

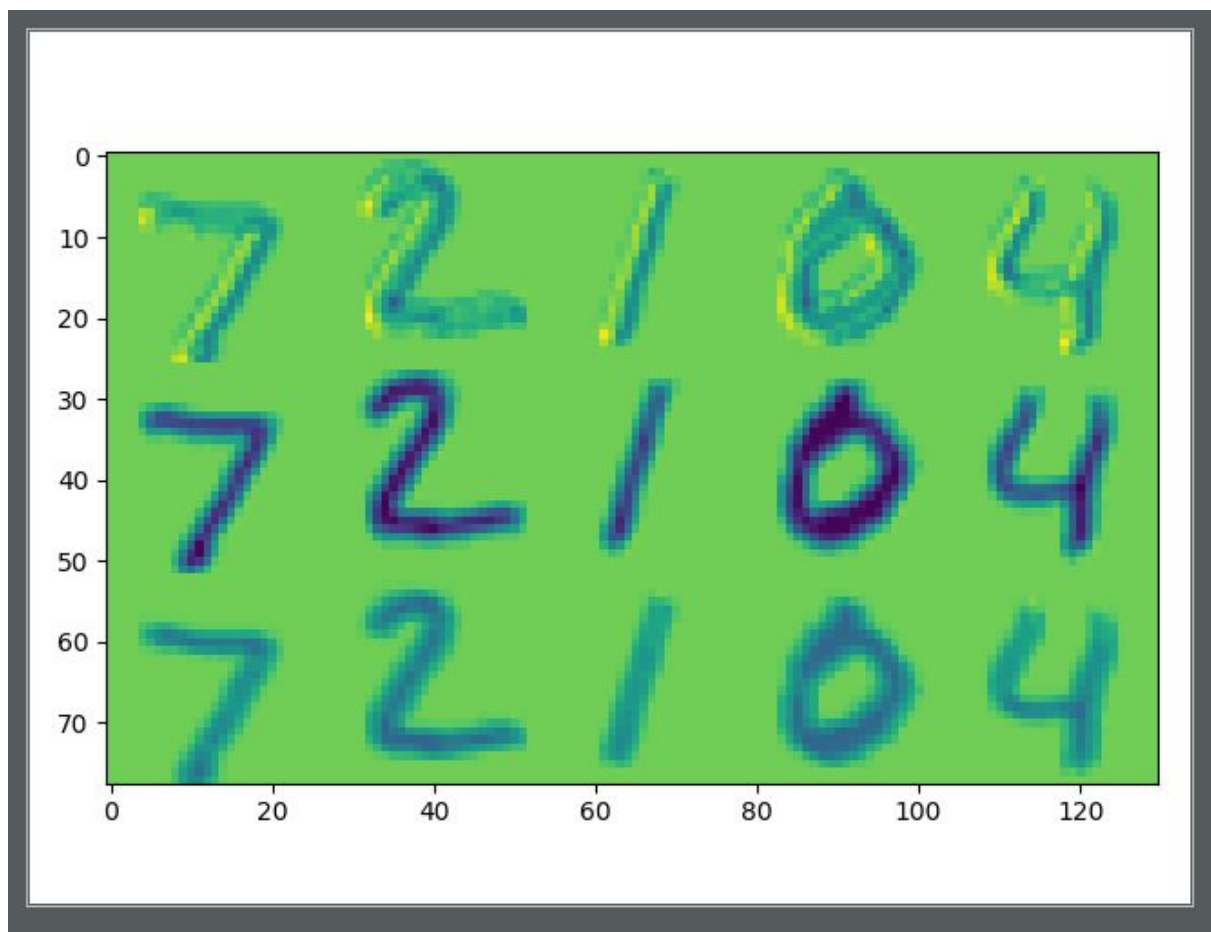
Sieci neuronowe - sprawozdanie

Józef Jasek

Zadanie 1.

Wizualizacja aktywacji warstw ukrytych

Tworzymy własną konwolucyjną sieć neuronową, a następnie trenujemy je na zbiorze mnist aż do otrzymania 90% skuteczności na zbiorze treningowym. Wykorzystano w tym celu framework PyTorch. Poniżej wyświetlamy output warstwy konwolucyjnej o rozmiarze liczba przykładów x liczba filtrów. Widzimy, że output nie różni się wiele od obrazu oryginalnego co świadczy o tym, że wytrenowana sieć decyduje o przynależności do odpowiednich klas przede wszystkim na podstawie późniejszych warstw. Wymagane byłoby dłuższe trenowanie.



