



# Głębokie uczenie wielozadaniowe dla semantycznej analizy wnętrz

---

Piotr Hondra

promotor: mgr inż. Maciej Stefańczyk  
kierunek: Automatyka i Robotyka  
Instytut Automatyki i Informatyki Stosowanej

# Cel pracy

# Cel pracy

## Cel pracy

Zbadanie problemu wspólnej segmentacji semantycznej i klasyfikacji scen we wnętrzach.

# Cel pracy

## Cel pracy

Zbadanie problemu wspólnej segmentacji semantycznej i klasyfikacji scen we wnętrzach.

- Opracowanie architektury wielozadaniowej

# Cel pracy

## Cel pracy

Zbadanie problemu wspólnej segmentacji semantycznej i klasyfikacji scen we wnętrzach.

- Opracowanie architektury wielozadaniowej
- Przeprowadzenie łącznego uczenia

# Cel pracy

## Cel pracy

Zbadanie problemu wspólnej segmentacji semantycznej i klasyfikacji scen we wnętrzach.

- Opracowanie architektury wielozadaniowej
- Przeprowadzenie łącznego uczenia
- Ewaluacja

# Segmentacja semantyczna oraz klasyfikacja sceny



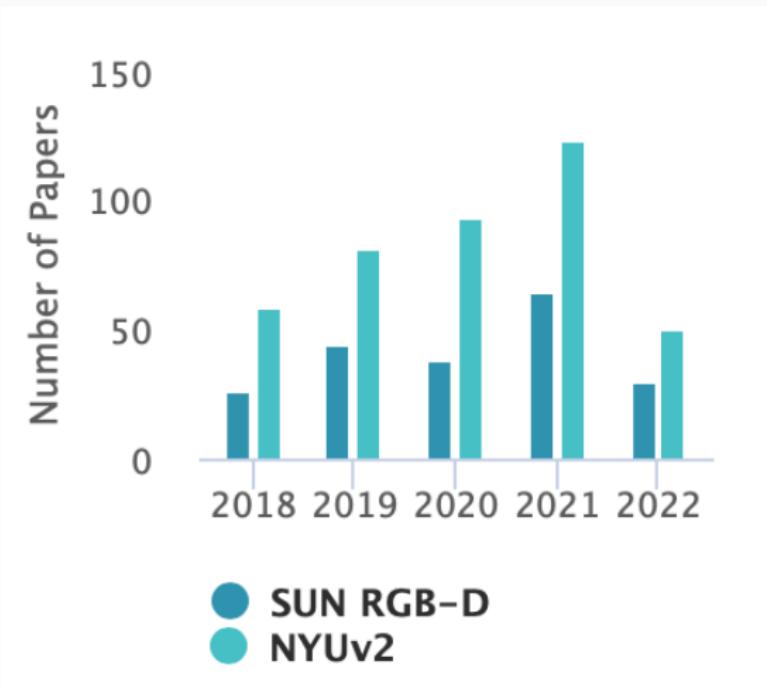
Figure 1: [1].

- Robotyka
  - nawigacja
  - manipulacja obiekttami
  - interakcja człowiek-robot
- Zarządzanie budynkiem
  - zarządzanie energią
  - bezpieczeństwo
- Rozszerzona rzeczywistość

