Next-Phase Project Proposal

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

Project Guide: Dr. R. Gunasekaran

Team Members:

Abhishek Manoharan	Nishanthini S	Vamsi Raju M
2019503502	2019503537	2019503568

Problem Statement:

Post-disaster scene understanding frameworks are becoming increasingly crucial in search and rescue operations and damage assessment initiatives. The use of 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 that 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.

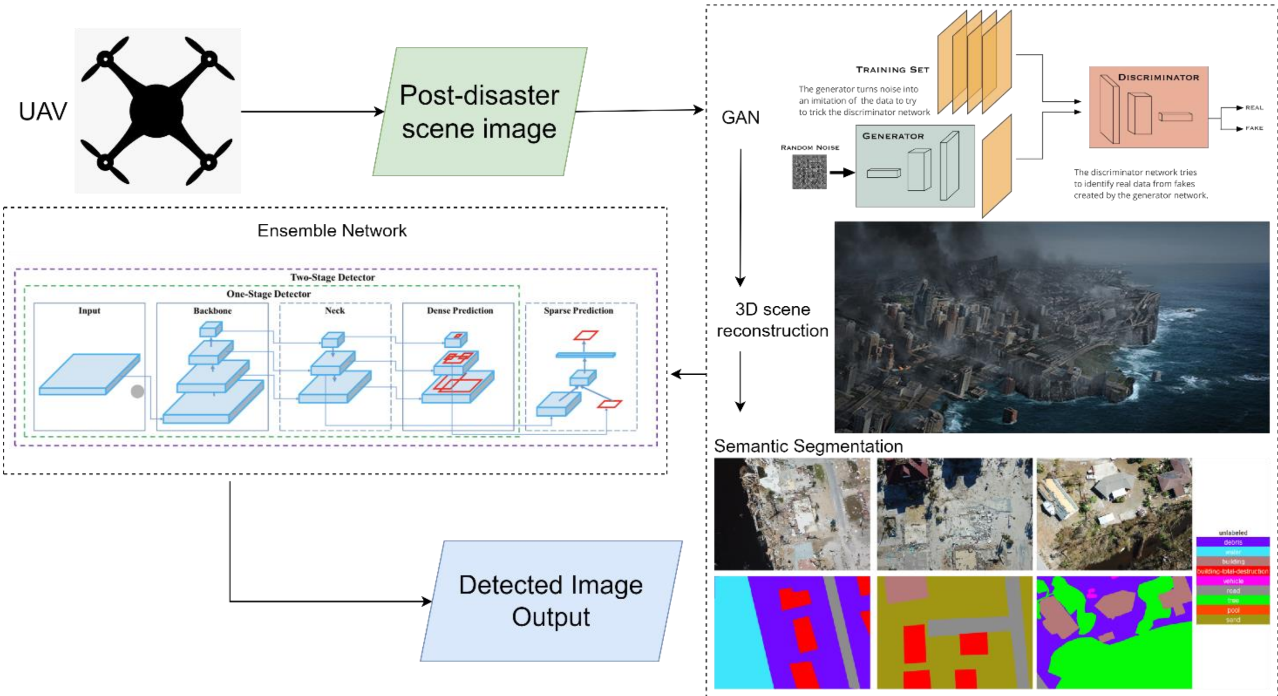
Identified CV mechanisms for the scene-understanding task:

An ensemble network comprising of the CenterNet and Cascade R-CNN mechanisms is to be deployed to overcome the limitations faced by current scene understanding topologies.

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 and decrease the high false negative rate of multi-stage mechanisms, improving the performance of single-stage detectors.
- To devise a 3D scene-reconstruction mechanism using thermal images obtained from the UAV 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 dense and small objects in post-disaster conditions.

