

# Analiza i eksperymenty na zbiorze Fashion-MNIST

## 1. Wstęp

Celem tego dokumentu jest przedstawienie modelu bazowego oraz prostych eksperymentów wykonanych na zbiorze **Fashion-MNIST**. Dokument stanowi uzupełnienie repozytorium i ma pokazać sposób myślenia przy doborze architektury oraz parametrów treningu.

Projekt ma charakter praktyczny i koncentruje się na:

- poprawności procesu uczenia,
- obserwacji przebiegu treningu,
- identyfikacji problemu przeuczenia,
- ocenie wpływu prostych modyfikacji modelu.

Nie celem projektu było osiągnięcie maksymalnej możliwej dokładności, lecz zrozumienie zachowania modelu i świadomy wybór parametrów.

## 2. Analiza wyników modelu bazowego

Model bazowy składa się z:

- dwóch warstw konwolucyjnych z funkcją aktywacji ReLU,
- warstw **MaxPooling** zmniejszających rozmiar obrazu,
- jednej warstwy w pełni połączonej,
- warstwy wyjściowej klasyfikującej do 10 klas.

Model trenowano z użyciem:

- funkcji straty **CrossEntropyLoss**,
- optymalizatora **Adam**,
- learning rate = 0.001,
- batch size = 64,
- 10 epok treningu.

Dane treningowe zostały podzielone na zbiór treningowy i walidacyjny w proporcji 80/20. Podczas treningu zapisywano model o najlepszej dokładności walidacyjnej.

### Wyniki modelu bazowego:

- Dokładność na walidacji: ~92%
- Dokładność na zbiorze testowym: ~91%

Po kilku epokach zaobserwowano oznaki przeuczenia, co stało się punktem wyjścia do dalszych eksperymentów.

```
Epoch [1/10]    Train Loss: 0.4588, Train Acc: 0.8339 | Val Loss: 0.3394, Val Acc: 0.8789
Epoch [2/10]    Train Loss: 0.2985, Train Acc: 0.8926 | Val Loss: 0.2878, Val Acc: 0.8957
Epoch [3/10]    Train Loss: 0.2499, Train Acc: 0.9082 | Val Loss: 0.2724, Val Acc: 0.9001
Epoch [4/10]    Train Loss: 0.2183, Train Acc: 0.9195 | Val Loss: 0.2564, Val Acc: 0.9077
Epoch [5/10]    Train Loss: 0.1947, Train Acc: 0.9288 | Val Loss: 0.2362, Val Acc: 0.9133
Epoch [6/10]    Train Loss: 0.1698, Train Acc: 0.9367 | Val Loss: 0.2273, Val Acc: 0.9200
Epoch [7/10]    Train Loss: 0.1493, Train Acc: 0.9436 | Val Loss: 0.2327, Val Acc: 0.9187
Epoch [8/10]    Train Loss: 0.1286, Train Acc: 0.9527 | Val Loss: 0.2557, Val Acc: 0.9157
Epoch [9/10]    Train Loss: 0.1119, Train Acc: 0.9586 | Val Loss: 0.2602, Val Acc: 0.9171
Epoch [10/10]   Train Loss: 0.0948, Train Acc: 0.9646 | Val Loss: 0.2873, Val Acc: 0.9139
Training Finished
```

