were analysed. The model robustness to weather conditions and motion blur were also investigated. The model that can be used in SAR operations because of the excellent results in detecting people in search and rescue scenarios. Alone with person detection an efficient aerial image classification from on-board UAV for emergency response/monitoring applications is done. Specifically, a dedicated Aerial Image Database for Emergency Response applications is introduced and a comparative analysis of existing approaches is performed. This method classify five aerial images such as car crash, building collapse, fire, flood ,traffic accident and normal images using atrous convolution feature fusion method. Through this analysis a lightweight convolutional neural network architecture is proposed, referred to as EmergencyNet, based on atrous convolutions to process multiresolution features and capable of running efficiently on low-power embedded platforms achieving upto 20× higher performance compared to existing models with minimal memory requirements with less than 1% accuracy drop compared to state-of-the-art models.

Keywords :-. **Convolutional neural networks, person detection, YOLOv4,search and rescue operation, deep learning, emergency monitoring, image processing, remote sensing**

I. INTRODUCTION

Due to a developing wide variety of human beings residing and wearing out divers sports inside the mountains and other inaccessible locations, and because of the very nature of these activities and the physical and intellectual loss of preparedness for such sports, there is more and more injuries, fractures and diverse accidents consisting of slipping, burying, etc. risks that increase the insecurity of hikers, climbers, and

different adrenaline athletes are, in addition to the prevalence of harm or illness, their capabilities and enjoy in dealing with feasible emergencies. Emergencies can arise, for example, due to incorrect assessment of the distance of the destination, incorrect assessment of the difficulty of the road, due to changes in weather conditions, inadequate clothing or equipment, non-compliance with information and warnings, or insufficient preparation and overestimation of one's capabilities or knowledge. Reports of missing persons due to disorientation, illness, or suicidal intentions are also common.

To useful resource and health care to the injured in those situations,it's far essential to prepare a search and rescue operation. the quest motion refers to a state of affairs whilst the position and circumstance of the lacking man or woman are unknown, so the intention of the motion is to find the placement of the missing person in nature. The rescue operation refers to a situation in which it's miles known that it is important to intrude and arrange someone's rescue. If the accident's area is unknown earlier, this motion consists of search elements, too.

over the past few years unmanned aerial automobiles (UAVs)/drones have received sizeable hobby as a faraway sensing platform for diverse practical packages, such as traffic tracking, search and rescue , precision agriculture , and satellite imagery processing . current technological advances inclusive of the integration of camera sensors provide the possibility for brand new UAV applications including stumble on, monitor, and examine passive and energetic threats and risks at incident scenes (e.g., fireplace spots in forested areas, flooding risk, street collisions, and landslide inclined regions). similarly, due to their small size UAVs provide fast deployment and can as a consequence be in-the-loop of venture important choices to higher manage the available assets and improve risk assessment, prevention, and mitigation.

A. Relevance of the project

This software allows one to quickly detect a person who trapped in mountains and inaccessible places and provide immediate health care to the person. It is also possible to identify the disaster the has occurred in that place from the ariel image captured by the drone.

B Scope of the Project

As new methods emerge, this software will be quite valuable. Due to a growing number of people living and carrying out various activities in the mountains and other inaccessible places, and because of the very nature of these activities and the physical and mental lack of preparedness for such activities, there is an increasing number of injuries, fractures and various accidents such as slipping, burying, etc. In some cases it is difficult to find the place or disaster where the person is trapped. This program aids in the detecting the person and the disaster where he is trapped from the aerial image captured using the drone. In the future, this system will be used to detect the person and classifying which disaster from a live video capture.

II. EXISTING SYSTEM

Today most object detectors consist of two elements, the backbone of the detector as a CNN community trained to extract functions and a head that predicts the elegance and boundary container of the detected objects. Networks including VGG, ResNet, ResNeXt or MobileNet pre-skilled on the ImageNet or Open Images dataset, are maximum normally used as backbones. the pinnacle of a detector can be divided into types: one-level and two-stage detectors. YOLOv3, SSD and RetinaNet are examples of the one-stage detector. The most consultant two-degree detectors are R-CNN detectors, which include speedy R-CNN, faster R-CNN and, R-FCN. -stage detectors are commonly greater accurate in terms of localization and class accuracy. alternatively, they're slower in processing than one-stage detectors. Many detectors add extra layers between the spine and head (neck), e.g., characteristic Pyramid network (FPN) typically used to acquire a couple of characteristic maps, each with a exclusive decision, that's useful for spotting items at one of a kind scales. we've got compared the overall performanc the CNN-based totally detectors: quicker R-CNN, YOLOv4, RetinaNet, and Cascade R-CNN. All selected detectors had been formerly educated at the MS COCO dataset. All detector fashions are in addition educated on fowl's eye view photographs from a part of the Vis Drone and a SARD custom dataset to improve their performances.

