

Czy można nauczyć komputer znajdować budynki na zdjęciach satelitarnych?

Andrzej Pyskir, Bydgoszcz, 24 kwietnia 2019









• Powiedzieć: "Znalazłem!" - klasyfikacja





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować
- Pokolorować segmentacja





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować
- Pokolorować segmentacja
- Pokolorować każdy obiekt danej klasy segmentacja obiektów





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować
- Pokolorować segmentacja
- Pokolorować każdy obiekt danej klasy segmentacja obiektów
- Wziąć w kółko





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować
- Pokolorować segmentacja
- Pokolorować każdy obiekt danej klasy segmentacja obiektów
- Wziąć w kółko
- Opisać prostokat na przedmiocie detekcja





- Powiedzieć: "Znalazłem!" klasyfikacja
- Pokazać palcem
- Opisać, np. "Indyk znajduje się w środku obrazka" - przypisywanie nagłówka
- Obrysować
- Pokolorować segmentacja
- Pokolorować każdy obiekt danej klasy segmentacja obiektów
- Wziąć w kółko
- Opisać prostokąt na przedmiocie detekcja
- ...

#### Nasze zadanie - segmentacja obiektów będących budynkami

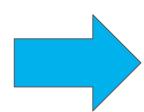














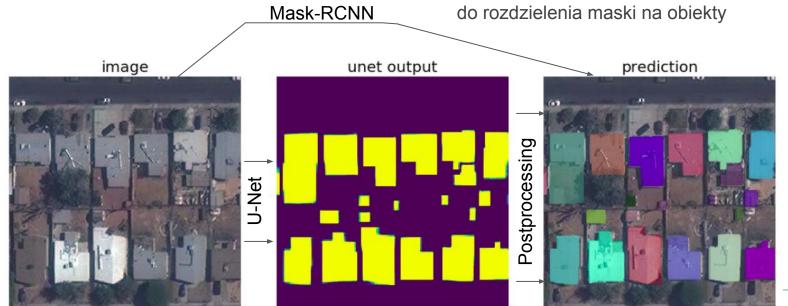
Link do rozwiązania: <a href="https://github.com/neptune-ml/open-solution-mapping-challenge">https://github.com/neptune-ml/open-solution-mapping-challenge</a>



#### Jak przeprowadzić segmentację obiektów?

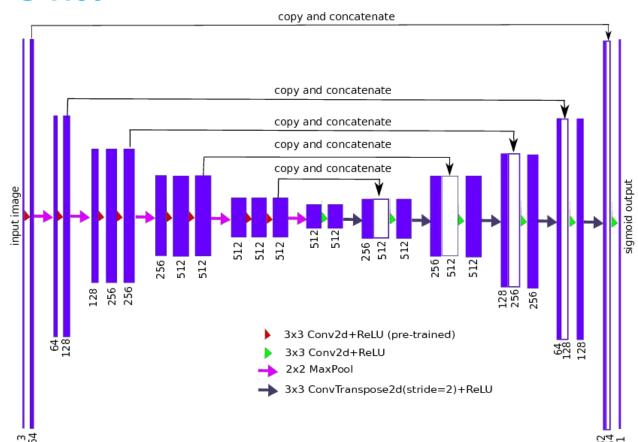
- Użyć modelu do segmentacji obiektów :-)
- Np. Mask-RCNN

- Użyć modelu do segmentacji otrzymujemy jedną maskę wspólną dla wszystkich obiektów
- Podział maski na poszczególne obiekty
- Np. U-Net do segmentacji + postprocessing do rozdzielenia maski na obiekty



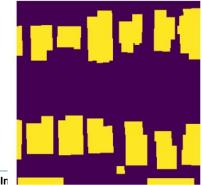
#### **U-Net**





Część "zstępująca", czyli zmniejszająca rozdzielczość, może być podmieniona. W naszym przypadku był to ResNet-152 pretrenowany na Image Net, co znacznie poprawia wyniki.

