

МІНІСТЕРСТВО ОСВІТИ І НАУКИ  
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЛЬВІВСЬКА ПОЛІТЕХНІКА»  
Кафедра систем штучного інтелекту



Лабораторна робота №2  
З курсу “Обробка зображень методами штучного інтелекту”

Виконала:  
студентка групи КН-408  
Жук Анастасія

Викладач:  
Пелешко Д. Д.

**Тема:** Класифікація зображень. Застосування нейромереж для пошуку подібних зображень.

**Мета:** набути практичних навиків у розв'язанні задачі пошуку подібних зображень на прикладі організації CNN класифікації.

### **Теоретичні відомості**

В основі класифікації (для пошуку подібних ) зображень пропонується використовувати Siamese networks. Ідея складається в тому щоб взяти випадково ініціалізовану мережу і застосувати її до зображень, щоб дізнатися наскільки вони схожі. Модель має значно полегшати виконання таких задач, як візуальний пошук по базі даних зображень, так як вона буде мати просту метрику подібності між 0 та 1 замість 2D масивів.

## Варіант 2

1. Побудувати CNN на основі AlexNet для класифікації зображень на основі датасету fashion-mnist. Зробити налаштування моделі для досягнення необхідної точності

На базі Siamese networks побудувати систему для пошуку подібних зображень в датасеті fashion-mnist. Візуалізувати отримані результати t-SNE.

## Хід роботи

1. Скачаємо датасет FashionMnist



рис. 1 Звізуалізовані картинки з датасету

Набір даних Fashion-MNIST складається із зображень одягу (футболок, штанів, суконь тощо), які походять із каталогу зображень Zalando. Zalando — європейська компанія електронної комерції, заснована у 2008 році.

## 2. Знайшли будову нейромережі AlexNet

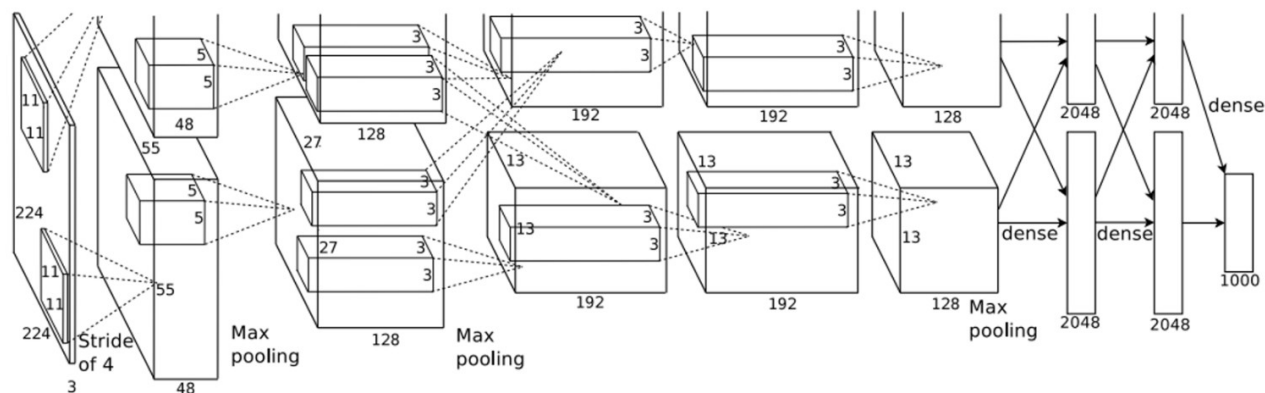


рис. 2 AlexNet

## 3. Збудувати модель

Перший шар в мережі моделі, `keras.layers.Flatten`, перетворює формат зображень із двовимірного масиву (28 на 28 пікселів) до одновимірного масиву ( $28 * 28 = 784$  пікселі). Цей шар розбирає рядки пікселів у зображенні та вирівнює їх у ряд і не має параметрів для вивчення; він лише переформатовує дані.

Подається в глибоку нейронну мережу з трьома прихованими шарами та вихідним шаром. У нейронних мережах із прямим зв'язком кількість одиниць зменшується, коли ми рухаємося вниз по шарах до вихідного шару, з кількох причин, головним чином через те, що ви вимагає більше нейронів для обробки вхідних даних на попередніх рівнях, а останній рівень відповідає кількості класів, необхідних для класифікації, у випадку набору даних fashionMNIST це буде 10.

