

Autor: Szymon Tokarz

Data: 26.11.2024 r. Godz. 8.00

Ćwiczenie: Głębokie sieci neuronowe

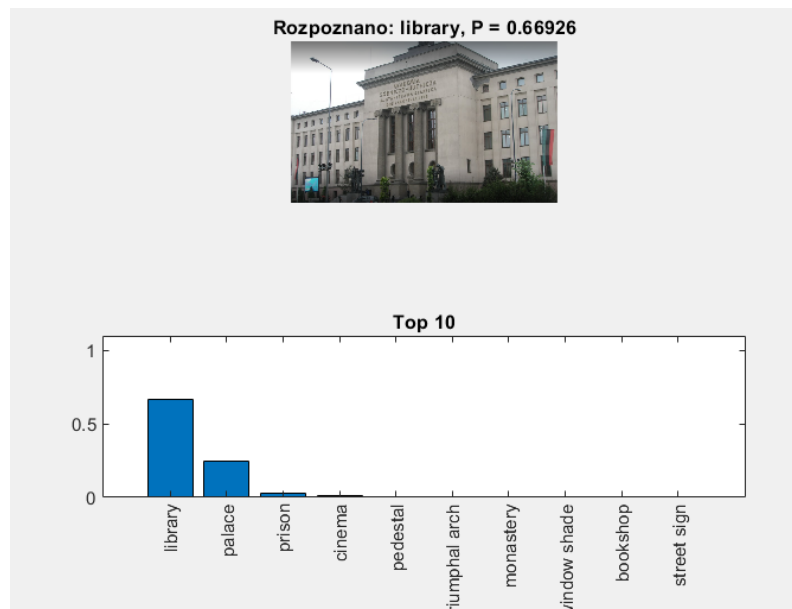
Rezultaty

Część 1

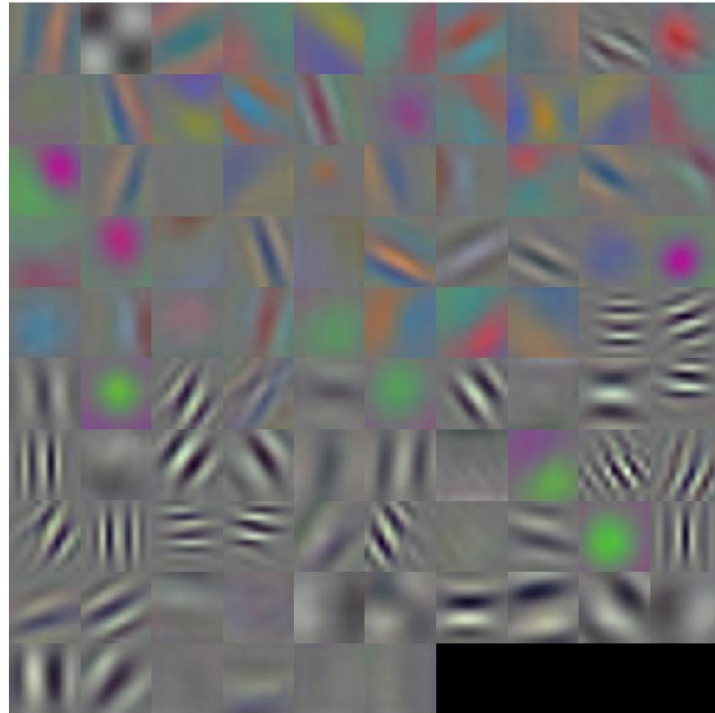
Sieć AlexNet wykorzystywana w ćwiczeniu ma 8 warstw. Na wejściu dostaje obraz o rozmiarach 227x227 3 – kanałowy i jest w stanie rozpoznać 1000 obiektów.

Rozmiar filtru w warstwie CONV1: 11,11. Ilość filtrów w warstwie CONV1: 96.