## Zbiór danych

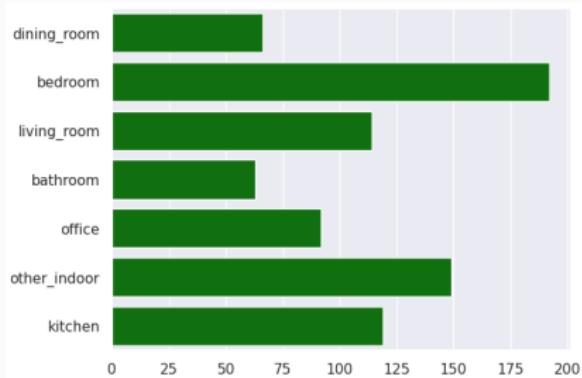
---

## Zbiór danych

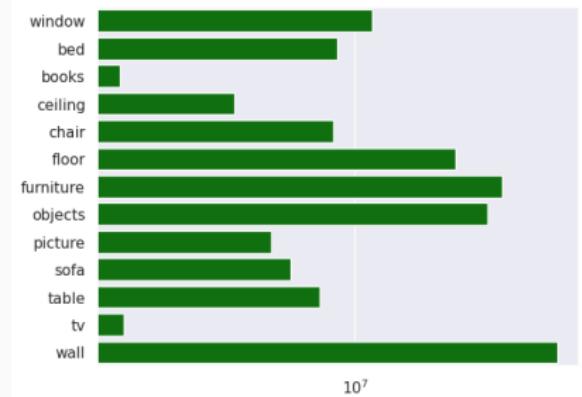


**Figure 2:** Szacowana liczba cytowań w latach 2018-2022 [paperswithcode.com]

# Zbiór danych



**Figure 3:** Histogram klas dla zadania klasyfikacji sceny.



**Figure 4:** Histogram pixeli klas dla zadania segmentacji semantycznej.

## Rozwiążanie problemu

---

# Rozwiążanie problemu - przedstawienie architektury

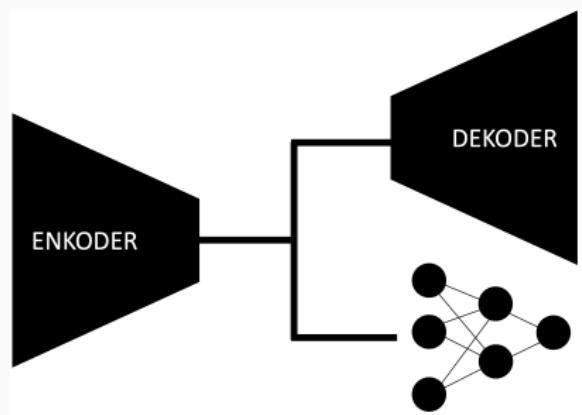


**Figure 5:** Architektura sieci zastosowana w pracy inżynierskiej jako model czarnej skrzynki.

## Opis eksperymentów

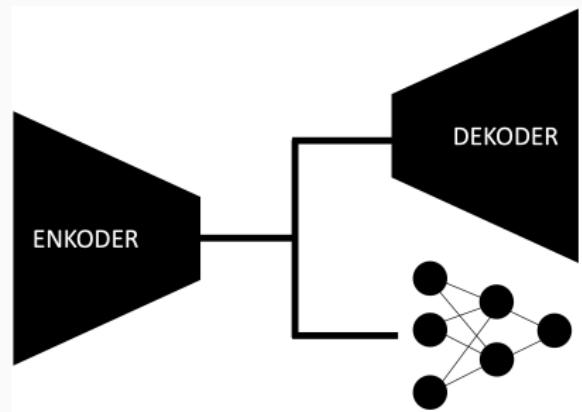
---

- Uczenie wielozadaniowe



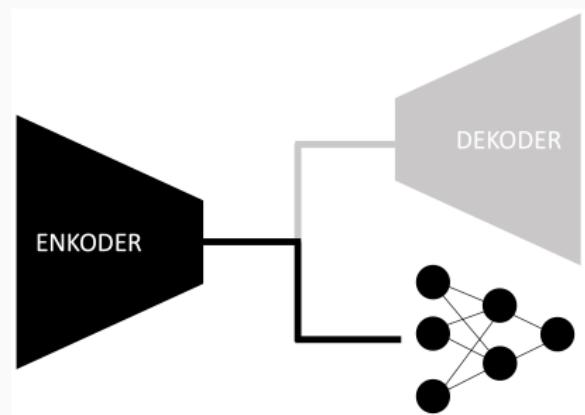