## Компоненти CNN

### Шар Conv2D (keras.layers.Conv2D)

Ми використовуємо клас `tf.keras.layers.Conv2D` для побудови згорткових шарів усередині згорткової нейронної мережі. У згортковому шарі виконується операція згортки між вхідним зображенням і ядром/фільтром шару перетворення. Вихід згорткового шару є результатом операції згортки між вхідними даними та значеннями ядра/фільтра.

### Шар пакетної нормалізації (keras.layers.BatchNormalization)

Пакетна нормалізація (BN) — це техніка, яка пом'якшує вплив нестабільних градієнтів у глибоких нейронних мережах. BN вводить додатковий шар до нейронної мережі, який виконує операції над входами попереднього шару. Операція стандартизує та нормалізує вхідні значення. Потім вхідні значення перетворюються за допомогою операцій масштабування та зсуву.

### Максимальний шар об'єднання (keras.layers.MaxPool2D)

Максимальний пул – це тип підвибірки, при якому максимальне значення пікселя набору пікселів, які потрапляють у сприйнятливий поле одиниці в межах шару підвибірки, приймається як вихідний результат.

### Зрівняти шар (keras.layers.Flatten)

Шар Flatten відомий як один із шарів зміни форми, які Keras надає для зміни розмірності вхідних даних. Клас Flatten діє на вхідні дані, зменшуючи розмірність вхідних даних до одиниці. Набори даних зображень є багатовимірними, і для того, щоб вхідні дані передавались через нейронну мережу, розміри вхідних даних потрібно зменшити до одного.

### Щільний шар (keras.layers.Dense)

Щільний шар містить нейрони всередині нейронної мережі. Усі нейрони/одиниці в

щільному шарі отримують вхідні дані від попереднього шару. Операція щільного шару на його вхідних даних є множенням матриці-вектора між вхідними даними, вагами шару, які можна вивчати, і зміщеннями.

Функції активації (`keras.activations.relu` / `keras.activations.softmax`)

Математична операція, яка перетворює результат або сигнали нейронів у нормований вихід. Функція активації — це компонент нейронної мережі, який вносить нелінійність усередину мережі. Включення функції активації дозволяє нейронній мережі мати більшу репрезентаційну потужність і вирішувати складні функції.

