

Detekcja i segmentacja obrazów

Wojciech Kozłowski

partner instytucjonalny:

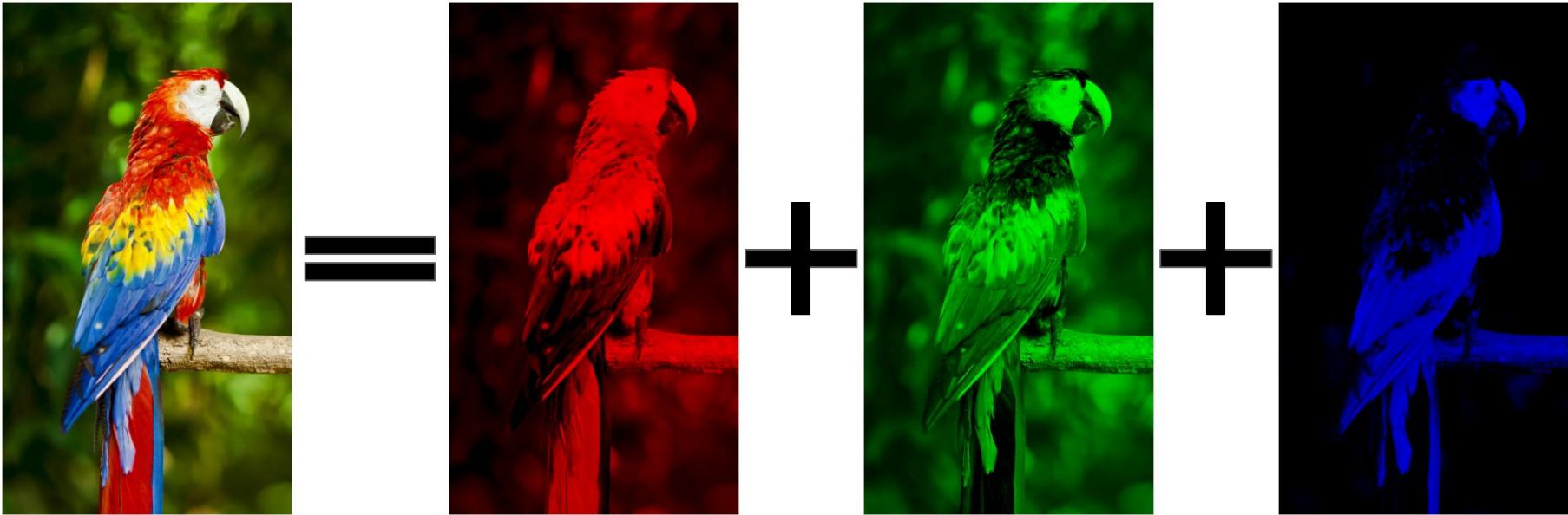


Wrocławskie
Centrum
Akademickie

Agenda

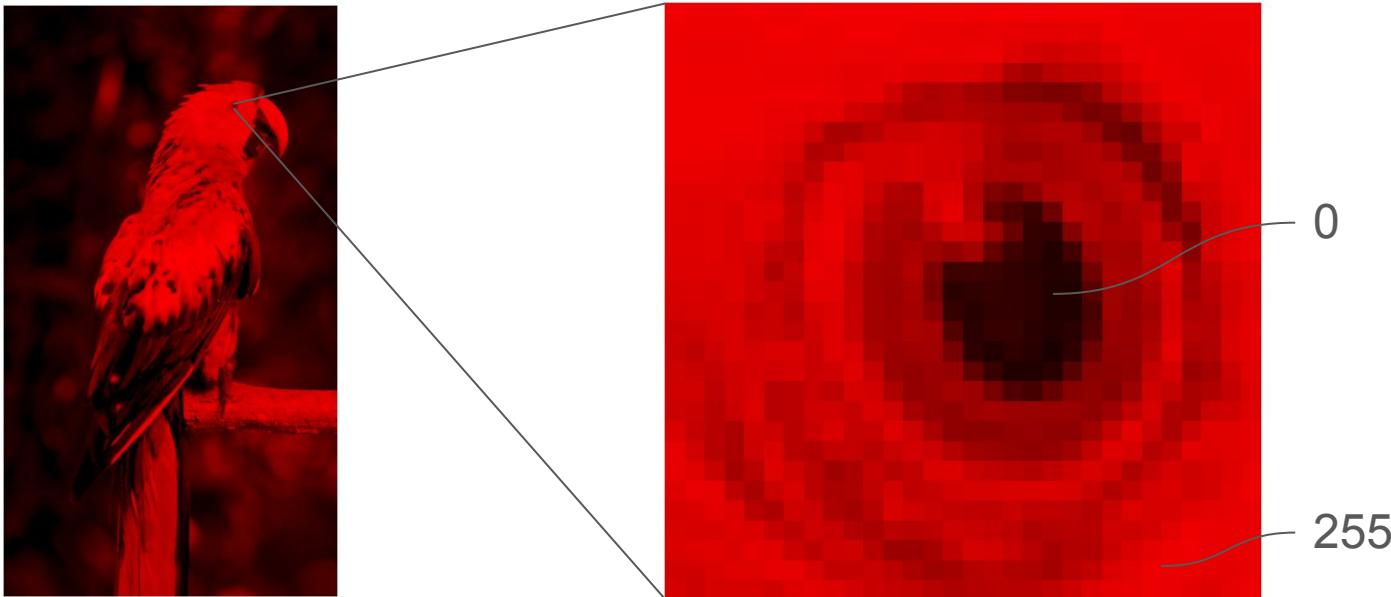
- Sieci konwolucyjne - przypomnienie
- Wstęp do widzenia komputerowego
- Problem detekcji obiektów na obrazach
- Problem segmentacji obrazów

Jak sieci widzą obraz



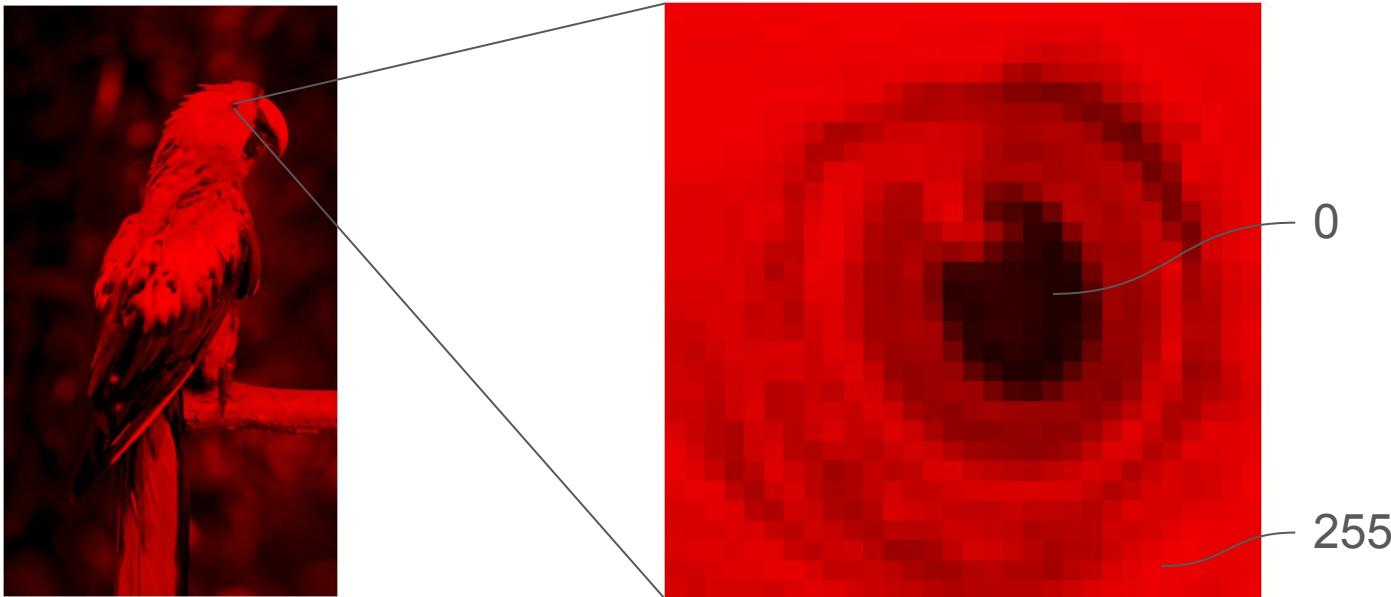
Kolorowy obraz zazwyczaj reprezentujemy jako połączenie trzech składowych kolorów (kanałów): czerwonego, zielonego i niebieskiego (**RGB**).

Jak sieci widzą obraz



W standardowych obrazach, każdy z kanałów składa się z pikseli o wartościach całkowitych (0, ..., 255).

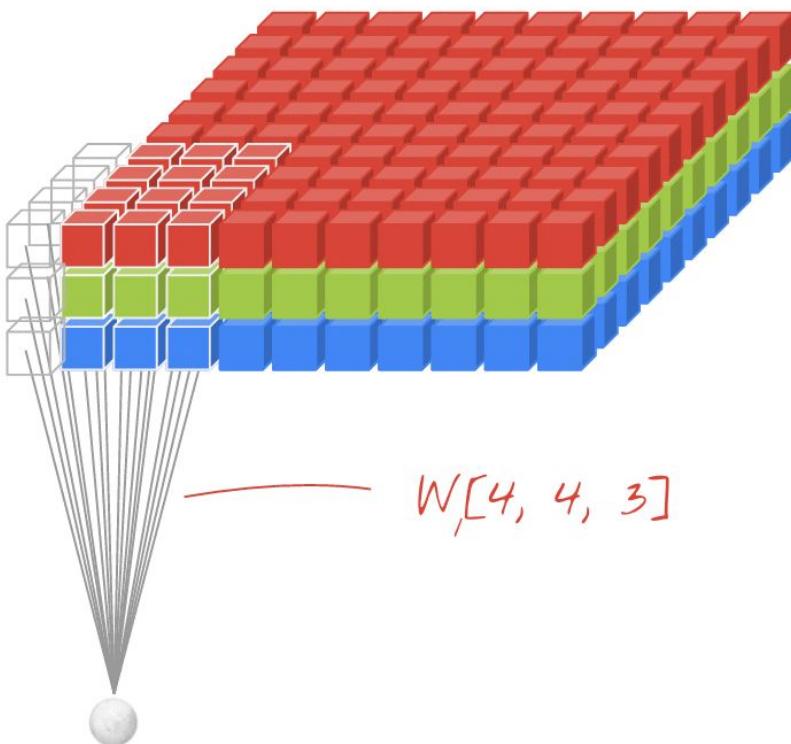
Jak sieci widzą obraz



W standardowych obrazach, każdy z kanałów składa się z pikseli o wartościach całkowitych (0, ..., 255).

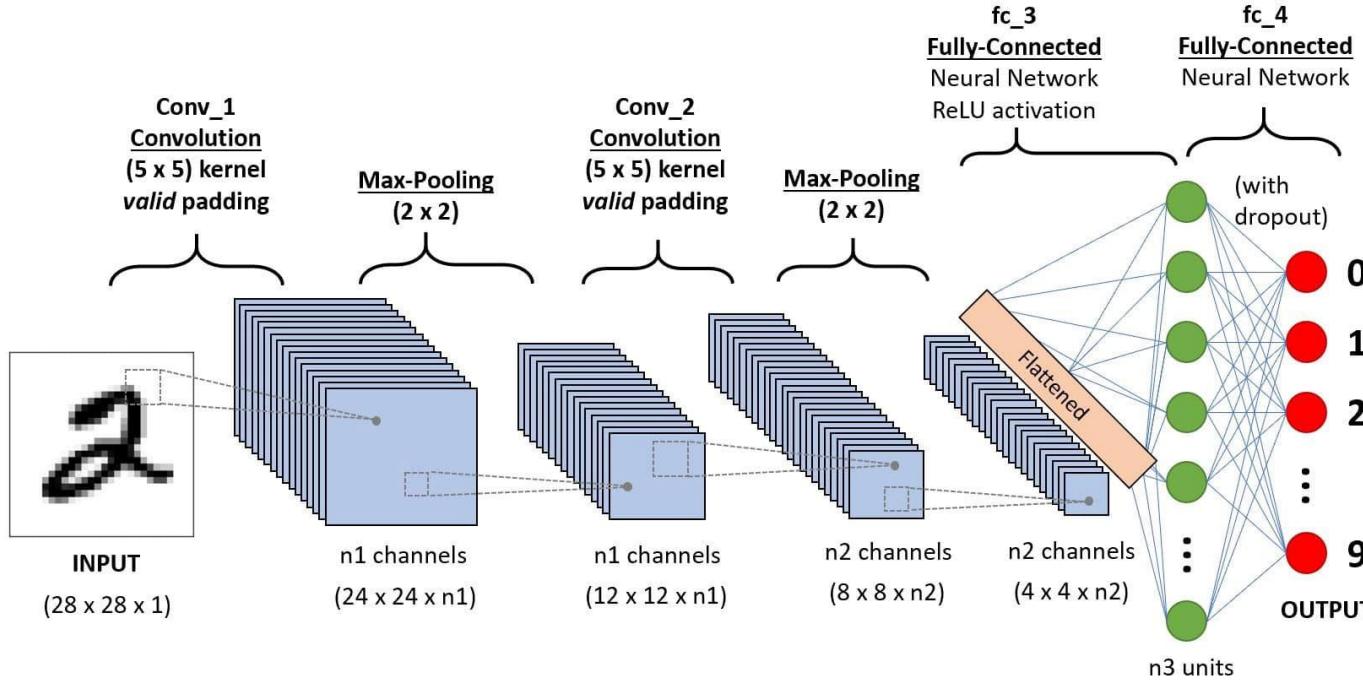
Przed przetworzeniem obrazu przez sieć neuronową warto dokonać normalizacji (np. dzieląc wszystkie wartości przez 255).

Jak sieci widzą obraz



- Zdjęcie możemy interpretować jako tensor 3D - trójwymiarową tablicę liczb.
- Taki tensor możemy przetwarzać z użyciem np. konwolucji.

Jak sieci widzą obraz



Sieci konwolucyjne składają się zazwyczaj z wielu operacji konwolucji oraz poolingu.

Widzenie komputerowe

Widzenie komputerowe

Widzenie komputerowe (ang. Computer Vision, CV) zajmuje się wyciąganiem informacji z obrazów.



Zdjęcie

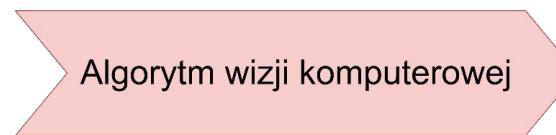


Widzenie komputerowe

Widzenie komputerowe (ang. Computer Vision, CV) zajmuje się wyciąganiem informacji z obrazów.



Zdjęcie



Zadania:

- **klasyfikacja**

Klasyfikacja



Zdjęcie



- 1 Kot
- 0 Pies
- 0 Krowa
- 0 Ryba

W zadaniu klasyfikacji chcemy przyporządkować obraz do jednej z wcześniej ustalonych **klas**.

Model stara się odpowiedzieć na pytanie: “**Co** przedstawia zdjęcie?”

Widzenie komputerowe

Widzenie komputerowe (ang. Computer Vision, CV) zajmuje się wyciąganiem informacji z obrazów.



Zdjęcie



Zadania:

- klasyfikacja
- detekcja

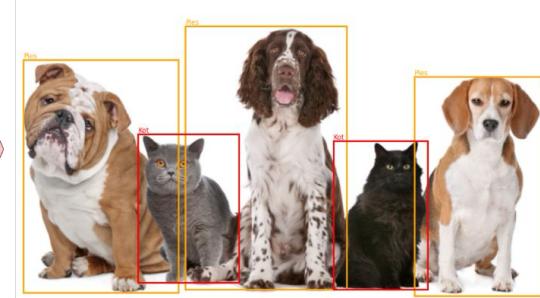
Detekcja



Zdjęcie



Detektor



Klasy i umiejscowienie obiektów na zdjęciu

W zadaniu detekcji chcemy **zlokalizować i zaklasyfikować wszystkie obiekty** znajdujące się na zdjęciu.
Model stara się odpowiedzieć na pytanie: "**Co i gdzie** znajduje się na zdjęciu?"

Widzenie komputerowe

Widzenie komputerowe (ang. Computer Vision, CV) zajmuje się wyciąganiem informacji z obrazów.



