# Object Classification for Drone Imagery using Attention Mechanism

Name of the Supervisor : Dr. Binod Kumar

Kethireddy Harshith Reddy (B20AI018)
Naikawadi Shivaprasad (B20AI057)

# Introduction

- Drone imagery has become a crucial source of data for many applications, ranging from agriculture and forestry to surveillance and disaster response.
- Object classification is one of the most important tasks in drone imagery analysis, as it allows us to identify and classify objects of interest in the imagery.
- In this presentation, we will explore the use of an attention mechanism for object classification in drone imagery, which can improve both the accuracy and efficiency of the classification process.

# Motivation

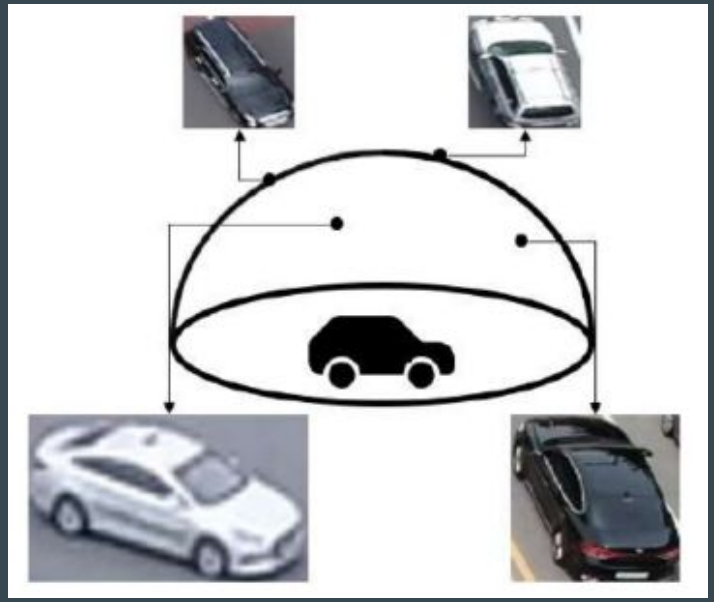
- Attention mechanisms can help to improve the accuracy and efficiency of object classification by focusing on the most relevant parts of the image, rather than processing the entire image.
- In the context of drone imagery, attention mechanisms can be particularly useful for identifying objects that are partially obscured or in difficult-to-see locations, such as cars parked under trees or street lamps in the shadows.
- This improved classification can have numerous practical applications, such as improving traffic monitoring and management, enhancing urban planning and development, and aiding in disaster response and recovery efforts.

# Objectives and Target Deliverables

- The objective of this presentation is to introduce the concept of attention mechanism in the object classification for drone imagery and explore the challenges associated with it.
- We will demonstrate how an attention mechanism can be used to improve the accuracy and efficiency of object classification in drone imagery.
- The key points we will cover include data preprocessing techniques, the architecture of the attention mechanism, and evaluation metrics for assessing the performance of the classification model.