#### 4. Тренування

#### 5. Візуалізація результатів

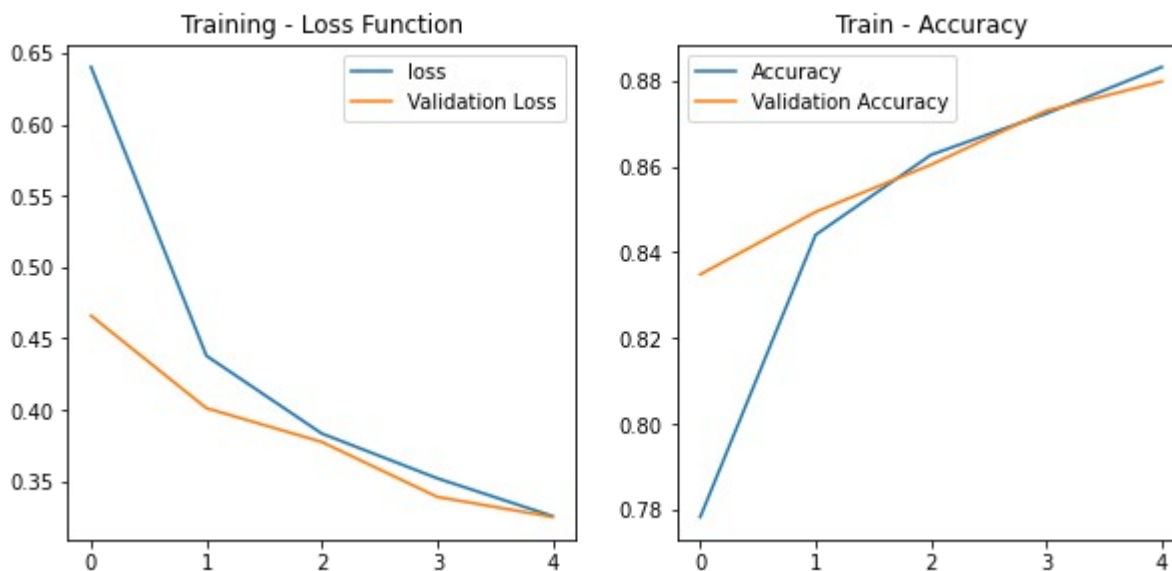


рис.3 Результати тренувань

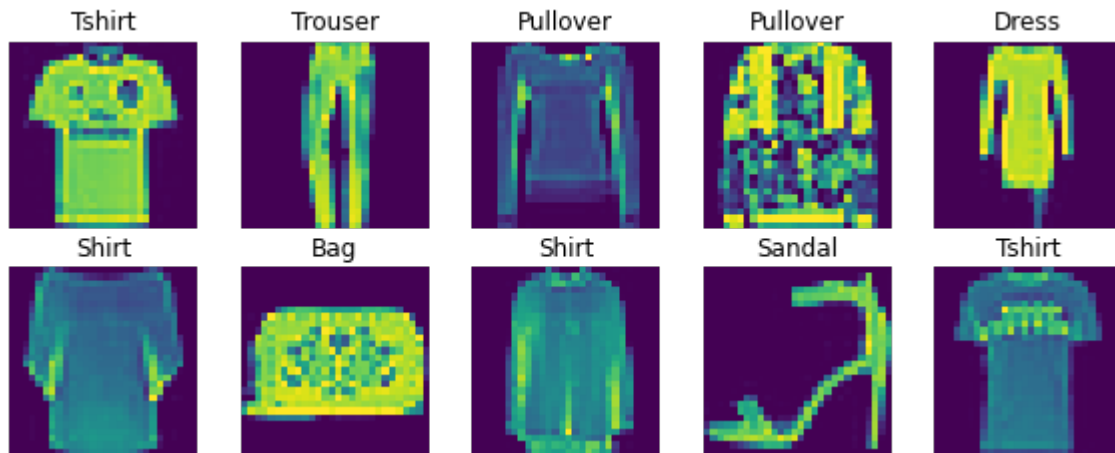


рис. 4 передбаченні тестові картинки

## 6. Будуємо сіамську модель

Кінцева ідея полягає в тому, що при отриманні нового зображення можна розрахувати для нього функціональний вектор за допомогою FeatureGenerationModel. Всі існуючі зображення були попередньо розраховані і збережені в базі даних векторів-фіч. Модель може бути застосована з використанням декількох векторних додавань і множень для визначення найбільш схожих растрів. Ці операції можуть бути реалізовані у вигляді збереженої процедури або всередині самої бази даних.

## 7. Тренуємо

## 8. Переглянемо передбачення моделі

Image A  
Actual: 100%

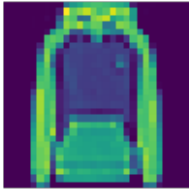


Image A  
Actual: 100%



Image A  
Actual: 100%

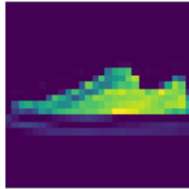


Image A  
Actual: 0%

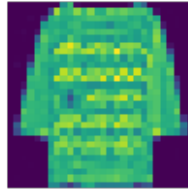


Image A  
Actual: 0%

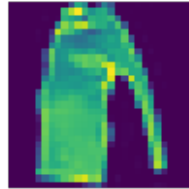


Image A  
Actual: 0%

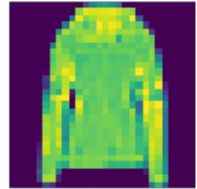


Image B  
Predicted: 68%

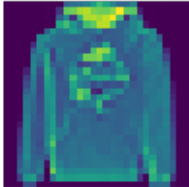


Image B  
Predicted: 68%



Image B  
Predicted: 72%

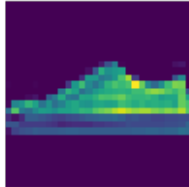


Image B  
Predicted: 68%

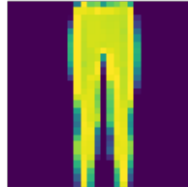
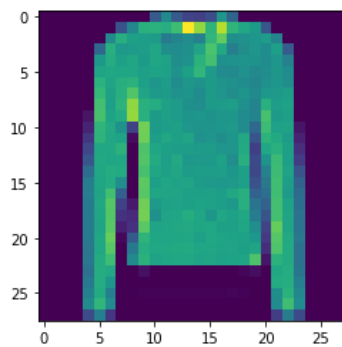
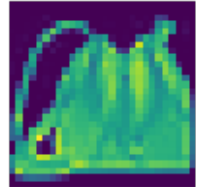