Zdjęcie



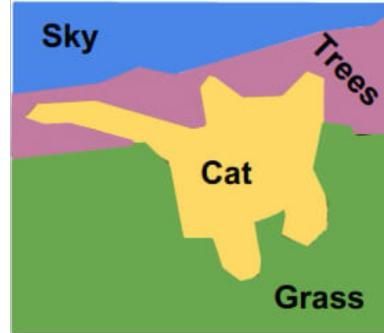
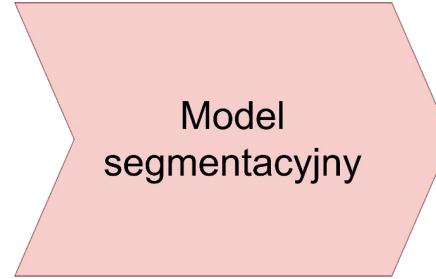
Zadania:

- klasyfikacja
- detekcja
- **segmentacja**

Segmentacja



Zdjęcie



Maska segmentacji

W zadaniu segmentacji chcemy przypisać każdemu pikselowi na zdjęciu **przynależność do** odpowiedniej **klasy**.

Model stara się odpowiedzieć na pytanie: "**Co, gdzie i o jakim kształcie** znajduje się na zdjęciu?"

Widzenie komputerowe

Widzenie komputerowe (ang. Computer Vision, CV) zajmuje się wyciąganiem informacji z obrazów.



Zdjęcie



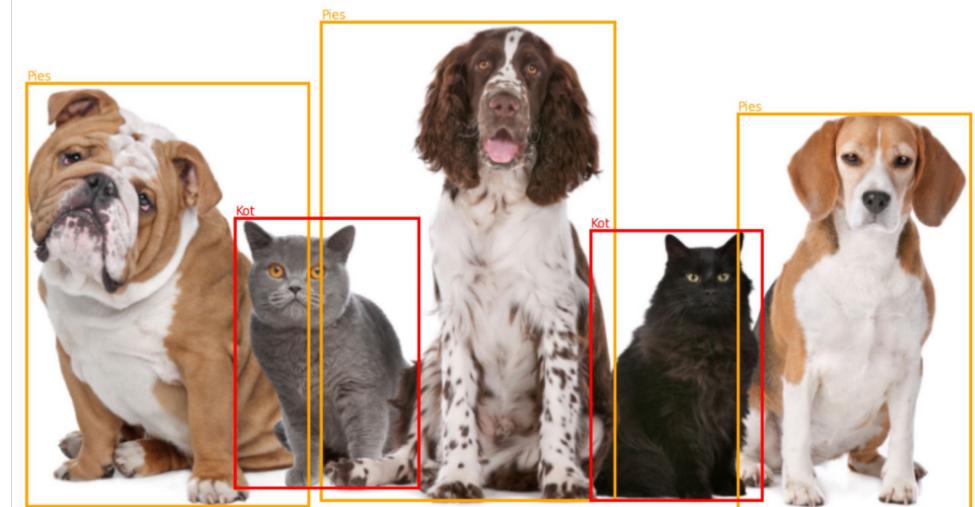
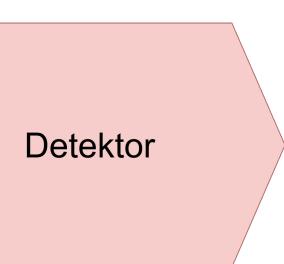
Zadania:

- klasyfikacja
- detekcja
- segmentacja
- inne (np. estymacja pozy, generowanie opisów zdjęć itp.)

Zadanie detekcji

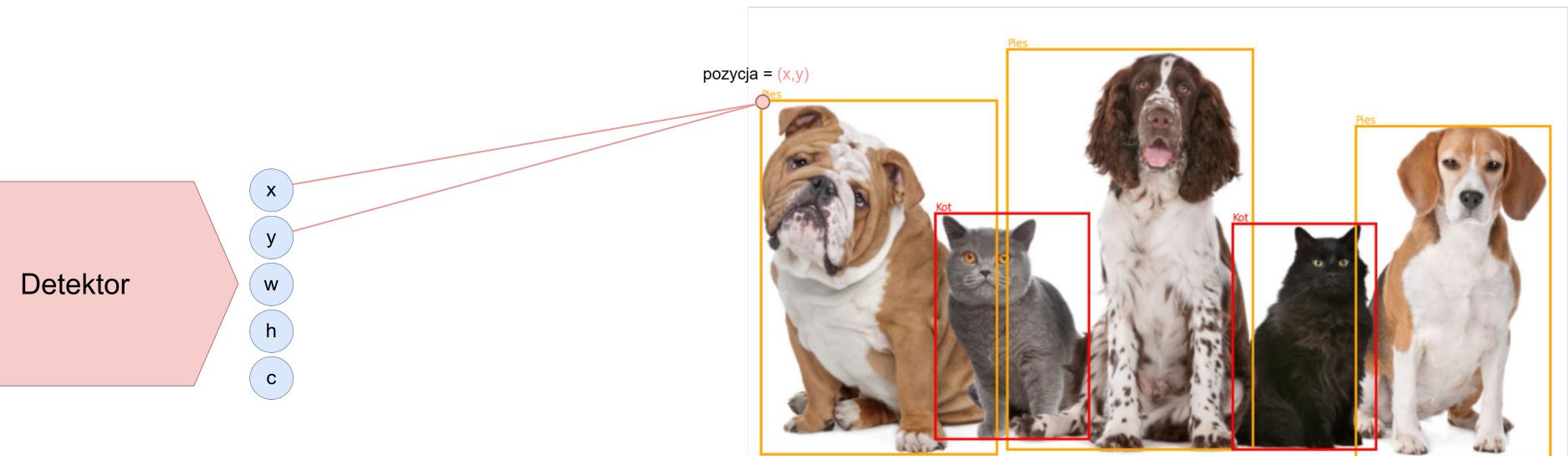
Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**



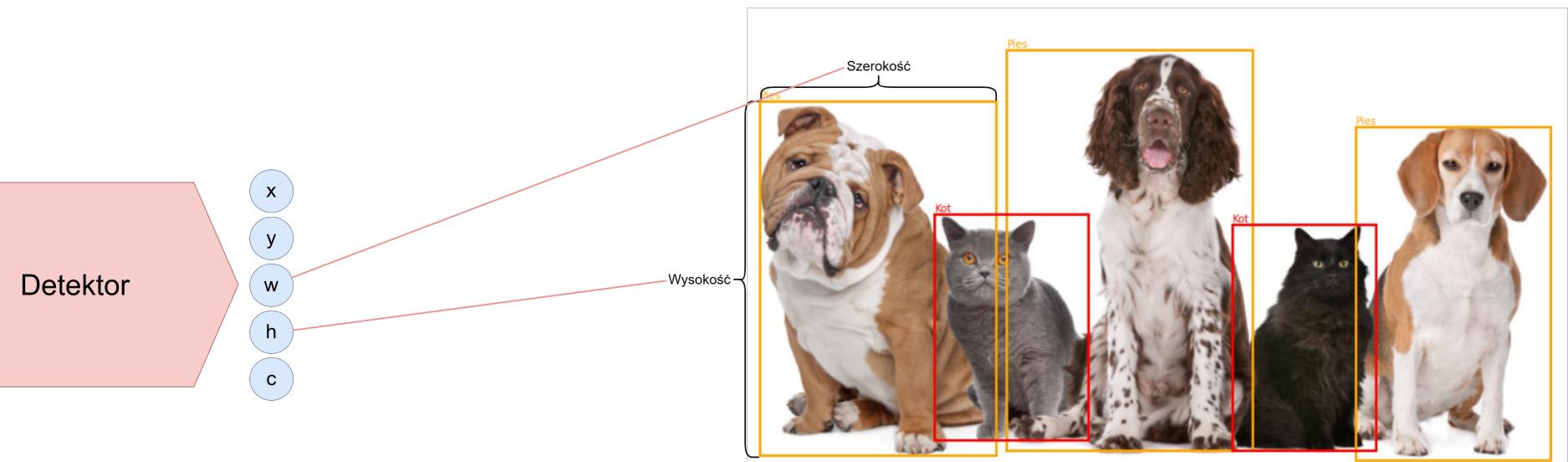
Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?



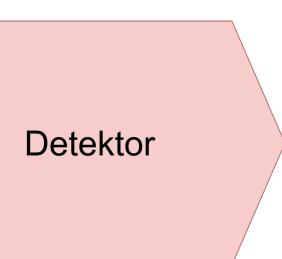
Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**

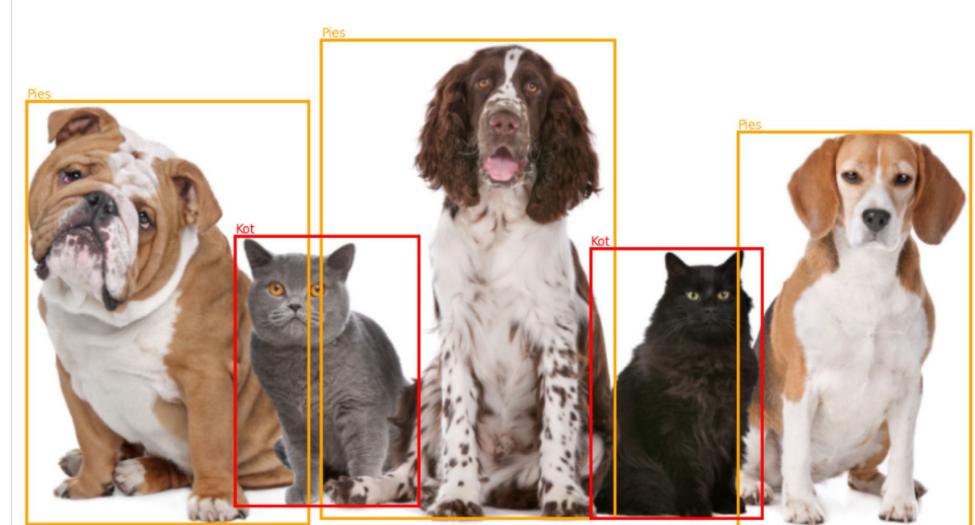


Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**

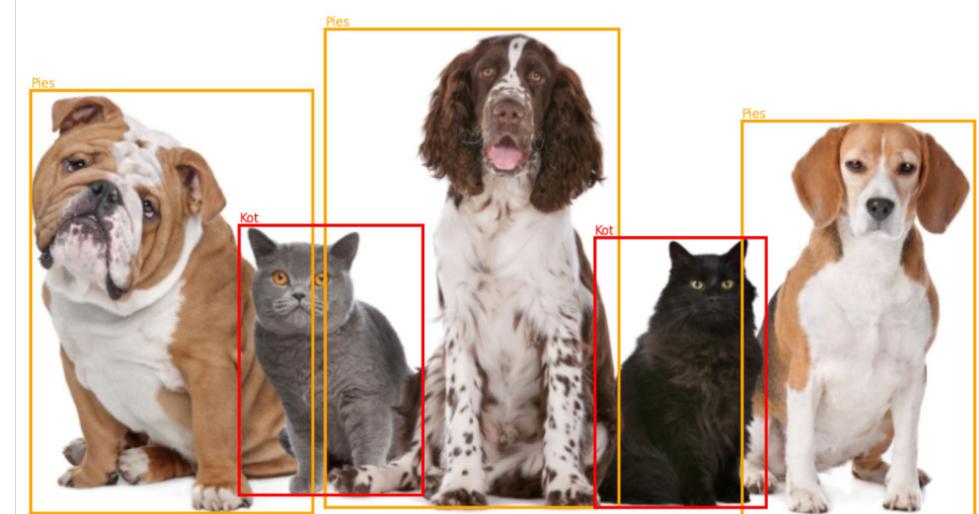
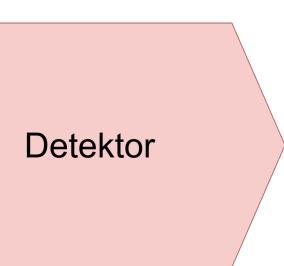


Klasa: Pies czy Kot?



Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**



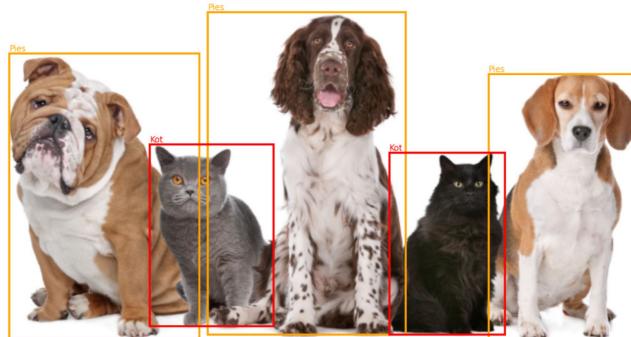
Co z resztą obiektów ?

Detekcja - problemy

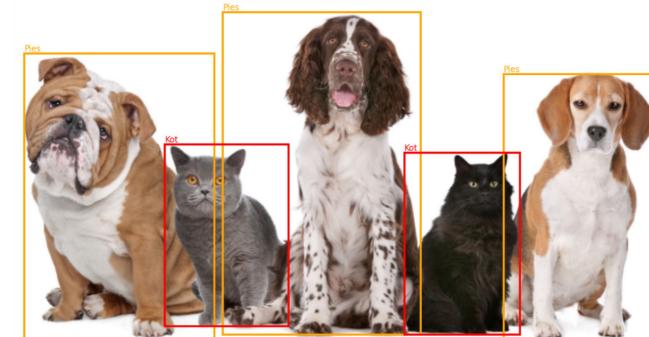
- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?

Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować wyjście sieci?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



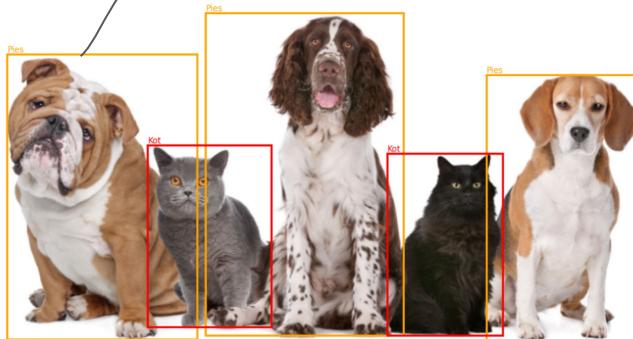
Predykcje modelu



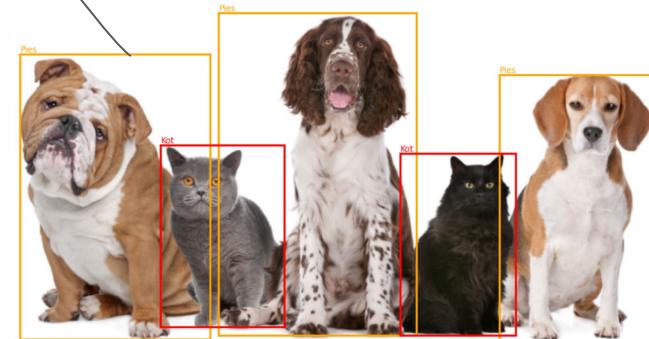
rozwiązanie docelowe

Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



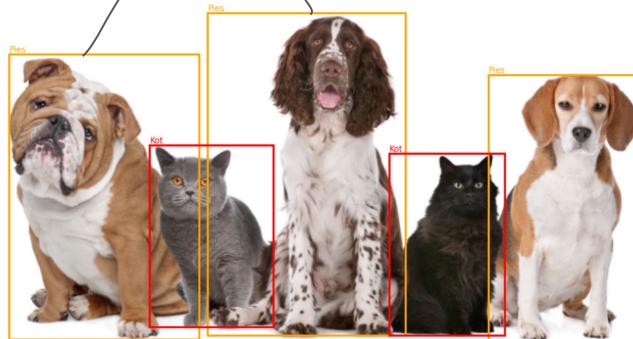
Predykcje modelu



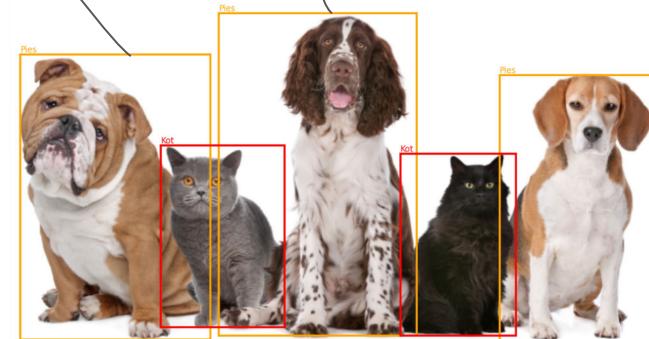
rozwiązanie docelowe

Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



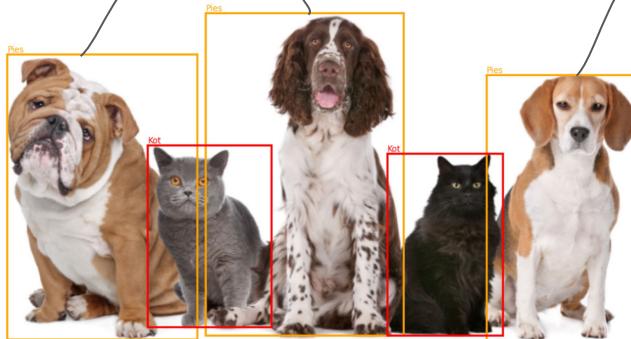
Predykcje modelu



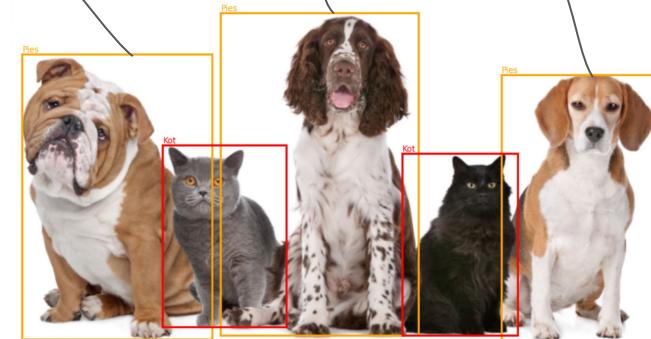
rozwiązanie docelowe

Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci?**
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