# Dataset and its characteristics

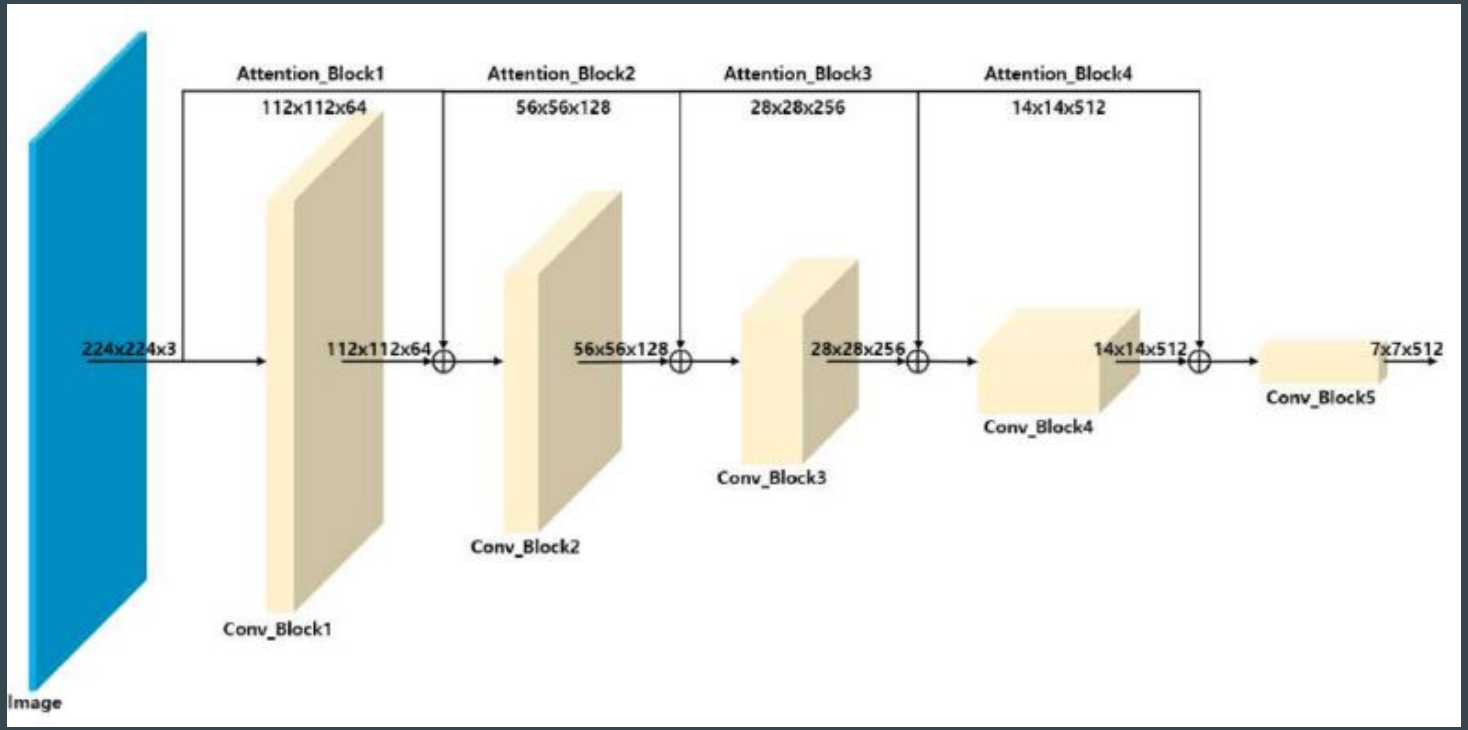
- We have used The VisDrone2019 dataset collected by the AISKYEYE team at Lab of Machine Learning and Data Mining which contains:
  - 6471 training samples
  - 548 validation samples
  - 1610 test samples
- When taking images with drones, it has a range of semi-spherical shapes based on cars. It would include more angle, especially bird's eye view and plane information than those in the traditional datasets.
- So, in this case object classification is difficult as we have images captured at different views and angles.



# Attention Mechanism

- Attention block is a proposed technique that complements lost features and extracts more diverse features through convolution layers.
- A feature map called attention block is add up to the output of the convolution layer before enter to next convolutional layers input.
- So, attention mechanism helps to highlight important features in the image that are relevant to the classification and improve the model's accuracy and efficiency.

