



Politechnika Opolska

**Wydział Elektrotechniki
Automatyki i Informatyki**

**Kierunek: Informatyka I st.
Semestr VI
Studia dzienne**

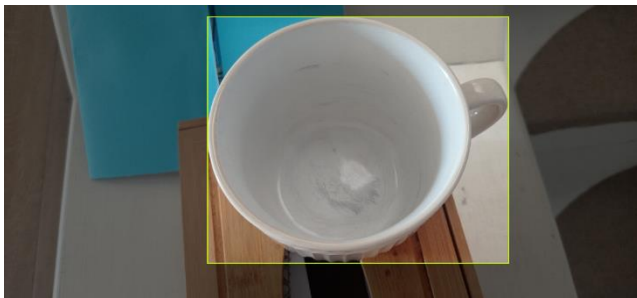
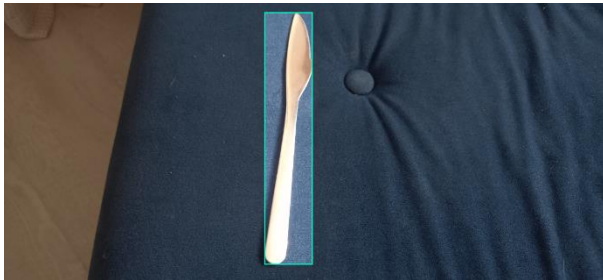
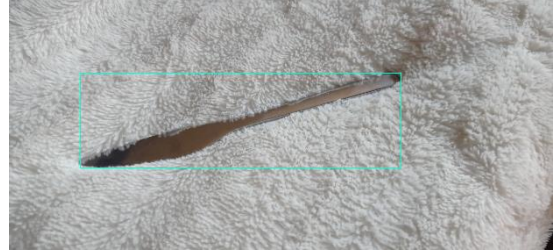
Narzędzia sztucznej inteligencji

Autor:

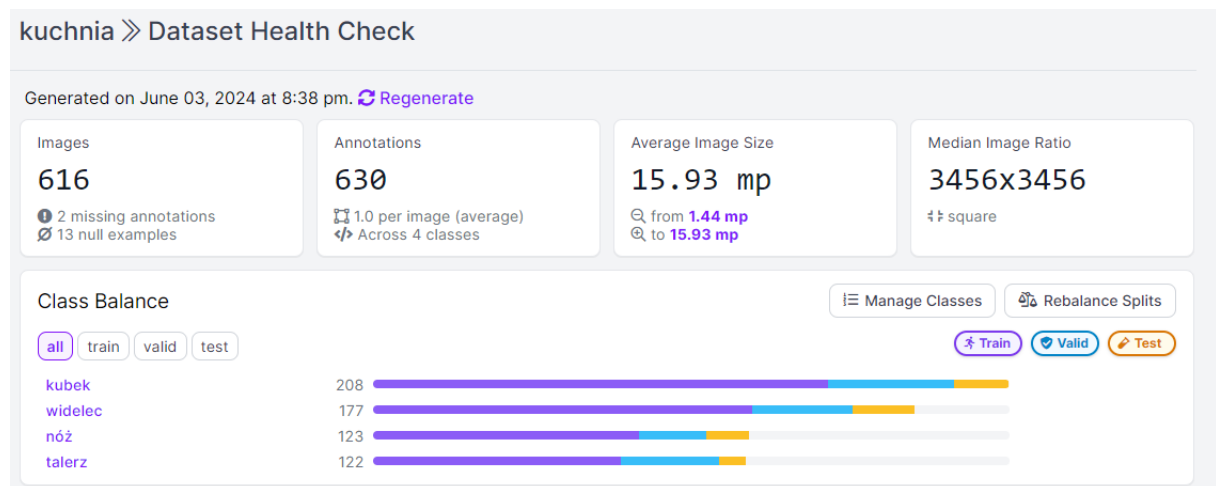
Dawid Tkacz s102584
Andrzej Szafrąński s102552

1. Dataset

Przykłady zdjęć w datasetcie:



Ocena datasetu:



Zbiór danych zawiera 616 obrazów. Sekcja "Class Balance" pokazuje rozkład klas: "kubek" (208), "widelec" (177), "nóż" (123) i "talerz" (122), z podziałem na zestawy treningowe, walidacyjne i testowe, co pozwala na ocenę równomierności rozkładu danych między tymi zestawami.