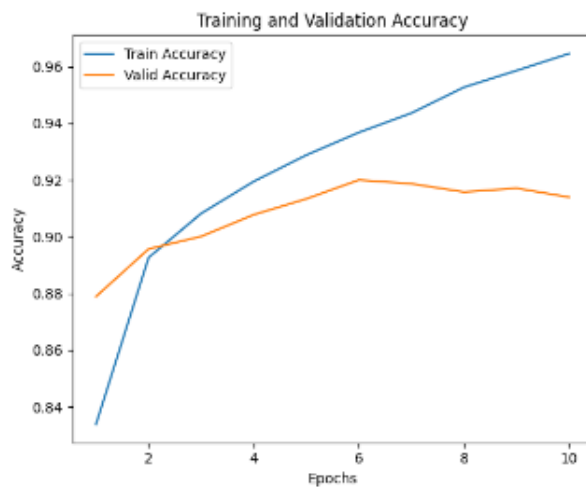
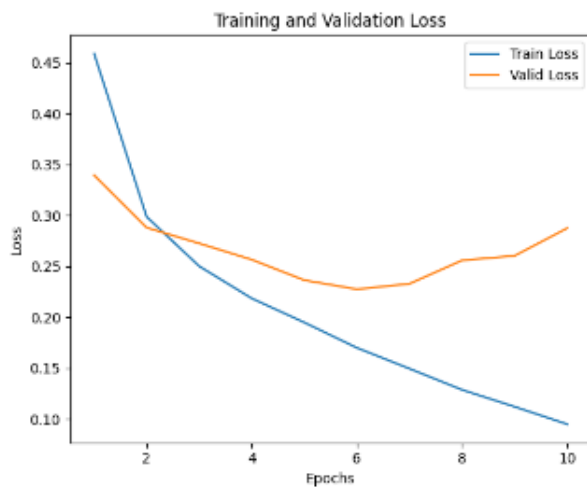
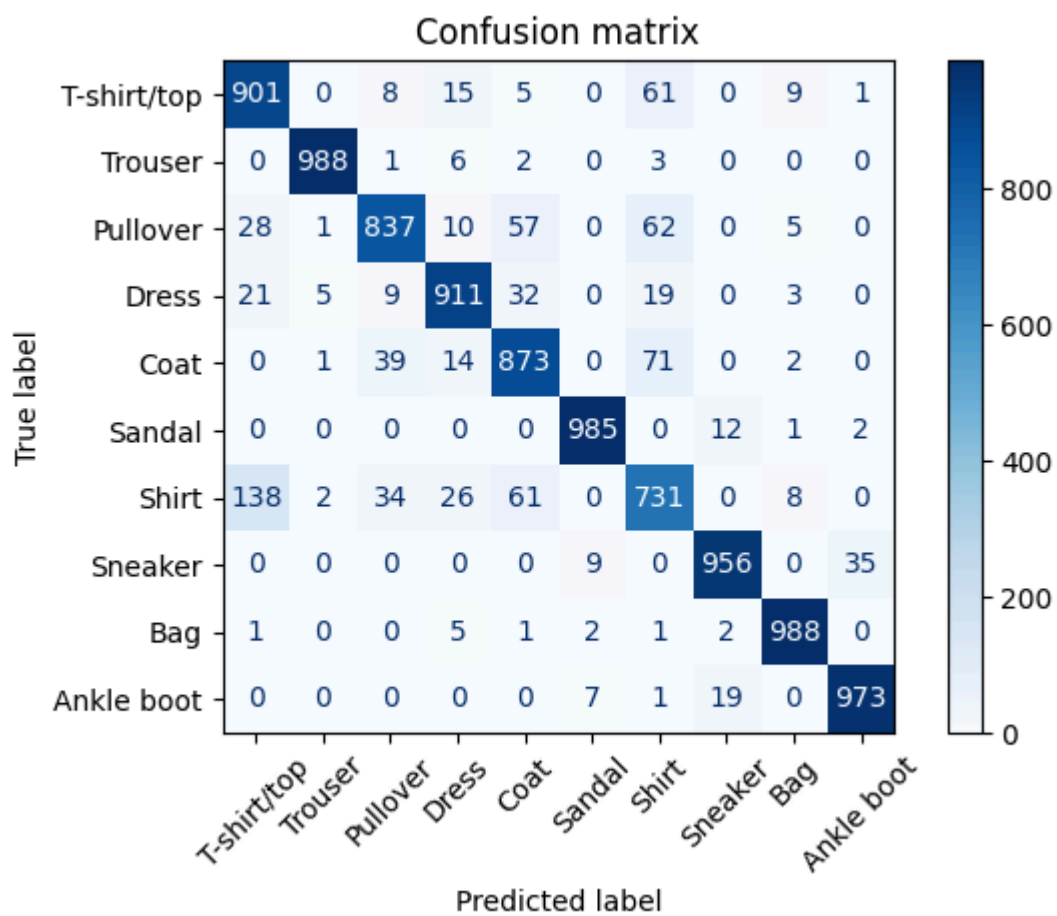
Start evaluating...