Zadanie 2:

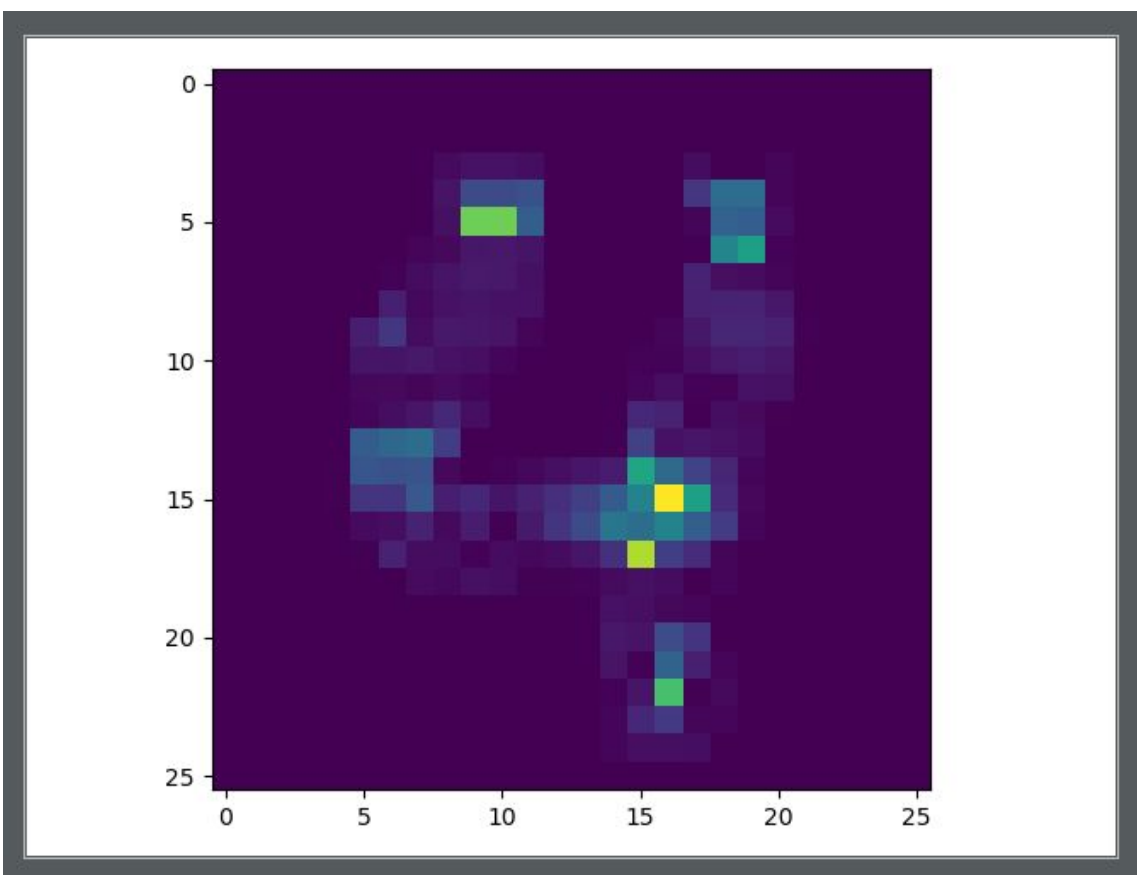
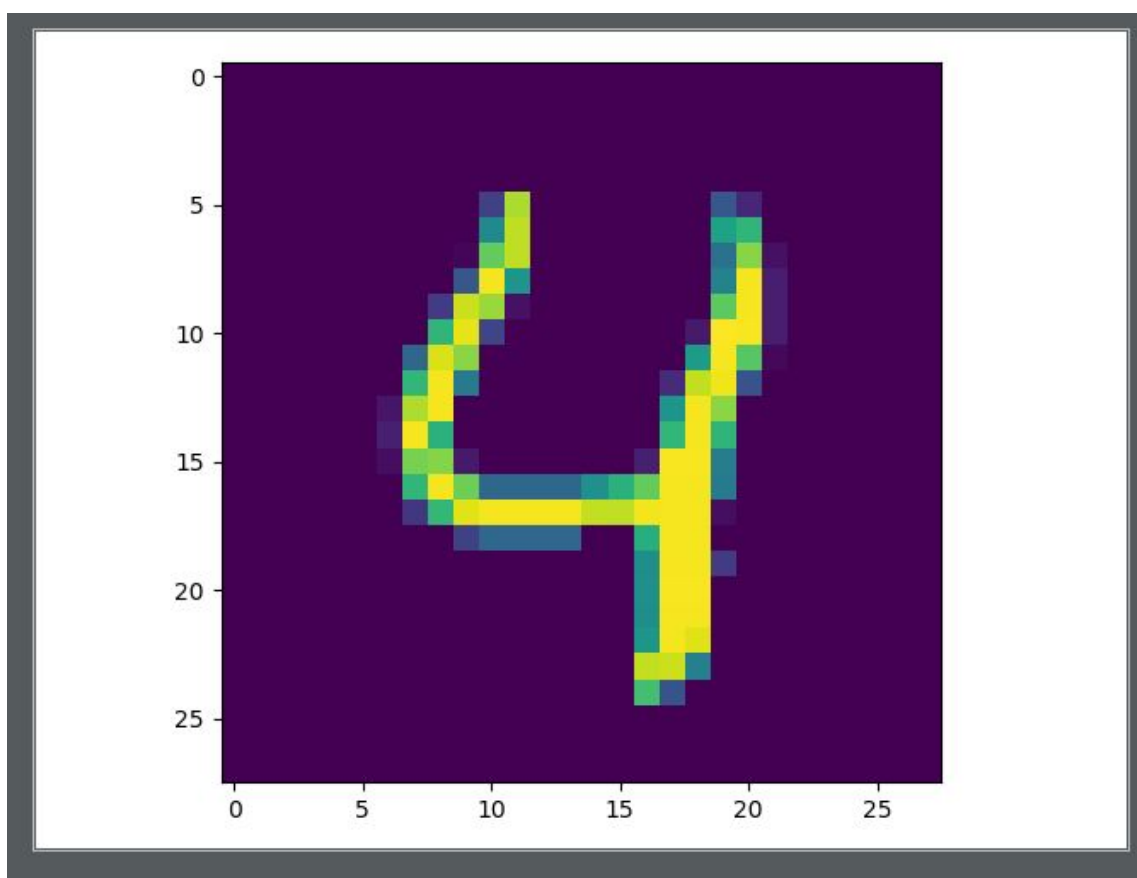
Heatmap