Dimension Insights

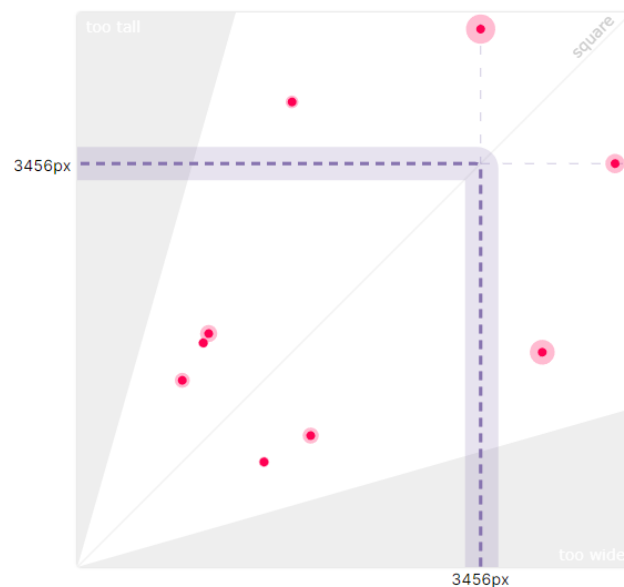
Size Distribution

The **purple box** indicates the median width by median height image (3456x3456).

jumbo

616

> 1024x1024



Aspect Ratio Distribution

tall

347

wide

269

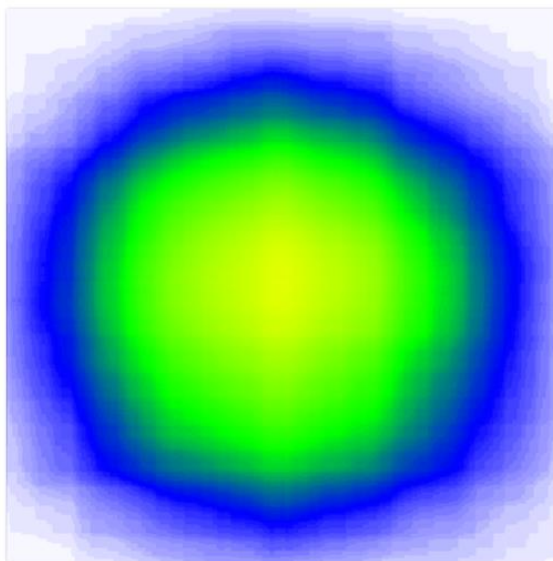
Zrzut ekranu przedstawia analizę wymiarów obrazów w zbiorze danych "kuchnia". Na górze widzimy rozkład rozmiarów obrazów, gdzie purpurowe pole wskazuje medianę szerokości i wysokości obrazu (3456x3456 pikseli). Widać, że wszystkie 616

obrazów znajdują się w kategorii "jumbo", czyli mają rozmiary większe niż 1024x1024 pikseli. Wykres punktowy pokazuje, że większość obrazów ma szerokość i wysokość zbliżoną do 3456 pikseli, a purpurowa linia odnosi się do tej mediany.

Na dole znajduje się rozkład proporcji wymiarów obrazów (Aspect Ratio Distribution). Obrazy są podzielone na dwie kategorie: "tall" (wysokie) i "wide" (szerokie). Rozkład wskazuje, że 347 obrazów jest wyższych niż szerszych, a 269 obrazów jest szerszych niż wyższych.

Annotation Heatmap

☒ all ☐ nóż (123) ☐ widelec (177) ☐ talerz (122) ☐ kubek (208)



Zrzut ekranu przedstawia mapę cieplną adnotacji w zbiorze danych "kuchnia". Mapa ta pokazuje rozkład wszystkich adnotacji dla czterech klas: "nóż" (123), "widelec" (177), "talerz" (122) i "kubek" (208). Kolory na mapie cieplnej reprezentują gęstość adnotacji, gdzie żółty kolor oznacza obszary o najwyższej gęstości, a niebieski obszary o najniższej gęstości.