Test accuracy: 0.9143

Test loss: 0.24765027594566344

#### Classification Report

	precision	recall	f1-score	support
T-shirt/top	0.8274	0.9010	0.8626	1000
Trouser	0.9910	0.9880	0.9895	1000
Pullover	0.9019	0.8370	0.8683	1000
Dress	0.9230	0.9110	0.9170	1000
Coat	0.8468	0.8730	0.8597	1000
Sandal	0.9821	0.9850	0.9835	1000
Shirt	0.7703	0.7310	0.7501	1000
Sneaker	0.9666	0.9560	0.9613	1000
Bag	0.9724	0.9880	0.9802	1000
Ankle boot	0.9624	0.9730	0.9677	1000
accuracy			0.9143	10000
macro avg	0.9144	0.9143	0.9140	10000
weighted avg	0.9144	0.9143	0.9140	10000



### 3. Wstęp do eksperymentów

Na podstawie wyników modelu bazowego zaplanowano proste eksperymenty mające na celu sprawdzenie, jak wybrane zmiany wpływają na jakość modelu i jego zdolność do generalizacji. Eksperymenty obejmują:

- dodanie warstwy **Dropout** w celu ograniczenia przeuczenia,
- zastosowanie **Batch Normalization**,
- porównanie wyników różnych wariantów modelu z modelem bazowym.

Każdy wariant modelu trenowany był na tych samych danych i z tymi samymi ustawieniami treningowymi, co pozwala na bezpośrednie porównanie wyników i ocenę wpływu wprowadzonych zmian.

### 4. Poprawa i eliminacja problemów

#### 4.1 Early Stopping

Chociaż trening prowadzono przez wszystkie epoki, zapis modelu o najlepszej dokładności walidacyjnej pozwolił na efektywne zastosowanie **Early Stopping**, co oszczędza czas i ogranicza przeuczenie.

## 4.2 Dropout i zwiększona liczba epok

Start training...

```
Epoch [1/15] Train Loss: 0.5938, Train Acc: 0.7847 | Val Loss: 0.3790, Val Acc: 0.8587
Epoch [2/15] Train Loss: 0.3950, Train Acc: 0.8607 | Val Loss: 0.3133, Val Acc: 0.8851
Epoch [3/15] Train Loss: 0.3371, Train Acc: 0.8787 | Val Loss: 0.2862, Val Acc: 0.8938
Epoch [4/15] Train Loss: 0.3016, Train Acc: 0.8923 | Val Loss: 0.2653, Val Acc: 0.9040
Epoch [5/15] Train Loss: 0.2788, Train Acc: 0.8997 | Val Loss: 0.2552, Val Acc: 0.9083
Epoch [6/15] Train Loss: 0.2557, Train Acc: 0.9077 | Val Loss: 0.2546, Val Acc: 0.9087
Epoch [7/15] Train Loss: 0.2391, Train Acc: 0.9130 | Val Loss: 0.2414, Val Acc: 0.9137
Epoch [8/15] Train Loss: 0.2230, Train Acc: 0.9171 | Val Loss: 0.2371, Val Acc: 0.9167
Epoch [9/15] Train Loss: 0.2101, Train Acc: 0.9215 | Val Loss: 0.2341, Val Acc: 0.9223
Epoch [10/15] Train Loss: 0.1990, Train Acc: 0.9255 | Val Loss: 0.2279, Val Acc: 0.9202
Epoch [11/15] Train Loss: 0.1861, Train Acc: 0.9301 | Val Loss: 0.2293, Val Acc: 0.9235
Epoch [12/15] Train Loss: 0.1747, Train Acc: 0.9339 | Val Loss: 0.2279, Val Acc: 0.9217
Epoch [13/15] Train Loss: 0.1667, Train Acc: 0.9361 | Val Loss: 0.2250, Val Acc: 0.9232
Epoch [14/15] Train Loss: 0.1593, Train Acc: 0.9396 | Val Loss: 0.2399, Val Acc: 0.9234
Early stopping at epoch 14
Training Finished
```