# Model Architecture



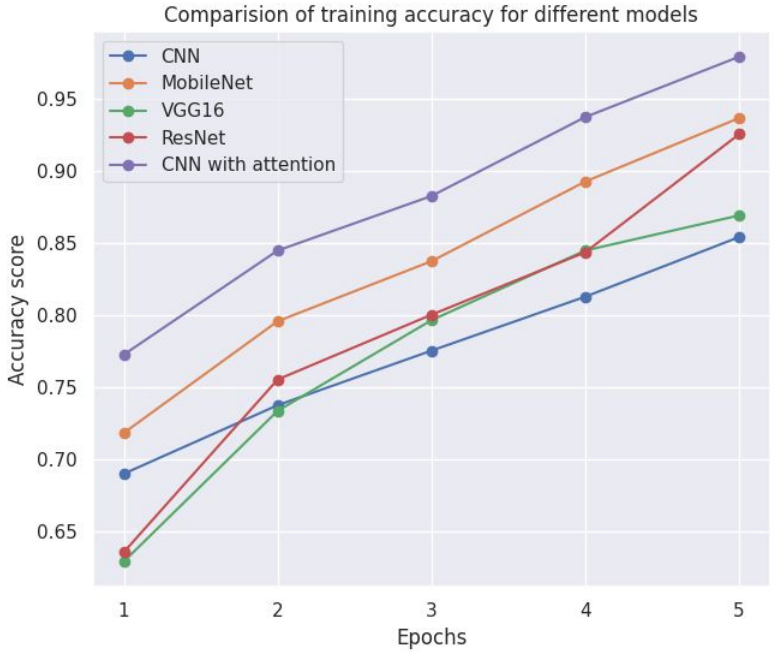
# Training procedure and evaluation metrics

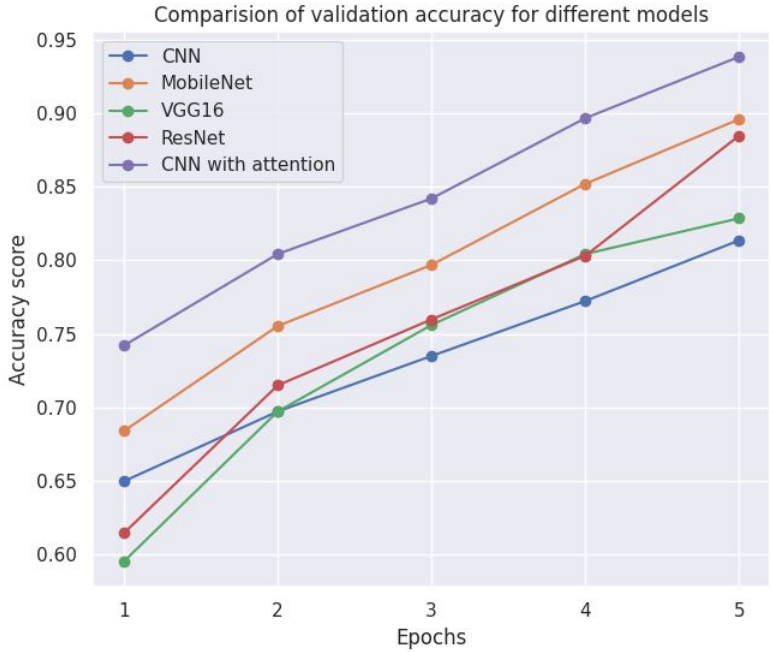
- We have implemented the proposed network consisting of convolution block and attention block using *Pytorch*.
- We have trained the proposed network using CUDA device, which is one of NVIDIA GPUs used to accelerates the training.
- We have evaluated the performance of the proposed network with different pre-trained models for image classification like *ResNet, VGG16* and *MobileNet* based on *accuracy and loss values, model parameters* and *model size*.

# Performance comparison of models

| Model | Test accuracy |
|---|---|
| CNN with Attention (proposed) | 93.87 % |
| MobileNet | 89.62% |
| VGG16 | 82.87 % |
| ResNet | 88.50% |
| CNN | 81.37% |

# Performance curves



# Comparison of Parameters and model size

| Model | Parameters | Model size |
|---|---|---|
| CNN with Attention | 11,177,538 | 45.43 MB |
| MobileNet | 8,251,397 | 42.71 MB |
| VGG16 | 134,268,738 | 512.23 MB |
| ResNet | 25,557,032 | 95.45 MB |
| CNN | 23,653,314 | 90.33 MB |

# Conclusion

- Evaluated Attention Mechanism on the VisDrone2019 dataset of drone images and compared the results to the traditional method of object classification.
- Attention mechanism achieved significant improvement in accuracy and robustness, and outperformed the baseline methods.
- On applying attention mechanism, the baseline CNN model's performance has improved significantly as it learns important feature maps.
- The proposed Attention based CNN has also less learnable parameters and occupies less memory space compared to other models.

# THANK YOU !