Mapa cieplna wskazuje, że większość adnotacji znajduje się w centralnej części obrazów, co sugeruje, że obiekty w tych klasach są zazwyczaj umieszczane w środku kadru. Taki rozkład może mieć wpływ na wyniki modelu sztucznej inteligencji, gdyż model będzie bardziej skłonny do przewidywania obecności obiektów w centrum obrazu.

2. Model YOLOV8n

```
import torch
from ultralytics import YOLO
from roboflow import Roboflow
import os

device = 'cuda' if torch.cuda.is_available() else 'cpu'

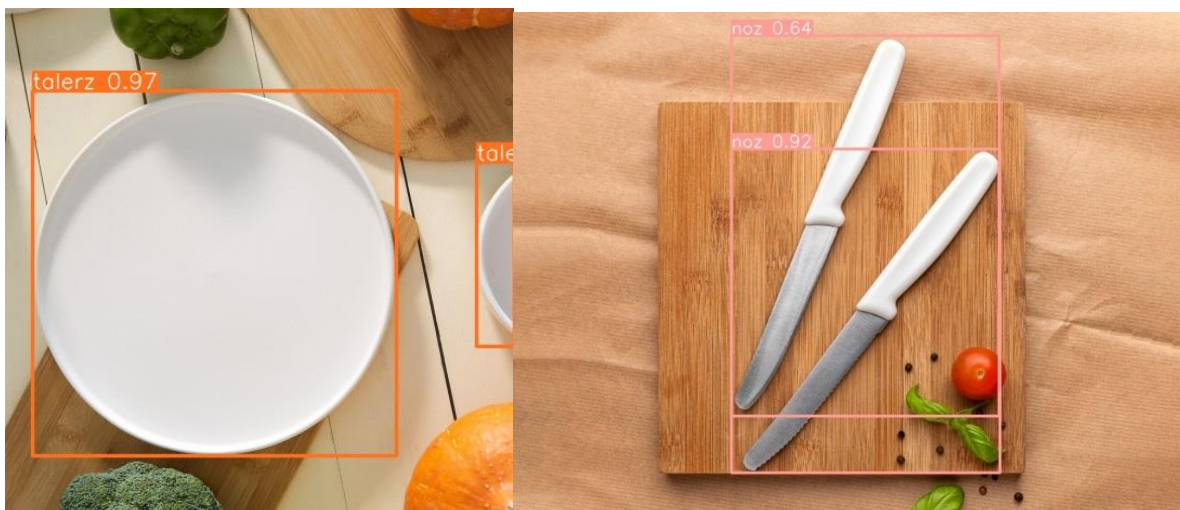
rf = Roboflow(api_key="ZArlQIkq4Z4P9CYIsjV1")
project = rf.workspace("kuchnia").project("kuchnia")
version = project.version(1)
dataset = version.download("yolov8")

if __name__ == "__main__":
    data_yaml = os.path.join(dataset.location, "data.yaml")
    model = YOLO('yolov8n.pt').to(device)
    model.train(data=data_yaml, epochs=100, imgsz=640, batch=16, name='sings')
```

Kod przedstawia proces trenowania modelu YOLOv8 przy użyciu biblioteki Ultralytics oraz narzędzia Roboflow. Model jest trenowany na pobranych danych z parametrami: 100 epok, rozmiar obrazu 640, rozmiar partii 16.

