The 12th conference of the Sensor Signal Processing for Defense (SSPD)- Edinburgh –Scotland, Sept-11 until 14, 2023

SSPD Conference

# A Lower Complexity Deep Learning Method for Drones Detection

**Mohamad Kassab**, Thales/SCAI Research Associate , *Mohamed bin Zayed University of AI*, Abu Dhabi, UAE.

**Amal El Fallah Seghrouchni**, *Mohammed VI Polytechnic University, Ai Movement*, Rabat, Morocco, *Sorbonne University*, Paris, France.

**Frederic Barbaresco**, *Thales Group*, Paris, France.

**Raed Abu Zitar**, Senior Research Scientist, Thales/SCAI, *Sorbonne University, Abu Dhabi*, UAE.

## Abstract

Detecting objects such as drones is a challenging task. Their relatively small size and maneuvering capabilities specially in complex environments can deceive machine learning models and cause them to mis detect the drones. In this work, we investigate proposing a lower complexity deep learning model and compare it with several similar techniques based on real data sets (benchmarks of flying drones). A Deep learning paradigm is proposed for the purpose of mitigating the complexity of those systems. The proposed paradigm consists of a hybrid between the AdderNet deep learning paradigm and the SSD paradigm. The goal was to minimize multiplication operation numbers within convolutional layers and, hence, reduce complexity. Some standard machine learning techniques such as SVM and RF are also tested and compared to other deep learning systems. The data sets used for training and validation were either complete or filtered in order to remove the images with small objects. The types of data were either RGB or IR data. Comparisons were made between all these types and conclusions are presented.

## Objectives

This study aims to investigate the use of deep learning/machine learning techniques for the task of airborne object detection. The objectives of this study also include examining traditional machine learning techniques such as SVM and RF to highlight the trade-offs between computational complexity and accuracy for real-time drone detection. To achieve lower complexity in deep learning, the convolutional filters in SSD were replaced with AdderNet filters. The pros and cons of the modified SSD were investigated. Different methods were compared the proposed SSD/AdderNet based on benchmarks and results were presented.

## Data sets

Below is description for the training data sets used in training and validation

TABLE I: Number of Samples in Drone-Vs-Bird

| Object Size | Training Samples | Validation Samples |
|---|---|---|
| Small | 63295 | 3578 |
| Medium | 21124 | 1104 |
| Large | 2413 | 63 |
| Background | 13034 | 444 |
| Total | 99866 | 5189 |

TABLE II: Number of Samples in Anti-UAV

| Object Size | Training Samples | Validation Samples |
|---|---|---|
| Small | 58362 | 2373 |
| Medium | 38660 | 2670 |
| Large | 94 | 0 |
| Background | 1560 | 197 |
| Total | 98676 | 5240 |

## SSD and AdderNet

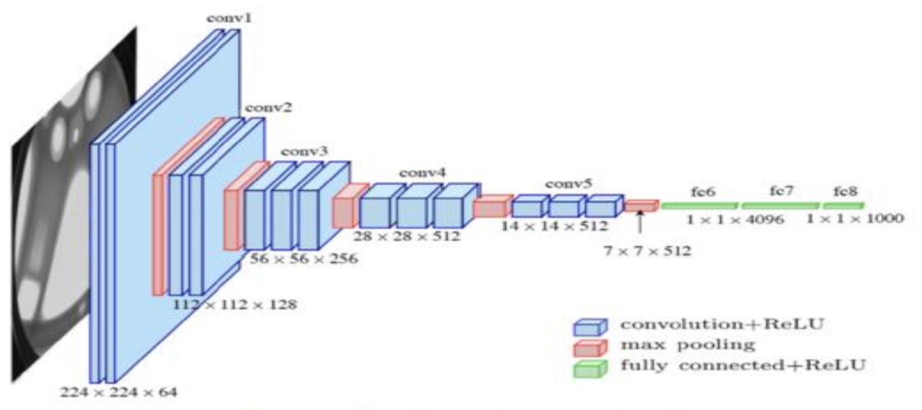



Fig. 2: SSD Backbone [17]

- **By implementing the proposed SSD + AdderNet, the number of multiplications has been reduced by 367.7785 M**