Do treningu modelu do segmentacji potrzeba maski z przyporządkowaniem: piksel ← klasa, np.:





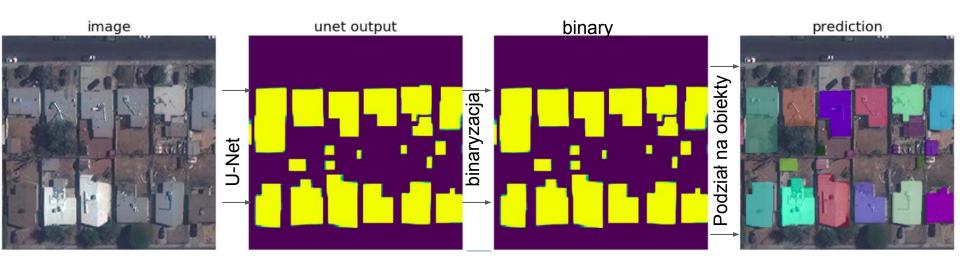
Jak dobre musi być wyjście z U-Neta, żeby było satysfakcjonujące?

Co w zasadzie zwraca U-Net?

A jaki będzie postprocessing?

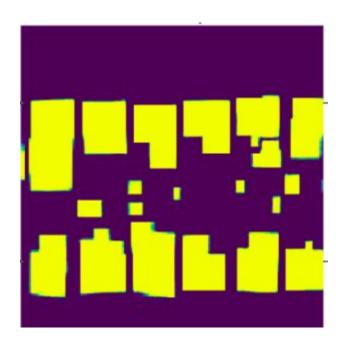
Macierz prawdopodobieństw przynależności piksela do klasy "budynek"

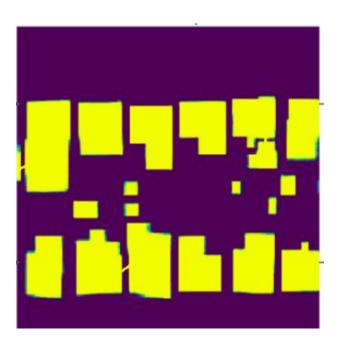
Binaryzacja + prosty podział na obiekty





## Jak dobre musi być wyjście z U-Neta, żeby było satysfakcjonujące?





deepsense.ai

Jak zmusić model do nauczenia się małych obiektów oraz rozdzielania obiektów leżących blisko siebie?

Funkcja straty:

BOB ~ 0.0 0.2 0.4 0.6 0.8 1.0 
$$\hat{y}_i$$

$$BCE = -\frac{1}{N} \sum_{i=0}^{N} y_i \cdot log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i)$$

BCE - entropia krzyżowa (binary cross entropy)

N - liczba pikseli na obrazku

y<sub>i</sub> - etykieta i-tego piksela (0 - tło, 1 - budynek)

 $\hat{y}_i$  - prawdopodobieństwo, że i-ty piksel należy do budynku



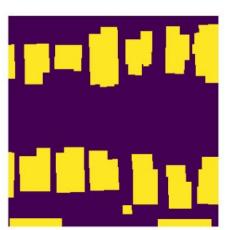
## Jak zmusić model do nauczenia się małych obiektów oraz rozdzielania obiektów leżących blisko siebie?

!!! Wprowadzamy różne wagi dla poszczególnych pikseli !!!

