

Jakub Tomaszewski 242555

Projekt II Zadanie II

## Architektury sieci Splotowych dla MNIST:

### SimpleCNN\_MNIST

**Opis:** Model składa się z dwóch warstw splotowych (Conv2d) z filtrami odpowiednio 16 i 32, każdej po nich następuje warstwa max-poolingu (MaxPool2d), która zmniejsza wymiary map cech. Po warstwach splotowych następuje warstwa spłaszczająca (Flatten) oraz dwie warstwy gęste (Linear), gdzie ostatnia odpowiada za klasyfikację na 10 klas. Parametry: learning rate= 0.001, output features= 256.

**Uzasadnienie wyboru:** Wybrano klasyczną, prostą architekturę CNN, która dobrze radzi sobie z prostymi zbiorami danych takimi jak MNIST. Dzięki temu można łatwo porównać wyniki z bardziej złożonymi rozwiązaniami.

### TwoFeatureCNN\_MNIST

**Opis:** Model również składa się z dwóch warstw splotowych (Conv2d) z filtrami 8 i 16, po których następuje max-pooling oraz warstwa spłaszczająca. Kluczową różnicą jest ostatnia warstwa gęsta przed klasyfikatorem, która redukuje wymiar przestrzeni cech do 2, co umożliwia wizualizację granic decyzyjnych. Parametr learning rate=0.01.

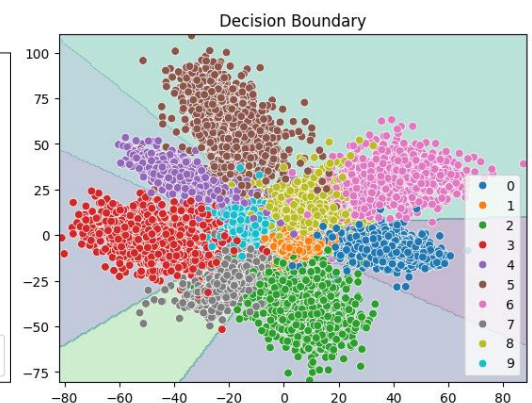
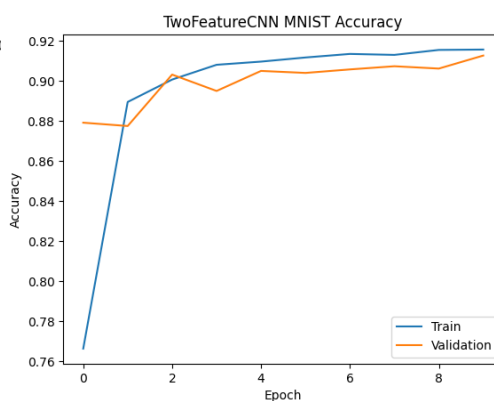
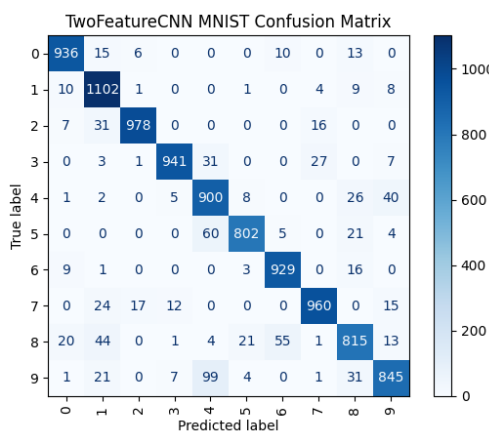
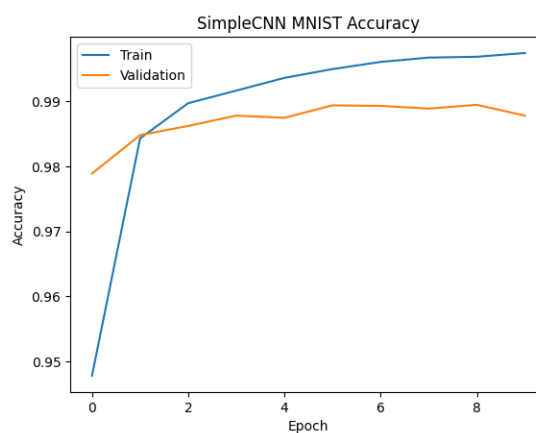
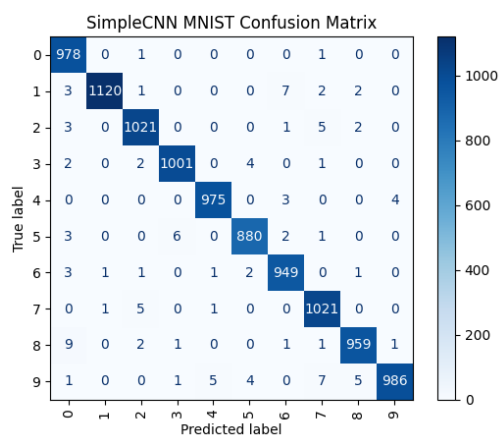
**Uzasadnienie wyboru:** Wybrano tę architekturę, aby umożliwić wizualną analizę cech wyuczonych przez model oraz ocenę separowalności klas w przestrzeni 2D.

