CS6811 – Project Work

Zeroth Review

UAV-based post-disaster scene understanding using a hybrid single-multi-stage ensemble network with GAN-aided semantic segmentation

Team Number: 17 (Batch 1)

Project Guide: Dr. R. Gunasekaran

Team Members:

Abhishek Manoharan	Nishanthini S	Vamsi Raju M
2019503502	2019503537	2019503568

Domains: Unmanned Aerial Vehicles, Computer Vision, Scene Understanding

Problem Statement:

Post-disaster scene understanding frameworks are becoming increasingly crucial in search and rescue operations and damage assessment initiatives. The use of Unmanned Aerial Vehicles (UAVs) provides an efficient method to complete the task of scene understanding. However, complex environments present in post-disaster scenarios make it difficult for UAVs to detect humans or objects accurately. Moreover, inefficient object detection mechanisms lead to low accuracy and a long time for object detection tasks. Hence, to mitigate these issues, we propose a UAV-based scene understanding scheme involving a GAN-aided semantic segmentation mechanism. This approach classifies objects present in the visual scope of the UAV using a 3D reconstruction from thermal images of the scene and pixel-level prediction. Furthermore, an ensemble network consisting of a combination of single-stage and multi-stage detectors is to be used to improve the performance of the detection model. This will help to reduce false positives and improve the overall accuracy of the system.

Objectives:

- To develop an efficient post-disaster scene understanding framework using UAVs for search and rescue operations.
- To implement a hybrid single-stage and multi-stage ensemble network comprising of the CenterNet and Cascade R-CNN mechanisms to combine the benefits of both, thereby decreasing the high false negative rate of

multi-stage mechanisms and improving the performance of single-stage detectors.

- ❖ To devise a 3D scene-reconstruction mechanism using thermal images obtained from a swarm of UAVs to map and extract useful information from the scene.
- ❖ To deploy a Generative Adversarial Network (GAN)-aided semantic segmentation framework to improve the detection of small and dense objects in post-disaster conditions. A GAN denoiser results in images having lower occlusion and optimal brightness, thereby highlighting the important features of the object. Semantic segmentation on these images leads to a pixel-level prediction of various entities or objects in the image.

Literature Survey:

Authors	Title	Proposed Work	Limitations
Jiong Dong , Kaoru Ota, and Mianxiong Dong	**UAV-Based Real-Time Survivor Detection System in Post-Disaster Search and Rescue Operations**	This paper proposes a new thermal image dataset consisting of 6447 thermal images designed for survivor detection using UAVs in post-disaster scenarios. The paper also describes optimal values to prune survivor detection models in order to reduce the complexity of the models and applies knowledge distillation techniques to fine-tune them and improve accuracy. The performance of several survivor detection models based on YOLOv3 and YOLOv3-MobileNetV1 were compared with and without pruning and fine-tuning.	Older and inferior detection models have been used for survivor detection, thereby resulting in models with high mean average precision (mAP) loss and low accuracy.
Hang Ren, Zhichao Sun , Jianyu Yang , Yuping Xiao, Hongyang An,	**Swarm UAV SAR for 3-D Imaging**	This paper implements a 3D imaging mechanism for 2D images obtained from a swarm of UAVs. The proposed work involves the	A considerable amount of data has to be transmitted from the UAV swarm, as images obtained

Zhongyu Li , and Junjie Wu		3D imaging of a scene by the usage of 2D images obtained from several UAVs present in the UAV Swarm at different perspectives with a few points of overlap. The point cloud obtained is then triangulated, and Bundle Adjustment is used to create the 3D rendering of the image.	from each node in the swarm are used to produce the 3D rendering. Multiple UAVs also need to exchange information in order to efficiently collect data of the scenario.
T. C. Bybee and S. E. Budge	**Method for 3-D Scene Reconstruction Using Fused LiDAR and Imagery From a Texel Camera**	In order to create greater fidelity terrain models, this study describes a bundle adjustment technique for aerial texel images that enables relatively low-accuracy navigation systems to be employed with inexpensive LiDAR and camera data.	Outliers present in the point cloud are not identified and mitigated, thereby leading to lower accuracy.
Albaba, Berat Mert, and Sedat Ozer	**SyNet: An ensemble network for object detection in UAV images**	With the goal of lowering the high false negative rate of multi-stage detectors and improving the quality of the single-stage detector proposals, the authors of this research propose an ensemble network called SyNet that combines a multi-stage method with a single-stage one.	According to the investigation, detecting objects in drone images is more challenging than detecting them in images that were taken from the ground, even with the most advanced object detection algorithms. Hence, the accuracy of the model trained on UAV images is still low compared to models trained on ground images.