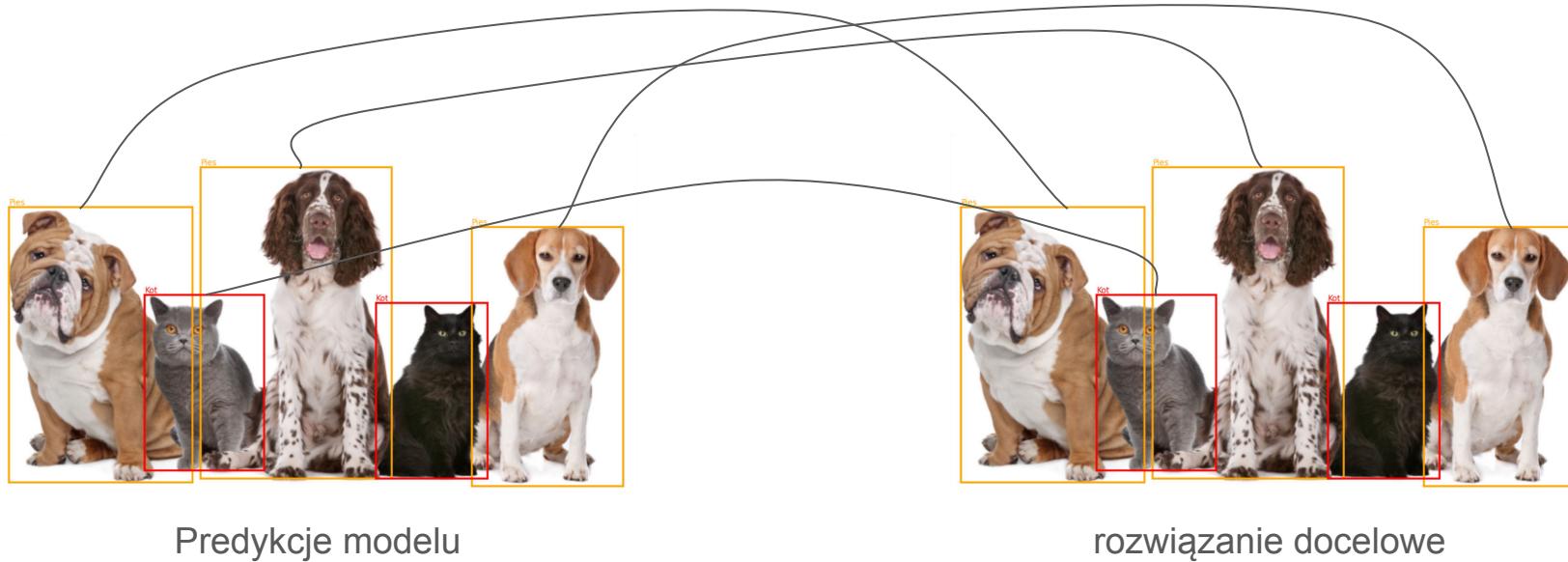
Predykcje modelu



rozwiązanie docelowe

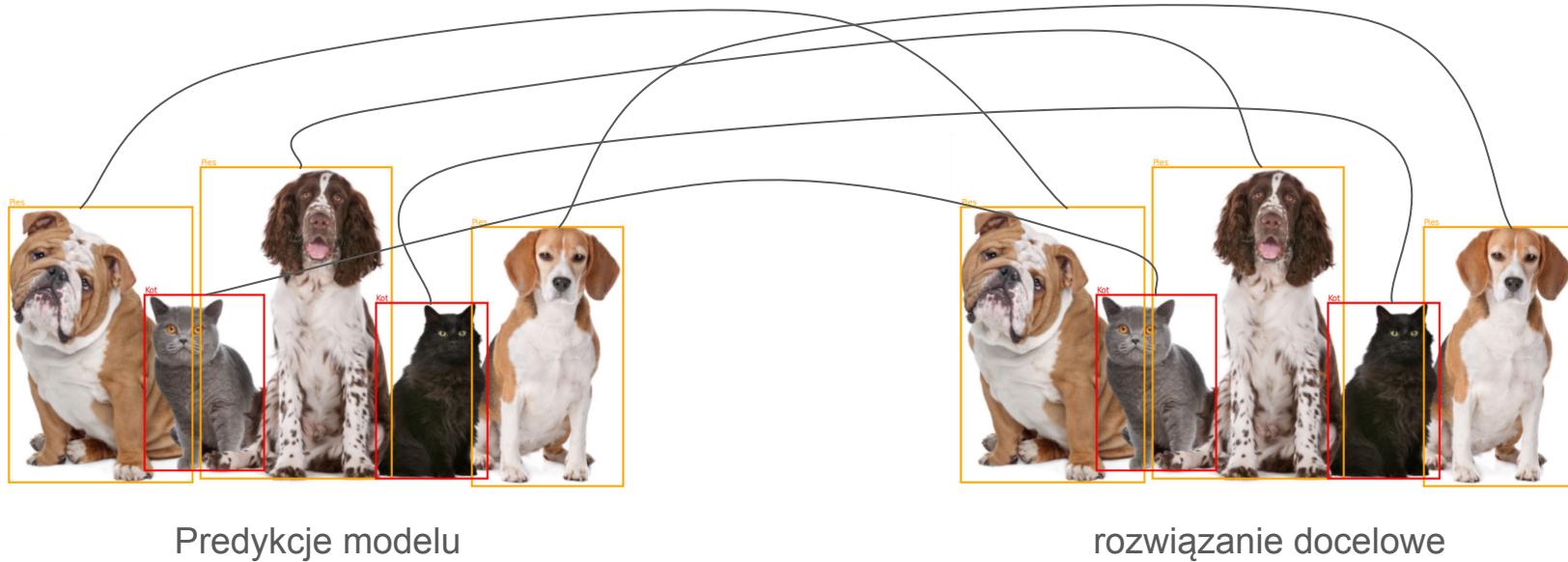
Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



Detekcja - problemy

- Na zdjęciu znajduje się **zmienna liczba obiektów**, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?



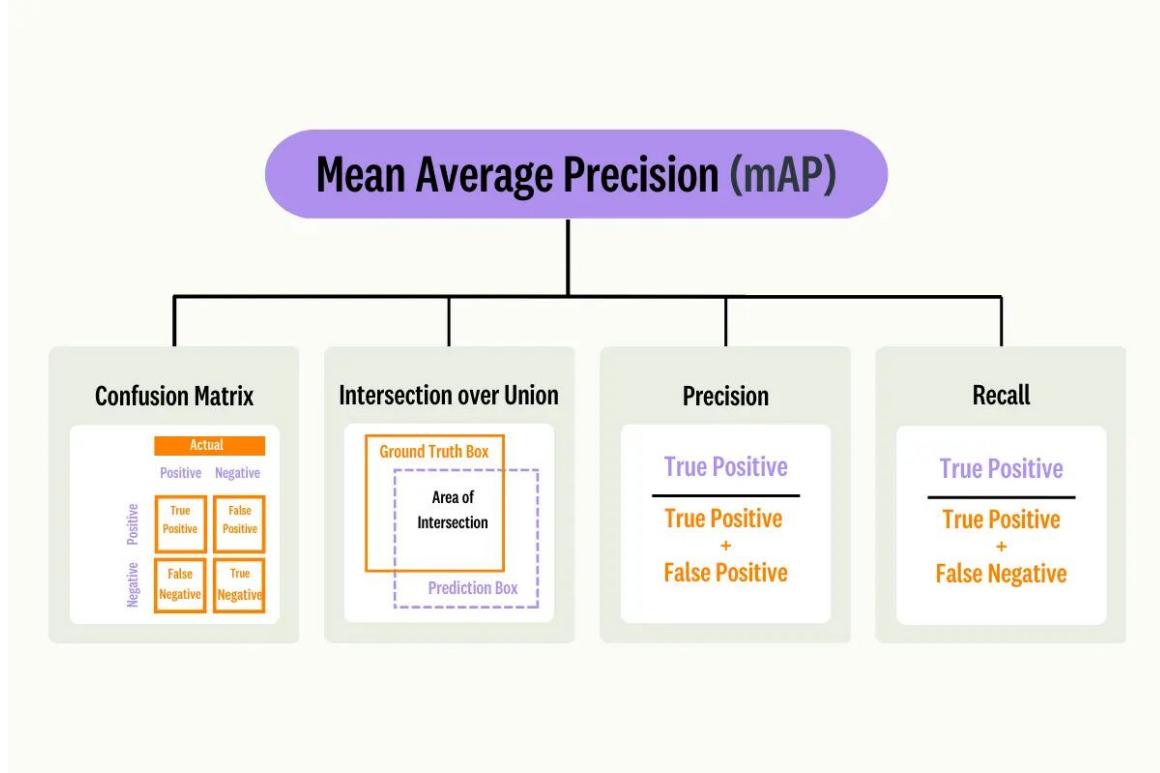
Detekcja - problemy

- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?
- W zadaniu detekcji rozwiązuje się dwa problemy jednocześnie:
 - znajdowanie lokalizacji obiektów - problem **regresji**
 - klasyfikacja zlokalizowanych obiektów - problem **klasyfikacji****Jak** powinny wyglądać **metryki** oraz **funkcja straty** modeli do detekcji?

Detekcja - problemy

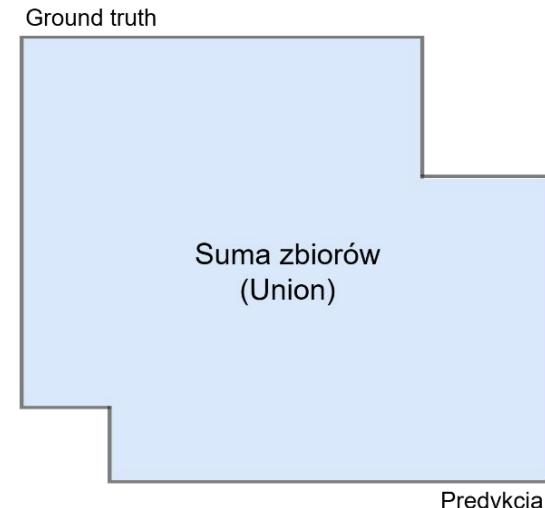
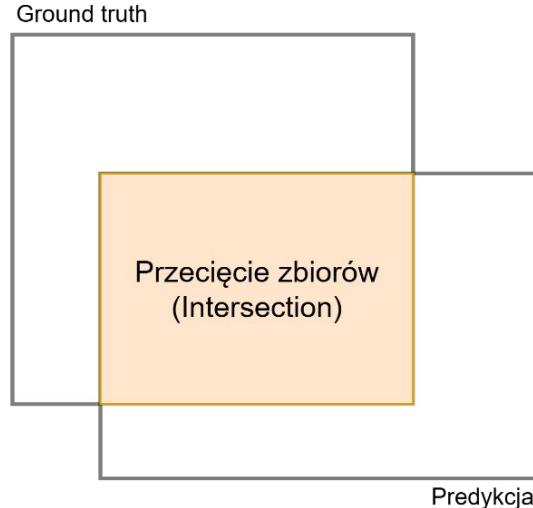
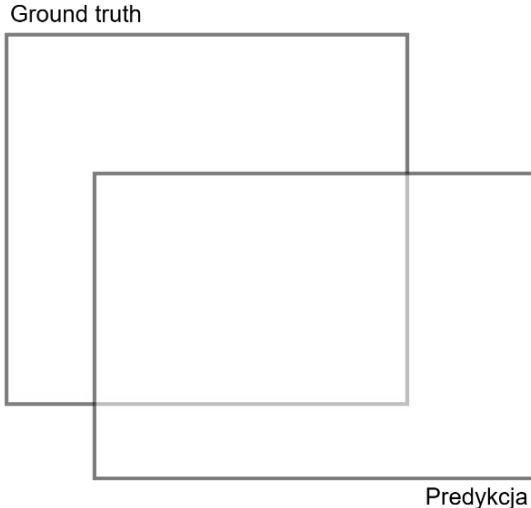
- Na zdjęciu znajduje się **zmienna** liczba obiektów, które chcemy wykryć.
Jak w takiej sytuacji zdefiniować **wyjście sieci**?
- **Kolejność** zwróconych **lokalizacji** (ang. bounding box) oraz lokalizacji docelowych (ground truth) może być **różna**. Jak zatem prawidłowo **parować** odpowiadające predykcje, gdy chcemy ocenić jakość modelu?
- W zadaniu detekcji rozwiązuje się dwa problemy jednocześnie:
 - znajdowanie lokalizacji obiektów - problem **regresji**
 - klasyfikacja zlokalizowanych obiektów - problem **klasyfikacji****Jak** powinny wyglądać **metryki** oraz **funkcja straty** modeli do detekcji?

Detekcja - Ewaluacja rozwiązań



Detekcja - Ewaluacja rozwiązań

- Do parowania predykcji z wartościami ground truth możemy wykorzystać metrykę IoU (ang. Intersection over Union).



$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Detekcja - Ewaluacja rozwiązań

Cechy mAP:

- Zakładamy konkretny **próg** wartości **IoU** w procesie parowania predykcji z docelowymi prostokątami.

Detekcja - Ewaluacja rozwiązań

Cechy mAP:

- Zakładamy konkretny **próg** wartości **IoU** w procesie parowania predykci z docelowymi prostokątami.
- Rozpatrujemy każdą z klas **osobno** licząc jej average precision (AP) i na koniec **uśredniamy** po wszystkich klasach (mean AP - mAP).

