Topic: Multimodal Learning
Project Title: Multimodal Learning for Detecting and Tracking Unmanned Aerial Vehicle
Description of the Project:
With the increasing utilization of Unmanned Aerial Vehicles (UAVs) across various sectors, the imperative for effective long-distance detection and tracking is evident. Radar is the mostly used technology in detecting UAVs. However, traditional radar systems are inadequate for detecting small UAVs due to their focus on larger airborne objects, requiring additional infrastructure and specialized equipment, which can be costly and may not be feasible for all setups, such as detecting UAVs from another UAV. While acoustic sensors can detect sound waves generated by UAVs, their range is typically shorter compared to other detection methods such as radar or optical sensors. Passive radio frequency (RF) signals emitted by UAVs present a promising avenue for detection, as they offer a non-intrusive and potentially cost-effective solution. However, their effectiveness can be impeded in urban environments characterized by various structures that may cause signal interference. Additionally, the presence of various electromagnetic sources in urban areas, such as Wi-Fi networks, cellular signals, and electronic devices, can further complicate RF detection by introducing background noise and interference. Furthermore, although it can detect UAVs, it is hard to localize the target using this technology. RGB vision camera can detect and localize the position of the UAVs yet are hindered by limitations in nighttime and low-light conditions and it is also needed to detect the target. Infrared cameras present a promising alternative to address the limitations of RGB cameras. However, they may face challenges in distinguishing between thermal emissions emitted by target objects and those emitted by surrounding obstacles. Therefore, there is a necessity to develop a solution capable of detecting and tracking small UAVs over long distances in various environmental conditions, including low lighting conditions, and potential interference, while remaining affordable and providing real-time functionality. Combining RGB and infrared cameras with a passive RF signal can offer a comprehensive solution to the challenges of UAV detection and tracking. Multimodal machine learning can be used to solve the issue. However, it is crucial to ensure that each sensor is individually effective in detecting UAVs before being integrated. It is also important to synchronize the data obtained from the sensors otherwise the extracted features will mislead the model. By integrating these sensors, a more robust and reliable system can be created to address all aspects of UAV detection and tracking, enhancing overall performance and accuracy.

Task for the Project:

1. **Overview of the Multimodal Learning:** Exploring contemporary algorithms and architectures.
2. **Requirement Analysis:** Assessing project objectives to determine model functional domain.
3. **Challenge Analysis:** Evaluating task difficulties and model limitations.
4. **Performance Metrics:** Establishing metrics to evaluate system suitability for identified requirements.
5. **Formulation of the pipeline:**
    - Develop appropriate algorithms for preprocessing sensor data and extracting features from each sensor.
    - Integrate the extracted features from multiple sensors.
    - Transmit the combined features to a network for drone detection and localization.
    - Fine-tuning the detection model for optimal performance.
6. **Validation:** Assess the soundness of the implemented system through the defined performance metrics and performing an ablation test of the system components.

Skills Needed:

1. Experience in Python programming, and Conda (anaconda or miniconda) environment.
2. Knowledge of deep learning and Convolution Neural Network.
3. Knowledge of image processing and signal processing is preferred.

Materials to be read:

[1] Jiang, N., Wang, K., Peng, X., Yu, X., Wang, Q., Xing, J., ... & Han, Z. Anti-UAV: A large multi-modal benchmark for UAV tracking. arXiv 2021. *arXiv preprint arXiv:2101.08466*.

[2] McCoy, J., Rawal, A., Rawat, D. B., & Sadler, B. M. (2022). Ensemble deep learning for sustainable multimodal uav classification. *IEEE Transactions on Intelligent Transportation Systems*.

[3] Svanström, F., Alonso-Fernandez, F., & Englund, C. (2021). A dataset for multi-sensor drone detection. *Data in Brief*, *39*, 107521.

[4] Jabeen, S., Li, X., Amin, M. S., Bourahla, O., Li, S., & Jabbar, A. (2023). A review on methods and applications in multimodal deep learning. *ACM Transactions on Multimedia Computing, Communications and Applications*, *19*(2s), 1-41.