

# Segmentacja zdjęć satelitarnych

---

## Autorzy

Witold Nowogórski, Szymon Jurecki, Arkadiusz Paterak

## Metody rozwiązania problemu

Dwie klasy - mapowanie terenów zabudowanych

- klasyfikacja binarna: teren niezabudowany (lasy, pola, woda), zabudowania (miasta, drogi, wsie)
- zbiór danych stworzony samodzielnie, wykorzystując zdjęcia satelitarne pobrane za pomocą oprogramowania geoinformacyjnego QGIS, które następnie zostały podzielone na fragmenty 50x50 pixeli i oznaczane odpowiednią klasą przy użyciu prostego interfejsu użytkownika w PyQt.
- model oparty na architekturze ResNet-34 z wykorzystaniem wag dostępnych w bibliotece PyTorch.

Segmentacja semantyczna przy pomocy modelu U-Net

- zbiór danych: [LandCover.ai v1](#); zdjęcia satelitarne z Polski wraz z adnotacjami w postaci masek
- klasy: inne (0), zabudowania (1), tereny leśne (2), woda (3), drogi (4)
- model oparty na architekturze U-Net, zaimplementowany samodzielnie w bibliotece PyTorch

## Napotkane trudności i ich rozwiązania

- dobór odpowiednich parametrów trenowania - ilość epok, rozmiar batch-a: model oparty na ResNet-34 po około 5 epokach, przy użyciu 3000 zdjęć i rozmiaru batcha 64, był w stanie osiągnąć dokładność 90% w klasyfikacji teren zabudowany / teren niezabudowany. Zastosowanie techniki early stopping pozwoliło uniknąć przetrenowania.
- stworzenie odpowiedniego zbioru danych do trenowania modelu klasyfikacji binarnej.
  - rozwiązanie: wystarczająca wielkość (~3000 zdjęć), równy rozkład klas,
- zbyt duży rozmiar zdjęć ze zbioru LandCover.ai spowalniał uczenie modelu U-Net
  - rozwiązanie: zdjęcia są skalowane do rozmiaru 64x64
- dobranie odpowiedniej funkcji straty dla modelu U-Net działającego dla wielu klas i określenie właściwych wymiarów wyjściowej warstwy (w wielu przykładach dokonywano segmentacji dla dwóch klas)
  - rozwiązanie: zastosowanie funkcji straty `nn.CrossEntropyLoss` oraz wyjściowej warstwy o wymiarach `[batch_size, n_classes, image_height, image_width]` (dla przypadku binarnego wystarczy jeden kanał w miejsce `n_classes`); inferencja dokonywana poprzez użycie funkcji `softmax`, a następnie wybór klasy o najwyższym prawdopodobieństwie

## Kontrybucje autorów

Witold Nowogórski był odpowiedzialny za implementację klasyfikatora binarnego ResNet34, od pobierania zdjęć satelitarnych przez QGIS, następnie przez ich przetwarzanie, do trenowania modelu. Szymon Jurecki przygotował zbiór danych i go zaetykietował. Arkadiusz Paterak zaproponował użycie modelu U-Net, który zaimplementował i wytrenował, stworzył także prosty interfejs do zaprezentowania działania modelu. Wszyscy autorzy wspólnie przygotowali raport.