CS6811 – Project Work

First Review

UAV-based post-disaster 3D scene reconstruction for efficient survivor detection

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, 3D Reconstruction

Problem Definition:

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 3D reconstruction mechanism. This approach deploys a Generative Adversarial Network (GAN)-based model to denoise and remove occlusion in the images obtained from the UAVs. The framework classifies objects present in the visual scope of the UAV using a 3D reconstruction of the images obtained from the UAV, followed by semantic segmentation, resulting in pixel-level prediction and classification of entities present in the 3D model. 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 survivor detection model. This will help reduce the false negative rate and improve the overall accuracy of the system.

Literature Survey:

S.No	Publication Venue and Year	Title	Proposed Work	Limitations
1.	*IEEE Journal on Miniaturization for Air and Space Systems*, vol. 2, no. 4, pp. 209-219 **(2021)**	**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. ❖ The model 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.
2.	*IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-16 **(2022)**	**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 3D imaging of a scene by the usage of 2D images obtained from several UAVs	❖ A considerable amount of data must be transmitted from the UAV swarm, as images obtained from each node in

			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.	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.
3.	*IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8879-8889 **(2019)**	**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. ❖ The model 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.
4.	*25th IEEE International Conference on Pattern Recognition (ICPR)*, pp. 10227-10234 **(2021)**	**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.
5.	*IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6047-6067 **(2022)**	**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 ➢ convolutional neural networks and their contributions to the challenge of improving vehicle detection. ❖ The paper then focuses on examining various vehicle detection techniques and	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.

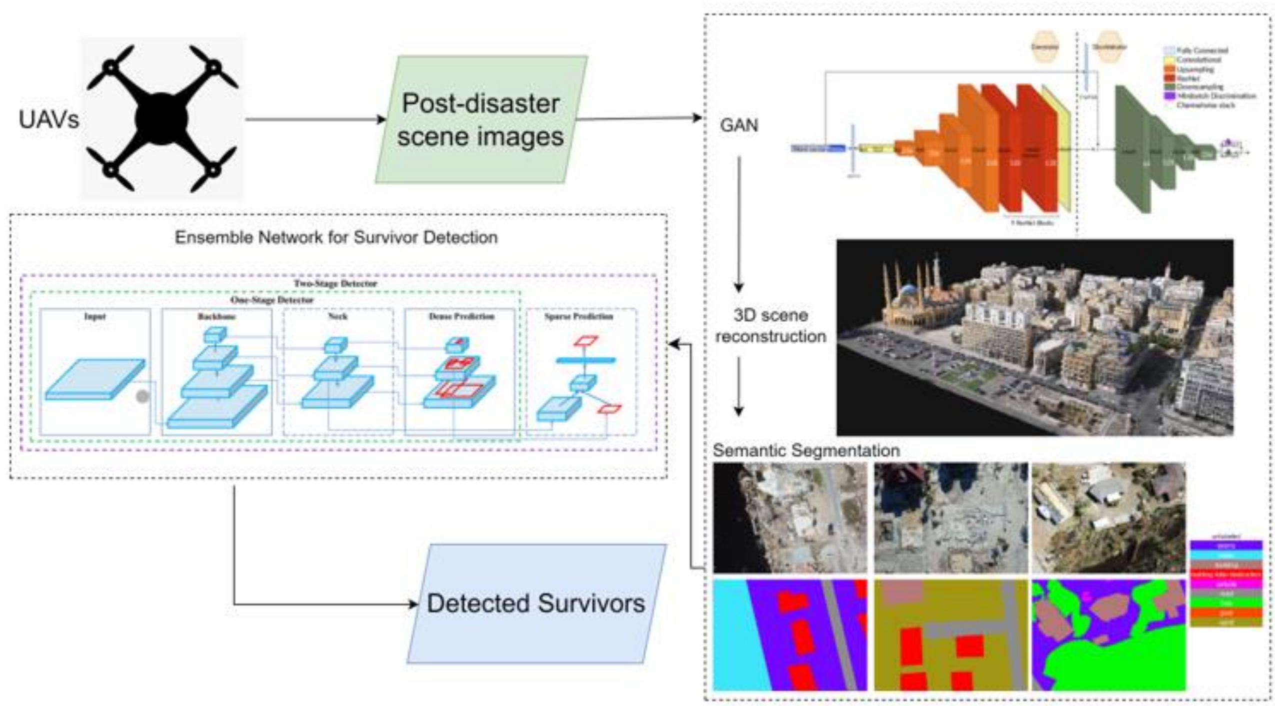
			presents different benchmark datasets and problems that have been discovered, along with possible remedies.	
6.	*IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-13 **(2022)**	**Dense and Small Object Detection in UAV-Vision Based on a Global-Local Feature Enhanced Network**	❖ This paper proposes a global-local feature-enhanced network (GLF-Net) to alleviate issues when detecting small and dense objects using UAVs. ❖ A feature-fusion module has been proposed to tackle the presence of numerous small objects. ❖ GLF-Net achieves 86.52% mean Average Precision (mAP) on the RO-UAV dataset.	The scalability of the framework is poor, and the application of GLF-Net on post-disaster UAV images leads to lower mAP, thereby requiring better frameworks.
7.	*IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1223-1232 **(2021)**	**Unmanned aerial vehicle visual detection and tracking using deep neural networks: A performance benchmark**	❖ This paper executes and compares various UAV detection mechanisms using air-borne UAVs that deploy deep neural networks. ❖ 4 datasets have been used and performance has been compared, namely MAV-VID, Drone-vs-Bird, Anti-UAV RGB, and Anti-UAV IR. ❖ The performance of 4 models was compared using the datasets mentioned, namely Faster RCNN, SSD512, YOLOv3, and DETR	❖ Long-distance detection of small UAVs was not taken into consideration. ❖ Deep neural networks for re-identification of UAVs were not considered as well.