## Najlepsze accuracy i parametry:

SimpleMNIST: test accuracy: 0.9890

TwoFeatureMNIST : test accuracy: 0.9208

## Wyniki klasyfikacji:



## Architektury sieci spłotowych dla Imagenette

### SimpleCNN\_Imagenette

**Opis:** Model składa się z dwóch warstw spłotowych (Conv2d) z filtrami 16 i 32, po których następuje max-pooling. Po warstwach spłotowych następuje warstwa spłaszczająca oraz dwie warstwy gęste, gdzie ostatnia odpowiada za klasyfikację na 10 klas. Parametry: learning rate= 0.001, output features= 64.

**Uzasadnienie wyboru:** Wybrano prostą architekturę, która pozwala na efektywne uczenie się cech z obrazów kolorowych i umożliwia porównanie z bardziej złożonymi modelami.

### TwoFeatureCNN\_Imagenette

**Opis:** Model analogiczny do SimpleCNN, ale ostatnia warstwa gęsta przed klasyfikatorem redukuje wymiar przestrzeni cech do 2, co umożliwia wizualizację granic decyzyjnych. Parametr learning rate=0.001.

**Uzasadnienie wyboru:** Wybrano tę architekturę, aby umożliwić wizualną analizę cech wyuczonych przez model oraz ocenę separowalności klas w przestrzeni 2D.

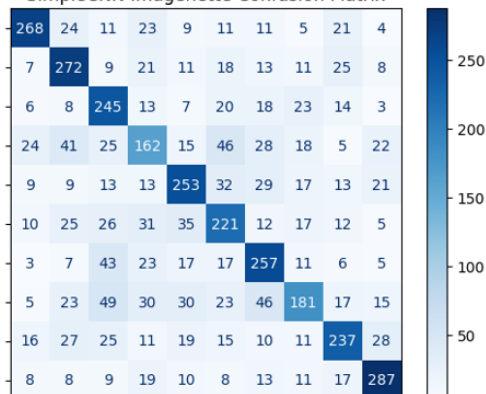
### Najlepsze accuracy i parametry:

SimpleIMAGENETTE: test accuracy: 0.6071

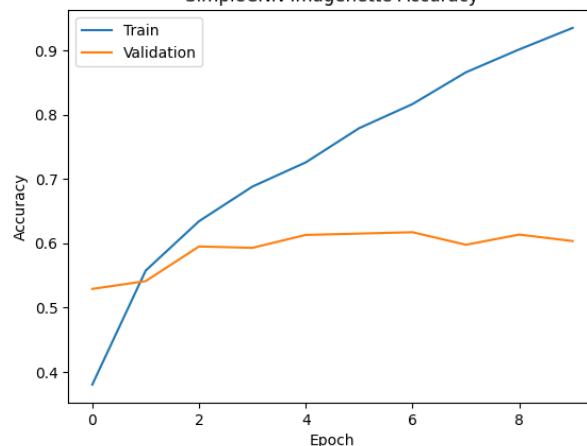
TwoFeatureIMAGENETTE : test accuracy: 0.3322