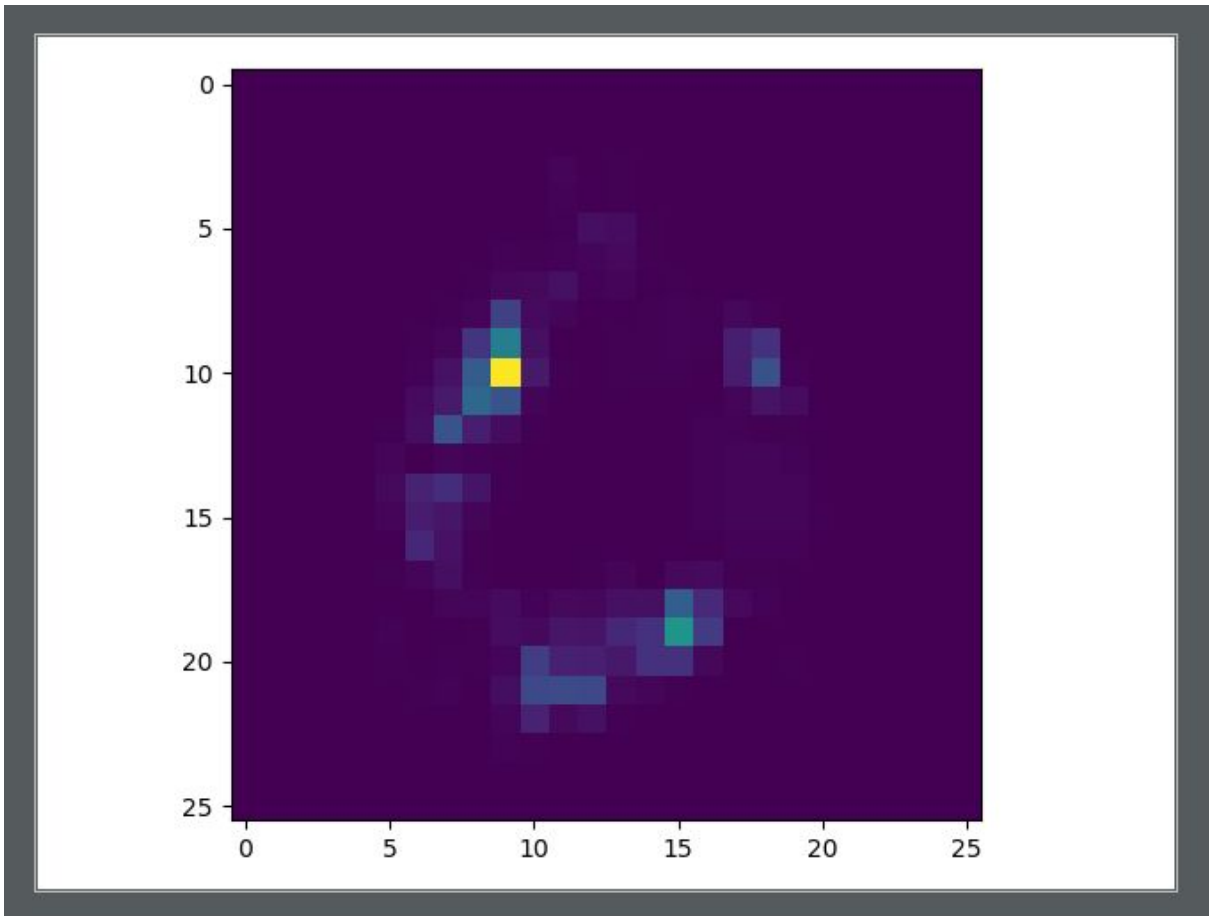
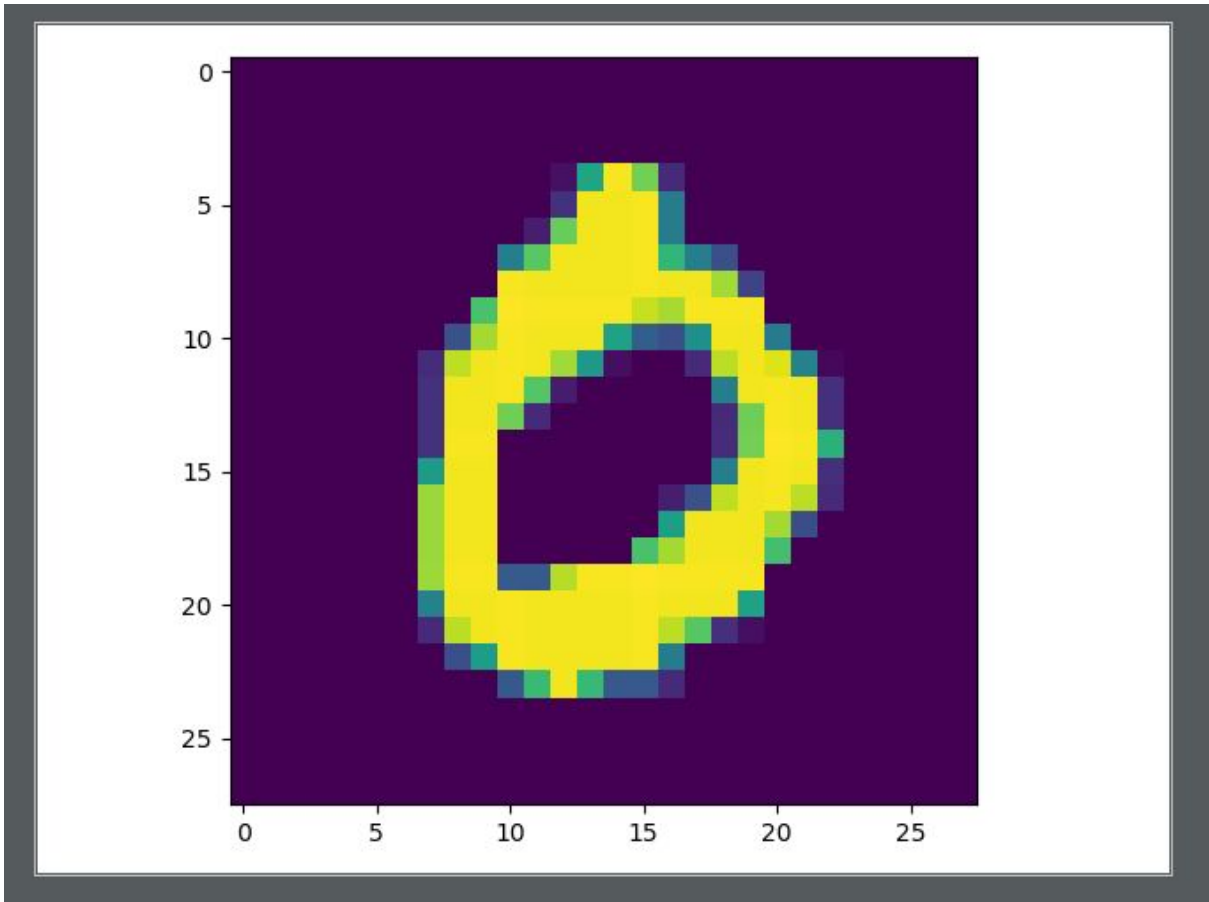
Tworzymy heatmapę przykładając czarny kwadrat 3x3 do każdego miejsca na obrazie i licząc dla niego $\text{abs}(x - y)$ gdzie

x - prawdopodobieństwo przynależności do klasy dla całego obrazka

y - prawdopodobieństwo przynależności do klasy dla obrazka z zasłoniętym fragmentem

Następnie plotujemy wszystkie otrzymane w ten sposób kwadraty.





Zgodnie z przypuszczeniami głównymi elementami odpowiedzialnymi za stwierdzanie przynależności do klasy są te piksele, gdzie następuje łączenie różnych linii oraz zakończenia cyfr.