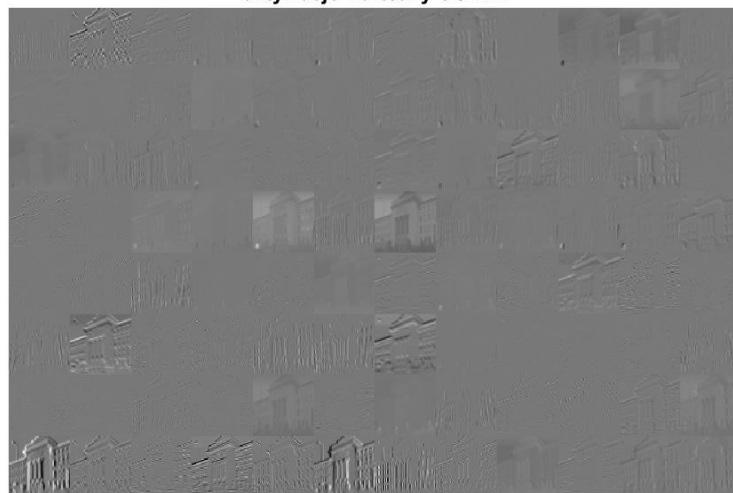
FilterSize to rozmiar filtru, NumChannels to liczba kanałów, a Stride to parametr, który ustala ruch kernela albo filtra.



tablica filtrów/wag warstwy CONV1)