LIMITATIONS:

- Comparatively low recall and more localization error
- Struggle to detect close objects because each grid can prose only 2 bounding boxes
- Struggles to detect small objects
- It depends on the time-consuming Selective Search Algorithm to generate region proposal.

III. PROPOSED SYSTEM

The YOLOv4 model was been selected for person detection in search and rescue operation by the proposed system and atrous convolution was used for the aerial image classification. The YOLOv4 model became decided on for similarly research due to accomplishing the best accuracy and detection velocity. to enhance the detection outcomes of the YOLOv4 version, we've got analyzed the impact of different internet paintings resolutions, detection accuracy, and switch getting to know settings on detection overall performance. The robustness of the YOLOv4 model to weather conditions and movement blur was additionally tested.

A dedicated Aerial Image Database for Emergency Response applications is introduced and a comparative analysis of existing approaches is performed. Through this analysis a lightweight convolutional neural network architecture is proposed, referred to as EmergencyNet, based on atrous convolutions to process multiresolution features and capable of running efficiently on low-power embedded platforms achieving upto 20× higher

performance compared to existing models with minimal memory requirements with less than 1% accuracy drop compared to state-of-the-art models.

The following modules make up the bulk of the proposed system:

1) Data preprocessing

- Loading dataset as data frame with images and corresponding labels,
- ImageDataGenerator is used for Pre-processing
- Resize images

2) Create CNN model

- Atrous Convolutional Feature Fusion Model is been created for the Classification.

3) Training

- Train Model using images and labels
- Save model

4) Prediction

- Choose image
- Load ACFF model
- Predict aerial image type

5) Person detection

- Pre-process the image
- Load yolo-v4 model
- Take detection and filter out person object
- Draw bounding box around person

IV. METHODOLOGY

c. Dataset Collection

The biggest challenge we faced during the experiments was to find a suitable and sufficient aerial image dataset. training a CNN for aerial image category for emergency reaction and disaster management applications first calls for gathering a suitable dataset for this mission. To the excellent of our know-how there's no broadly used and publicly available dataset for emergency reaction programs. As such, a devoted database for this task is built called AIDER (Aerial photograph Dataset for Emergency reaction applications). The dataset creation worried manually accumulating all pics for four catastrophe events, specifically fire/Smoke, Flood, Collapsed building/Rubble, and visitors injuries, in addition to one class for the normal case.

D. Image Preprocessing

Training the raw images as it is might lead to poor performance. Thus, simple image processing algorithms can be implemented to achieve maximum accuracy. Image processing algorithms such ImageDataGenerator is used to reduce the training time and power consumption. The ImageDataGenerator function is used for resizing, rescaling, adjust the height, width and length. The noise from the images can be eliminated. Positions of persons in the images range from standard (standing position, sitting, lying, walking, running)

to positions typical of exhausted or injured persons reconstructed by actors at their discretion.

E. Training

All the networks are evolved and examined through the same education framework so as to have the same situations and a honest assessment during the inference phase. The Keras deep studying framework is used which has available all fashions with Tensorflow walking because the backend. The same photograph size is used for all networks where feasible (except for the mobileNet V1 and V2 models which mainly require a smaller image length). consequently, earlier than augmenting and including an picture to the batch it's miles first resized to the correct photo size depending at the community (default is 240 × 240 pixels which is an ordinary length for education CNNs). It need to be noted that it's far possible to use large photo sizes at a cost of slower inference time, however, on this paintings the picture length space isn't explored but alternatively focus is at the community design.