Detekcja - Ewaluacja rozwiązań

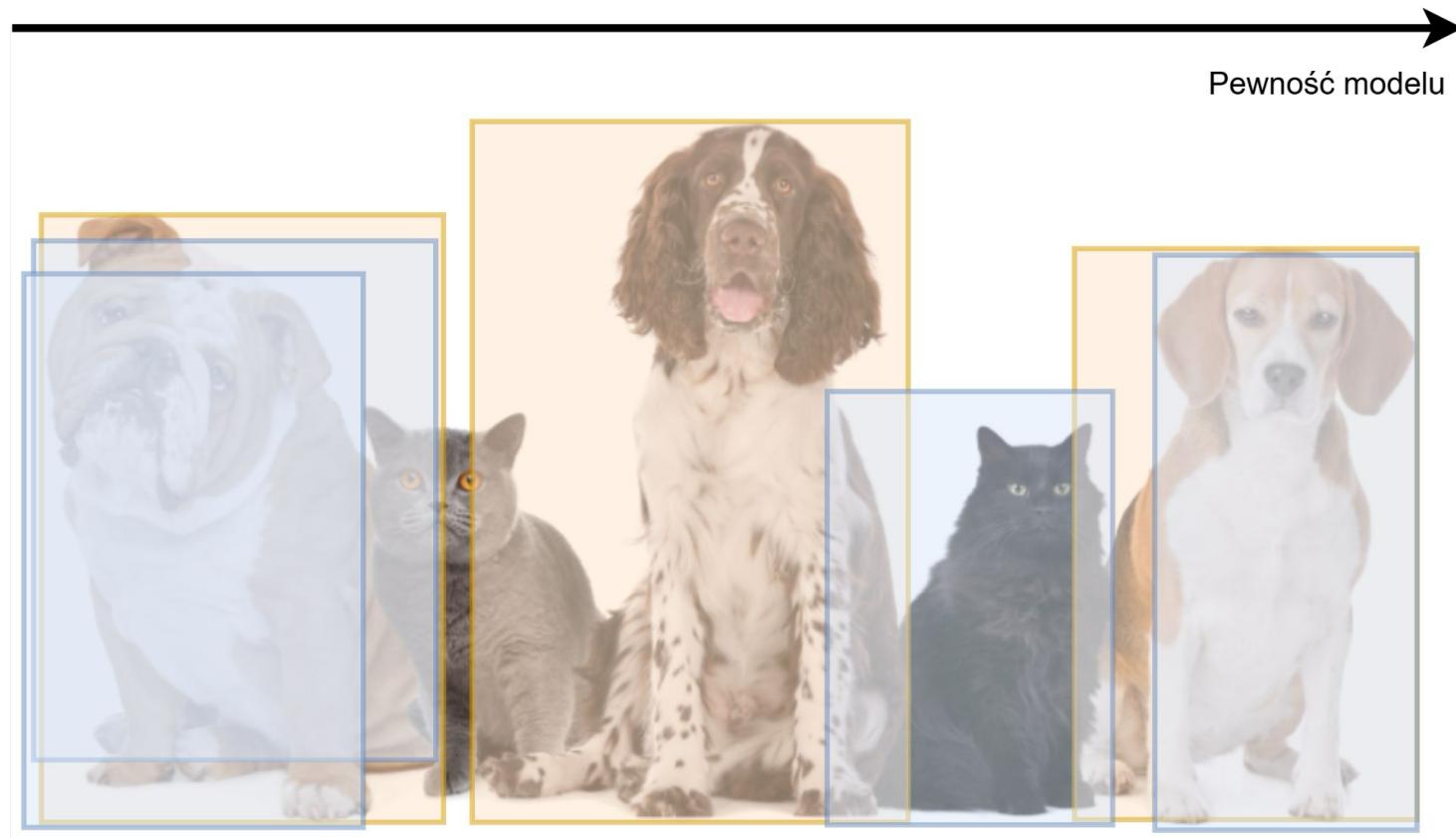
Cechy mAP:

- Zakładamy konkretny **próg** wartości **IoU** w procesie parowania predykcji z docelowymi prostokątami.
- Rozpatrujemy każdą z klas **osobno** licząc jej average precision (AP) i na koniec **uśredniamy** po wszystkich klasach (mean AP - mAP).
- Dodatkowo, oprócz pozycji prostokąta i jego klasy, wykorzystujemy **pewność modelu** co do konkretnej predykcji.

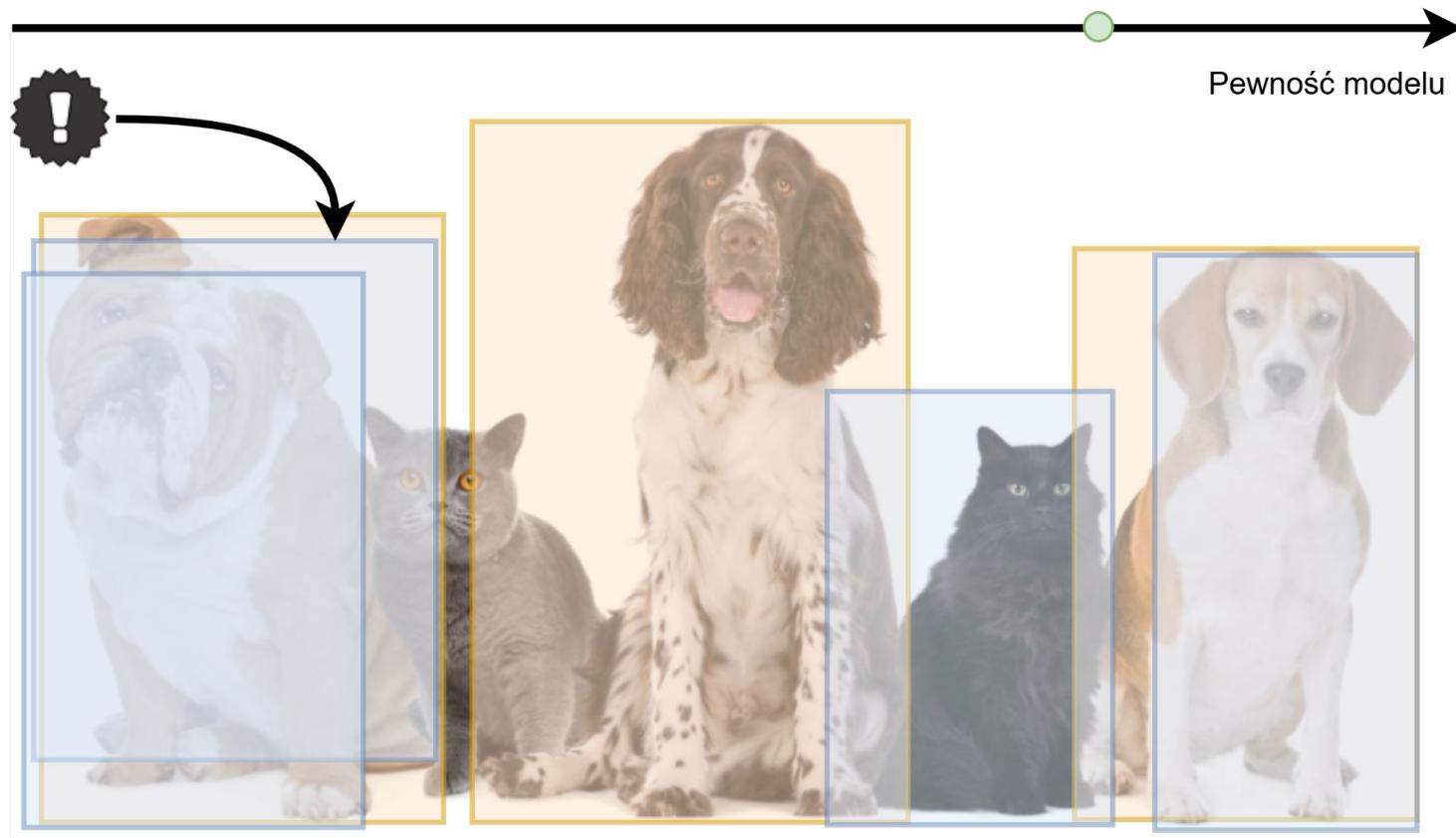
Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



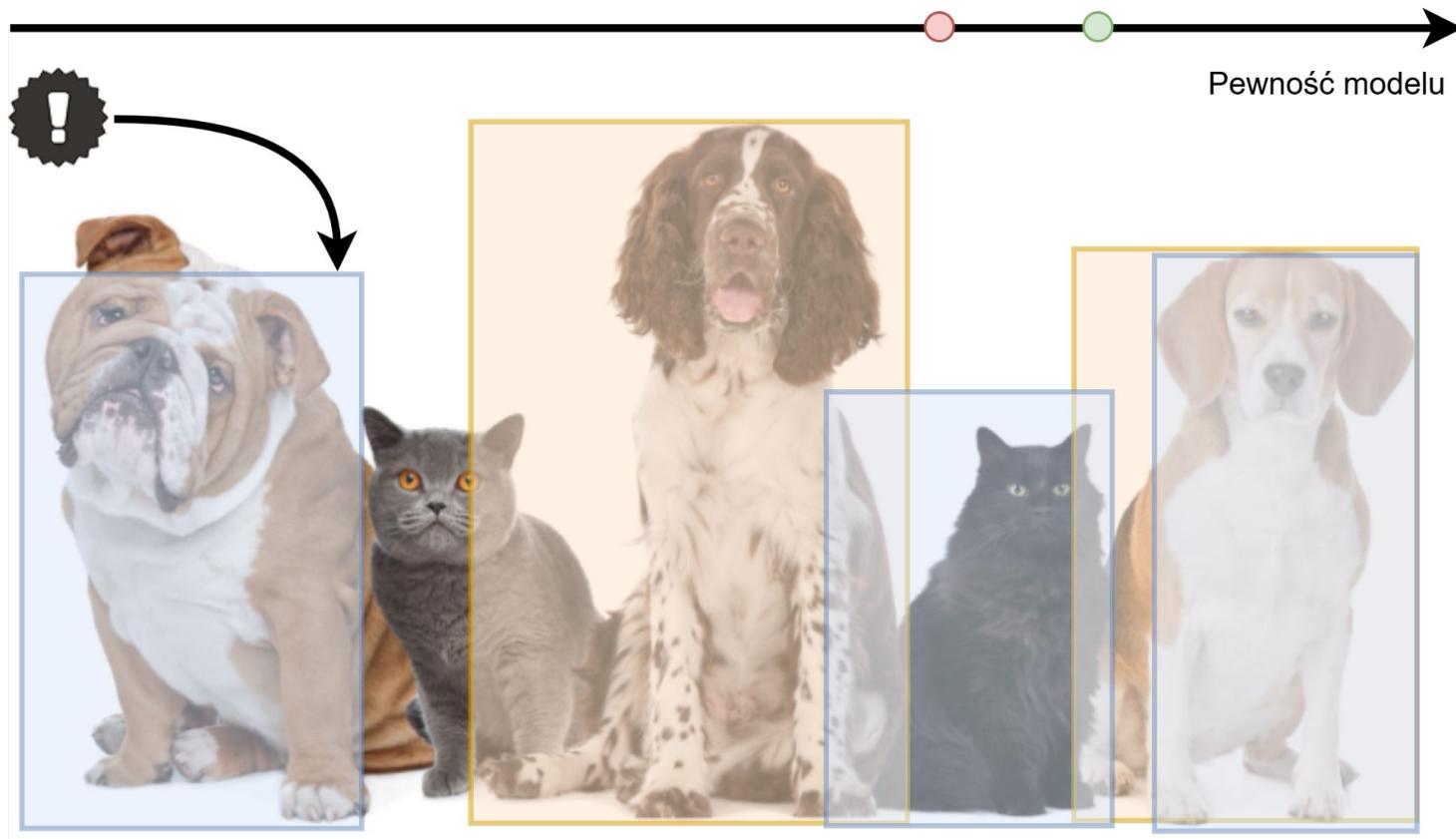
Detekcja - wizualizacja metryki mAP



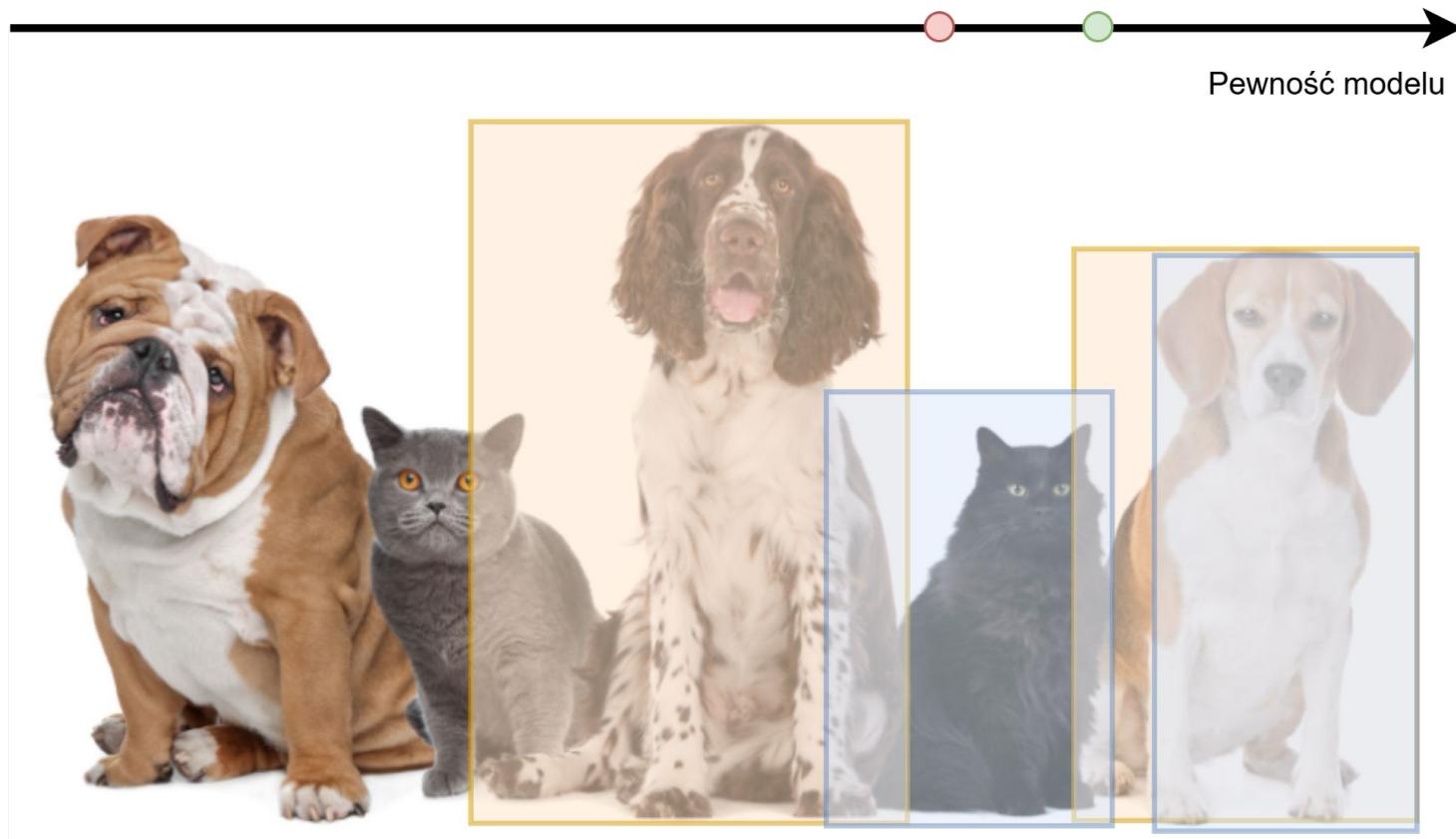
Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP



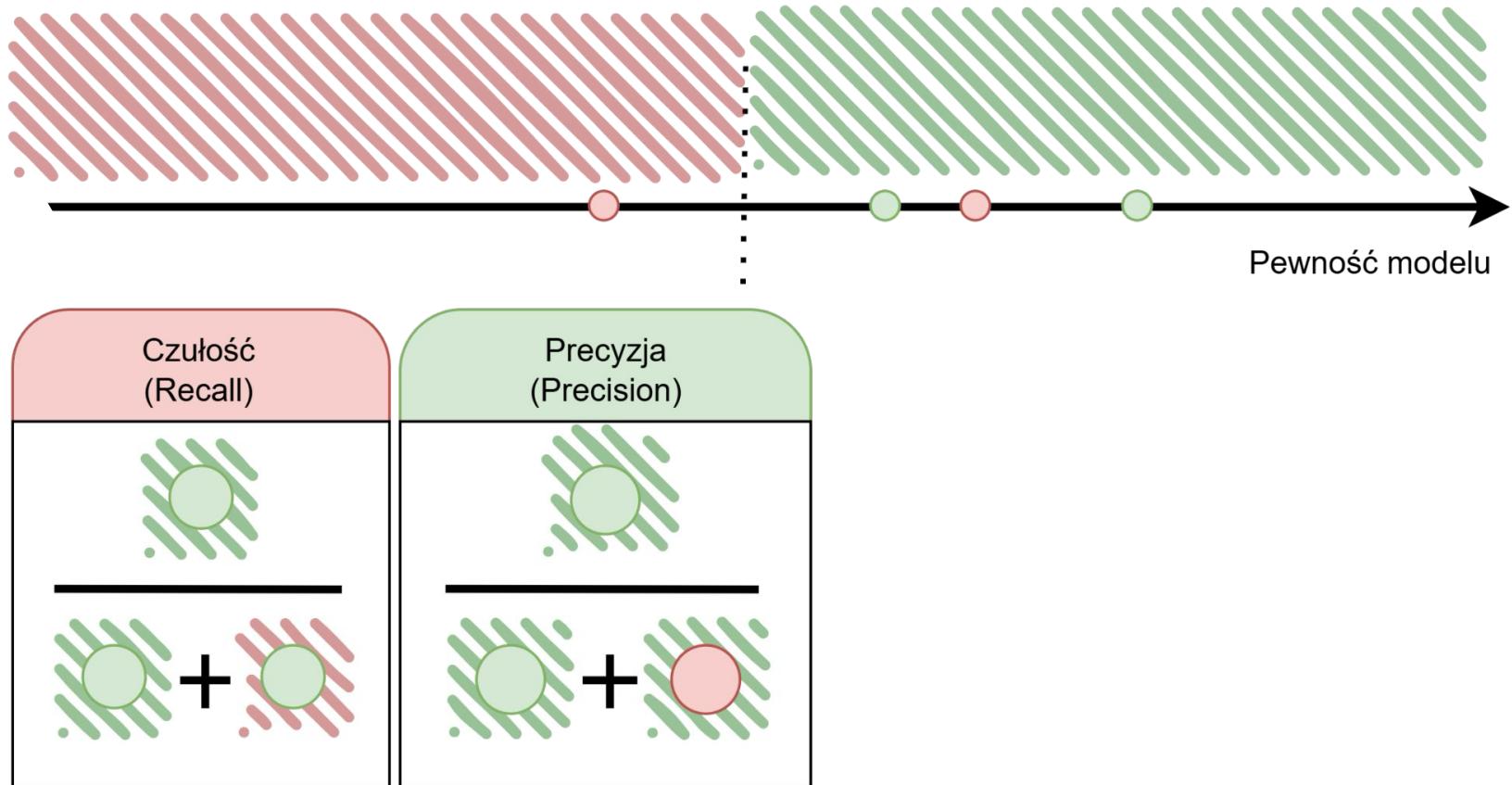
Detekcja - wizualizacja metryki mAP



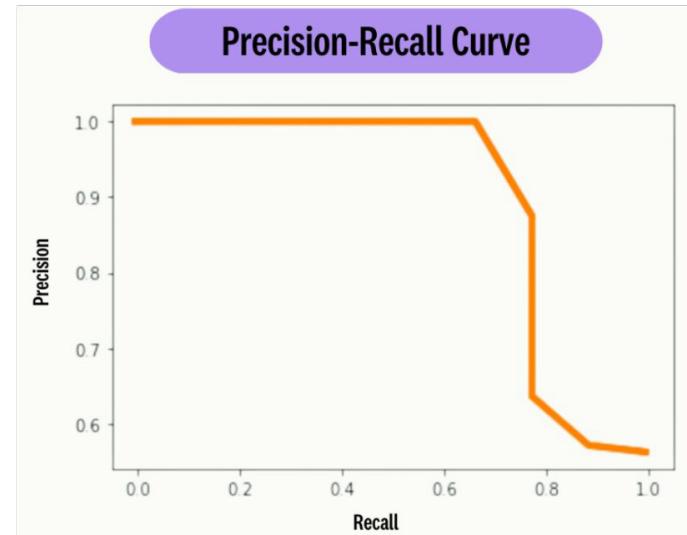
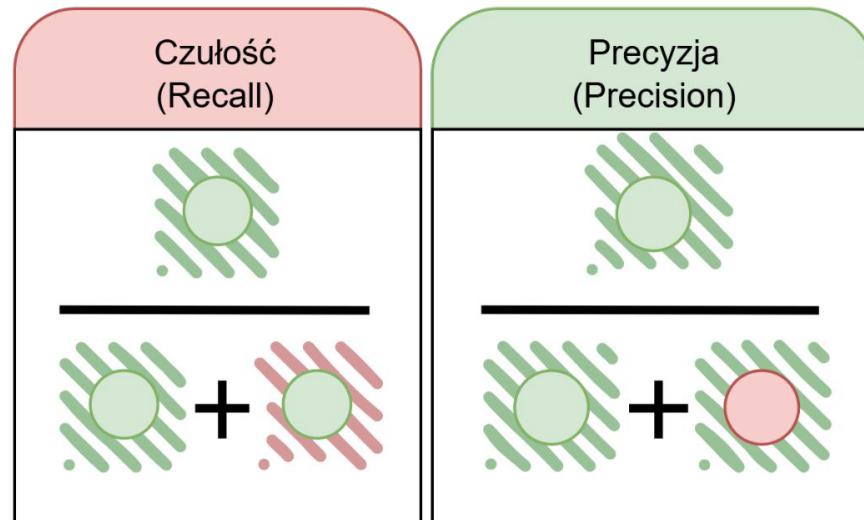
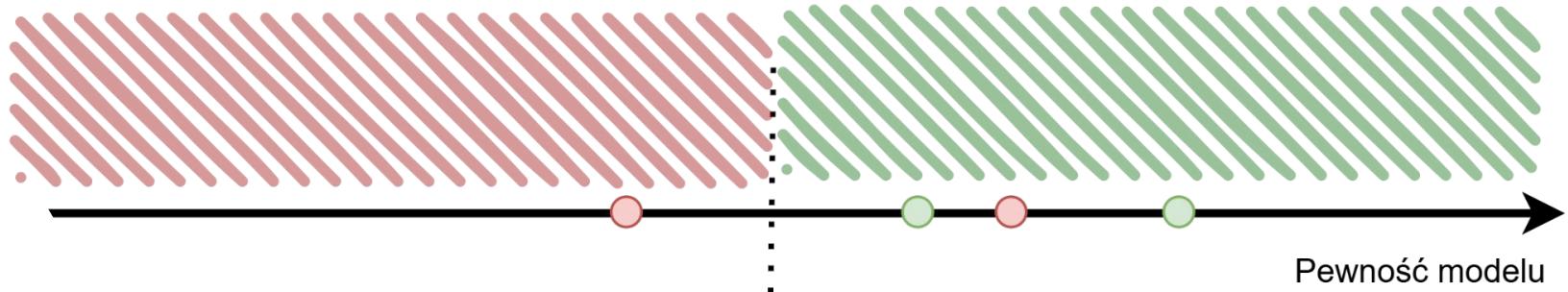
Detekcja - wizualizacja metryki mAP



Detekcja - wizualizacja metryki mAP

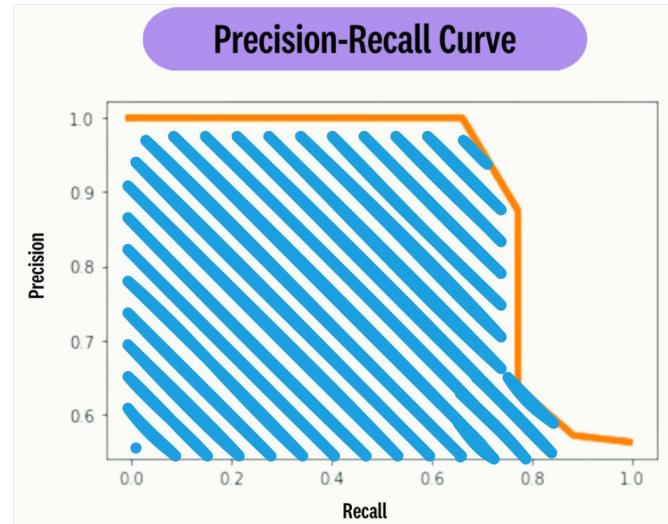


Detekcja - wizualizacja metryki mAP



Detekcja - podsumowanie mAP

- Pole pod krzywą precision-recall to average precision **AP**.

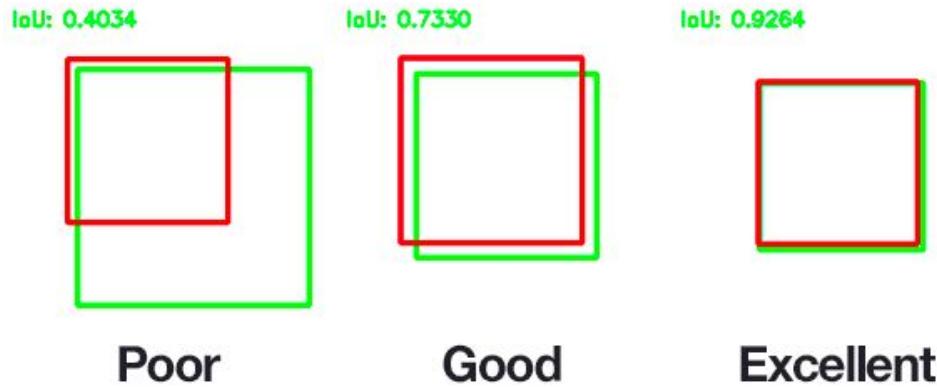


Detekcja - podsumowanie mAP

- Pole pod krzywą precision-recall to average precision **AP**.
- Wartości AP uśredniamy dla wszystkich klas - mean AP (mAP)

Detekcja - podsumowanie mAP

- Pole pod krzywą precision-recall to average precision **AP**.
- Wartości AP **uśredniamy** dla wszystkich klas - mean AP (mAP)
- Ale jakiego progu IoU użyć, żeby metryka mAP nie była ani zbyt surowa ani zbyt łaskawa w parowaniu prostokątów? ...



Detekcja - podsumowanie mAP

- Pole pod krzywą precision-recall to average precision **AP**.
- Wartości AP **uśredniamy** dla wszystkich klas - mean AP (mAP)
- Ale jakiego progu IoU użyć, żeby metryka mAP nie była ani zbyt surowa ani zbyt łaskawa w parowaniu prostokątów? ...
- Najlepiej użyć różnych progów IoU i na końcu również uśrednić po wszystkich wartościach. Standardowo stosuje się progi od **0.5** do **0.95** z interwałem **0.05**.

Detekcja - podsumowanie mAP

- **Pole pod krzywą** precision-recall to average precision **AP**.
- Wartości AP **uśredniamy** dla wszystkich klas - mean AP (mAP)
- Ale jakiego progu IoU użyć, żeby metryka mAP nie była ani zbyt surowa ani zbyt łaskawa w parowaniu prostokątów? ...
- Najlepiej użyć różnych progów IoU i na końcu również uśrednić po wszystkich wartościach. Standardowo stosuje się progi od **0.5** do **0.95** z interwałem **0.05**. Zazwyczaj oznaczamy je w nazwie jako mAP:0.5:0.95:0.05.
 - Wartości progu IoU bliżej 0.5 sprawdzają czy model mniej więcej potrafi coś zlokalizować i poprawnie zaklasyfikować
 - Wartości progu IoU bliżej 0.95 sprawdzają czy model potrafi zwracać niemal idealnie zlokalizowane predykcje.

Detekcja - istniejące metody

- Metody klasyczne (nie neuronowe), np. detektor SIFT, Viola Johns itp.

Detekcja - istniejące metody

- Metody klasyczne (nie neuronowe), np. detektor SIFT, Viola Johns itp.
- Metoda okna przesuwnego (ang. sliding window)

Detekcja - istniejące metody

- Metody klasyczne (nie neuronowe), np. detektor SIFT, Viola Johns itp.
- Metoda okna przesuwnego (ang. sliding window)
- Region based CNN (RCNN) i pochodne

Detekcja - istniejące metody

- Metody klasyczne (nie neuronowe), np. detektor SIFT, Viola Johns itp.
- Metoda okna przesuwnego (ang. sliding window)
- Region based CNN (RCNN) i pochodne
- YOLO, ..., YOLOv9

Detekcja - istniejące metody

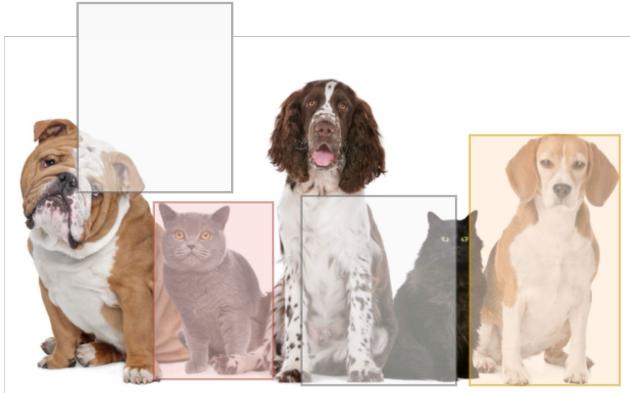
- Metody klasyczne (nie neuronowe), np. detektor SIFT, Viola Johns itp.
- Metoda okna przesuwnego (ang. sliding window)
- Region based CNN (RCNN) i pochodne
- YOLO, ..., YOLOv9
- Wiele innych podejść, np. RetinaNet, EfficientDet, SSD itp...

Detekcja - Technika okna przesuwnego



Trening:

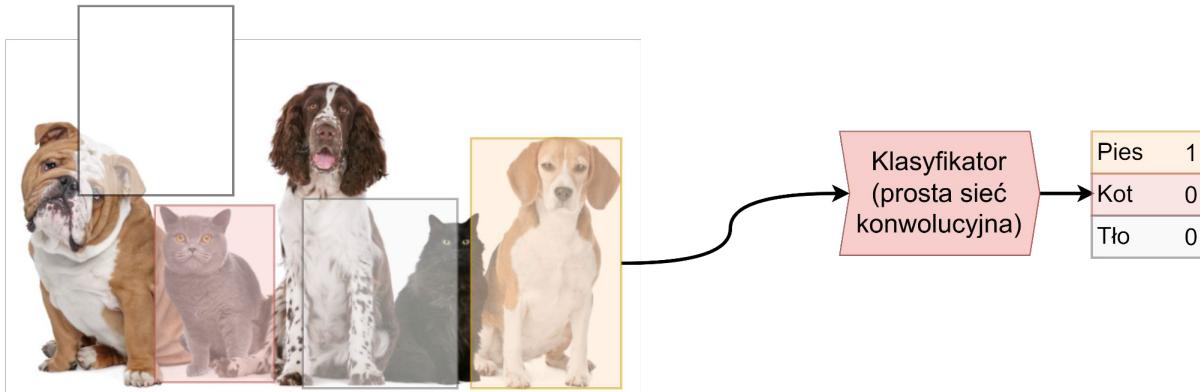
Detekcja - Technika okna przesuwnego



Trening:

- Do treningu używamy wycinków **zawierających klasy obiektów** do detekcji oraz **losowe wycinki z tłem**.

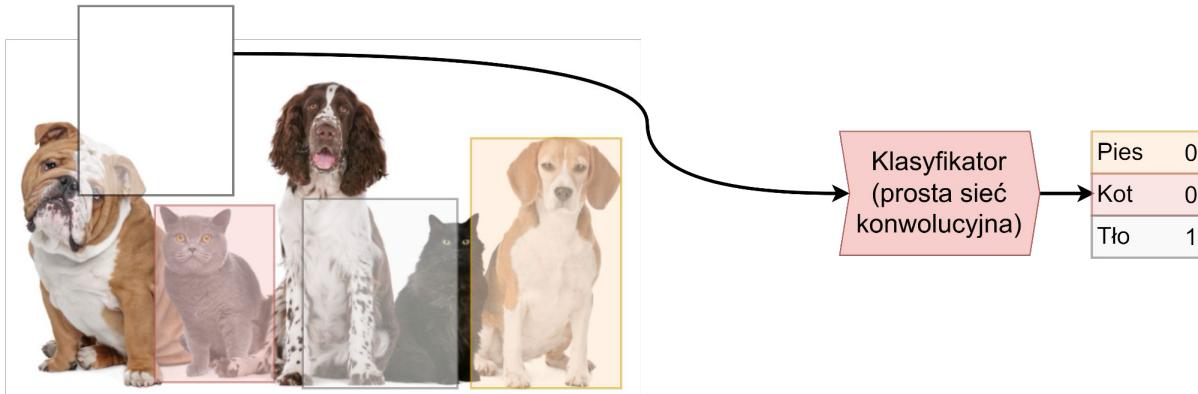
Detekcja - Technika okna przesuwnego



Trening:

- Do treningu używamy wycinków **zawierających klasy obiektów** do detekcji oraz **losowe wycinki z tłem**.
- Trenujemy sieć neuronową na zadaniu **klasyfikacji** wycinków.