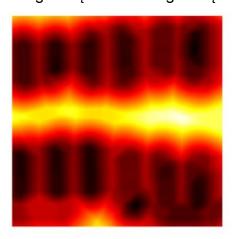
Obraz



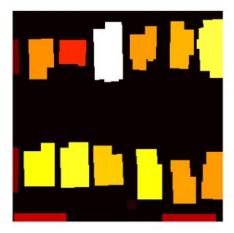
Maska



Wagi związane z odległością

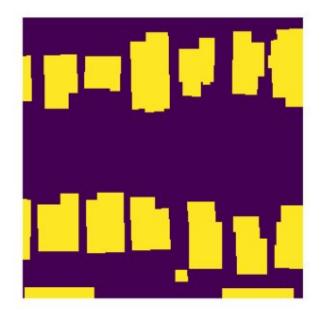


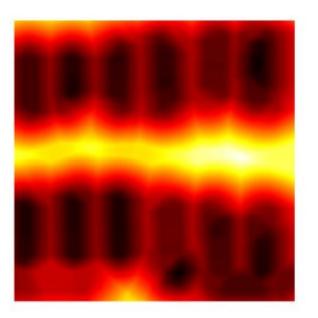
Wagi związane z wielkością





#### Wagi vs maska

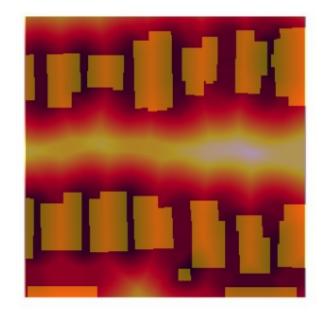




\_

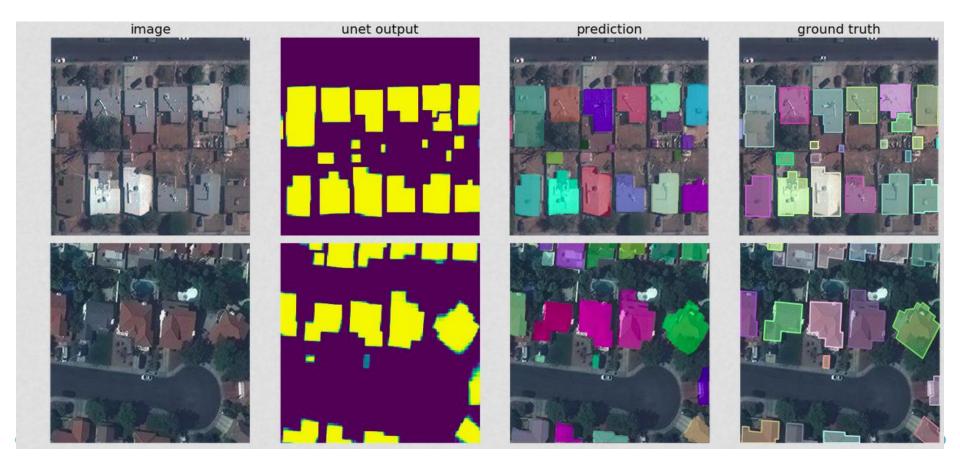


#### Wagi vs maska



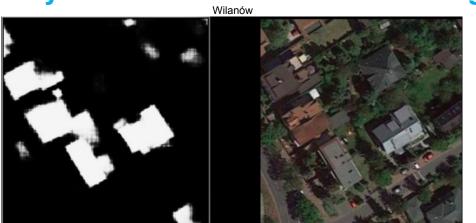
### Wyniki





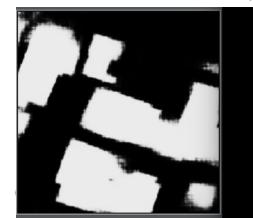


### Czy ten model dobrze działa na Google Maps?





Włochy (warszawskie) Aleppo











#### Dlaczego tak źle? Czy da się to poprawić?







Imagery ©2019 , DigitalGlobe, U.S. Geological Survey, USDA Farm Service Ager



