KIOS
Research and Innovation Center of Excellence

Design Exploration of CNN Parameters for Multi-Altitude UAV Object Detection

Michalis Piponidis, Theocharis Theocharides
KIOS Research and Innovation Center of Excellence
Department of Electrical and Computer Engineering
University of Cyprus

University of Cyprus
Imperial College London

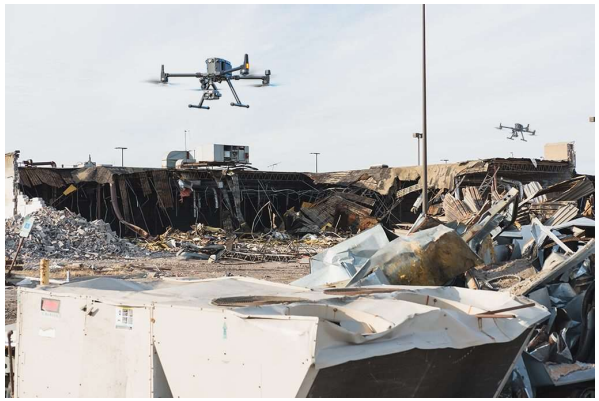
CPS Summer School 2024, September 16-20 2024, Alghero, Sardinia, Italy

16/09/2024

Funded by:

Introduction

- **Unmanned Aerial Vehicles (UAVs) are increasingly used in various fields**
  - Search and rescue
  - Emergency management
  - Infrastructure inspection
- **High versatility**
  - Fast, cheap and easy deployment
  - Aerial image capture from various angles/altitudes
  - Unmanned ➔ Ideal in dangerous situations!
- **Effective object detection is crucial for UAVs**
  - Identify objects of interest
  - Situational awareness ➔ Navigation/Obstacle avoidance



Images sources: https://www.flytbase.com/blog/drone-disaster-relief
https://www.nbcnews.com/mach/science/drones-are-fighting-wildfires-some-very-surprising-ways-ncna820966
https://haztech.com/uav-inspection/