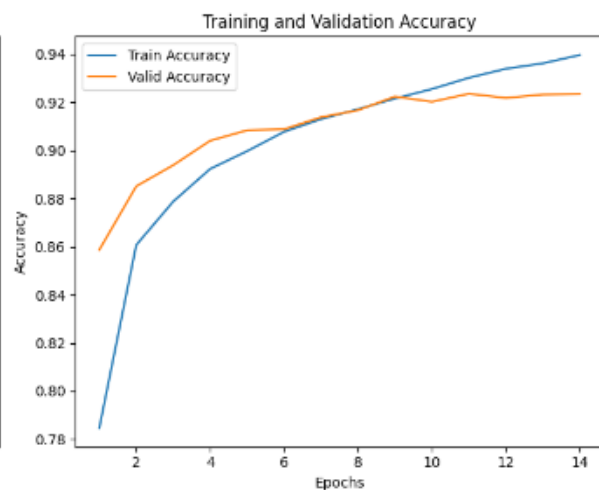
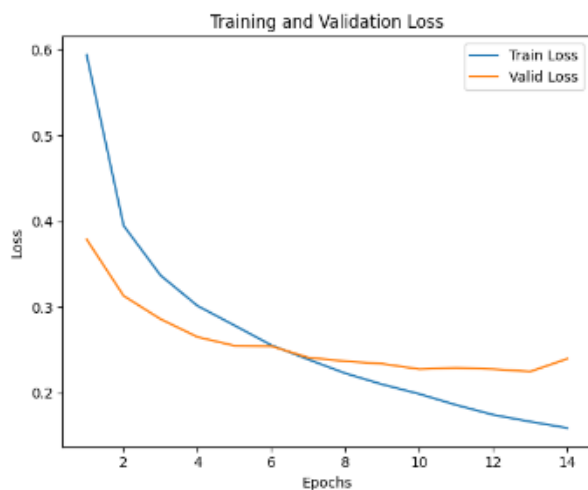
Start evaluating...