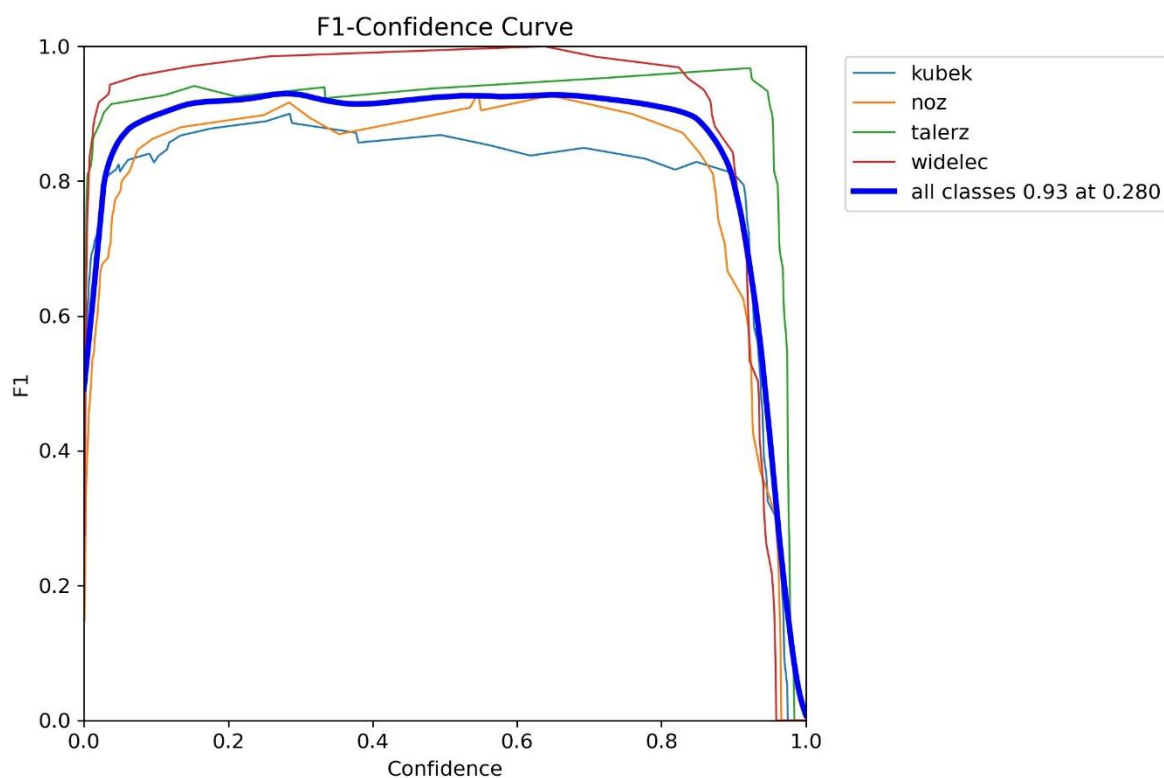
W tym kodzie użyto modelu YOLO z końcówką **n** (nano) jest to najlżejsza wersja, zaprojektowana do szybkiego działania na urządzeniach o ograniczonej mocy obliczeniowej.

Przykładowe wyniki wytrenowanego modelu:



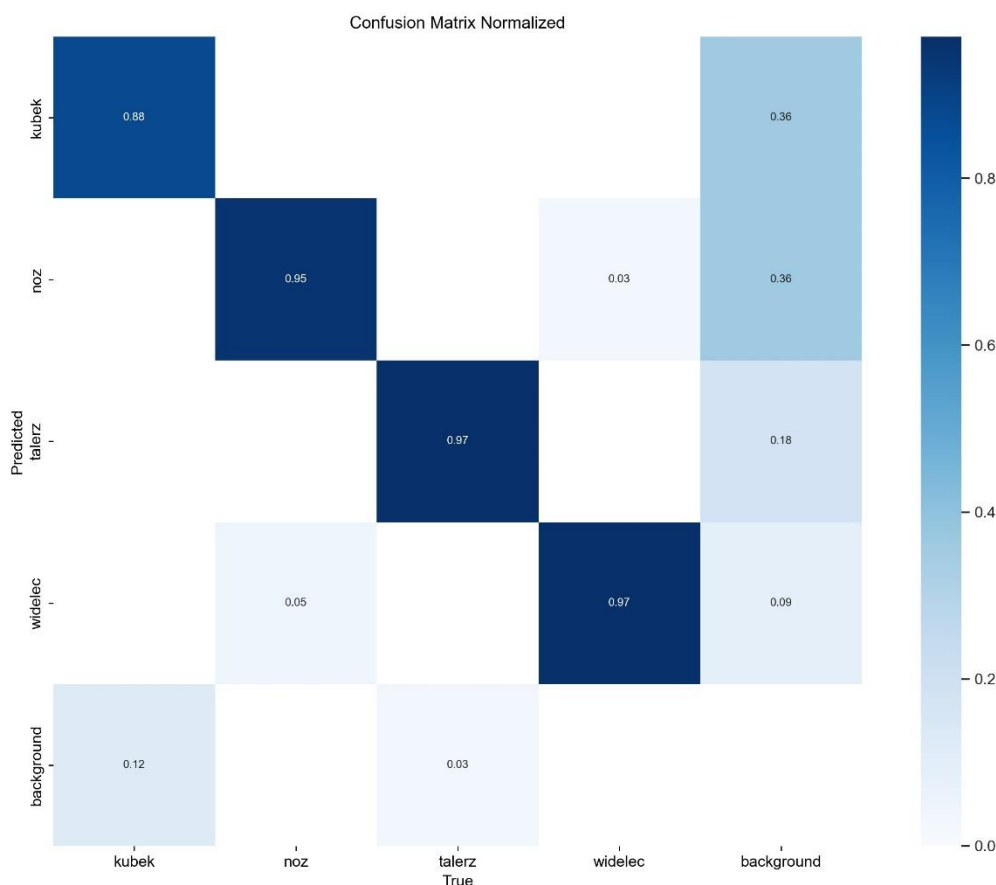


Ocena modelu:



Powyższy wykres przedstawia krzywą F1 w zależności od poziomu pewności dla różnych klas w modelu YOLOv8, który został wytrenowany na zbiorze danych

"kuchnia". Krzywe F1 pokazują wydajność modelu dla każdej z czterech klas: "kubek", "nóż", "talerz" i "widelec". Oś pozioma (Confidence) przedstawia poziom pewności predykcji modelu, natomiast oś pionowa (F1) przedstawia wartość miary F1, która jest harmoniczną średnią precyzji i czułości. Gruba niebieska linia reprezentuje uśrednioną wydajność modelu dla wszystkich klas, która osiąga wartość F1 równą 0.93 przy poziomie pewności 0.280.



Powyższy wykres przedstawia znormalizowaną macierz konfuzji dla modelu YOLOv8, który został wytrenowany na zbiorze danych "kuchnia". Macierz konfuzji pokazuje, jak dobrze model klasyfikuje obrazy do odpowiednich klas: "kubek", "nóż", "talerz" i "widelec", wraz z dodatkową kategorią "background" dla błędnych klasyfikacji.

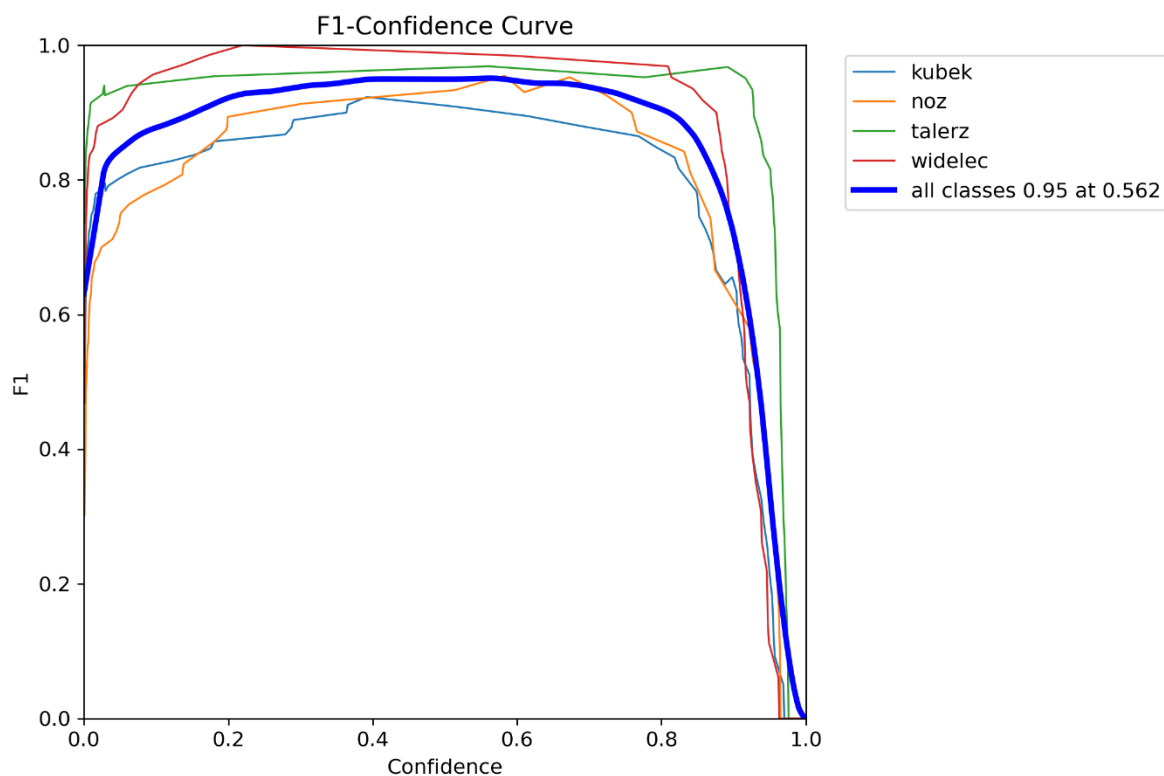
Klasa "kubek" ma najwyższy odsetek błędnych klasyfikacji do tła (12%), co sugeruje, że model może mieć trudności z rozróżnianiem kubków od tła w niektórych przypadkach.