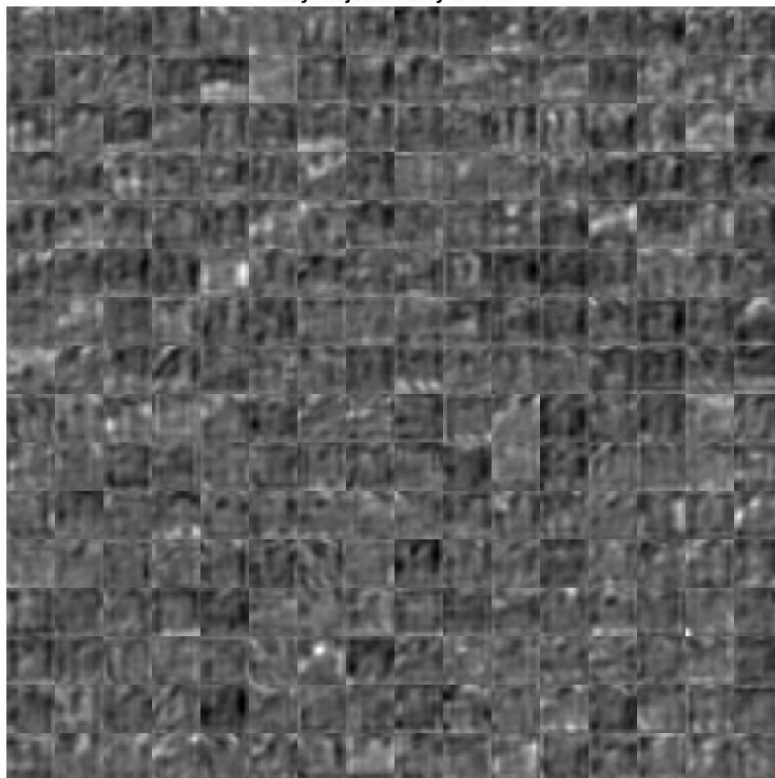
aktywacje warstwy CONV1



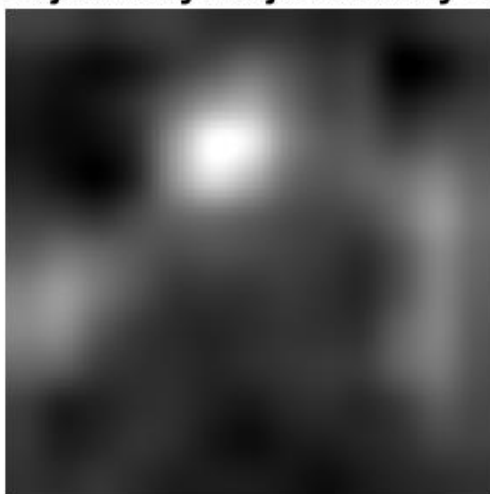
najsilniejsza aktywacja warstwy CONV1

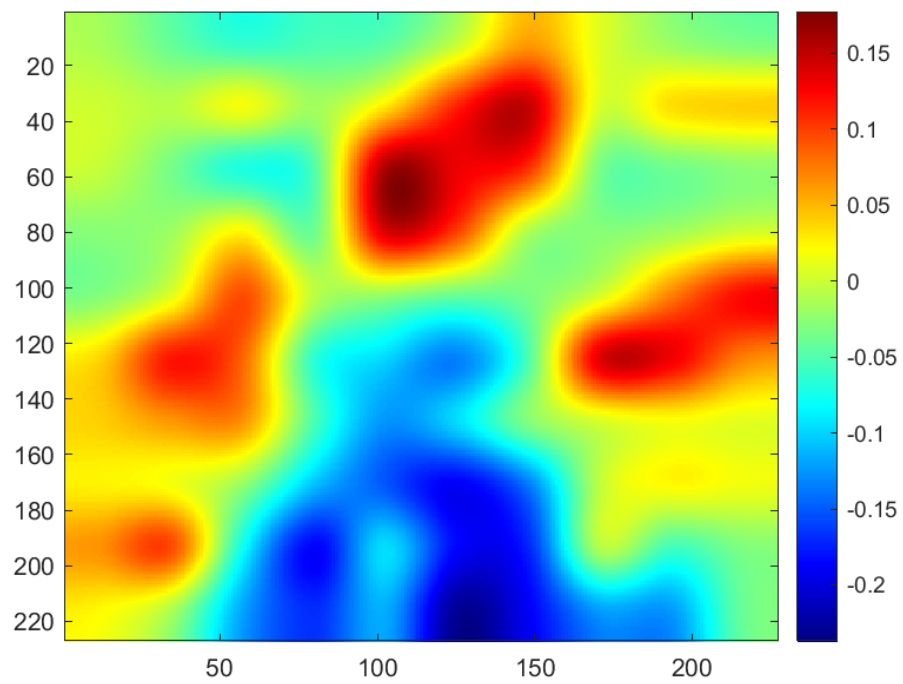
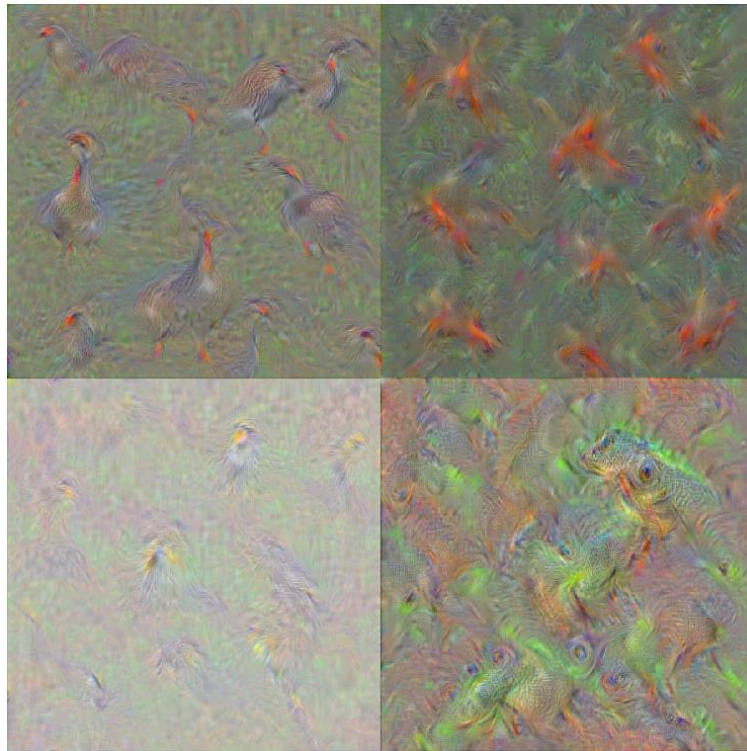