# Opis eksperymentów

- Uczenie wielozadaniowe
- Finetuning



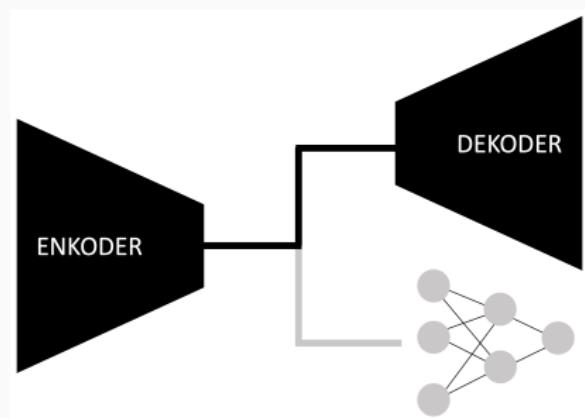
# Opis eksperymentów

- Uczenie wielozadaniowe
- Finetuning
- Wyłącznie klasyfikacja



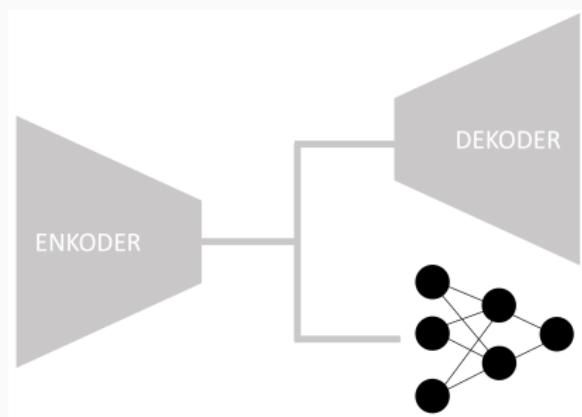
# Opis eksperymentów

- Uczenie wielozadaniowe
- Finetuning
- Wyłącznie klasyfikacja
- Wyłącznie segmentacja



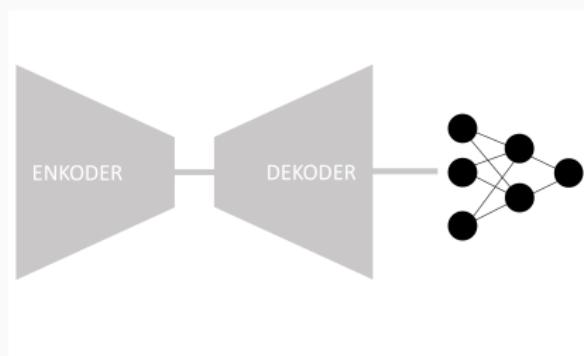
# Opis eksperymentów

- Uczenie wielozadaniowe
- Finetuning
- Wyłącznie klasyfikacja
- Wyłącznie segmentacja
- Pośrednia klasyfikacja z segmentacji



# Opis eksperymentów

- Uczenie wielozadaniowe
- Finetuning
- Wyłącznie klasyfikacja
- Wyłącznie segmentacja
- Pośrednia klasyfikacja z segmentacji
- Bezpośrednia klasyfikacja z segmentacji