3. Model YOLOV8s

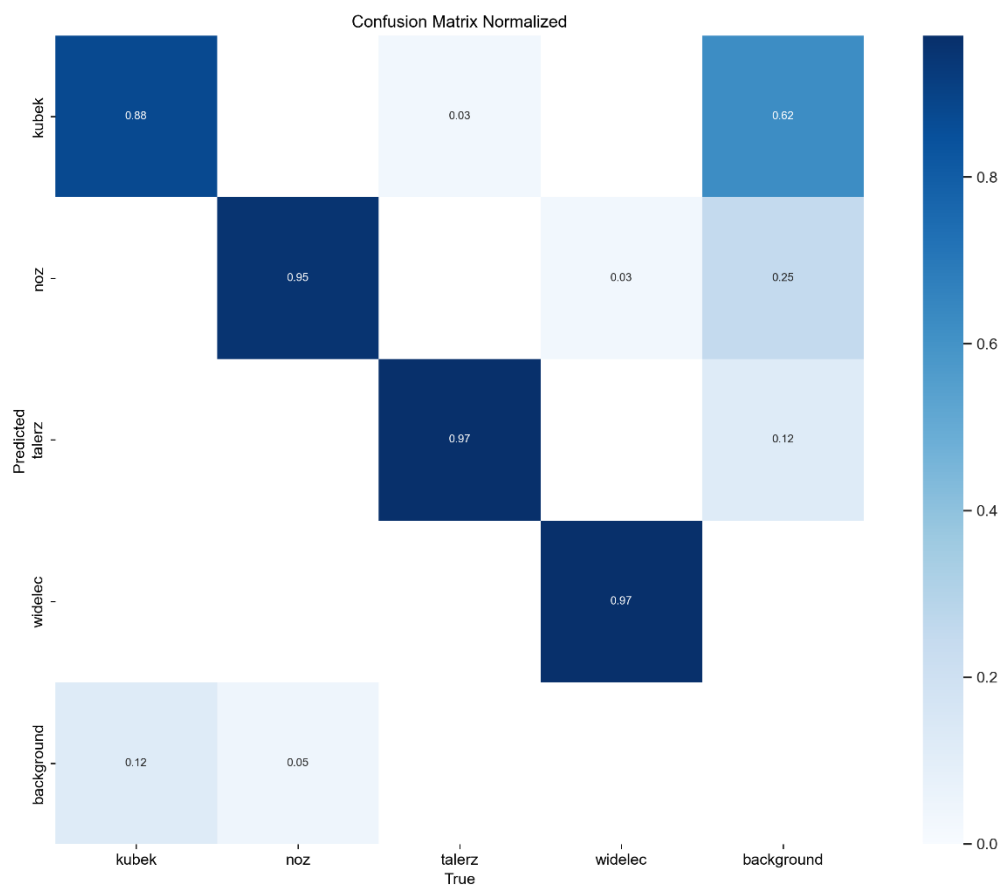
Przykładowe wyniki wytrenowanego modelu:



Ocena modelu:



Powyższy wykres przedstawia krzywą F1 w zależności od poziomu pewności dla różnych klas w modelu YOLOv8s, wytrenowanym na zbiorze danych "kuchnia". Podobnie jak w przypadku modelu YOLOv8n, krzywe F1 pokazują wydajność modelu dla każdej z czterech klas: "kubek", "nóż", "talerz" i "widelec".

