

Raport projektu - detekcja dronów

Michał Bober, Antoni Jaszcz

Streszczenie

W naszym projekcie zajęliśmy się problemem detekcji dronów na obrazach/klatkach filmu. W naszych eksperymentach testowaliśmy klasyczne konwolucyjne sieci neuronowe, VGG16 i ResNet50 oraz różniące się wielkością warianty architektur YOLOv5 i YOLOv8. Założenie projektu dotyczyło porównania modeli i wyznaczenia najbardziej optymalnego (pod względem złożoności, wykorzystania zasobów w procesie uczenia i predykcji oraz uzyskanej dokładności) modelu. Wszystkie modele były dotrenowywane przez 10 epok i korzystały z wcześniej wytrenowanych wag na zbiorach ImageNet oraz COCO. Uzyskane wyniki pokazują, że model YOLOv8n jest najlepszym wyborem do wybranego przez nas zadania, osiągając najmniejsze zużycie zasobów pamięciowych i uzyskując dokładność mAP50 wynoszącą 91.48%.

I. EKSPERYMENTY

A. Dane i ustawienia eksperymentalne

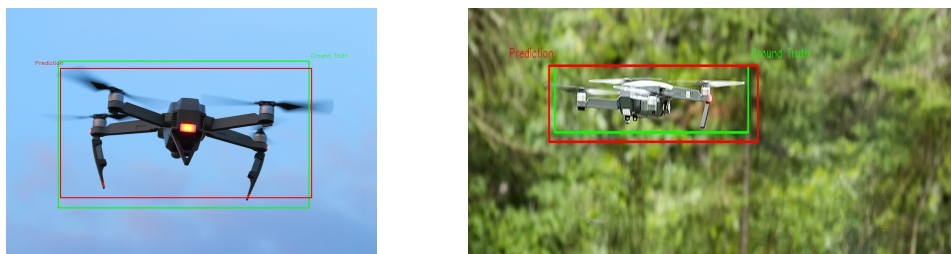
W wykonanych eksperymentach, posłużyliśmy otwarto-źródłową bazą zdjęć dronów, prezentującą jednostki latające w różnych sytuacjach (platforma Kaggle). Dane zostały podzielone w proporcjach 7:2:1 na zbiór treningowy, walidacyjny i testowy.

Wszystkie obliczenia były prowadzone na komputerze z GPU nvidia rtx 2060, która była wykorzystywana w obliczeniach. Trening odbywał się mini grupami, liczącymi 16 próbek. Każdy z testowanych modeli był uczony przez 10 epok i korzystał z wcześniej wytrenowanych wag na zbiorach ImageNet lub COCO. Wybrany algorytm optymalizacji był Adam, z maksymalnym wskaźnikiem uczenia ustawionym na $1e-3$. Kod źródłowy dostępny jest platformie GitHub.

B. Wyniki

Model	mean memory usage (MB)	mean gpu usage (MB)	runtime (s)	Ilość parametrów (mln.)	Acc. (mAP 50)
VGG16	4540	5642	270	18.9	68.38
ResNet50	4860	5686	240	40.3	77.94
YOLOv5n	1664	2290	403	1.9	90.69
YOLOv5s	1810	3364	383	7.2	91.60
YOLOv5m	3578	4305	1353	21.2	90.42
YOLOv8n	1635	2118	451	3.2	91.48
YOLOv8s	1721	3073	391	11.2	90.55

Tabela I: Porównanie zużycia zasobów, złożoności i uzyskanej dokładności dla różnych modeli



Rysunek 1: Przykładowe predykcje modelu YOLOv8n na zbiorze testowym

II. WNIOSKI

Z uzyskanych wyników widać, że wszystkie architektury typu YOLO okazały się lepsze, od klasycznych CNN, osiągając ponad 90-procentową dokładność w każdym przypadku. Choć model YOLOv8n ma niższą dokładność, niż YOLOv5s, jest to różnica zaledwie 0.12 punktów procentowych. Jest to niewielka strata, w porównania z ewidentną przewagą YOLOv8n w zakresie ilości parametrów oraz użycia karty graficznej. Z uzyskanych wyników wynika, że YOLOv8n jest najbardziej optymalnym wyborem do zadania.

Wydział Matematyki Stosowanej, Politechnika Śląska
Kasubka 28, 44-100 Gliwice
M. Bober, mb300327@student.polsl.pl
A. Jaszcz, aj303181@student.polsl.pl