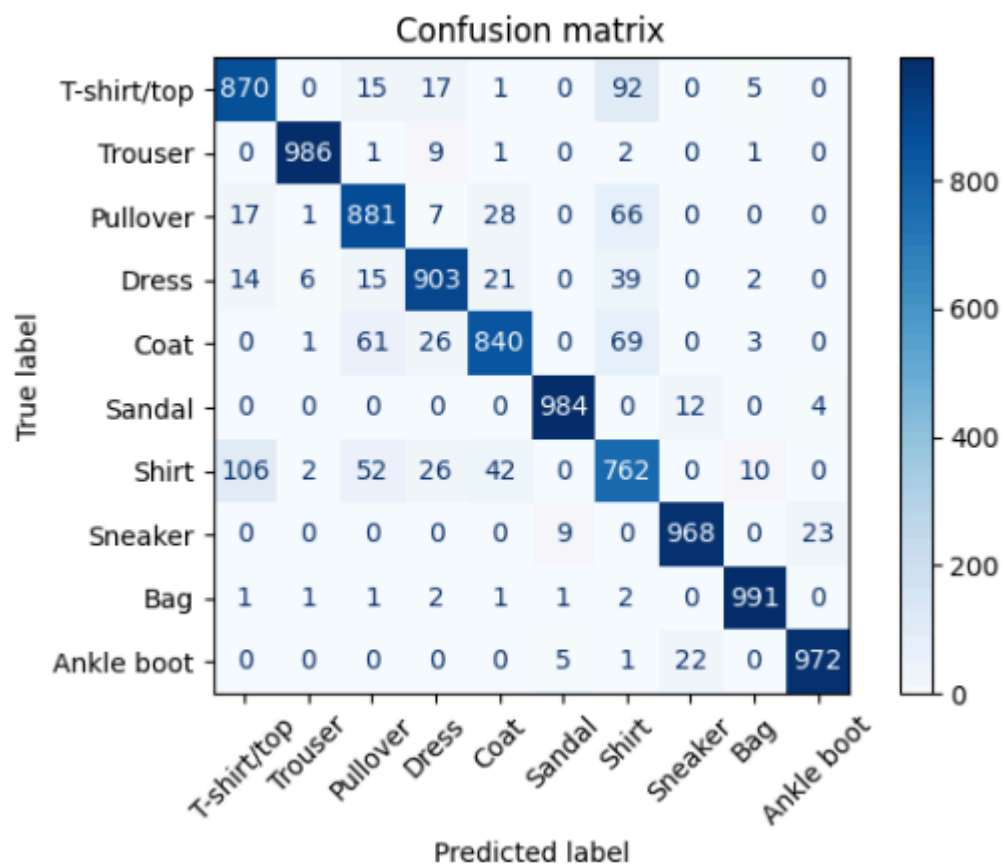
Test accuracy: 0.9157

Test loss: 0.23939590423703194

Classification Report

	precision	recall	f1-score	support
T-shirt/top	0.8631	0.8700	0.8665	1000
Trouser	0.9890	0.9860	0.9875	1000
Pullover	0.8587	0.8810	0.8697	1000
Dress	0.9121	0.9030	0.9075	1000
Coat	0.8994	0.8400	0.8687	1000
Sandal	0.9850	0.9840	0.9845	1000
Shirt	0.7377	0.7620	0.7496	1000
Sneaker	0.9661	0.9680	0.9670	1000
Bag	0.9792	0.9910	0.9851	1000
Ankle boot	0.9730	0.9720	0.9725	1000
accuracy			0.9157	10000
macro avg	0.9163	0.9157	0.9159	10000
weighted avg	0.9163	0.9157	0.9159	10000





Dodanie warstwy **Dropout** zmniejszało overfitting i stabilizowało dokładność walidacyjną. Nie zwiększyło jednak znacząco dokładności testowej w przypadku prostego CNN na Fashion-MNIST. Jest to typowy efekt regularizacji – mniejsze ryzyko przeuczenia kosztem wolniejszego treningu.

## 4.3 Batch Normalization

Start training...

```
Epoch [1/15]   Train Loss: 0.3964, Train Acc: 0.8542 | Val Loss: 0.3047, Val Acc: 0.8923
Epoch [2/15]   Train Loss: 0.2705, Train Acc: 0.9022 | Val Loss: 0.2581, Val Acc: 0.9077
Epoch [3/15]   Train Loss: 0.2298, Train Acc: 0.9146 | Val Loss: 0.2552, Val Acc: 0.9067
Epoch [4/15]   Train Loss: 0.2032, Train Acc: 0.9232 | Val Loss: 0.2341, Val Acc: 0.9143
Epoch [5/15]   Train Loss: 0.1814, Train Acc: 0.9331 | Val Loss: 0.2250, Val Acc: 0.9187
Epoch [6/15]   Train Loss: 0.1592, Train Acc: 0.9400 | Val Loss: 0.2325, Val Acc: 0.9211
Epoch [7/15]   Train Loss: 0.1403, Train Acc: 0.9464 | Val Loss: 0.2348, Val Acc: 0.9187
Epoch [8/15]   Train Loss: 0.1195, Train Acc: 0.9558 | Val Loss: 0.2504, Val Acc: 0.9147
Epoch [9/15]   Train Loss: 0.1060, Train Acc: 0.9615 | Val Loss: 0.2533, Val Acc: 0.9206
```