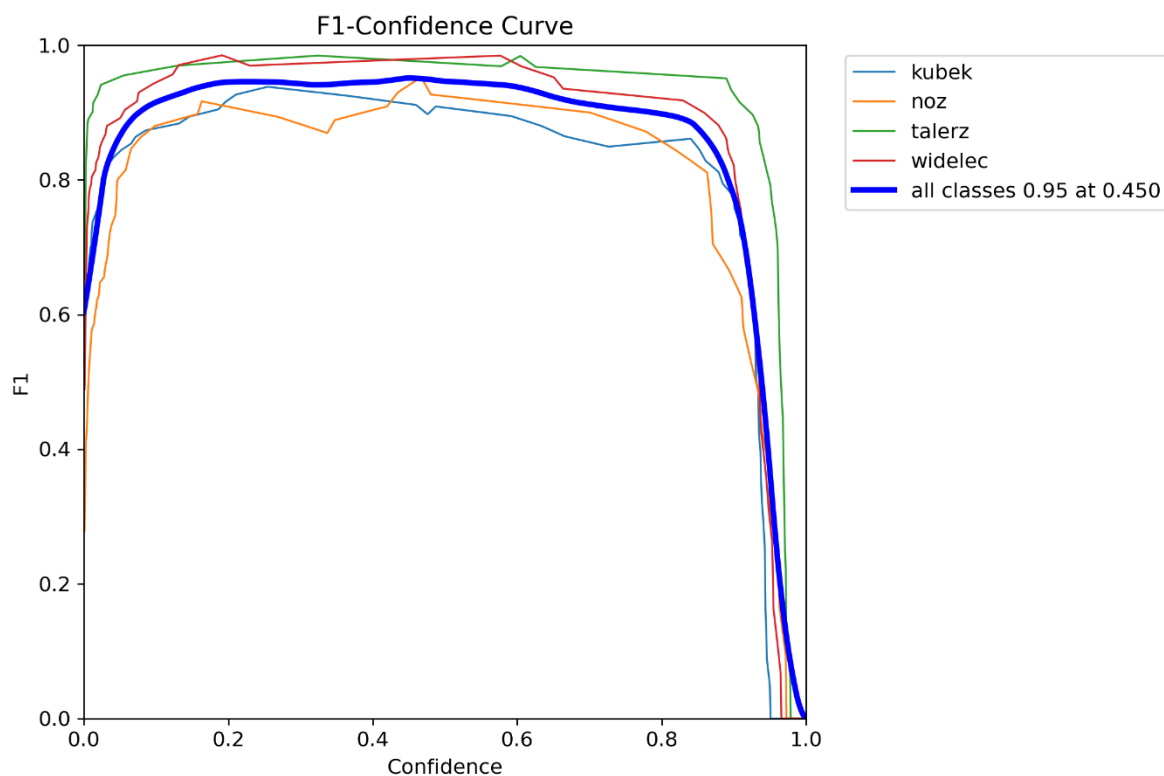


Model wykazuje wysoką skuteczność w klasyfikacji większości klas: "kubek" (88%), "nóż" (95%), "talerz" (97%) i "widelec" (97%). Jednak w przypadku "kubka" 9% przykładów zostało błędnie sklasyfikowanych jako tło, a w przypadku "noża" 2% zostało sklasyfikowanych jako "kubek".