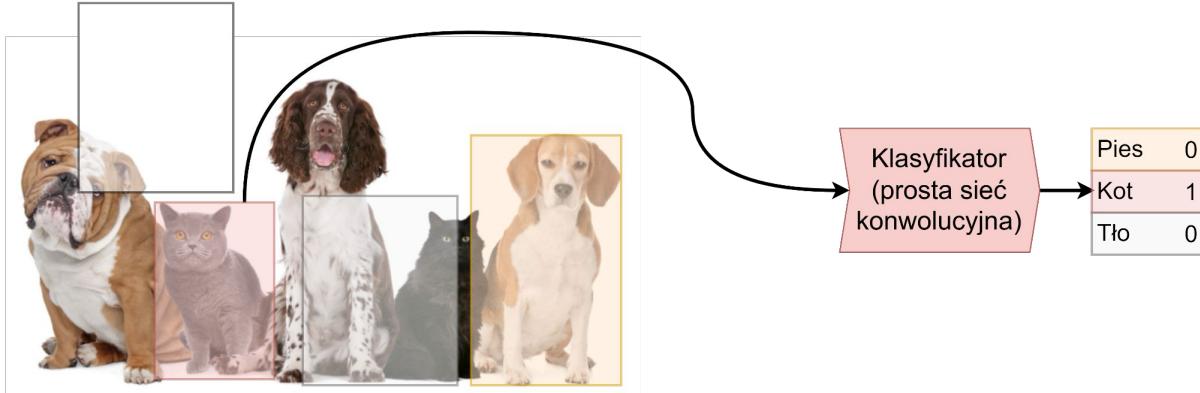
Detekcja - Technika okna przesuwnego



Trening:

- Do treningu używamy wycinków **zawierających klasy obiektów** do detekcji oraz **losowe wycinki z tłem**.
- Trenujemy sieć neuronową na zadaniu **klasyfikacji** wycinków.

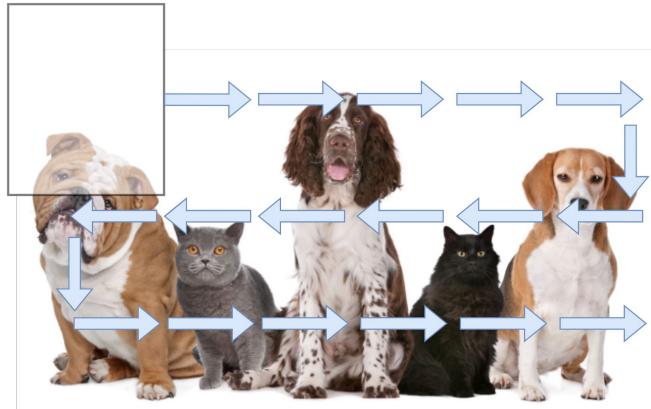
Detekcja - Technika okna przesuwnego



Trening:

- Do treningu używamy wycinków **zawierających klasy obiektów** do detekcji oraz **losowe wycinki z tłem**.
- Trenujemy sieć neuronową na zadaniu **klasyfikacji** wycinków.

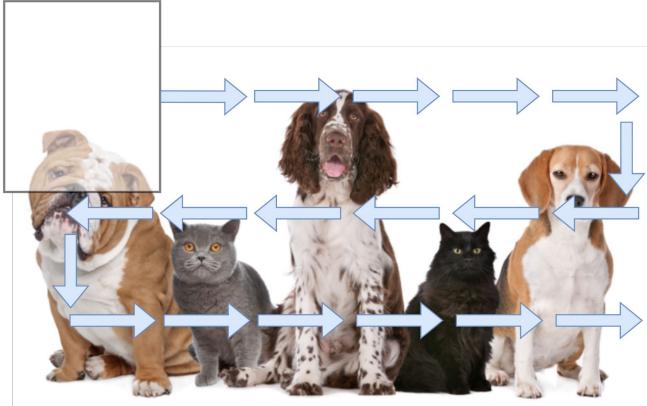
Detekcja - Technika okna przesuwnego



Inferencja (faza testowania):

- W celu znalezienia wszystkich obiektów na zdjęciu, **przesuwamy** okno, które trafia do klasyfikatora. Gdy klasyfikator zwróci klasę inną niż tło, pozycja okna zostanie zapisana.
- Proces można powtarzać dla **różnych** wielkości prostokątów.

Detekcja - Technika okna przesuwnego



Zalety:

- Intuicyjne i proste do zaimplementowania rozwiązanie.

Wady:

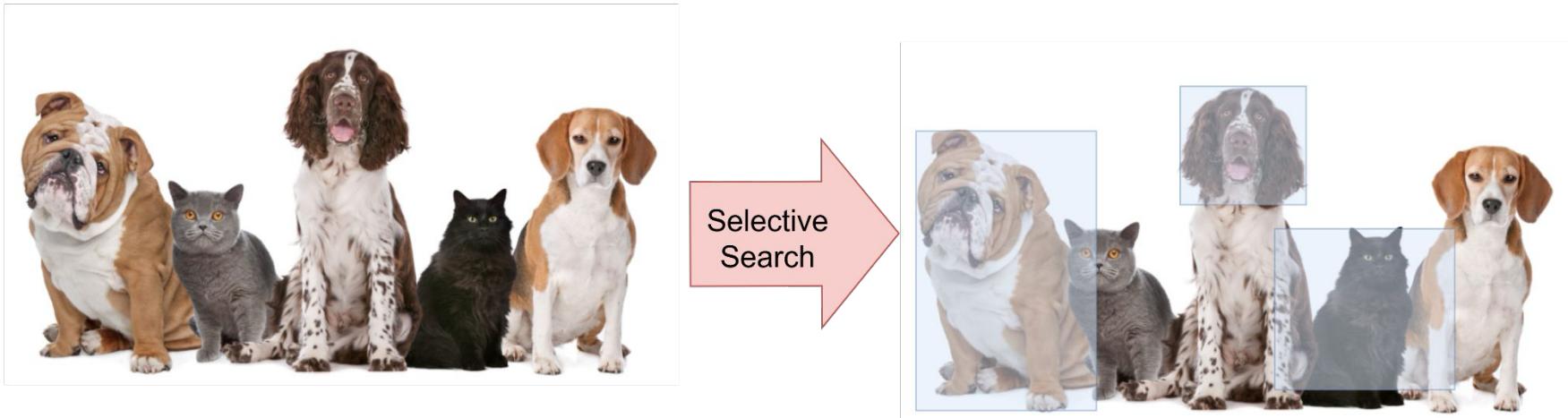
- Niedokładne lokalizacje
- Wolna inferencja
- Brak możliwości dopasowania rozmiaru okna do faktycznego obiektu.

Inferencja (faza testowania):

- W celu znalezienia wszystkich obiektów na zdjęciu, **przesuwamy** okno, które trafia do klasyfikatora. Gdy klasyfikator zwróci klasę inną niż tło, pozycja okna zostanie zapisana.
- Proces można powtarzać dla różnych **wielkości** prostokątów.

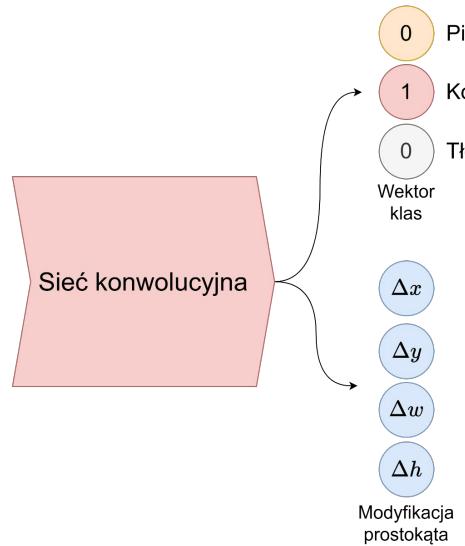
Detekcja - Region based CNN

- Pierwsza wersja modelu RCNN starała się ulepszyć technikę okna przesuwnego poprzez sprytniejszy wybór potencjalnych lokalizacji (ang. region proposal) dla modelu klasyfikatora. Zostało to uzyskane za pomocą algorytmu **selective search**.



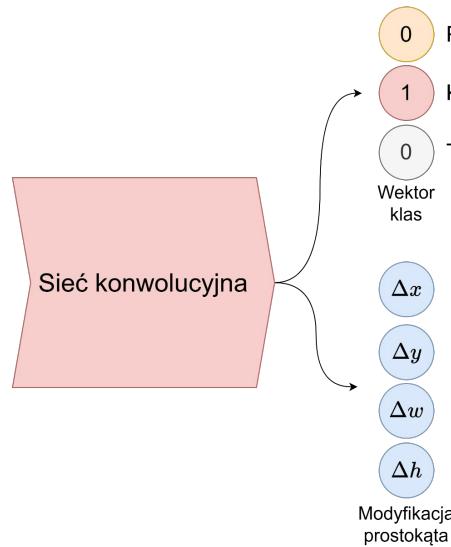
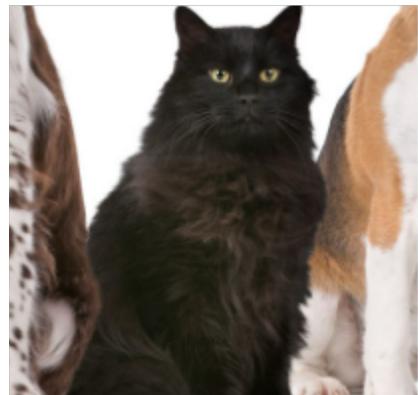
Detekcja - Region based CNN

- Pierwsza wersja modelu RCNN starała się ulepszyć technikę okna przesuwnego poprzez sprytniejszy wybór potencjalnych lokalizacji (ang. region proposal) dla modelu klasyfikatora. Zostało to uzyskane za pomocą algorytmu **selective search**.
- Dodatkowo, zamiast przewidywać jedynie odpowiednią klasę regionu, sieć była także trenowana do poprawy lokalizacji i rozmiaru.



Detekcja - Region based CNN

- Pierwsza wersja modelu RCNN starała się ulepszyć technikę okna przesuwnego poprzez sprytniejszy wybór potencjalnych lokalizacji (ang. region proposal) dla modelu klasyfikatora. Zostało to uzyskane za pomocą algorytmu **selective search**.
- Dodatkowo, zamiast przewidywać jedynie odpowiednią klasę regionu, sieć była także trenowana do poprawy lokalizacji i rozmiaru.



Zalety:

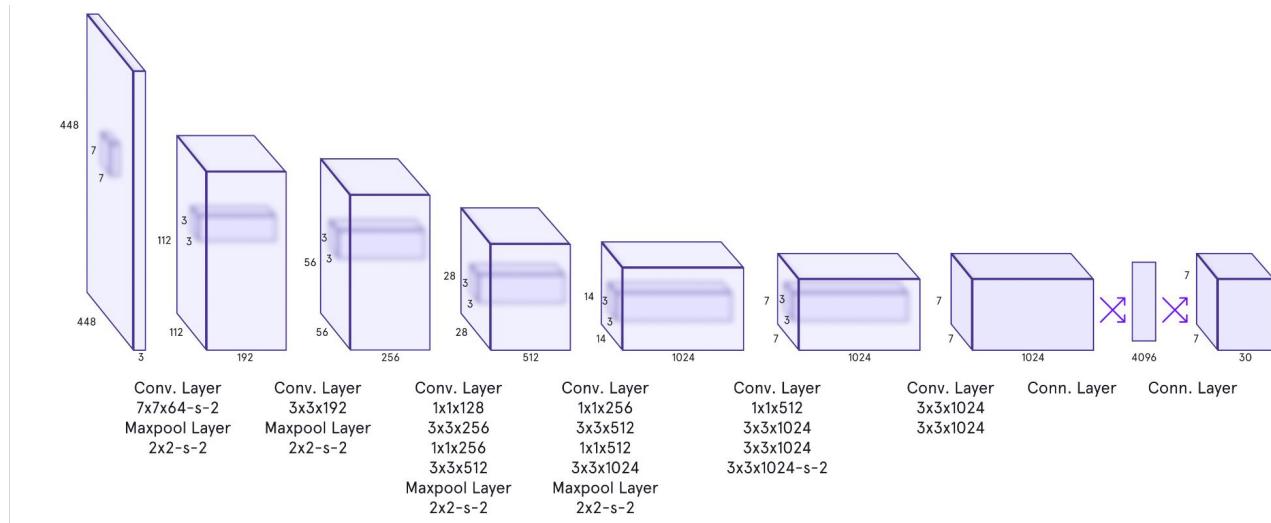
- Możliwość manipulacji pozycji i rozmiaru bounding box'ów.
- Dokładność.
- Szybsza inferencja niż w technice okna przesuwnego.

Wady:

- Wolna inferencja (wciąż bardzo dużo wycinków do klasyfikacji)

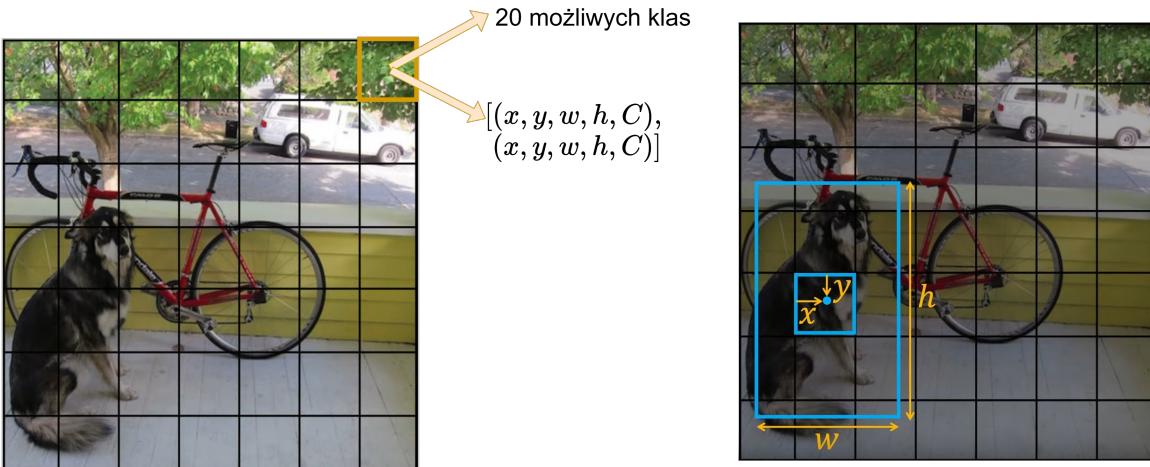
Detekcja - YOLO

- Nazwa YOLO pochodzi od “You Only Look Once”, co odnosi się do faktu, że modele YOLO wykorzystują tylko **jedno przejście** przez sieć dla konkretnego obrazka. Ta strategia sprawia, że model ten jest wyjątkowo szybki w inferencji.
- Na wyjściu z sieci dostajemy tensor 7×7 o 30 kanałach



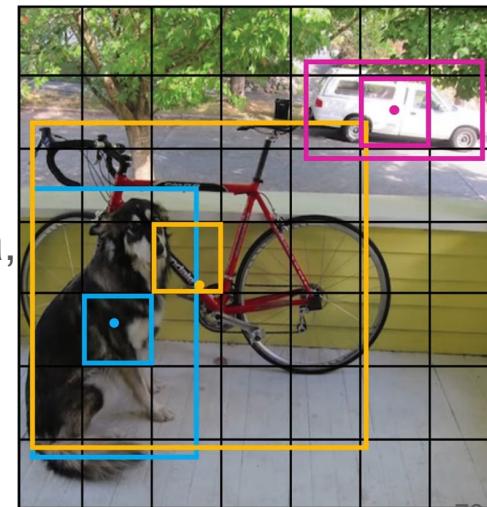
Detekcja - YOLO

- Nazwa YOLO pochodzi od “You Only Look Once”, co odnosi się do faktu, że modele YOLO wykorzystują tylko **jedno przejście** przez sieć dla konkretnego obrazka. Ta strategia sprawia, że model ten jest wyjątkowo szybki w inferencji.
- Na wyjściu z sieci dostajemy tensor 7×7 o 30 kanałach.
- Każdy z 49 obszarów koduje informacje na temat klasy i pozycji prostokąta, wewnątrz którego się znajduje.