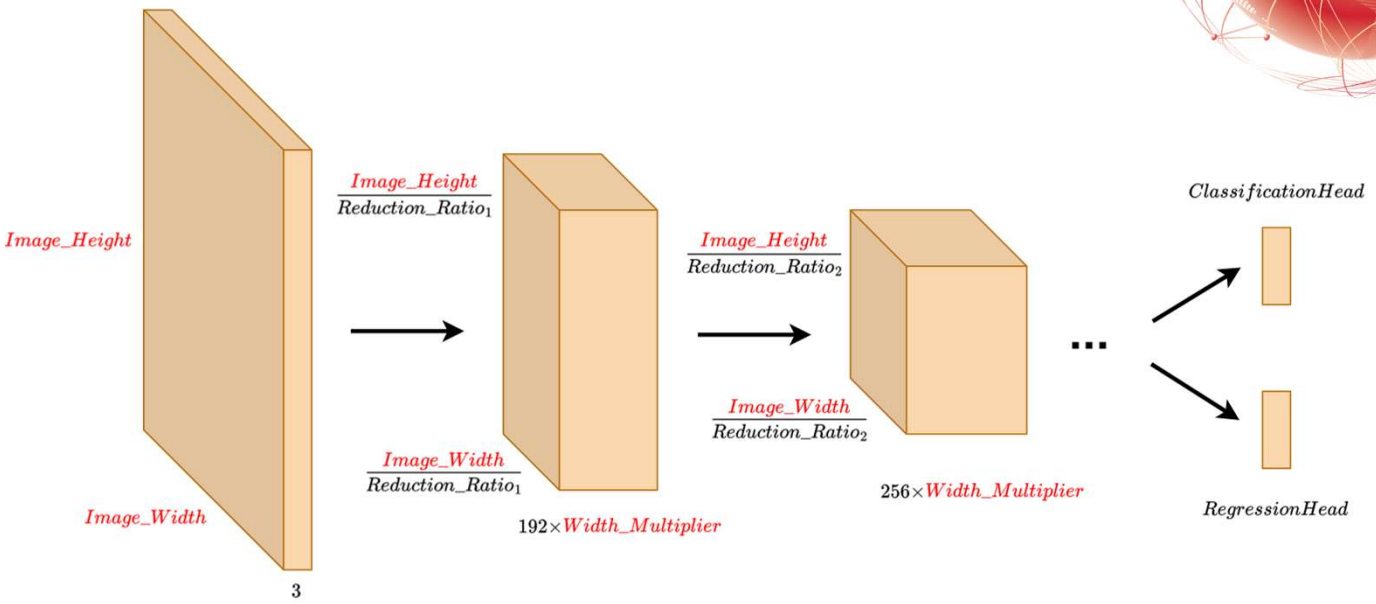
Challenges in UAV Object Detection

- **Real-time processing needs**
  - Limited on-board computational and power budget
  - Models must be computationally efficient while maintaining the desired accuracy
- **Large variety angles and altitudes**
  - Traditional models are typically trained to generalize on images taken from similar angles/distances
- **Diverse environmental conditions**
  - Reflections, smoke, and occlusions can obscure objects
  - Weather conditions (rain, fog, sun rays etc.) can reduce image clarity

Objective

Examine the effect of CNN parameters for UAV object detection at various altitudes

1. Input image resolution
2. Network width (number of channels)

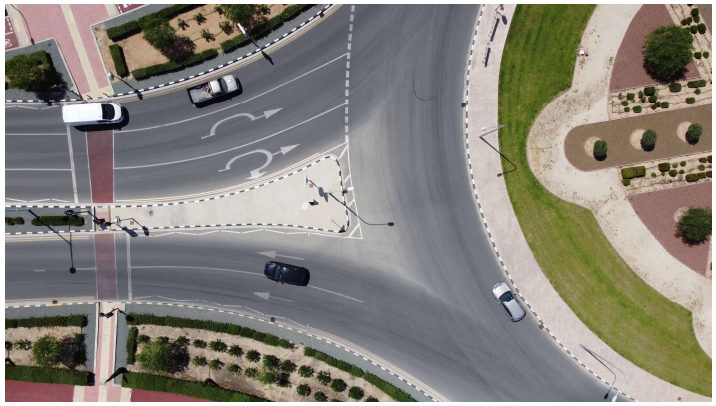


- **Higher resolutions:**
  - Capture more detail
  - Increase computational load
- **Wider networks:**
  - Learn more features
  - Increase computational load

- **Goal:** Find optimal parameter values for various altitude ranges
  - Optimize **accuracy-performance trade-off**

Dataset

- **Multi-Altitude Aerial Vehicles Dataset [1]**
  - Single-class object detection (Car)
  - Images taken from 50 to 500 meters with a 50-meter step
  - Each altitude contains:
    - ≈250 training images
    - ≈60 test images
    - ≈60 validation images