aktywacje warstwy CONV5



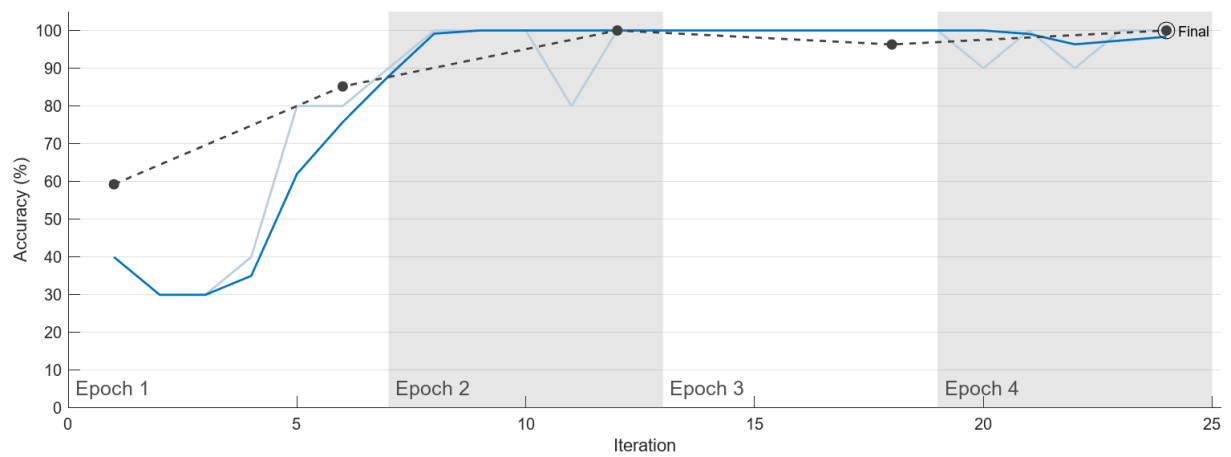
najsilniejsza aktywacja warstwy CONV5







Część 2



| | | | | |
|------------|--------|-----------------|--------|--------|
| True Class | obektA | 30 | | |
| | obektB | | 30 | |
| | obektC | | 3 | 27 |
| | | obektA | obektB | obektC |
| | | Predicted Class | | |

Dokładność na zbiorze walidacyjnym: 100%

Dokładność na zbiorze testowym: 96.7%

Najczęściej mylone klasy to C i B (telefon i ręka).

Analiza

W trakcie analizy skryptu *deep_learning_1.m* zidentyfikowano poszczególne elementy i parametry opisujące strukturę sieci neuronowej. Wizualizacja filtrów i wag sugeruje, że pierwsze warstwy sieci są odpowiedzialne za wyodrębnianie podstawowych cech, takich jak kolory grup pikseli, plamy czy krawędzie znajdujące się w obrębie obrazu. W głębszych warstwach sieci można zauważyć, że filtry zaczynają rozpoznawać bardziej złożone wzorce, takie jak konkretne plamy lub struktury na obrazie. Jednak trudno jednoznacznie określić, jakie dokładnie cechy są istotne dla rozpoznawania przez sieć.

Pytania

Wymień rodzaje warstw sieci konwolucyjnych i ich role

Sieci konwolucyjne składają się z różnych typów warstw, z których każda pełni określoną funkcję w przetwarzaniu danych wejściowych. Przykładowe warstwy to:

- **Warstwa konwolucyjna:** zawiera filtry o niewielkich rozmiarach (zwykle kilka na kilka pikseli), których parametry są optymalizowane podczas procesu uczenia. Filtry te

wykonują operacje splotu (iloczyn skalarny) na fragmentach danych wejściowych, generując dwuwymiarowe mapy aktywacji. Pozwala to na wyodrębnianie różnych cech, takich jak krawędzie, tekstury czy inne elementarne wzory.

- **Warstwa ReLU:** wprowadza nieliniowość do działania sieci za pomocą funkcji aktywacji $\text{ReLU}(x) = \max(0, x)$. Funkcja ta eliminuje wartości ujemne i zwiększa zdolność modelu do uczenia się bardziej złożonych zależności w danych.
- **Warstwa poolingowa (POOL):** realizuje proces subsamplingu (np. max-pooling), co zmniejsza wymiarowość danych, pozwalając na kompresję informacji i redukcję ryzyka przeuczenia. Warstwa ta umożliwia grupowanie cech i wprowadza odporność na niewielkie przesunięcia w danych wejściowych.
- **Warstwa w pełni połączona (fully connected):** znajduje się na końcu sieci i odpowiada za ostateczną interpretację wyników. Każdy neuron w tej warstwie jest połączony z wszystkimi neuronami poprzedniej warstwy, co pozwala na kompleksową analizę wyodrębnionych cech i ich klasyfikację.

Wyjaśnij pojęcie transfer learningu

Transfer Learning jest techniką wykorzystywaną w machine learningu polegającą na ponownym użyciu modelu do zwiększenia jakości lub szybkości rozwiązania dla podobnego zadania. Definicja brzmi następująco:

Dla dziedziny źródła D_s i zadania uczącego T_s , dziedzina celu D_t i zadania uczącego T_t , transfer learning ma na celu poprawę efektów uczenia funkcji przewidywania celu $f_t()$ w dziedzinie D_t wykorzystując wiedzę w dziedzinie D_s i zadaniu T_s , gdzie $D_s \neq D_t$ oraz $T_s \neq T_t$.