Detekcja - YOLO

- Nazwa YOLO pochodzi od “You Only Look Once”, co odnosi się do faktu, że modele YOLO wykorzystują tylko **jedno przejście** przez sieć dla konkretnego obrazka. Ta strategia sprawia, że model ten jest wyjątkowo szybki w inferencji.
- Na wyjściu z sieci dostajemy tensor 7×7 o 30 kanałach.
- Każdy z 49 obszarów koduje informacje na temat klasy i pozycji prostokąta, wewnątrz którego się znajduje.
- W procesie uczenia:
 - Dla obszarów, gdzie **nie** ma środka żadnego z bounding box'ów, **minimalizujemy** wartość **pewności** modelu.
 - Dla obszarów, gdzie **znajduje** się środek prostokąta, **minimalizujemy entropię krzyżową** dla klas, **MSE** dla predykcji pozycji i rozmiarów prostokąta oraz **maksymalizujemy pewność** (max 1).



Detekcja - YOLO

Zalety:

- bardzo szybka inferencja (zazwyczaj w czasie rzeczywistym).
- Duża dokładność predykcji zarówno co do klas jak i położenia.

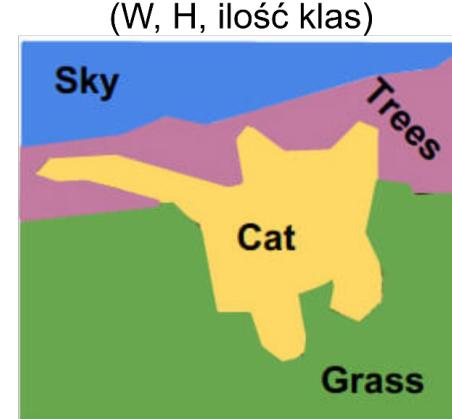
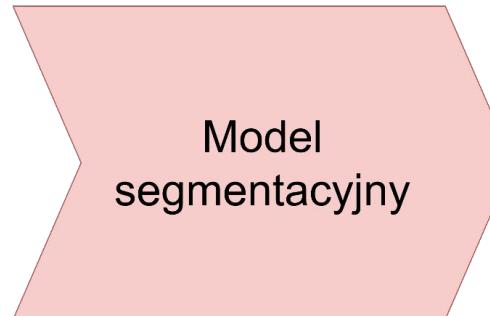
Wady:

- Zazwyczaj nieco gorsze rezultaty niż metody bazujące o RCNN.

Zadanie segmentacji

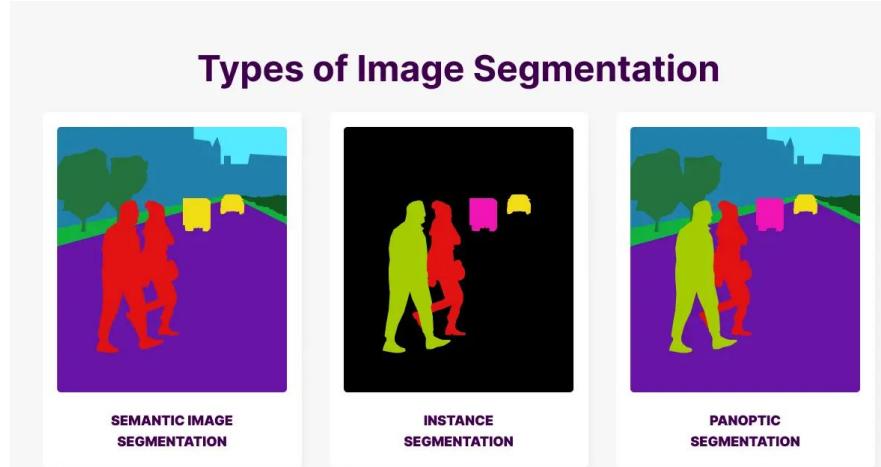
Segmentacja

- W zadaniu segmentacji mapujemy zdjęcie na mapę semantyczną, mówiącą do jakiego obiektu należy każdy piksel.
- Możemy zatem myśleć o segmentacji jako o **klasyfikacji każdego piksela**.



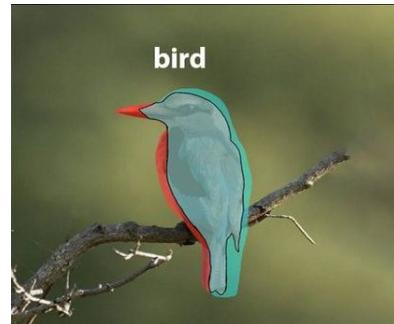
Segmentacja

- W zadaniu segmentacji mapujemy zdjęcie na mapę semantyczną, mówiącą do jakiego obiektu należy każdy piksel.
- Możemy zatem myśleć o segmentacji jako o **klasyfikacji każdego piksela**.
- Rozróżniamy trzy typy segmentacji:
 - segmentacja semantyczna
 - segmentacja instancyjna
 - segmentacja panoptyczna



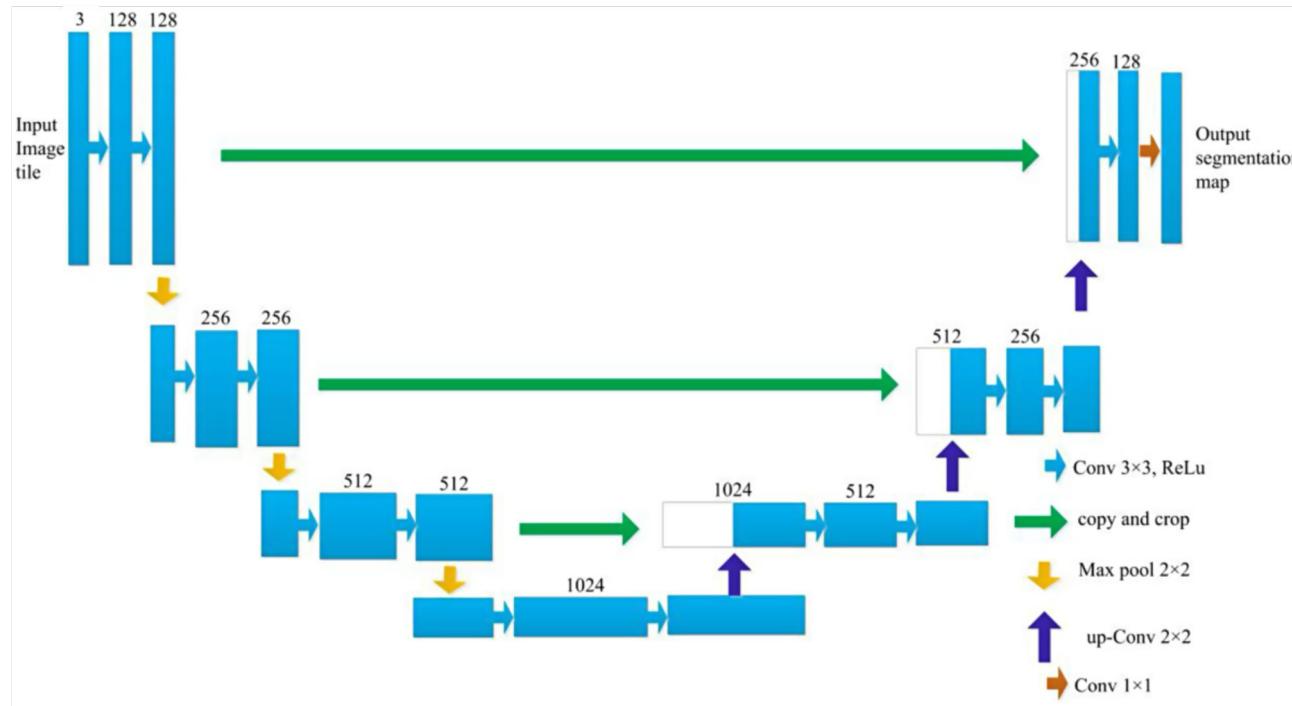
Segmentacja

- W zadaniu segmentacji mapujemy zdjęcie na mapę semantyczną, mówiącą do jakiego obiektu należy każdy piksel.
- Możemy zatem myśleć o segmentacji jako o **klasyfikacji każdego piksela**.
- Rozróżniamy trzy typy segmentacji:
 - segmentacja semantyczna
 - segmentacja instancyjna
 - segmentacja panoptyczna
- Jedną z podstawowych metryk w zadaniu segmentacji to Intersection over Union (IoU) uśrednione po wszystkich klasach.

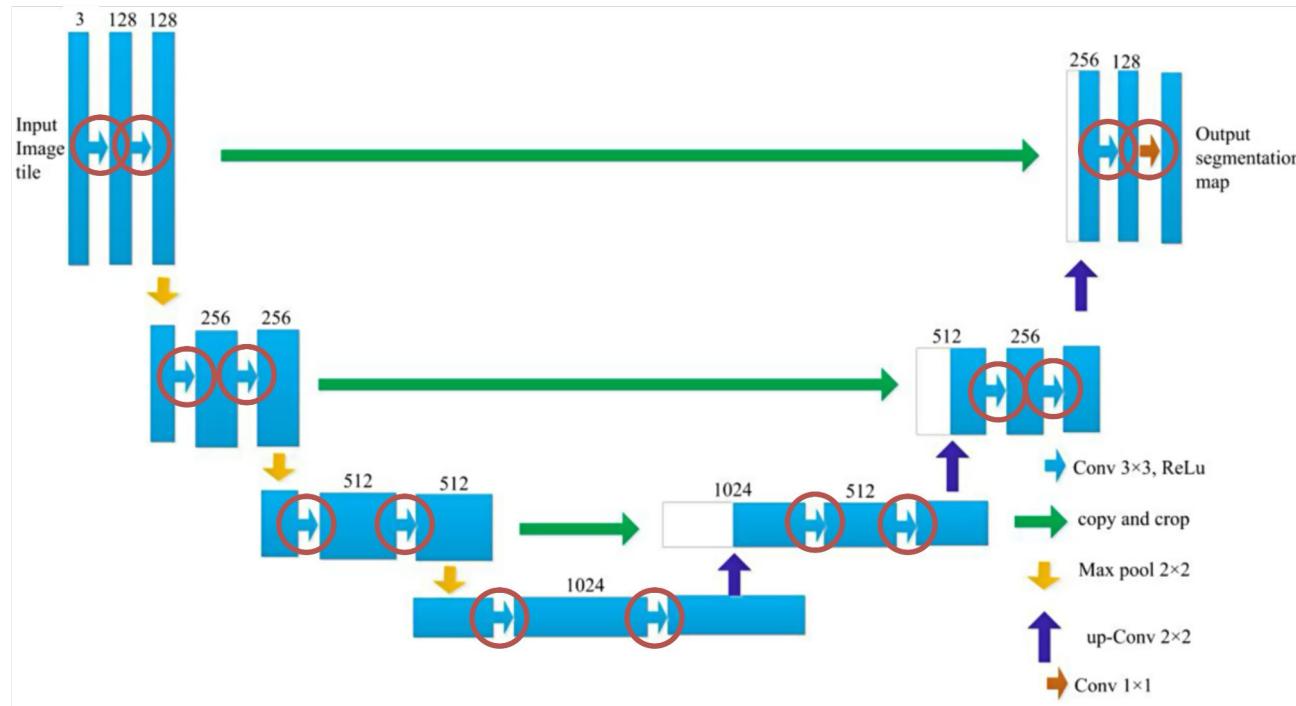


UNet

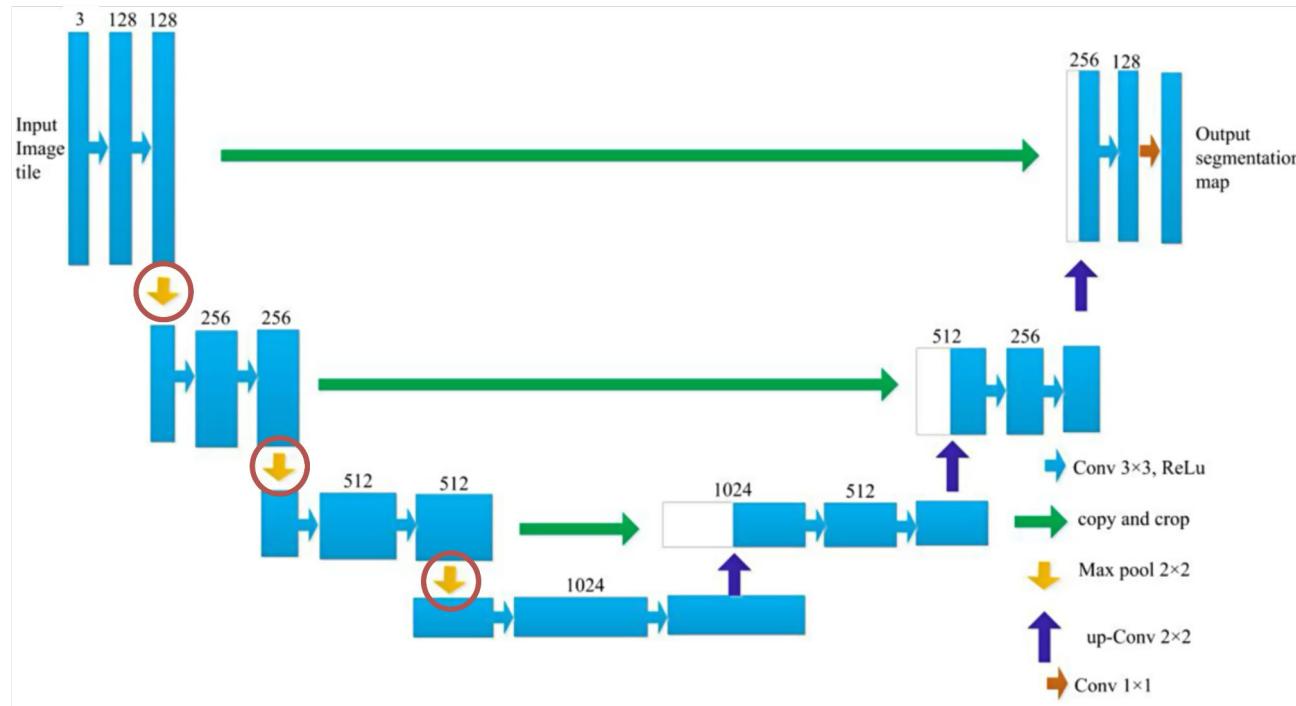
- Najbardziej popularną architekturą do zadania segmentacji jest sieć UNet.