F. YOLOv4 Model

YOLO is short for You Only Look Once. It is a real-time object recognition system that can recognize multiple objects in a single frame. YOLO recognizes objects more precisely and faster than other recognition systems. It can predict up to 9000 classes and even unseen classes. The real-time recognition system will recognize multiple objects from an image and also make a boundary box around the object. It can be easily trained and deployed in a production system.YoloV4 is an important improvement of YoloV3, the implementation of a new architecture in the backbone and the modifications in the neck have improved the mAP**(** mean Average Precision) by 10% and the number of FPS(Frame per Second) by In addition, it has become easier to train this neural network on a single GPU. The main objective of the backbone is to extract the essential features**,** the selection of the backbone is a key step it will improve the performance of object detection. Often pre-trained neural network are used to train the backbone. The essential role of the neck is to collect feature maps from different stages**.** Usually, a neck is composed of several bottom-up paths and several top-down paths. The role of the head in the case of a one stage detector is to perform dense prediction**.** The dense prediction is the final prediction which is composed of a vector containing the coordinates of the predicted bounding box (center, height, width), the confidence score of the prediction and the label**.**

G. Atrous Convolution Feature Fusion Model

Atrous (also called dilated) convolutions can capture and transform images at different resolutions depending on the dilation rate which determines the spacing between the kernel points, effectively increasing their receptive field without increasing the parameter count. Hence, it can be used to incorporate larger context to the model. The proposed block computes multiple such atrous convolutional features (Ud) for the same input map across

different dilation rates d. Each atrous convolution is factored into depth-wise convolution that performs light-weight filtering by applying a single convolutional kernel per input channel to reduce the computational complexity. Then, some form of fusion takes place to merge the different features together. The intuition is to take advantage of the different dilation rates since one path may peek up features that another may have missed due to changes in object/region resolution. It is important to note that the weights are not shared between paths and each learns different weights wd that may be more useful. Another advantage of using atrous convolutions stems from the fact that the same number of parameters and computations are needed regardless of the resolution. Each atrous convolution acts on the same feature map but at a different spatial resolution; starting from a dilation rate of 1 and a filter size of 3 × 3 (i.e., no spacing) and going up to 3 which is equivalent to 7 × 7 receptive field effectively.

V. SYSTEM ARCHITECTURE

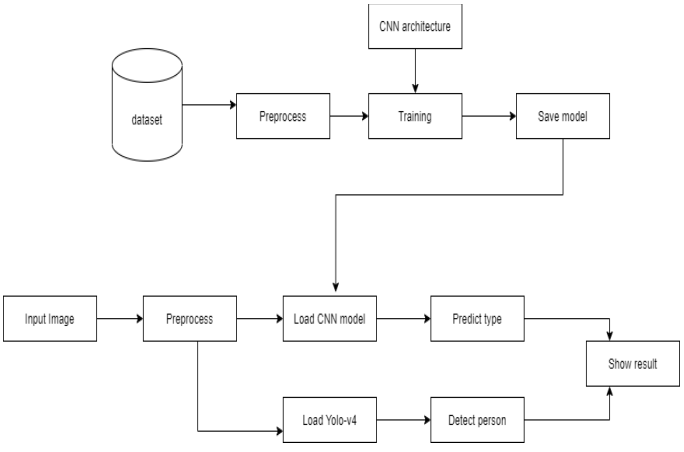

Fig 1. System architecture of the model

VI. LITERATURE REVIEW

H. Faster R-Cnn: Towards Real-Time Object Detection With Region Proposal Networks (Shaoqing Ren, Kaiming He, Ross Girshick, And Jian)2015

An RPN is a completely convolutional community that concurrently predicts object bounds and objectness rankings at each role. The RPN is skilled quit-to-stop to generate location proposals, that are utilized by fast R-CNN for detection. We in addition merge RPN and fast R-CNN right into a unmarried network by sharing their convolutional capabilities—the use of the these days famous terminology today's neural networks with 'interest' mechanisms, the RPN issue tells the unified community wherein to appearance. For the very deep VGG-sixteen version , our detection device has a frame rate modern day 5fps (including all steps) on a GPU, whilst reaching object detection accuracy on PASCAL VOC 2007, 2012, and MS COCO datasets with simplest 300 proposals consistent with photograph. In ILSVRC and COCO 2015 competitions, quicker R-CNN and RPN are the rules modern day the 1st-place prevailing entries in numerous tracks. Code has been made publicly to be had.

I. Cascade R-Cnn: Delving Into High Quality Object Detection (Zhaowei Cai Uc San Diego, Nuno Vasconcelos Uc San Diego)2018

A multi-degree object detection architecture, the Cascade R-CNN, is

proposed to deal with these issues. It includes a chain of detectors trained with growing thresholds, to be sequentially extra selective in opposition to close fake positives. The detectors are educated degree by way of stage, leveraging the remark that the output of a detector is a good distribution for schooling the following better pleasant detector. The resampling of steadily progressed hypotheses ensures that every one detectors have a effective set of examples of equivalent size, reducing the overfitting hassle. The equal cascade procedure is applied at inference, permitting a closer fit among the hypotheses and the detector exceptional of each stage. A simple implementation of the Cascade R-CNN is proven to surpass all single-version object detectors at the challenging COCO dataset. Experiments additionally show that the Cascade R-CNN is extensively relevant throughout detector architectures, attaining steady profits independently of the baseline detector power.