### Wyniki klasyfikacji:

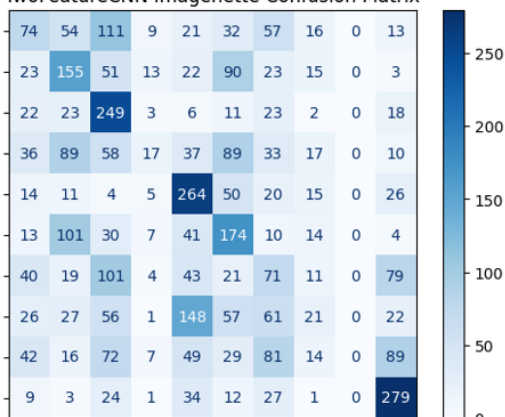
SimpleCNN Imagenette Confusion Matrix



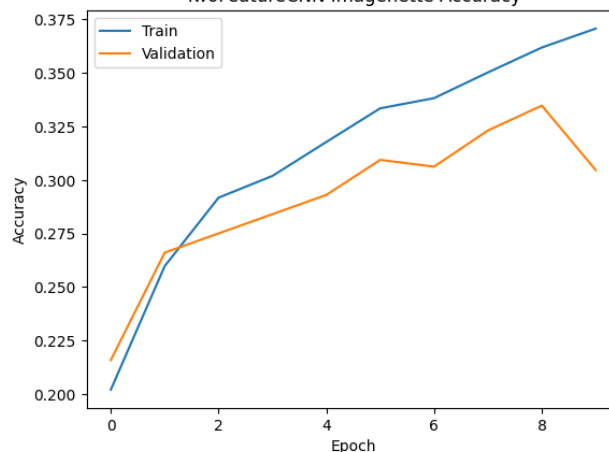
SimpleCNN Imagenette Accuracy



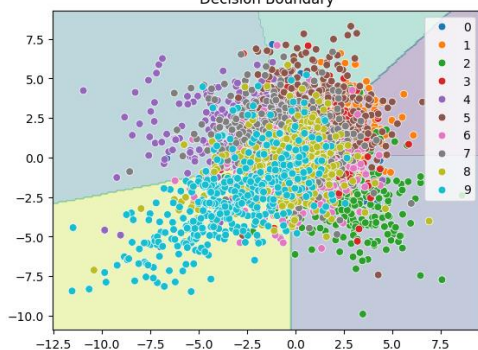
TwoFeatureCNN Imagenette Confusion Matrix



TwoFeatureCNN Imagenette Accuracy



Decision Boundary



## Opis metod augmentacji dla MNIST

- **Brak augmentacji:** Obrazy nie są modyfikowane.
- **Augmentacja 1:** Losowy obrót o 10 stopni.
- **Augmentacja 2:** Losowy obrót o 20 stopni oraz losowy crop (przycinanie) z paddingu 2 piksele.
- **Uzasadnienie wyboru:** Wybrano te transformacje, ponieważ losowy obrót i crop są typowe dla cyfr ręcznie pisanych i nie zniekształcają znacząco obrazu.

## Wyniki eksperymentu

### Simple MNIST

	60000	1000	200	100
<b>brak</b>	0.9902 ± 0.0011	0.9220 ± 0.0057	0.8175 ± 0.0261	0.7162 ± 0.0248
<b>aug1</b>	0.9914 ± 0.0014	0.9276 ± 0.0046	0.8100 ± 0.0288	0.7499 ± 0.0255
<b>aug2</b>	0.9916 ± 0.0008	0.9332 ± 0.0109	0.8369 ± 0.0250	0.7443 ± 0.0460

### Two feature MNIST

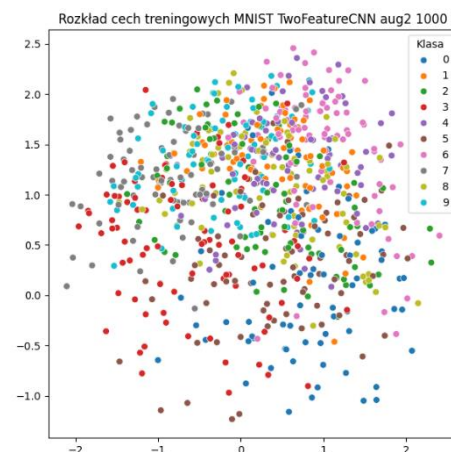
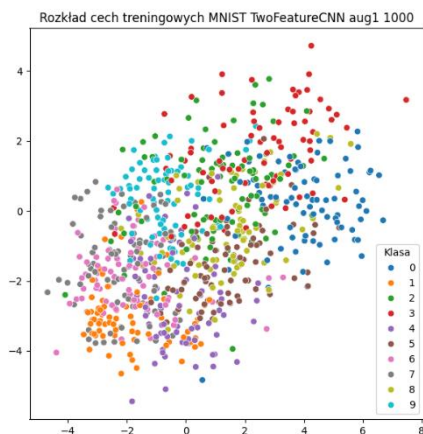
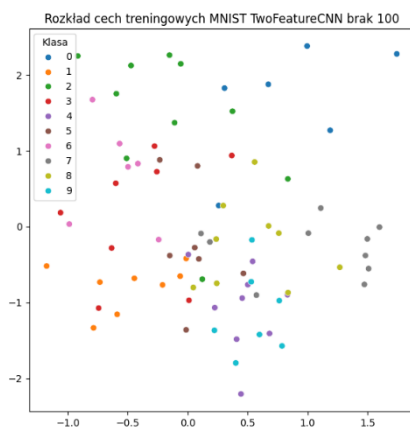
	60000	1000	200	100
<b>brak</b>	0.8737 ± 0.0118	0.4294 ± 0.0427	0.2700 ± 0.0394	0.2165 ± 0.0638
<b>aug1</b>	0.8822 ± 0.0192	0.4260 ± 0.0426	0.2906 ± 0.0765	0.1550 ± 0.0667
<b>aug2</b>	0.8634 ± 0.0170	0.3780 ± 0.0548	0.2323 ± 0.0565	0.1553 ± 0.0627