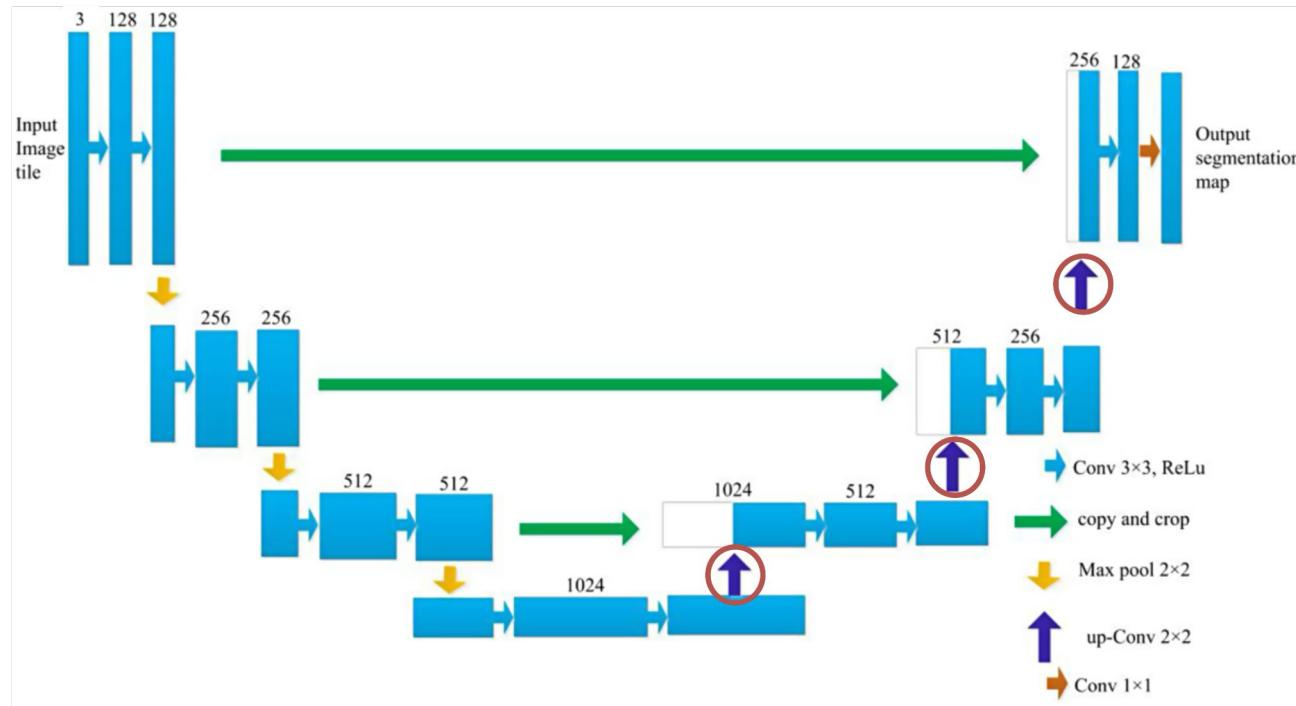
UNet



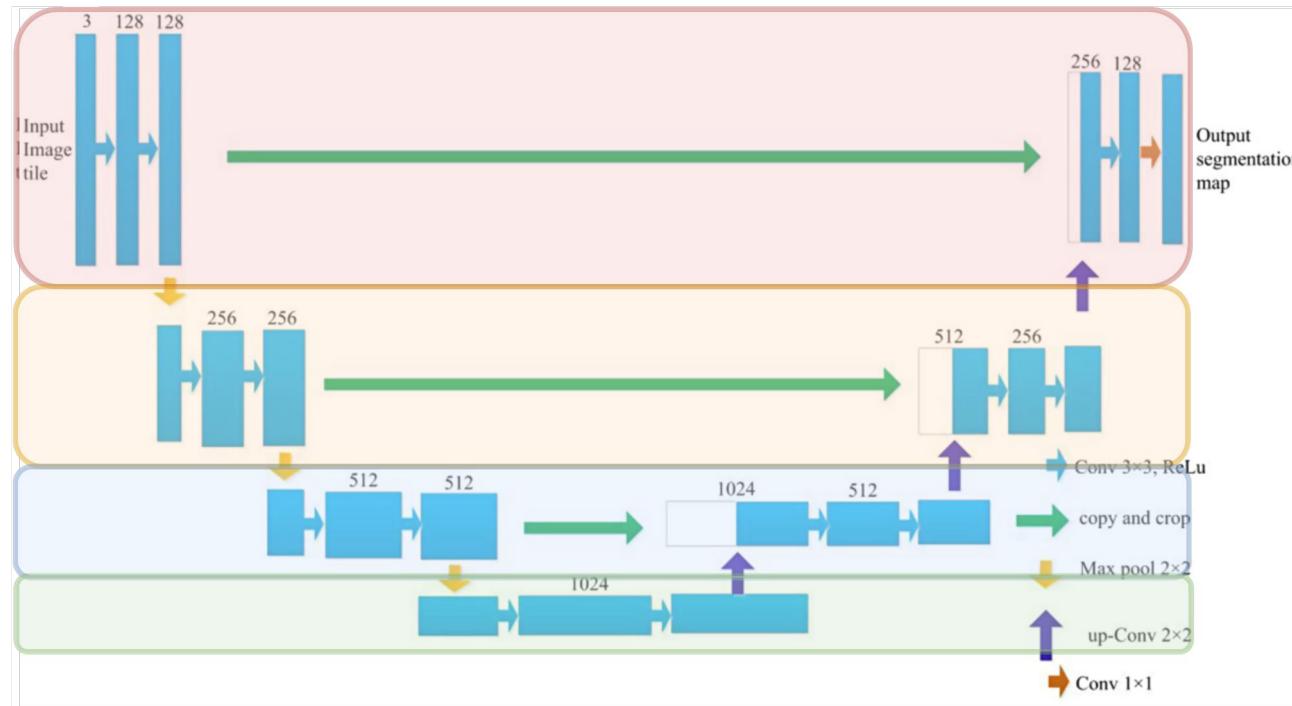
UNet



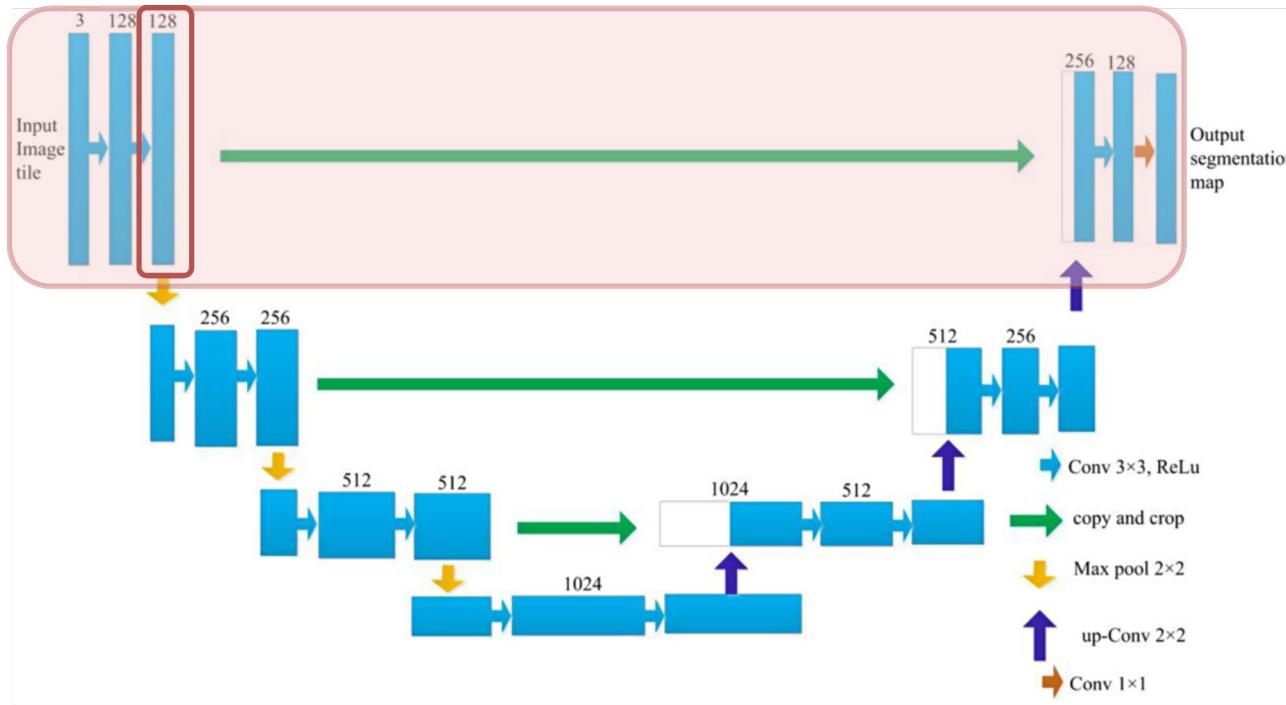
UNet



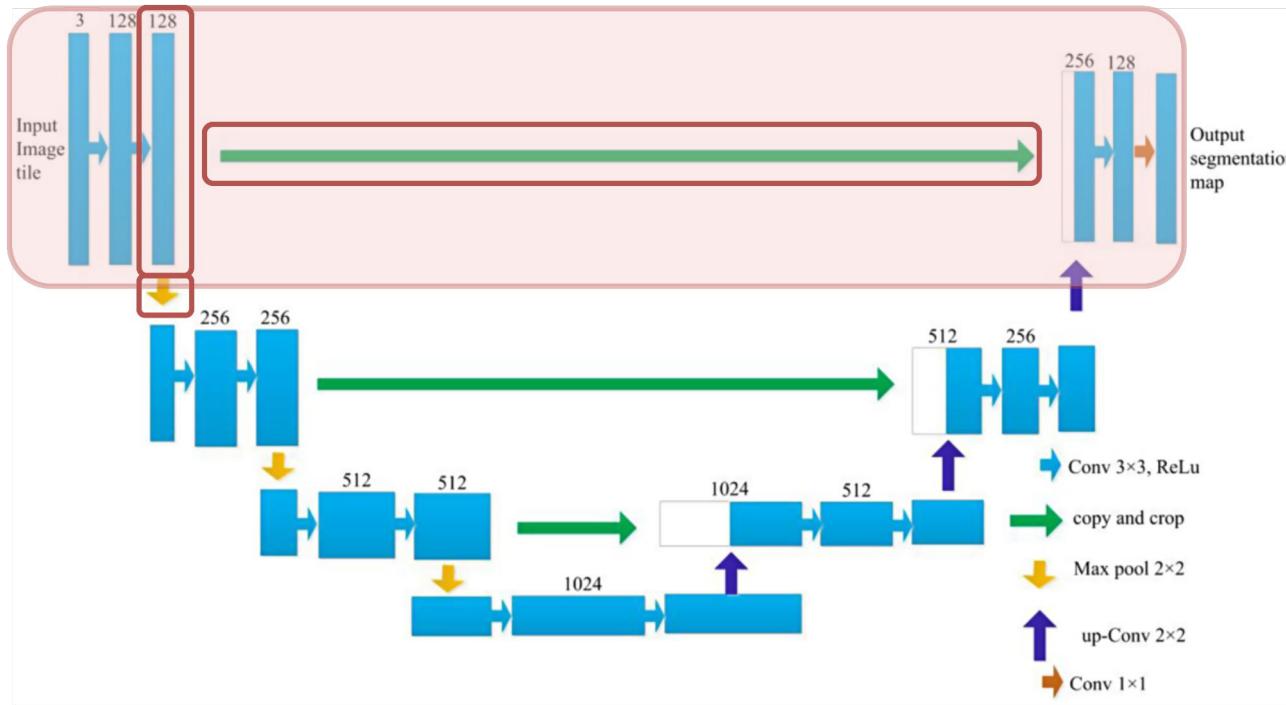
UNet



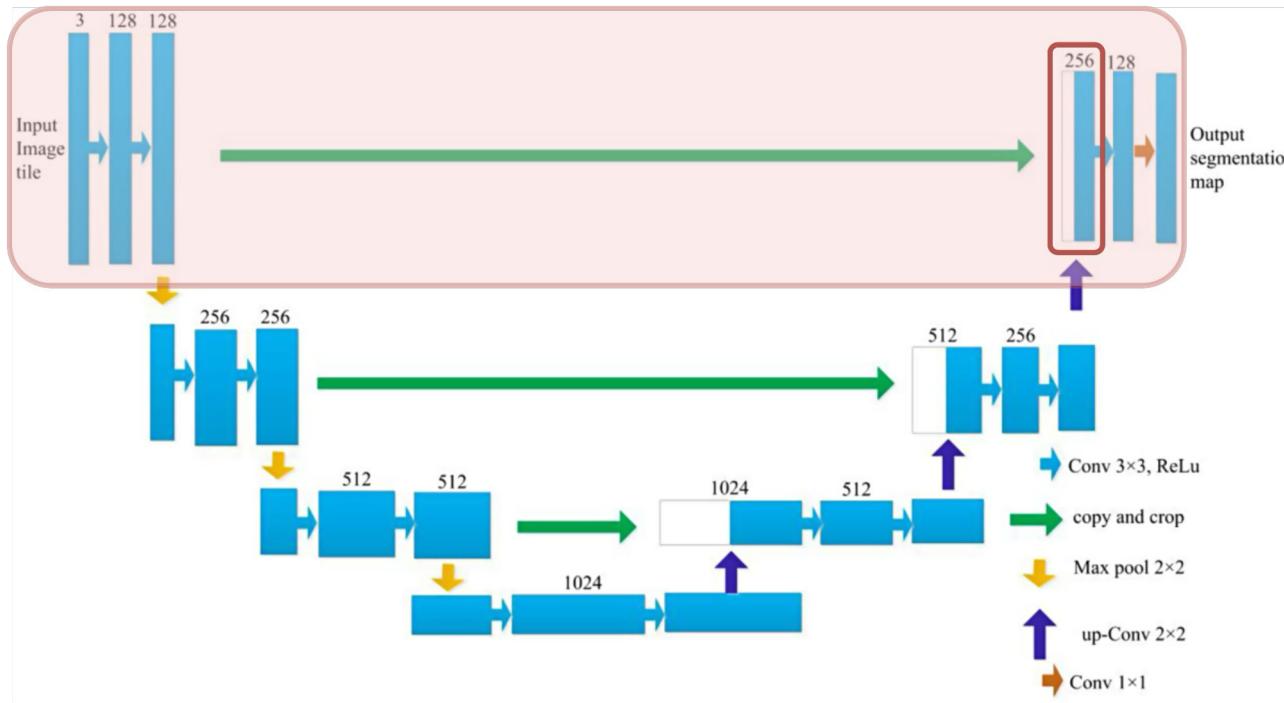
UNet



UNet



UNet



UNet

- Zmiany rozdzielczości ułatwiają propagację informacji za pomocą konwolucji - im mniejsza rozdzielcość map cech, tym mniej konwolucji potrzeba, żeby łączyć informacje z różnych miejsc na zdjęciu.
- Połączenia rezydualne (skip connections) poprawiają stabilność modelu usuwając problem zanikającego i eksplodującego gradientu.
- Do zmniejszania rozdzielczości zazwyczaj wykorzystuje się Max Pooling, natomiast do powrotu do wyższej rozdzielczości stosujemy np. konwolucję transponowaną.
- W przypadku segmentacji semantycznej, funkcją straty jest **entropia krzyżowa dla każdego piksela**.

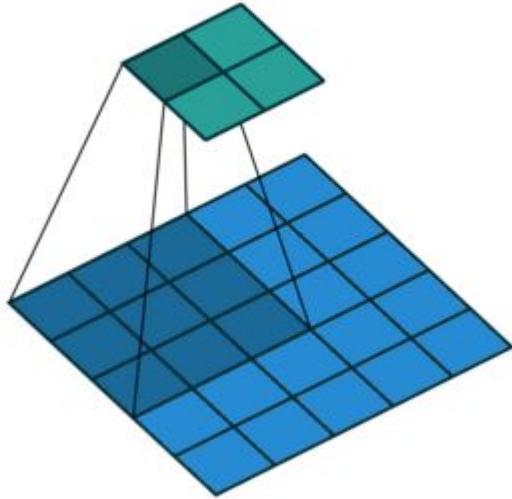
Polecane materiały

- Publikacje:
 - R-CNN <https://arxiv.org/abs/1311.2524>
 - Fast R-CNN <https://arxiv.org/abs/1504.08083>
 - Faster R-CNN <https://arxiv.org/abs/1506.01497>
 - YOLO <https://arxiv.org/abs/1506.02640>
 - UNet <https://arxiv.org/abs/1505.04597>
- Tutoriale online
 - metryka mAP - <https://kili-technology.com/data-labeling/machine-learning/mean-average-precision-map-a-complete-guide>
 - metryka IoU - <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>
 - zadanie detekcji - <https://www.v7labs.com/blog/object-detection-guide>
 - zadanie segmentacji - <https://medium.com/@raj.pulapakura/image-segmentation-a-beginners-guide-0ede91052db7>
- Materiały wideo
 - YOLO <https://www.youtube.com/watch?v=svn9-xV7wjk>
 - RCNN <https://www.youtube.com/watch?v=nJzQDpppFj0>
- Książki
 - Deep Learning for Vision Systems, Mohamed Elgendy
 - Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras, Rajalingappa Shanmugamani

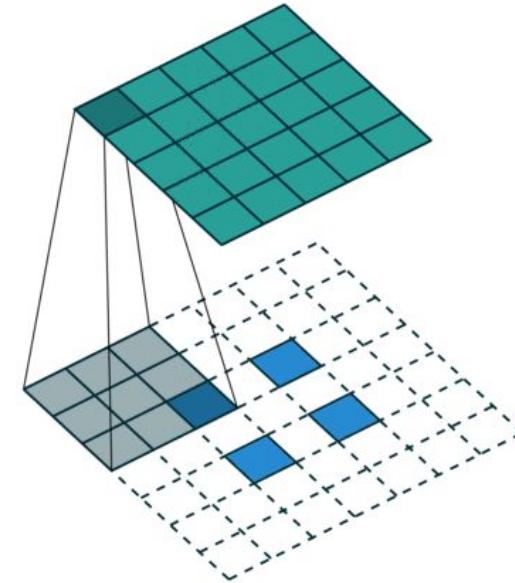
Dziękuję za uwagę.

Zapraszam do zadawania pytań.

Konwolucje - materiał dodatkowy



konwolucja 3x3



transponowana
konwolucja 3x3