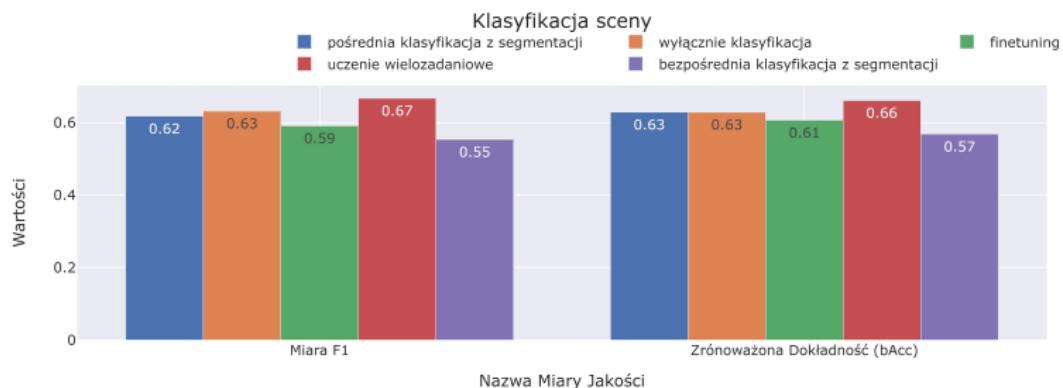
# Wyniki

---

## Analiza jakościowa

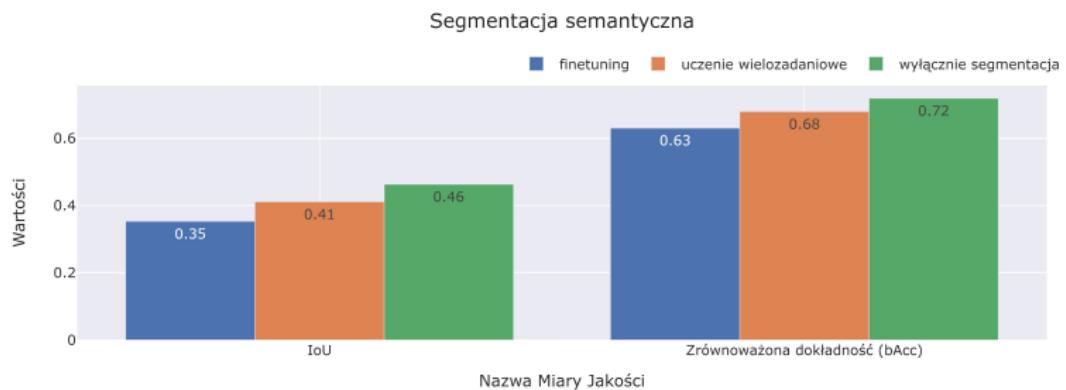
---

# Klasyfikacja sceny



**Figure 6:** Porównanie miar F1 oraz dokładności dla klasyfikacji sceny.

# Segmentacja semantyczna



**Figure 7:** Porównanie miar IoU oraz dokładności dla segmentacji sceny.

## Analiza wydajnościowa

---

## Analiza czasu uczenia

nazwa zadania	$\approx$ czas[s]
wyłącznie segmentacja + wyłącznie klasyfikacja	360
wyłącznie segmetnacja + pośrednia klasyfikacja	260
wyłącznie segmetnacja + bezpośrednia klasyfikacja	275
uczenie wielozadaniowe	211
finetuning	160

**Table 1:** Porównanie czasu uczenia względem całości.

## Analiza czasu wnioskowania

rodzaj sieci	czas[s]
dwie szeregowe sieci	15.7
jedna architektura	8.6

**Table 2:** Porównanie czasu wnioskowania.

# Wyniki

---

## Wady

- Pogorszenie segmentacji

## Zalety

- Poprawa klasyfikacji
- Skrócenie czasu uczenia
- Dwukrotne skrócenie czasu wnioskowania

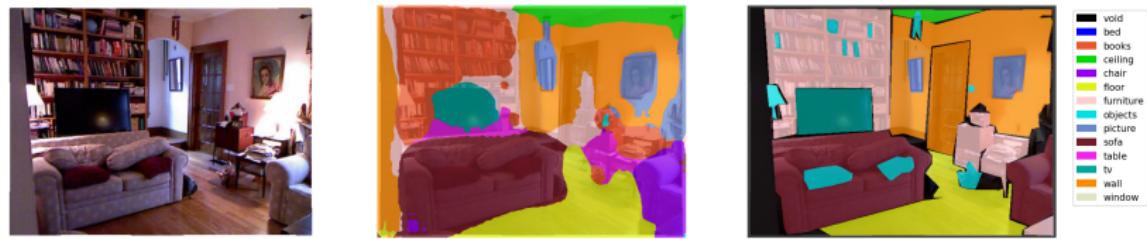
# Podsumowanie

Cel pracy został spełniony, dodatkowo odpowiadając na inne pytania badawcze.

- Jak można zaprojektować model oparty na głębokim uczeniu do wspólnej segmentacji semantycznej i klasyfikacji scen w środowiskach wewnętrznych?
- Czy przestrzeń reprezentacji po wytrenowaniu na zadaniu segmentacji semantycznej może być użyta do zadania klasyfikacji sceny?
- Jak dobrze proponowany model radzi sobie na dużym zbiorze danych scen wewnętrznych i jak wypada w porównaniu z aktualnymi metodami segmentacji semantycznej i klasyfikacji scen osobno?
- Jak proponowany model może być wykorzystany do poprawy wydajności w robotyce mobilnej?

-  H. Zhang, K. Dana, J. Shi, Z. Zhang, X. Wang, A. Tyagi, and A. Agrawal.  
**Context encoding for semantic segmentation.**  
In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 7151–7160, 2018.