TABLE V: Number of Multiplications

| | Conv 1 | Conv 2 | Conv 3 |
|---|---|---|---|
| **Block 1** | Number of filters : 64<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 224x224x3<br>Multiplications = 86.704 M | Number of filters : 64<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 224x224x3<br>Multiplications = 86.704 M | - |
| **Block 2** | Number of filters : 128<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 112x112x3<br>Multiplications = 44.35 M | Number of filters : 128<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 112x112x3<br>Multiplications = 44.35 M | - |
| **Block 3** | Number of filters: 256<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 56x56x3<br>Multiplications = 21.676 M | Number of filters: 256<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 56x56x3<br>Multiplications = 21.676 M | Number of filters: 256<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 56x56x3<br>Multiplications = 21.676 M |
| **Block 4** | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 28x28x3<br>Multiplications = 10.838 M | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 28x28x3<br>Multiplications = 10.838 M | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 28x28x3<br>Multiplications = 10.838 M |
| **Block 5** | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 14x14x3<br>Multiplications = 2.7095 M | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 14x14x3<br>Multiplications = 2.7095 M | Number of filters : 512<br>Size of filter : 3x3<br>Padding : same<br>Stride : 1<br>Input size : 14x14x3<br>Multiplications = 2.7095 M |

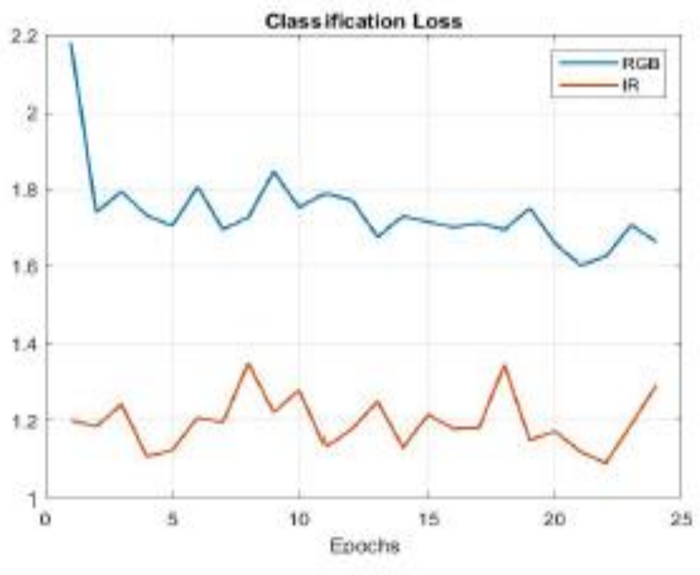
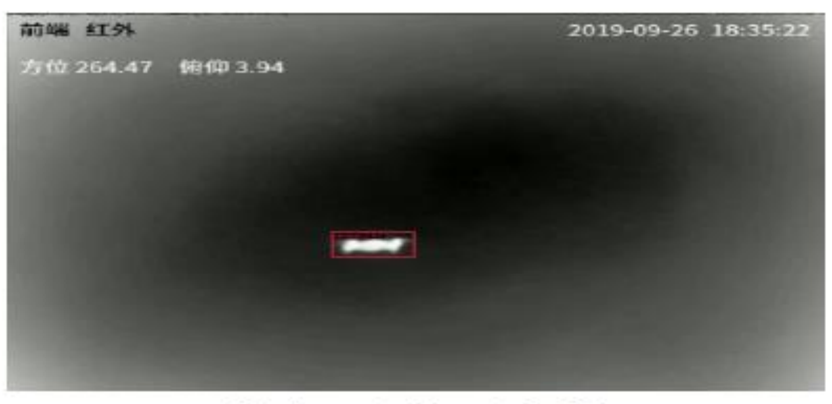



Fig. 3: Hybrid SSD Detection Learning Loss

## Simulations ans Results

(a) Sample Result 1: RGB




(b) Sample Result 2: IR

TABLE II: Inference Time (seconds/image)

| Dataset | Model | Inference Time (Second/Image) |
|---|---|---|
| Drone-Vs-Bird | Faster-RCNN | 0.0493 |
| | SSD | 0.0297 |
| | YOLOv3 | 0.0240 |
| | DETR | 0.0521 |
| | SSD/AdderNet | 0.0285 |
| Anti-UAV-IR | Faster-RCNN | 0.0459 |
| | SSD | 0.0291 |
| | YOLOv3 | 0.0253 |
| | DETR | 0.0483 |
| | SSD/AdderNet | 0.0275 |
| Anti-UAV-RGB | Faster-RCNN | 0.0481 |
| | SSD | 0.0293 |
| | YOLOv3 | 0.0242 |
| | DETR | 0.0524 |
| | SSD/AdderNet | 0.0281 |