50m

200m

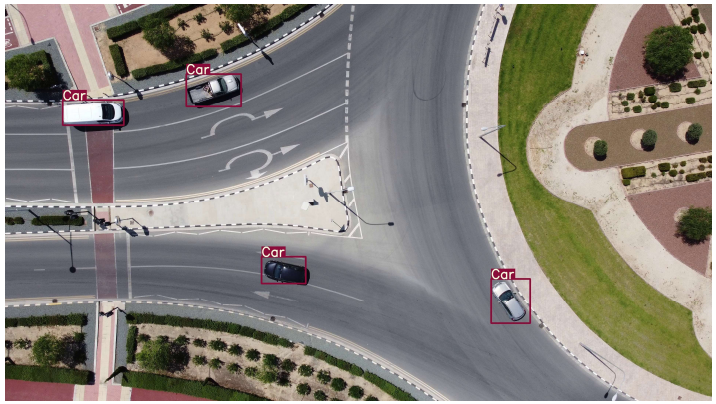
350m

500m

[1] https://www.kios.ucy.ac.cy/evai/datasets/multi-altitude-aerial-vehicles/

Dataset

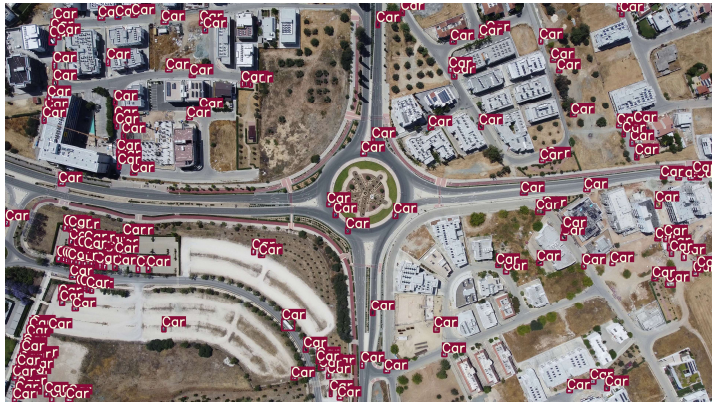
- **Multi-Altitude Aerial Vehicles Dataset** [1]
  - Single-class object detection (Car)
  - Images taken from 50 to 500 meters with a 50-meter step
  - Each altitude contains:
    - ≈250 training images
    - ≈60 test images
    - ≈60 validation images



50m

200m

350m



500m

[1] https://www.kios.ucy.ac.cy/evai/datasets/multi-altitude-aerial-vehicles/

Methodology

- **Model:** tiny YOLOv7
- **Resolutions:**
  - 1088 x 1088
  - 896 x 896
  - 768 x 768
  - 640 x 640
  - 512 x 512
  - 384 x 384
  - 256 x 256
- **Width Multipliers:**
  - 1x (Default)
  - 0.75x
  - 0.5x
  - 0.25x
  - 0.1x
  - 0.01x
- **Evaluation Metrics:**
  - Accuracy
    - mAP@.5
  - Computational Efficiency
    - MAC Operations
    - Network size (MB)

Methodology

- 5 variable models for altitude ranges:
  - 50-100m
  - 150-200m
  - 250-300m
  - 350-400m
  - 450-500m
- 1 fixed model including the whole dataset
- Each model trained for all parameter combinations
- Trained in python using PyTorch

Parameter	Value
Epochs	300
Batch size	16
Optimizer	SGD
Initial learning rate	0.01
Final learning rate	0.1
Momentum	0.937
Learning rate scheduler	One Cycle

Results – Accuracy

50-100m

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,996	0,997	0,996	0,994	0,994	0,992
896	0,997	0,996	0,995	0,995	0,991	0,973
768	0,998	0,994	0,994	0,994	0,989	0,969
640	0,997	0,996	0,995	0,992	0,961	0,936
512	0,997	0,991	0,995	0,985	0,984	0,898
384	0,998	0,991	0,982	0,954	0,928	0,971
256	0,996	0,954	0,926	0,879	0,688	0,731

Width Multiplier

350-400m

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,951	0,917	0,83	0,739	0,584	0,372
896	0,935	0,892	0,849	0,649	0,383	0,143
768	0,922	0,845	0,766	0,314	0,45	0,0168
640	0,571	0,017	0,111	0,139	0,0427	0,157
512	0,774	0,00547	0,0569	0,0054	0,00205	0,0113
384	0,314	0,0109	4,45E-05	0,000342	7,72E-05	0,000101
256	5,73E-06	6,54E-06	2,69E-05	1,15E-05	2,11E-05	0,000834