			(Detection Transformer). Overall, Faster RCNN performed best.	
8.	*UMBC Student Collection* **(2021)**	**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.	❖ 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.
9.	IEEE/CVF Conference on Computer	**Uav-human: A large benchmark for**	❖ This paper proposes a UAV-Human dataset for human action, pose,	❖ The UAV-Human dataset poses

	Vision and Pattern Recognition (CVPR), pp. 16266-16275 **(2021)**	**human behavior understanding with unmanned aerial vehicles**	and behavior understanding. ❖ The proposed UAV-Human contains ➢ 67,428 multi-modal video sequences ➢ 119 subjects for action recognition ➢ 22,476 frames for pose estimation ➢ 41,290 frames ➢ 1,144 identities for person re-identification ➢ 22,263 frames for attribute recognition which encourages the exploration and deployment of various data-intensive learning models for UAV-based human behavior understanding.	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.
10.	*2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 2325-2328 **(2021)**	**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

				destruction damage make a great impact on the segmentation performance of the evaluated network models.

System Architecture:



Novelty:

A GAN denoiser results in images having lower occlusion and optimal brightness, thereby highlighting the important features of the object. Furthermore, using GAN improves the detection of small and dense objects, which is the case of survivors in images obtained from a UAV. A 3D reconstruction of the scene using the enhanced images obtained from the GAN-based model will be used to map and extract useful information from the scene. The GAN-based 3D reconstruction framework incorporating the Structure-From-Motion (SFM) and bundle adjustment algorithms enhances occluded survivors in the 3D model, thereby resulting in a more efficient survivor detection mechanism compared to existing frameworks. Semantic segmentation on the 3D model leads to a pixel-level prediction of various entities or objects present in the image. The ensemble model, a hybrid architecture consisting of single-stage and multi-stage detectors, overcomes the disadvantages of both frameworks. Deploying an ensemble network comprising the CenterNet and Cascade R-CNN frameworks improves the performance and efficiency of survivor detection. The overall framework will increase the accuracy and performance of the survivor detection task, thereby resulting in efficient Search-And-Rescue operations.

Expected Outcomes:

- To develop an efficient post-disaster scene understanding framework using UAVs for survivor detection and Search-And-Rescue (SAR) operations.
- To deploy a Generative Adversarial Network (GAN) framework to improve the detection of survivors, who are generally present as small and dense objects in post-disaster UAV images. A GAN denoiser will result in images having lower occlusion and optimal brightness, thereby highlighting the important features of the object.
- To devise a 3D scene-reconstruction mechanism based on the Structure-From-Motion (SFM) and bundle adjustment algorithms using images obtained from a swarm of UAVs to map and extract useful information from the scene.
- To deploy a semantic segmentation mechanism on the 3D model, leading to a pixel-level prediction of various entities or objects in the 3D model. This will improve the detection of survivors present in the post-disaster scene.
- To implement a hybrid single-stage and multi-stage ensemble network for survivor detection on the previously obtained semantic entities. The model will comprise 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.

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. T. Ye, W. Qin, Y. Li, S. Wang, J. Zhang and Z. Zhao, "**Dense and Small Object Detection in UAV-Vision Based on a Global-Local Feature Enhanced Network**," in *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-13, 2022
7. Rahnemoonfar, Maryam, Tashnim Chowdhury, and Robin Murphy. "**RescueNet: A High-Resolution Post Disaster UAV Dataset for Semantic Segmentation**." *UMBC Student Collection*, 2021
8. 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
9. 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
10. Isaac-Medina, Brian KS, Matt Poyser, Daniel Organisciak, Chris G. Willcocks, Toby P. Breckon, and Hubert PH Shum. "**Unmanned aerial**

vehicle visual detection and tracking using deep neural networks: A performance benchmark." In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1223-1232, 2021.