### Model dotrenowany na Google Maps + Open Street Map

**San Francisco** 

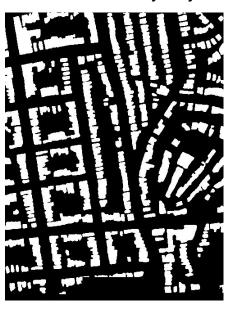




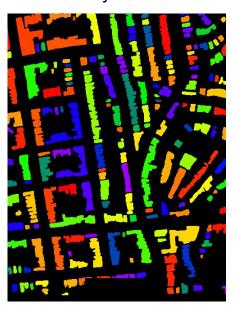
Wyjście U-Neta

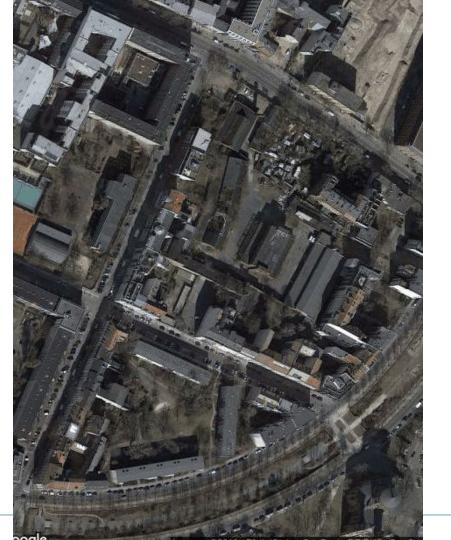


Po binaryzacji



Wynik









## Model dotrenowany na Google Maps + Open Street Map Berlin

Obraz



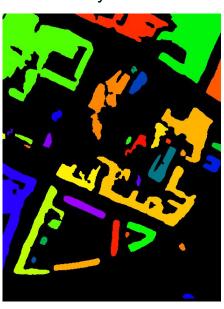
Wyjście U-Neta



Po binaryzacji



Wynik





## deepsense.ai



## Model dotrenowany na Google Maps + Open Street Map Bydgoszcz

Obraz



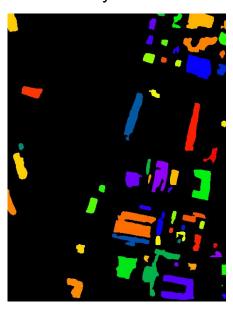
Wyjście U-Neta

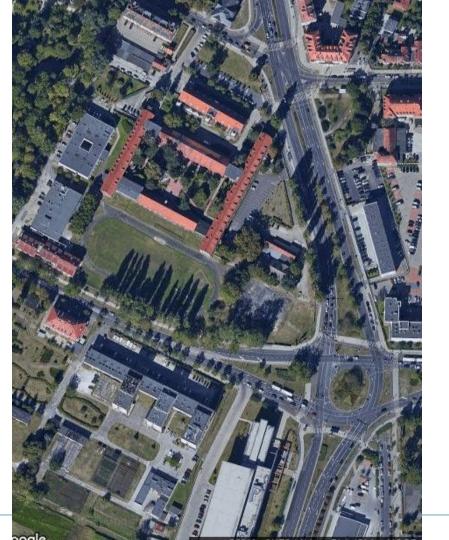


Po binaryzacji



Wynik





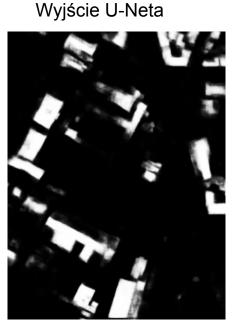
## deepsense.ai



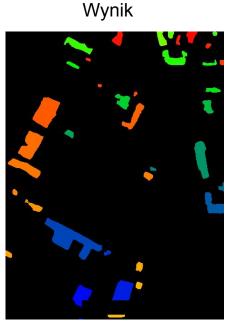
## Model dotrenowany na Google Maps + Open Street Map