Pierwsza augmentacja



Druga augmentacja



## Opis metod augmentacji dla Imagenette

- **Brak augmentacji:** Obrazy nie są modyfikowane.
- **Augmentacja 1:** Losowy obrót o 15 stopni oraz losowe odbicie poziome.
- **Augmentacja 2:** Losowy obrót o 30 stopni, losowe odbicie poziome oraz losowy crop z paddingu 8 pikseli.
- **Uzasadnienie wyboru:** Te transformacje są typowe dla obrazów naturalnych i pozwalają zwiększyć różnorodność danych treningowych.

## Wyniki eksperymentu

### Simple Imagenette

	9469	1000	200	100
<b>brak</b>	$0.6137 \pm 0.0095$	$0.4102 \pm 0.0125$	$0.3004 \pm 0.0189$	$0.2591 \pm 0.0097$
<b>aug1</b>	$0.6104 \pm 0.0115$	$0.4220 \pm 0.0231$	$0.2917 \pm 0.0138$	$0.2327 \pm 0.0119$
<b>aug2</b>	$0.5819 \pm 0.0064$	$0.3519 \pm 0.0372$	$0.2664 \pm 0.0222$	$0.2332 \pm 0.0129$

### TwoFeature Imagenette

	9469	1000	200	100
<b>brak</b>	$0.3105 \pm 0.0097$	$0.1845 \pm 0.0207$	$0.1335 \pm 0.0220$	$0.1273 \pm 0.0180$
<b>aug1</b>	$0.3161 \pm 0.0175$	$0.1759 \pm 0.0252$	$0.1279 \pm 0.0115$	$0.1329 \pm 0.0171$
<b>aug2</b>	$0.2733 \pm 0.0165$	$0.1542 \pm 0.0246$	$0.1236 \pm 0.0221$	$0.1323 \pm 0.0273$