Width Multiplier

150-200m

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,997	0,995	0,994	0,989	0,991	0,978
896	0,997	0,996	0,994	0,984	0,976	0,983
768	0,998	0,994	0,982	0,985	0,967	0,957
640	0,997	0,973	0,984	0,969	0,932	0,943
512	0,996	0,99	0,967	0,929	0,869	0,848
384	0,969	0,144	0,0239	0,00934	0,00826	0,000339
256	0,223	0,00145	0,000104	0,00142	3,53E-05	9,12E-05

Width Multiplier

450-500m

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,935	0,878	0,801	0,628	0,454	0,39
896	0,897	0,714	0,69	0,296	0,25	0,271
768	0,807	0,494	0,511	0,35	0,189	0,288
640	0,417	0,463	0,121	0,339	0,116	0,176
512	0,0954	0,232	0,00675	0,00652	0,0777	0,0422
384	0,00102	0,001	0,000114	0,00905	0,0123	0,00513
256	4,97E-06	0,000209	3,66E-05	1,47E-05	2,16E-05	9,17E-06

Width Multiplier

250-300m

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,992	0,986	0,987	0,967	0,883	0,834
896	0,991	0,983	0,988	0,938	0,794	0,685
768	0,989	0,978	0,959	0,807	0,689	0,408
640	0,981	0,956	0,911	0,704	0,438	0,00322
512	0,938	0,00962	0,499	0,465	0,384	0,0185
384	0,483	0,4	0,294	0,00335	0,000375	0,000369
256	0,00547	0,133	7,62E-05	8,03E-05	4,55E-05	7,07E-05

Width Multiplier

mix_alt

Resolution \ mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,961	0,935	0,916	0,808	0,652	0,551
896	0,939	0,826	0,845	0,677	0,559	0,441
768	0,943	0,859	0,881	0,708	0,594	0,453
640	0,919	0,879	0,513	0,533	0,482	0,264
512	0,802	0,318	0,26	0,0591	0,0569	0,0701
384	0,0211	0,274	0,0324	0,00188	0,000165	0,0308
256	0,0154	0,0174	0,0157	0,000104	0,0025	0,00687

Width Multiplier

Results – Computational Efficiency

MAC Operations (M)

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	45,42	34,13	22,84	11,55	5,81	2,38
896	31,99	24,04	16,09	8,14	4,09	1,68
768	22,71	17,07	11,42	5,78	2,9	1,19
640	16,35	12,29	8,22	4,16	2,09	0,86
512	11,03	8,29	5,55	2,81	1,41	0,58
384	6,75	5,07	3,39	1,72	0,86	0,35
256	2,92	2,19	1,47	0,74	0,37	0,15

Size (MB)

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	17,41	12,15	8,39	6,13	5,55	5,4
896	15,83	10,57	6,81	4,55	3,97	3,82
768	14,73	9,47	5,71	3,45	2,87	2,72
640	13,98	8,72	4,96	2,7	2,12	1,97
512	13,35	8,09	4,33	2,07	1,49	1,34
384	12,85	7,59	3,83	1,57	0,99	0,84
256	12,4	7,14	3,38	1,12	0,54	0,39

Model	Parameters	Parameter Size (MB)
1	12.015.192	12,05
0.75	6.768.120	6,79
0.5	3.016.600	3,03
0.25	760.632	0,77
0.1	190.680	0,19
0.01	36.488	0,04

Case Study

Target Accuracy: **90%**

50-100m

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,996	0,997	0,996	0,994	0,994	0,992
896	0,997	0,996	0,995	0,995	0,991	0,973
768	0,998	0,994	0,994	0,994	0,989	0,969
640	0,997	0,996	0,995	0,992	0,961	0,936
512	0,997	0,991	0,995	0,985	0,984	0,898
384	0,998	0,991	0,982	0,954	0,928	0,971
256	0,996	0,954	0,926	0,879	0,688	0,731