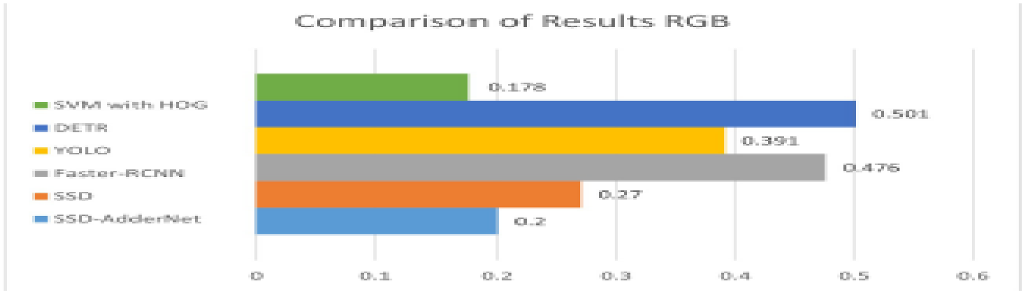



Fig. 4: Comparison of Results RGB

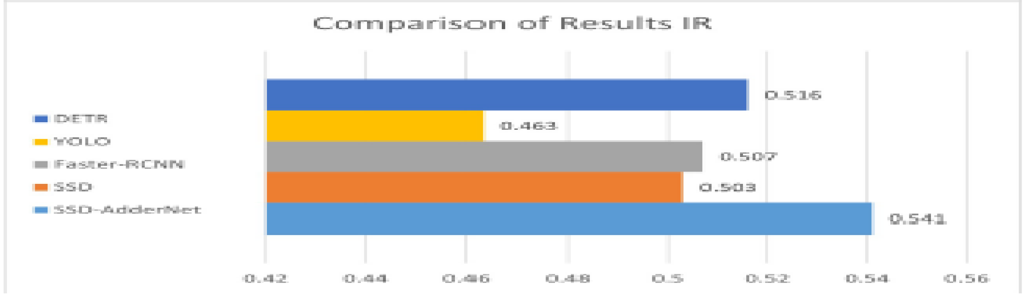



Fig. 5: Comparison of Results IR

## Conclusions

The work summarizes efforts to propose a new deep learning paradigm of the one-stage detector with much less complexity in the filtering layers. The reduced complexity is based on minimizing the number of multiplications in the convolutional layer in an SSD-AdderNet architecture. The training/testing data was extracted from real videos with moderate resolution and a mixture of small/medium/large objects (drones) sizes. The goal was to detect the presence of a drone in the image. Despite the low precision achieved by our proposed SSD-AdderNet when trained/tested on RGB images compared to other well-known techniques, the reduction in the complexity was remarkable. However, the performance of the SSD-AdderNet outperformed other models when trained/tested on IR images. The proposed method is recommended to be used for large-to-medium-sized objects when dealing with RGB images where the results are acceptable. Furthermore, the SSD-AdderNet showed good performance when dealing with small images on IR data . This is considered to be promising. The users now have the privilege to switch between high precision detectors such as the DETR or Faster R-CNN for small RGB objects or our less complex and, hence, faster during inference, SSD-AdderNet for larger and higher resolution RGB objects that usually require more computations and time. SSD/AdderNet is more appropriate for applications that require faster response and shorter decision time.

## References (selected)

1- H. Chen, Y. Wang, C. Xu, B. Shi, C. Xu, Q. Tian, C. Xu, Addernet: Do we really need multiplications in deep learning?, in *the Proceedings IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1468–1477.

2- W.Liu, D.Anguelov, D.Erhan, C.Szegedy, S.Reed,C.-Y.Fu, A.C.Berg, Ssd: Single shot multibox detector, in: *European conference on computer vision*, Springer, 2016, pp.21–37.

## Acknowledgements

THALES

15 YEARS OF EXCELLENCE SORBONNE UNIVERSITY ABU DHABI

MOHAMED BIN ZAYED UNIVERSITY OF ARTIFICIAL INTELLIGENCE

scai SORBONNE CENTER FOR ARTIFICIAL INTELLIGENCE