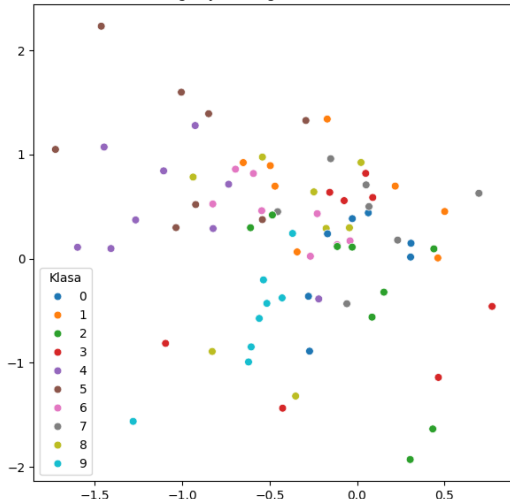
Pierwsza augmentacja



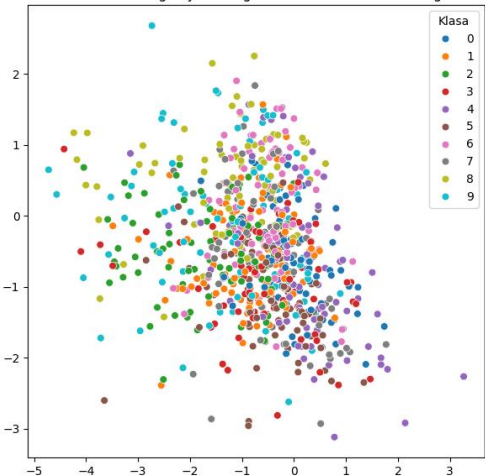
Druga augmentacja



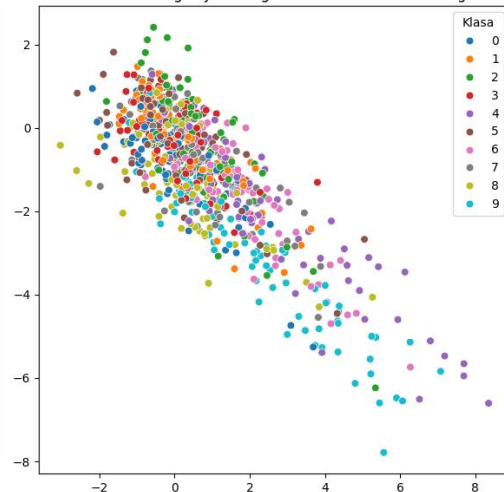
Rozkład cech treningowych Imagenette TwoFeatureCNN brak 100



Rozkład cech treningowych Imagenette TwoFeatureCNN aug1 1000



Rozkład cech treningowych Imagenette TwoFeatureCNN aug2 1000



## **Wnioski końcowe na podstawie przeprowadzonych eksperymentów:**

### **Wpływ architektury sieci splotowych:**

Prostsze modele (np. SimpleCNN) osiągają wyższą dokładność klasyfikacji na zbiorze testowym niż modele, w których ekstraktor cech redukuje wymiar do 2 (TwoFeatureCNN). Modele z dwuwymiarowym wektorem cech są jednak nieocenione w analizie wizualnej – pozwalają na ocenę separowalności klas i zrozumienie mechanizmu działania modelu na poziomie cech.

### **Wpływ liczby danych treningowych i augmentacji:**

Zwiększenie liczby przykładów treningowych oraz zastosowanie augmentacji (np. losowe obroty, przycinanie, odbicia) znacząco poprawia zdolność generalizacji modeli. Najlepsze rezultaty uzyskano przy pełnym zbiorze danych i rozsądnie dobranej augmentacji, co potwierdza, że różnorodność danych treningowych jest kluczowa dla skuteczności klasyfikacji. Jednakże dla zbioru ImageNet augmentacja nie przyniosła poprawy, możliwe że wykorzystane metody augmentacji nie były bardzo dobrze dopasowane do charakterystyki obrazków.

### **Wizualizacja cech i granic decyzyjnych:**

Dla modeli z dwuwymiarowym ekstraktorem cech możliwa była wizualizacja rozkładu klas oraz granic decyzyjnych. Pokazało to, że augmentacja pozwala lepiej wypełnić przestrzeń cech, co przekłada się na większą rozdzielczość granic decyzyjnych i lepsze wyniki klasyfikacji, zwłaszcza przy ograniczonej liczbie danych.

### **Zastosowanie dla innych zadań:**

Wyniki eksperymentów potwierdzają, że nawet proste architektury CNN są skuteczne w klasyfikacji obrazów, a augmentacja danych jest szczególnie istotna w sytuacjach, gdy dostępnych jest niewiele przykładów treningowych. Wizualizacja cech i analiza rozkładu klas mogą być pomocne przy projektowaniu modeli dla innych problemów klasyfikacyjnych.

### **Podsumowując:**

Dobór architektury, odpowiednia liczba danych treningowych oraz zastosowanie augmentacji mają kluczowy wpływ na skuteczność klasyfikacji obrazów. Uzyskane wyniki są uniwersalne i mogą być z powodzeniem wykorzystane w innych zadaniach klasyfikacyjnych, zwłaszcza tam, gdzie dostęp do dużych zbiorów danych jest ograniczony. Eksperymenty wykazały także, jak ważna jest możliwość wizualizacji cech i granic decyzyjnych dla pełniejszego zrozumienia działania modeli głębokiego uczenia.

Eksperymenty potwierdzają, że dobór architektury sieci, liczba danych treningowych oraz zastosowanie augmentacji mają istotny wpływ na zdolność modelu do klasyfikacji obrazów. Wyniki można przenieść na inne zadania klasyfikacji obrazów, zwłaszcza tam, gdzie dostęp do dużych zbiorów danych jest ograniczony.