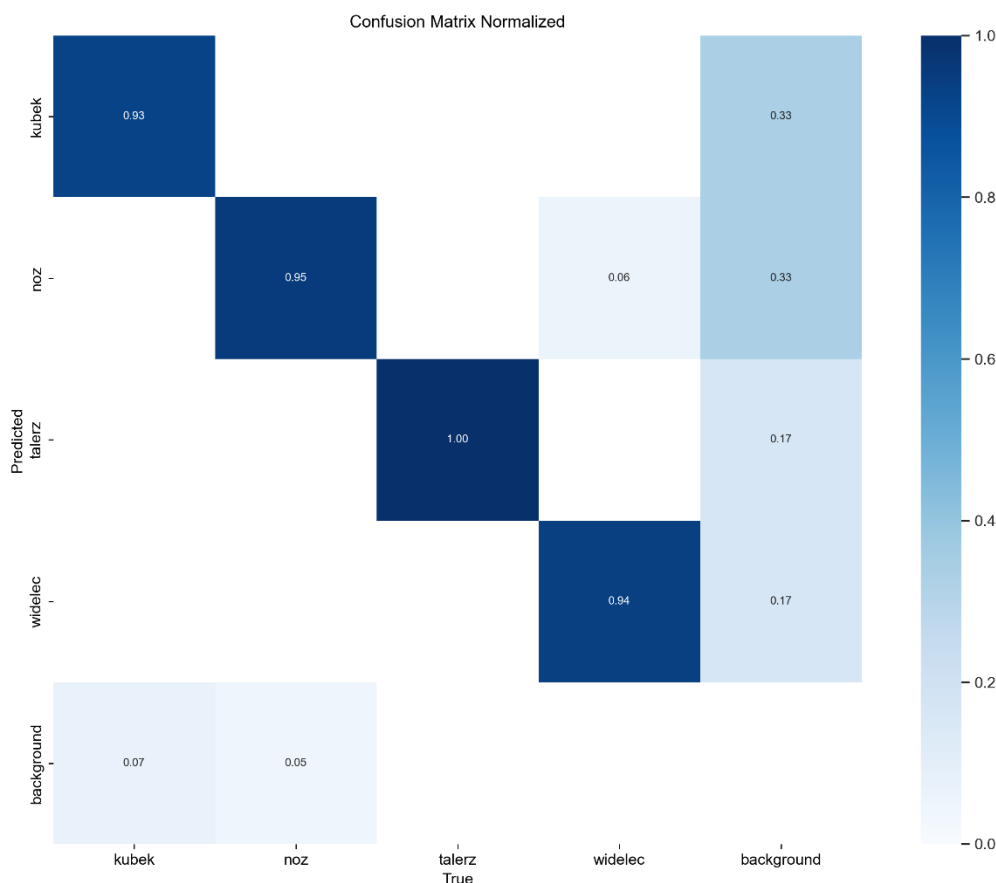
4. Model YOLOV8m



Ocena modelu:



Uśredniona krzywa dla wszystkich klas (gruba niebieska linia) osiąga wartość F1 równą 0.95 przy poziomie pewności 0.450, co jest porównywalne z modelem YOLOv8s. Poszczególne klasy również osiągają wysokie wartości F1, z "talerzem" i "widelec" przewodzącymi, a "nóż" wykazuje nieco niższą, ale nadal wysoką skuteczność.



Model wykazuje wysoką skuteczność w klasyfikacji większości klas: "kubek" (93%), "nóż" (95%), "talerz" (100%) i "widelec" (94%). Klasa "kubek" ma 7% błędnych klasyfikacji do tła, a klasa "nóż" ma 5% błędnych klasyfikacji do tła. Ponadto, 33% przypadków klasy "kubek" zostało błędnie sklasyfikowanych jako tło, a 6% przypadków klasy "nóż" zostało błędnie sklasyfikowanych jako "kubek".

Wnioski:

Na podstawie przeprowadzonej analizy trzech modeli YOLOv8n, YOLOv8s i YOLOv8m, można wyciągnąć następujące wnioski dotyczące ich skuteczności w klasyfikacji obiektów w zbiorze danych "kuchnia". Model YOLOv8n osiągnął wartość F1 równą 0.93 przy poziomie pewności 0.280, natomiast YOLOv8s i YOLOv8m osiągnęły wyższe wartości F1, odpowiednio 0.95 przy poziomie pewności 0.562 i 0.450, co wskazuje na ogólną poprawę wydajności w miarę wzrostu złożoności modelu. Macierze konfuzji dla modeli YOLOv8s i YOLOv8m pokazują, że oba modele wykazują wysoką skuteczność w klasyfikacji większości klas, jednak YOLOv8m osiąga najwyższą dokładność, szczególnie w przypadku klasy "talerz" z 100% poprawnych klasyfikacji. Mimo to, wszystkie modele miały pewne trudności z rozróżnieniem "kubków" od tła. Podsumowując, model YOLOv8m okazał się najskuteczniejszy, oferując najlepszą kombinację precyzji i czułości, co czyni go najbardziej efektywnym narzędziem do rozpoznawania obiektów w analizowanym zbiorze danych.