High-level Architecture:



Novelty:

The implementation of a combined GAN and Semantic Segmentation pre-processing mechanism is used for the images passed as input to the ensemble model. 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. A 3D reconstruction of the scene using thermal images obtained from a Thermal Infrared Radiation (TIR) sensor fitted on the UAV will map and extract useful information from the scene. The ensemble model, being a hybrid architecture consisting of single-stage and multi-stage detectors, overcomes the disadvantages of both frameworks. Deploying an ensemble network comprising of CenterNet and Cascade R-CNN frameworks improves the performance and efficiency of the scene understanding task. The overall framework will increase the accuracy and performance of the system.

Modules:

Module 1: Data collection, initial image pre-processing techniques, and data augmentation.

Module 2: Image Processing module consisting of the GAN + Sematic segmentation framework and 3D scene reconstruction mechanism.

Module 3: Selection of optimal single-stage and multi-stage detection frameworks. Implementation of the hybrid single-stage and multi-stage object detection mechanism.

Module 4: Simulation after integration of both mechanisms, real-time simulation using a web app.

Mathematical Modelling Scope:

Algorithm 1 High-Level Mathematical Modelling

Input: Images (λ) of post-disaster scene obtained from UAV
Output: Minimized Loss of Ensemble Network (L_N)

procedure GAN_DENOISING(λ)
    Pre-processing of UAV images λ
    $\lambda_{\text{epochs}} \leftarrow$ Number of iterations to train GAN
    $\lambda_{\text{batchsize}} \leftarrow$ Number of images to train per epoch
    for β `in` λ **do**
        $\vartheta \leftarrow TrainGAN(\beta, \lambda_{\text{epochs}}, \lambda_{\text{batchsize}})$
    end for
    $\varpi \leftarrow GAN\_denoiser(\vartheta)$
    return ϖ
end procedure
procedure SEMANTIC_SEGMENTATION(ϖ)
    $\Phi[] \leftarrow$ entities present in ϖ
    $\nu \leftarrow 0$
    for α `in` ϖ **do**
        $\Phi[\nu] = Classify(\alpha)$
        $\nu = \nu + 1$
    end for
    return Φ
end procedure
procedure ENSEMBLE_NETWORK(Φ)
    $L_C \leftarrow$ Loss of Cascade RCNN
    $L_N \leftarrow$ Loss of CenterNet
    $y_n \leftarrow$ actual label of entity x_n
    $y \leftarrow$ predicted label of entity x_n
    $p(\psi) \leftarrow$ probability of occurrence of ψ
    $t_c \leftarrow$ true width of entity x_n
    for x_n `in` Φ **do**
        $L_C \leftarrow -log(p(y = y_n|x_n))$
        Applying Ensemble Loss Minimization:
        $L_N(L_C) \leftarrow -log(|y - y_n t_c|)$
    end for
    return L_N
end procedure

Tentative Implementation Timeline:

Task	Tentative Timeline
Data collection and initial tuning	Jan 23 to Feb 10
GAN + Semantic segmentation framework, 3D scene reconstruction mechanism	Feb 10 to Feb 25
Selection of optimal detection frameworks	Feb 25 to Mar 5
Implementation of a hybrid mechanism of the chosen single-stage and multi-stage detection frameworks	Mar 5 to Mar 20
Simulation of the system after integration of modules, performance testing, and evaluation	Mar 20 to Mar 30

References:

1. 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
2. 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
3. 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
4. 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

5. 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
6. M. Żarski, B. Wójcik, J. A. Miszczak, B. Blachowski and M. Ostrowski, "**Computer Vision Based Inspection on Post-Earthquake With UAV Synthetic Dataset**," in *IEEE Access*, vol. 10, pp. 108134-108144, 2022
7. 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.
8. 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
9. M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari and R. R. Murphy, "**FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding**," in *IEEE Access*, vol. 9, pp. 89644-89654, 2021
10. Rahnemoonfar, Maryam, Tashnim Chowdhury, and Robin Murphy. "**RescueNet: A High Resolution Post Disaster UAV Dataset for Semantic Segmentation**." *UMBC Student Collection* (2021).
11. 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
12. 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