# Obraz



**Figure 8:** Porównanie jakości segmentacji dla klasy salon.

# Architektura wielozadaniowej

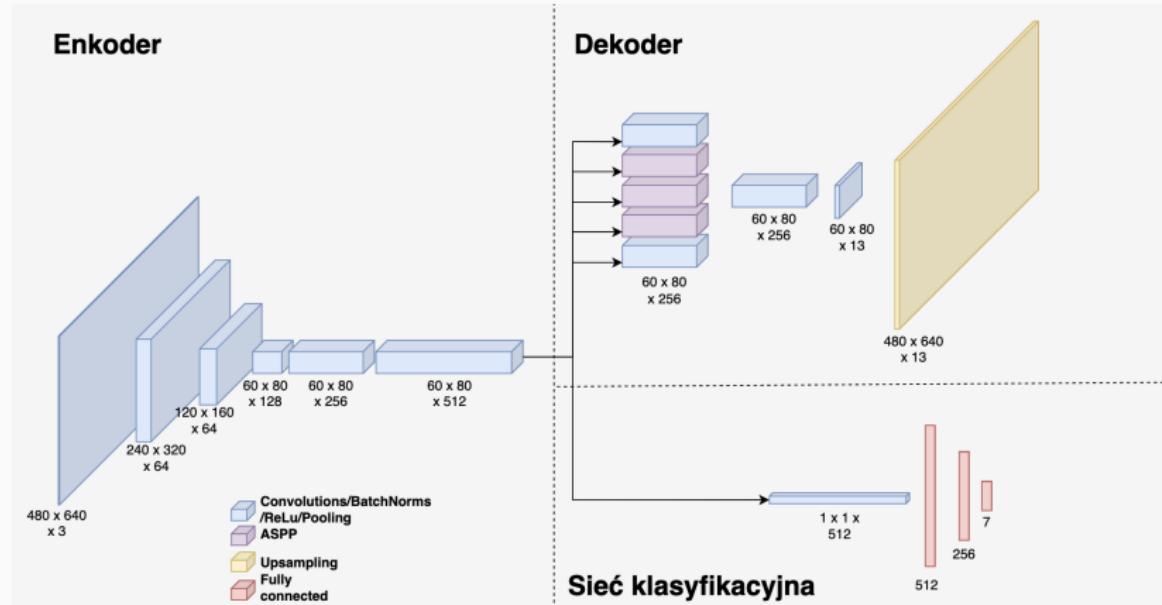
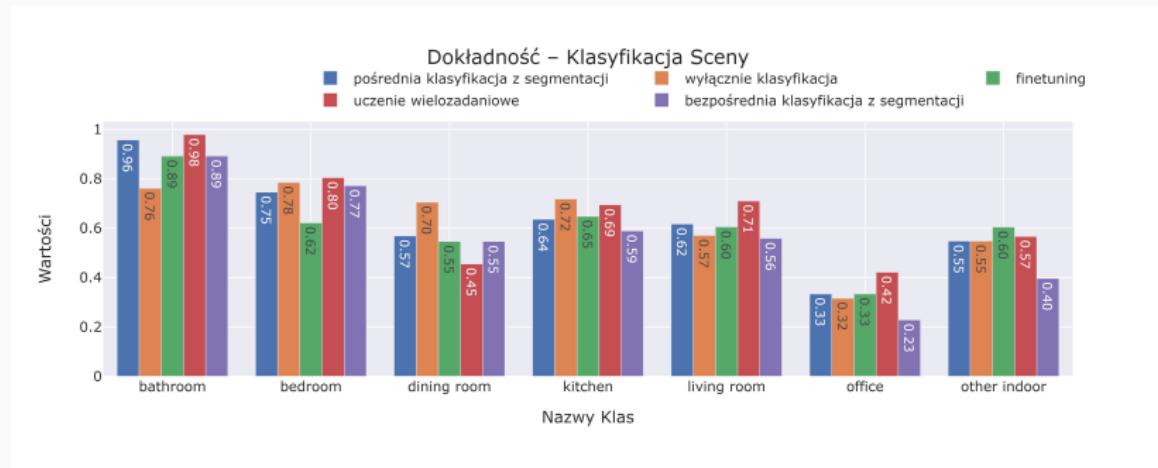


Figure 9: Architektura wielozadaniowej sieci.