Early stopping at epoch 9

Training Finished

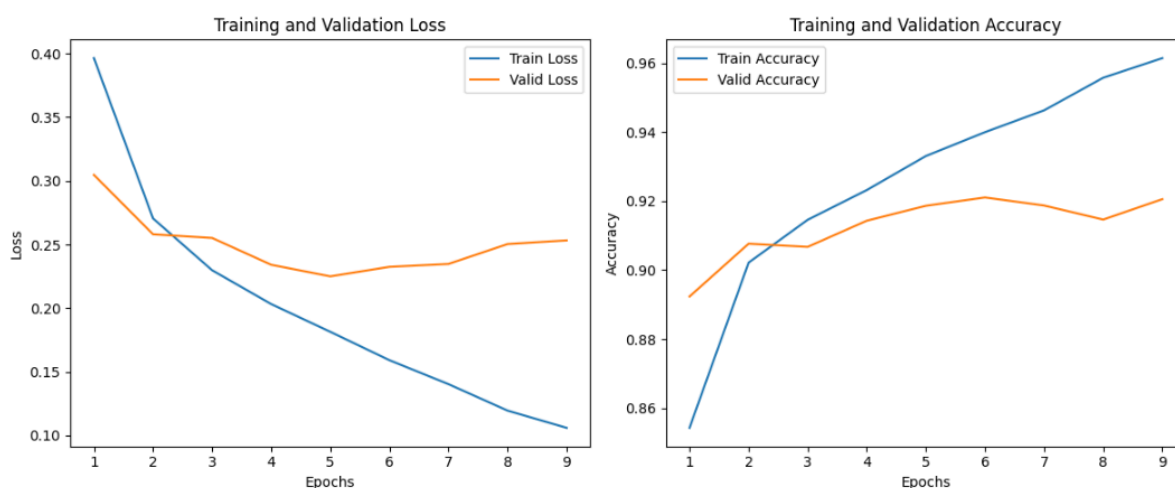
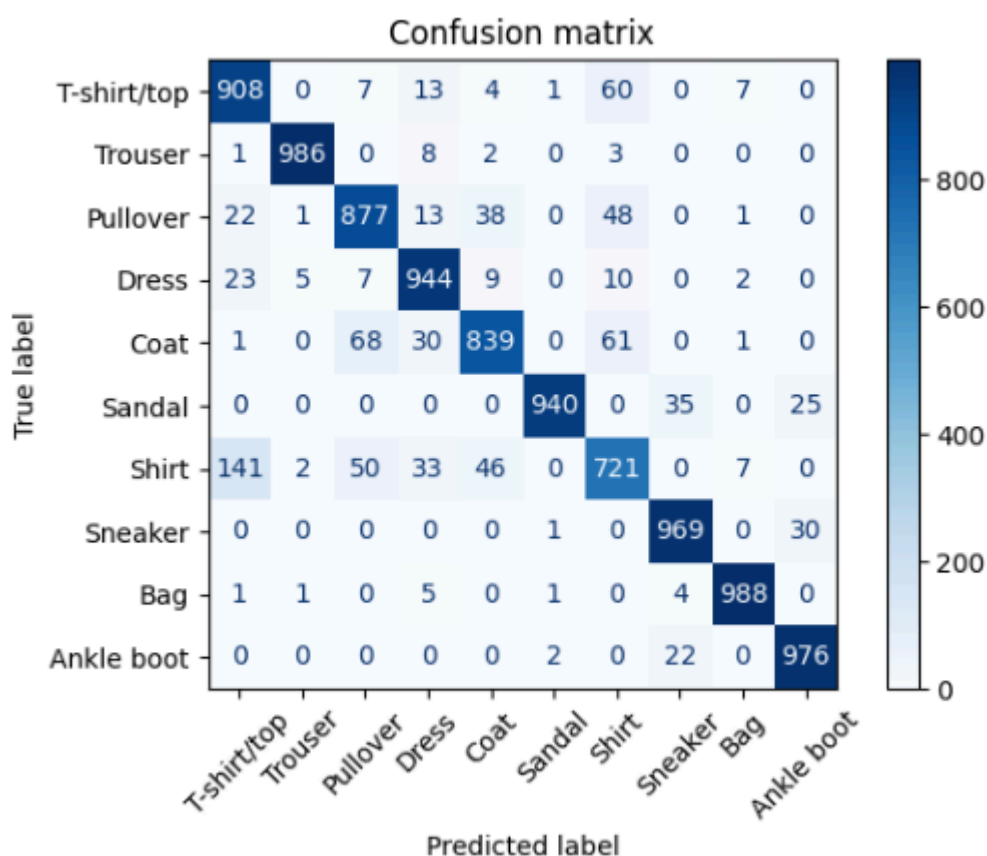
Start evaluating...

Test accuracy: 0.9148

Test loss: 0.24508068487644197

Classification Report

	precision	recall	f1-score	support
T-shirt/top	0.8277	0.9080	0.8660	1000
Trouser	0.9910	0.9860	0.9885	1000
Pullover	0.8692	0.8770	0.8731	1000
Dress	0.9025	0.9440	0.9228	1000
Coat	0.8945	0.8390	0.8658	1000
Sandal	0.9947	0.9400	0.9666	1000
Shirt	0.7984	0.7210	0.7578	1000
Sneaker	0.9408	0.9690	0.9547	1000
Bag	0.9821	0.9880	0.9850	1000
Ankle boot	0.9467	0.9760	0.9611	1000
accuracy			0.9148	10000
macro avg	0.9147	0.9148	0.9141	10000
weighted avg	0.9147	0.9148	0.9141	10000



Zastosowanie **Batch Normalization** poprawiło stabilność treningu. Model osiągnął dokładność walidacyjną ~92% przy wcześniejszym zatrzymaniu (Early Stopping po 9 epokach). Dokładność testowa wyniosła ~91.5%, co pokazuje, że BatchNorm poprawia stabilność uczenia, ale nie wpływa znacząco na końcową jakość modelu.

Najtrudniejsze do sklasyfikowania pozostały klasy wizualnie podobne, takie jak **Shirt** i **T-shirt/top**, natomiast łatwe klasy (**Trouser**, **Bag**, **Sandal**) osiągnęły wysokie metryki.