150-200m

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,997	0,995	0,994	0,989	0,991	0,978
896	0,997	0,996	0,994	0,984	0,976	0,983
768	0,998	0,994	0,982	0,985	0,967	0,957
640	0,997	0,973	0,984	0,969	0,932	0,943
512	0,996	0,99	0,967	0,929	0,869	0,848
384	0,969	0,144	0,0239	0,00934	0,00826	0,000339
256	0,223	0,00145	0,000104	0,00142	3,53E-05	9,12E-05

250-300m

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,992	0,986	0,987	0,967	0,883	0,834
896	0,991	0,983	0,988	0,938	0,794	0,685
768	0,989	0,978	0,959	0,807	0,689	0,408
640	0,981	0,956	0,911	0,704	0,438	0,00322
512	0,938	0,00962	0,499	0,465	0,384	0,0185
384	0,483	0,4	0,294	0,00335	0,000375	0,000369
256	0,00547	0,133	7,62E-05	8,03E-05	4,55E-05	7,07E-05

350-400m

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,951	0,917	0,83	0,739	0,584	0,372
896	0,935	0,892	0,849	0,649	0,383	0,143
768	0,922	0,845	0,766	0,314	0,45	0,0168
640	0,571	0,017	0,111	0,139	0,0427	0,157
512	0,774	0,00547	0,0569	0,0054	0,00205	0,0113
384	0,314	0,0109	4,45E-05	0,000342	7,72E-05	0,000101
256	5,73E-06	6,54E-06	2,69E-05	1,15E-05	2,11E-05	0,000834

450-500m

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,935	0,878	0,801	0,628	0,454	0,39
896	0,897	0,714	0,69	0,296	0,25	0,271
768	0,807	0,494	0,511	0,35	0,189	0,288
640	0,417	0,463	0,121	0,339	0,116	0,176
512	0,0954	0,232	0,00675	0,00652	0,0777	0,0422
384	0,00102	0,001	0,000114	0,00905	0,0123	0,00513
256	4,97E-06	0,000209	3,66E-05	1,47E-05	2,16E-05	9,17E-06

mix_alt

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	0,961	0,935	0,916	0,808	0,652	0,551
896	0,939	0,826	0,845	0,677	0,559	0,441
768	0,943	0,859	0,881	0,708	0,594	0,453
640	0,919	0,879	0,513	0,533	0,482	0,264
512	0,802	0,318	0,26	0,0591	0,0569	0,0701
384	0,0211	0,274	0,0324	0,00188	0,000165	0,0308
256	0,0154	0,0174	0,0157	0,000104	0,0025	0,00687

MAC Operations (M)

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	45,42	34,13	22,84	11,55	5,81	2,38
896	31,99	24,04	16,09	8,14	4,09	1,68
768	22,71	17,07	11,42	5,78	2,9	1,19
640	16,35	12,29	8,22	4,16	2,09	0,86
512	11,03	8,29	5,55	2,81	1,41	0,58
384	6,75	5,07	3,39	1,72	0,86	0,35
256	2,92	2,19	1,47	0,74	0,37	0,15

Model	Resolution	Width Multiplier	MAC Operations (M)
50-100m	640p	0.01x	0,86
150-200m	640p	0.01x	0,86
250-300m	896p	0.25x	8,14
350-400m	768p	1x	22,71
450-500m	896p	1x	31,99
mix_alt	640p	1x	16,35

Case Study – Dataset

MAC Operations (M)

Resolution \ Width Multiplier	1	0.75	0.5	0.25	0.1	0.01
1088	45,42	34,13	22,84	11,55	5,81	2,38
896	31,99	24,04	16,09	8,14	4,09	1,68
768	22,71	17,07	11,42	5,78	2,9	1,19
640	16,35	12,29	8,22	4,16	2,09	0,86
512	11,03	8,29	5,55	2,81	1,41	0,58
384	6,75	5,07	3,39	1,72	0,86	0,35
256	2,92	2,19	1,47	0,74	0,37	0,15

Dataset test set images:

- 50-100m: 93
- 150-200m: 122
- 250-300m: 122
- 350-400m: 122
- 450-500m: 128

Selected models:

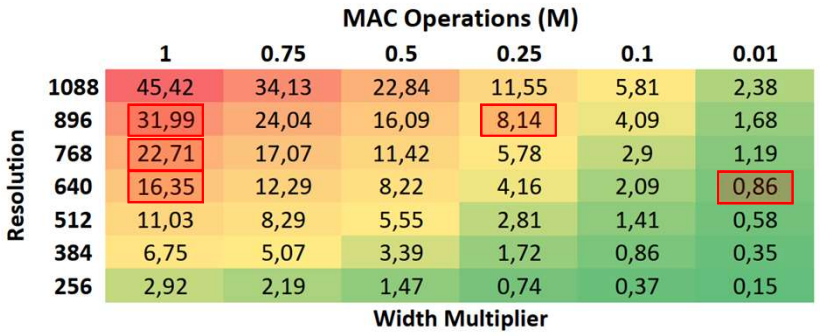
- 50-100m: 0.01@640p
- 150-200m: 0.01@640p
- 250-300m: 0.25@896p
- 350-400m: 1@768p
- 450-500m: 1@896p
- mix_alt: 1@640p

MAC Operations using:

- **Variable models:**
  - **MAC Operations:** (93 x 0,86) + (122 x 0,96) + (122 x 8,14) + (122 x 22,71) + (128 x 31,99) = **8055,52M**
  - **Accuracy: 90,5%**
- **Fixed model:**
  - **MAC Operations:** 587 x 16,35 = **9597,45M**
  - **Accuracy: 91,9%**

- **+19.1% Speed**
- **-1.4% Accuracy**

Case Study – Theoretical Lower Altitudes



MAC Operations (M)

Resolution	1	0.75	0.5	0.25	0.1	0.01
1088	45,42	34,13	22,84	11,55	5,81	2,38
896	31,99	24,04	16,09	8,14	4,09	1,68
768	22,71	17,07	11,42	5,78	2,9	1,19
640	16,35	12,29	8,22	4,16	2,09	0,86
512	11,03	8,29	5,55	2,81	1,41	0,58
384	6,75	5,07	3,39	1,72	0,86	0,35
256	2,92	2,19	1,47	0,74	0,37	0,15

Width Multiplier

Theoretical image distribution:

- 50-100m: 35%
- 150-200m: 30%
- 250-300m: 20%
- 350-400m: 10%
- 450-500m: 5%

Selected models:

- 50-100m: 0.01@640p
- 150-200m: 0.01@640p
- 250-300m: 0.25@896p
- 350-400m: 1@768p
- 450-500m: 1@896p
- mix_alt: 1@640p

MAC Operations using:

- **Variable models:**
  - **MAC Operations:** (35 x 0,86) + (30 x 0,96) + (20 x 8,14) + (10 x 22,71) + (5 x 31,99) = **608,75M**
  - **Expected Accuracy:** (35% x 93,6%) + (30% x 94,3%) + (20% x 93,8%) + (10% x 92,2%) + (5% x 89,7%) = **93,5%**
- **Fixed model:**
  - **MAC Operations:** 100 x 16,35 = **1635M**
  - **Accuracy: 91,9%**

- **+168.6% Speed**
- **+1.6% Accuracy**

Conclusions

- Different altitude ranges require distinct parameter configurations to achieve target accuracy
- Dynamic network structures that adapt parameters based on altitude data can significantly enhance the efficiency and performance of UAV object detection
- Flexible and efficient object detection for UAV applications

Challenges – Next steps

- **Implementing seamless switching between models**
  - Dynamic networks
  - Reconfigurable hardware
- **Determining optimal parameters for each altitude range**
  - Testing everything is time/computationally intensive
- **Selecting the appropriate model for each image**
  - Addressing scenarios with unknown current altitude
- **Creating a comprehensive dataset with:**
  - Non-fixed altitudes
  - Diverse locations and environmental conditions

Thank you for your attention!