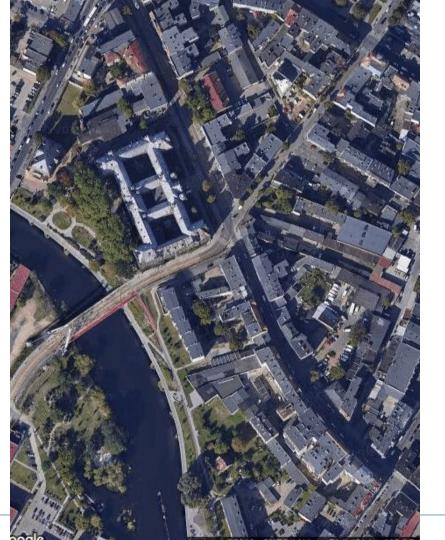
**Bydgoszcz** 









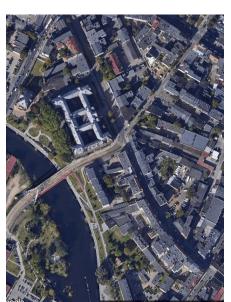






#### Model dotrenowany na Google Maps + Open Street Map **Bydgoszcz**

Wyjście U-Neta Po binaryzacji Obraz

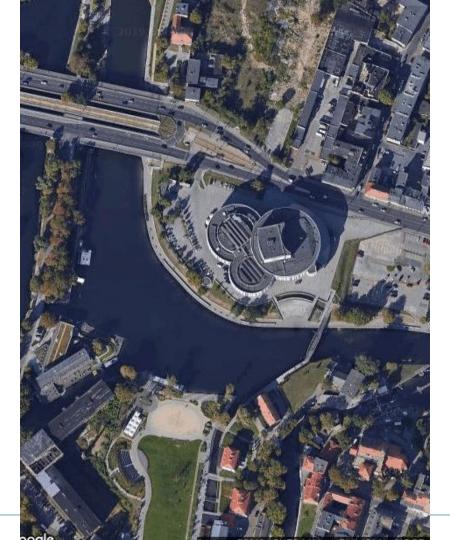














## Model dotrenowany na Google Maps + Open Street Map

**Bydgoszcz** 



Wyjście U-Neta











## Model dotrenowany na Google Maps + Open Street Map Mount Everest

Wyjście U-Neta Po binaryzacji Obraz Wynik



## Model nauczy się tego, co dostanie w treningu!



#### Dziękuję!

Andrzej Pyskir

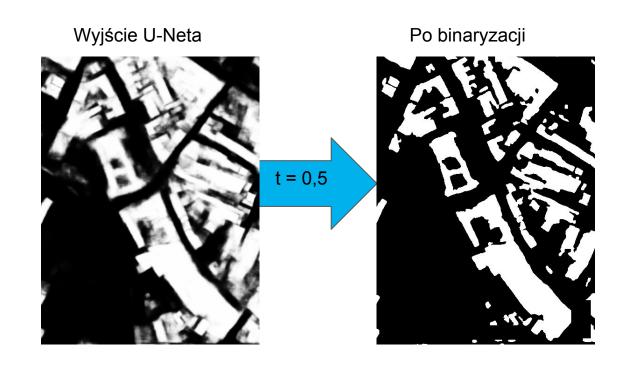
andrzej.pyskir@deepsense.ai



# Backup - Model drugiego rzędu

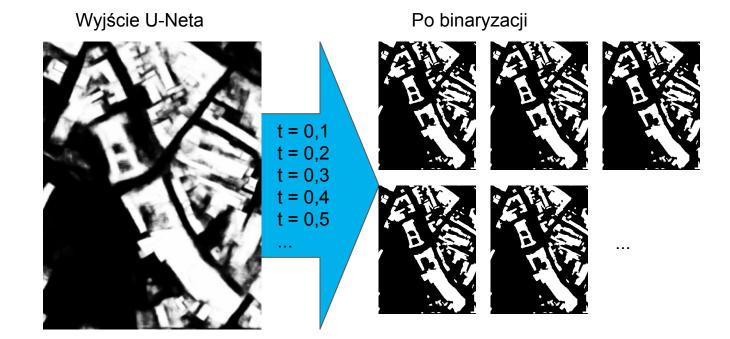


#### Jaka binaryzacja jest najlepsza?





#### Jaka binaryzacja jest najlepsza? - niech zdecyduje model!



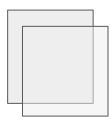
\_



## Jaka binaryzacja jest najlepsza? - niech zdecyduje model!

#### Procedura:

- Wybieramy n progów prawdopodobieństwa: t<sub>1</sub>, t<sub>2</sub>, ..., t<sub>n</sub>
- Otrzymujemy n binarnych masek, każda dla innego progu prawdopodobieństwa
- Każdą maskę dzielimy na predykcje budynków, dla i-tego progu mamy m, predykcji
- Dla każdej predykcji sprawdzamy jak jest dobra jakie ma loU z prawdziwą maską



IoU = 75 %

- Trenujemy model, który dla każdej predykcji przewiduje IoU na podstawie cech takich jak:
  - Pole
  - Obwód
  - Średnie, minimalne, maksymalne prawdopodobieństwo na masce
  - Stosunek pola do obwodu
  - Itp...
- Otrzymujemy zestaw predykcji wraz z ich przewidywanymi IoU z prawdziwym budynkiem
- Dla pokrywających się predykcji wybieramy tą z najwyższym przewidywanym IoU