## 4.4 Minimalna augmentacja danych

Start training...

Epoch [1/15]	Train Loss: 0.4588, Train Acc: 0.8339	Val Loss: 0.3394, Val Acc: 0.8789
Epoch [2/15]	Train Loss: 0.2985, Train Acc: 0.8926	Val Loss: 0.2878, Val Acc: 0.8957
Epoch [3/15]	Train Loss: 0.2499, Train Acc: 0.9082	Val Loss: 0.2724, Val Acc: 0.9001
Epoch [4/15]	Train Loss: 0.2183, Train Acc: 0.9195	Val Loss: 0.2564, Val Acc: 0.9077
Epoch [5/15]	Train Loss: 0.1947, Train Acc: 0.9288	Val Loss: 0.2362, Val Acc: 0.9133
Epoch [6/15]	Train Loss: 0.1698, Train Acc: 0.9367	Val Loss: 0.2273, Val Acc: 0.9200
Epoch [7/15]	Train Loss: 0.1493, Train Acc: 0.9436	Val Loss: 0.2327, Val Acc: 0.9187
Epoch [8/15]	Train Loss: 0.1286, Train Acc: 0.9527	Val Loss: 0.2557, Val Acc: 0.9157
Epoch [9/15]	Train Loss: 0.1119, Train Acc: 0.9586	Val Loss: 0.2602, Val Acc: 0.9171

Early stopping at epoch 9

Training Finished

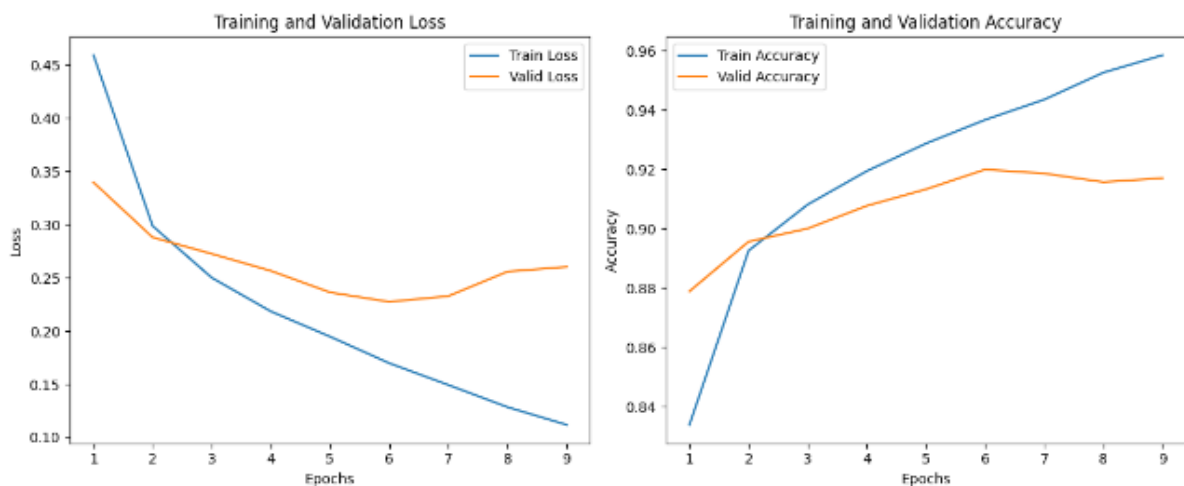
Start evaluating...