J. Thermal Object Detection In Difficult Weather Conditions Using Yolov3(Mate Krišto, Marina Ivasic-Kos, (Member, Ieee), And Miran Pobar)2020

The performance modern the standard item detectors inclusive of faster R-CNN, SSD, Cascade R-CNN, and YOLOv3, that were retrained on a dataset trendy thermal pictures extracted from videos that simulate illegal actions across the border and in included regions. videos are recorded at night time in clear climate, rain, and in the fog, at exceptional tiers, and with exceptional movement kinds. YOLOv3 became appreciably faster than different detectors even as attaining overall performance similar with the first-rate, so it become utilized in in addition experiments. We experimented with special education dataset settings with a purpose to decide the minimum wide variety latest pictures needed to acquire properly detection effects on take a look at datasets. We accomplished first-rate detection effects with appreciate to common accuracy for all check eventualities despite the fact that a modest set modern thermal photographs become used for education.

K. Deep-Learning-Based Aerial Image Classification For Emergency Response Applications Using Unmanned Aerial Vehicles(Christos Kyrkou)2020

This Method focuses on the automated aerial scene class of disaster occasions from on-board a UAV. in particular, a committed Aerial photograph Database for Emergency response (AIDER) applications is added and a comparative analysis of existing strategies is carried out. thru this evaluation a light-weight convolutional neural community (CNN) structure is evolved, capable of walking correctly on an embedded platform achieving ~ three× better overall performance as compared to existing fashions with minimum reminiscence requirements with much less than 2% accuracy drop in comparison to the nation-ofthe-art. these initial consequences provide a solid basis for further experimentation toward actual-time aerial photograph classification for emergency response applications the usage of UAVs.

an embedded platform and deep mastering UAVs can autonomously reveal a catastrophe troubled region, analyze the image in real-time.

VII. CONCLUSION

The capacity to discover people on drone pix the usage of laptop vision methods automatically is a good-sized help in SAR operations. YOLOv4 has executed the high-quality detection performances on the SARD dataset in terms of common precision (AP) considering IoU precision and the object length in addition to the least false detection (FP), so it changed into similarly used in the test, known as YOLOv4 (SARD). when the version become educated on 512 × 512 picture resolution, the first-class AP of 60% changed into finished for a network resolution of 832 × 832 In attempting to find a missing person, the most critical component is that the detector locates that person, and it's far much less critical how accurate the detection is. We experimentally decided on parameters as a exchange-off between accuracy and keep in mind in order that the version can be helpful in SAR actions. The results confirmed that the YOLOv4 (SARD) model in a community decision of 832 × 832, IoU = 0.1, done the excellent outcomes for thresh of zero.4, namely AP of ninety-seven.15% (TP: 2538, FP: 46). An analysis at the design and implementation of an efficient deep gaining knowledge of machine has been carried out to routinely recognize and classify catastrophe activities in actual-time from on-board a UAV. The proposed solution affords an adequate tradeoff between accuracy, inference velocity, and complexity and that it may be used as a building block in the direction of use-instances with similar constraints

REFERENCE

1. S. Ren, K. He, R. Girshick, and J. Sun, ‘‘Faster R-CNN: Towards real-time object detection with region proposal networks,’’ in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 91–99.

2. Automatic Person Detection in Search and Rescue Operations Using Deep CNN Detectors, Sasa Sambolek1 And Marina Ivasic-Kos,Member IEEE, Feb 2021.

3. EmergencyNet: Efficient Aerial Image Classification for Drone-Based Emergency Monitoring Using Atrous Convolutional Feature Fusion, Christos Kyrkou , Member, IEEE, and Theocharis Theocharides , Senior Member, IEEE ,2020.

4. Z. Cai and N. Vasconcelos, ‘‘Cascade R-CNN: Delving into high-quality object detection,’’ in Proc. IEEE Conf. Comput. Vis. pattern Recognit., Jun. 2018, pp. 6154–6162.

5. M. Kristo, M. Ivasic-Kos, and M. Pobar, ‘‘Thermal object detection in difficult weather conditions using YOLO,’’ IEEE Access, vol. 8, pp. 125459–125476, 2020.