A. Bouguettaya, H. Zarzour, A. Kechida and A. M. Taberkit,	**Vehicle Detection From UAV Imagery With Deep Learning**	The article provides a review of vehicle detection from UAV imagery using deep learning techniques. It begins by outlining the various deep learning architectures, including generative adversarial networks, autoencoders, recurrent neural networks, and convolutional neural networks, and their contributions to the challenge of improving vehicle detection. The paper then focuses on examining various vehicle detection techniques and presents different benchmark datasets and problems that have been discovered, along with possible remedies.	Videos captured in the UAVs are sent to on-ground workstations or to the cloud for processing rather than being implemented on the UAV itself, thereby leading to the absence of a lightweight system for vehicle detection.
M. Rahnemoonfar, Maryam, Tashnim Chowdhury, and Robin Murphy	**RescueNet: A High-Resolution Post Disaster UAV Dataset for Semantic Segmentation**	This paper introduces a high-resolution post-disaster UAV dataset named RescueNet, which contains comprehensive pixel-level annotation of 11 classes for semantic segmentation to assess damage after a natural disaster. The dataset collection and annotation process are discussed, along with the challenges it poses. Four state-of-art semantic segmentation methods have been evaluated on RescueNet, and the results are discussed.	RescueNet contains a small number of classes. As a result, smaller objects like “vehicles” and “pools” make it difficult to get a good segmentation compared to larger objects like buildings and roads. Besides that, since UAV images include only the top view of a scene, it is difficult to assess the actual damage since the

			horizontal view also brings information regarding all sides of a building.
T. Chowdhury and M. Rahnemoonfar	**Attention-Based Semantic Segmentation on UAV Dataset for Natural Disaster Damage Assessment**	This paper proposes and evaluates a novel self-attention segmentation model named ReDNet on a new high-resolution UAV natural disaster dataset named HRUD. The challenges of semantic segmentation on the HRUD dataset are discussed, along with the excellent performance of the proposed model.	HRUD is a very challenging dataset due to its variable-sized classes along with similar textures among different classes. Debris, textures of debris, sand, and building with total destruction damage make a great impact on the segmentation performance of the evaluated network models.
Li, Tianjiao, Jun Liu, Wei Zhang, Yun Ni, Wenqian Wang, and Zhiheng Li	**Uav-human: A large benchmark for human behavior understanding with unmanned aerial vehicles**	This paper proposes a UAV-Human dataset for human action, pose, and behaviour understanding. The proposed UAV-Human contains 67,428 multi-modal video sequences and 119 subjects for action recognition, 22,476 frames for pose estimation, 41,290 frames and 1,144 identities for person re-identification, and 22,263 frames for attribute recognition which encourages the exploration and deployment of various data-intensive learning models for UAV-based human behaviour understanding.	The UAV-Human dataset poses a limitation for attribute recognition because the dataset is captured over a relatively long period of time. As a result, the subjects have been diversified with different dressing types and large variations of viewpoints caused by multiple UAV altitudes.

References:

1. J. Dong, K. Ota and M. Dong, "**UAV-Based Real-Time Survivor Detection System in Post-Disaster Search and Rescue Operations**," in *IEEE Journal on Miniaturization for Air and Space Systems*, vol. 2, no. 4, pp. 209-219, 2021
2. H. Ren et al., "**Swarm UAV SAR for 3-D Imaging: System Analysis and Sensing Matrix Design**," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-16, 2022
3. T. C. Bybee and S. E. Budge, "**Method for 3-D Scene Reconstruction Using Fused LiDAR and Imagery From a Texel Camera**," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 11, pp. 8879-8889, Nov. 2019
4. Albaba, Berat Mert, and Sedat Ozer, "**SyNet: An ensemble network for object detection in UAV images**," in 25th IEEE International Conference on Pattern Recognition (ICPR), pp. 10227-10234, 2021
5. A. Bouguettaya, H. Zarzour, A. Kechida and A. M. Taberkit, "**Vehicle Detection From UAV Imagery With Deep Learning: A Review**," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6047-6067, Nov. 2022
6. Rahnemoonfar, Maryam, Tashnim Chowdhury, and Robin Murphy. "**RescueNet: A High Resolution Post Disaster UAV Dataset for Semantic Segmentation**." *UMBC Student Collection*, 2021
7. T. Chowdhury and M. Rahnemoonfar, "**Attention Based Semantic Segmentation on UAV Dataset for Natural Disaster Damage Assessment**," *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 2325-2328, 2021
8. Li, Tianjiao, Jun Liu, Wei Zhang, Yun Ni, Wenqian Wang, and Zhiheng Li. "**Uav-human: A large benchmark for human behavior understanding with unmanned aerial vehicles**." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 16266-16275, 2021