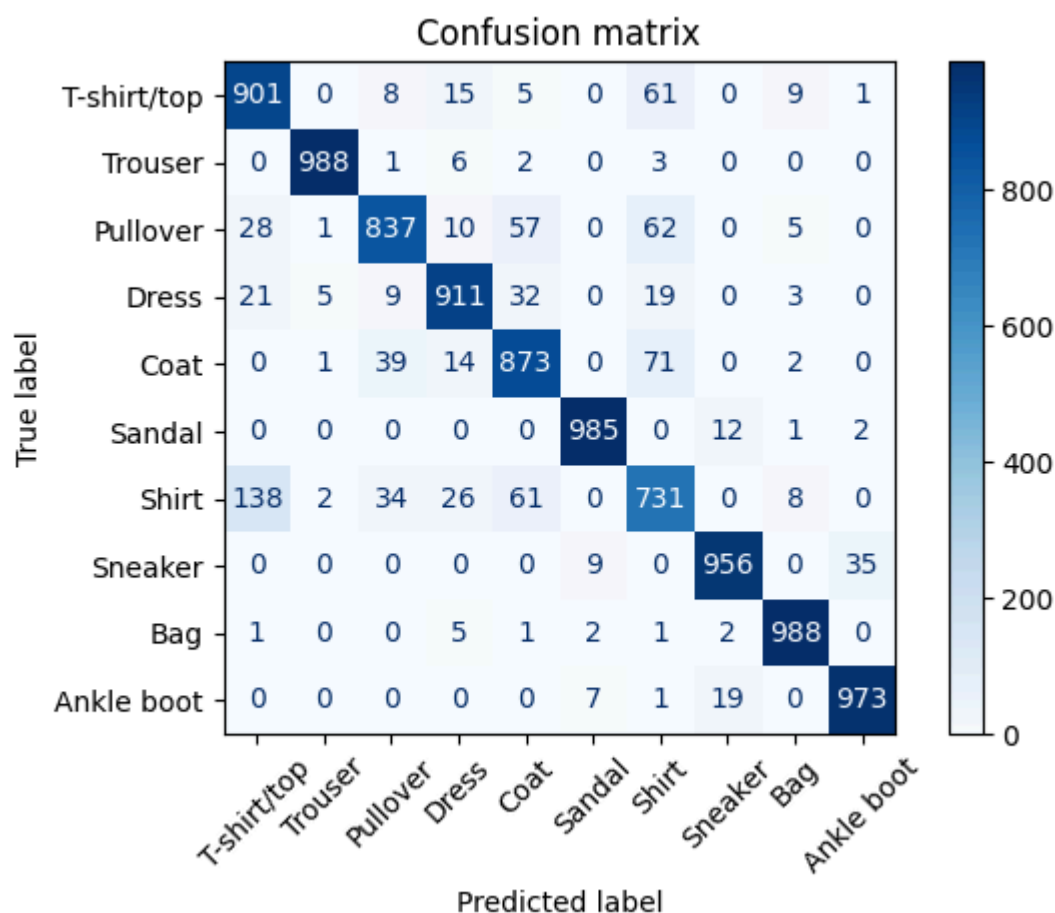
Test accuracy: 0.9143

Test loss: 0.24765027594566344

Classification Report

	precision	recall	f1-score	support
T-shirt/top	0.8274	0.9010	0.8626	1000
Trouser	0.9910	0.9880	0.9895	1000
Pullover	0.9019	0.8370	0.8683	1000
Dress	0.9230	0.9110	0.9170	1000
Coat	0.8468	0.8730	0.8597	1000
Sandal	0.9821	0.9850	0.9835	1000
Shirt	0.7703	0.7310	0.7501	1000
Sneaker	0.9666	0.9560	0.9613	1000
Bag	0.9724	0.9880	0.9802	1000
Ankle boot	0.9624	0.9730	0.9677	1000
accuracy			0.9143	10000
macro avg	0.9144	0.9143	0.9140	10000
weighted avg	0.9144	0.9143	0.9140	10000





Po zastosowaniu minimalnej augmentacji danych model osiągnął podobną dokładność testową (~91.4%) jak w baseline. Dokładność walidacyjna rosła stabilnie do ~92%, a Early Stopping zatrzymał trening po 9 epokach, co ograniczyło przeuczenie.

Augmentacja nie zmieniła znacząco ogólnej dokładności, lecz mogła nieco poprawić wyniki dla trudniejszych klas wizualnie podobnych (**T-shirt/top**, **Shirt**).

## 5. Wnioski ogólne

1. Model bazowy osiągnął wysoką dokładność (~91–92%), ale szybko pojawiły się oznaki przeuczenia.
2. **Dropout** stabilizuje walidację i ogranicza overfitting, ale nie poprawia znacząco końcowej dokładności testowej.
3. **Batch Normalization** zwiększa stabilność treningu i pozwala na wcześniejsze zatrzymanie (Early Stopping), zachowując wysoką dokładność.
4. Minimalna augmentacja danych ma ograniczony wpływ na ogólną dokładność, lecz może poprawić klasy trudniejsze do rozróżnienia.
5. Proste modyfikacje architektury i regularizacji pozwalają lepiej zrozumieć zachowanie modelu i świadomie dobierać parametry treningowe.