# Miary jakości klasyfikacji



**Figure 10:** Porównanie dokładności klasyfikacji sceny z rozróżnieniem konkretnych klas.

# Miary jakości klasyfikacji

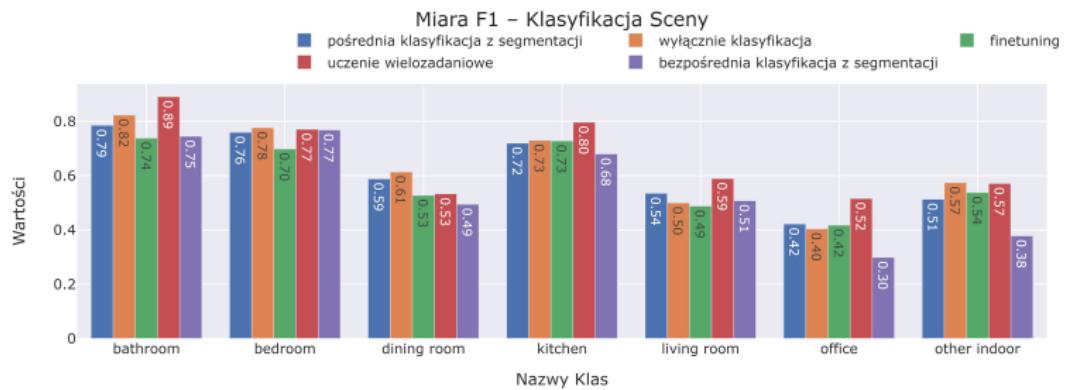


Figure 11: Porównanie miary F1 dla klasyfikacji sceny z rozróżnieniem konkretnych klas.

# Miary jakości segmentacji

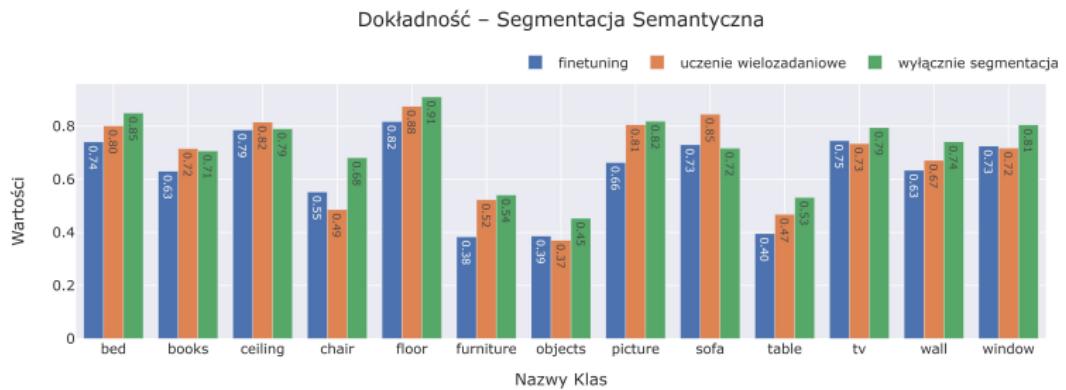


Figure 12: Porównanie dokładności segmentacji z rozróżnieniem konkretnych klas.

# Miary jakości segmentacji

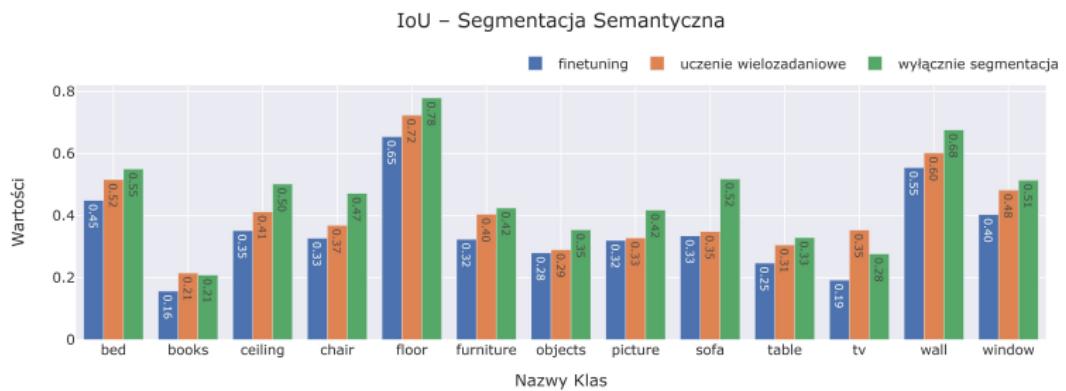


Figure 13: Porównanie miary IoU segmentacji z rozróżnieniem konkretnych klas.