Image B  
Predicted: 68%



Image B  
Predicted: 11%

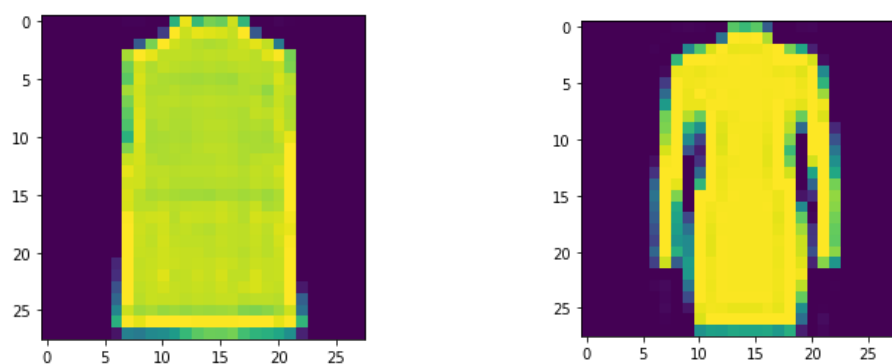


Знайдені схожі:

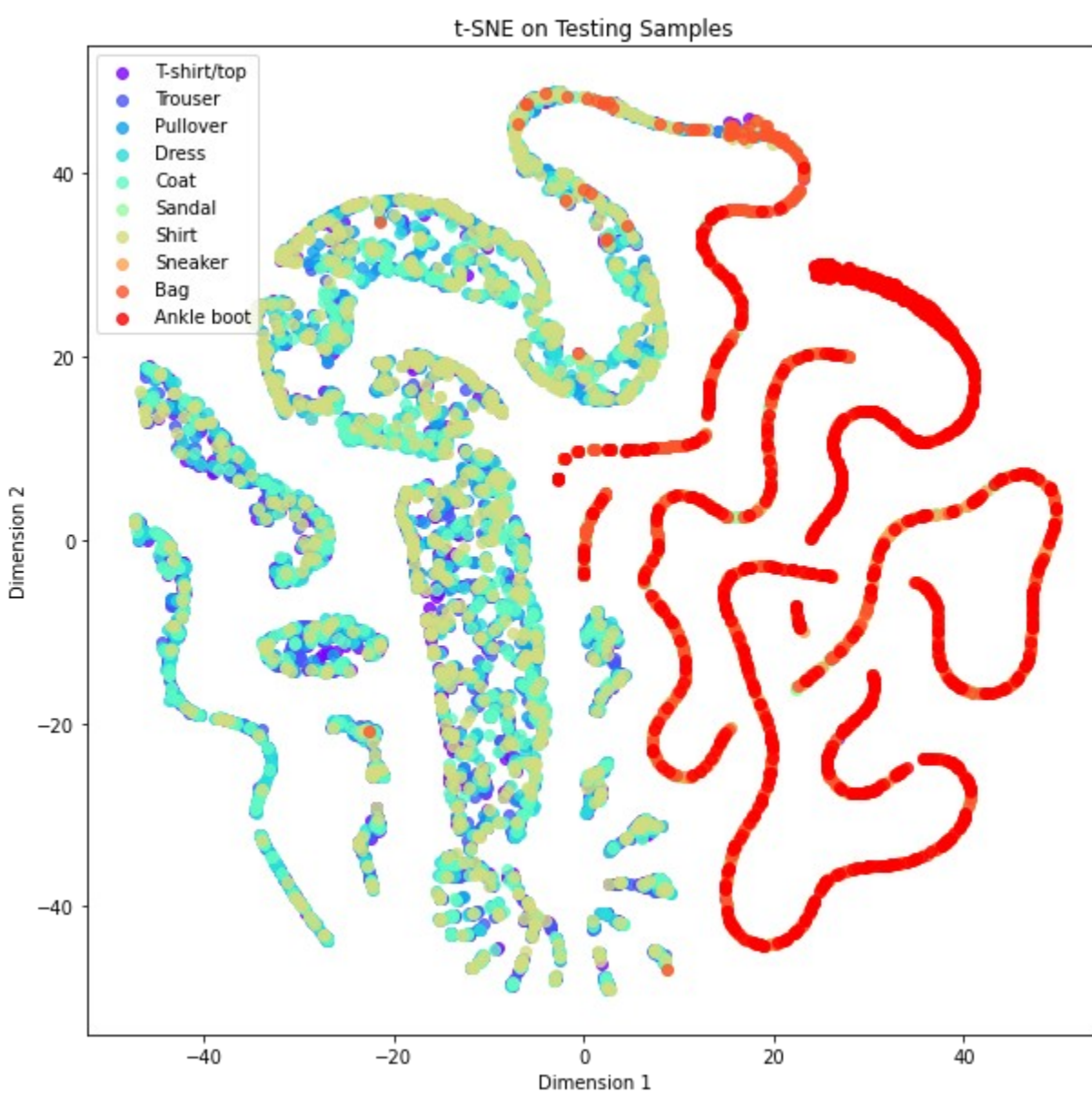
0.67854476

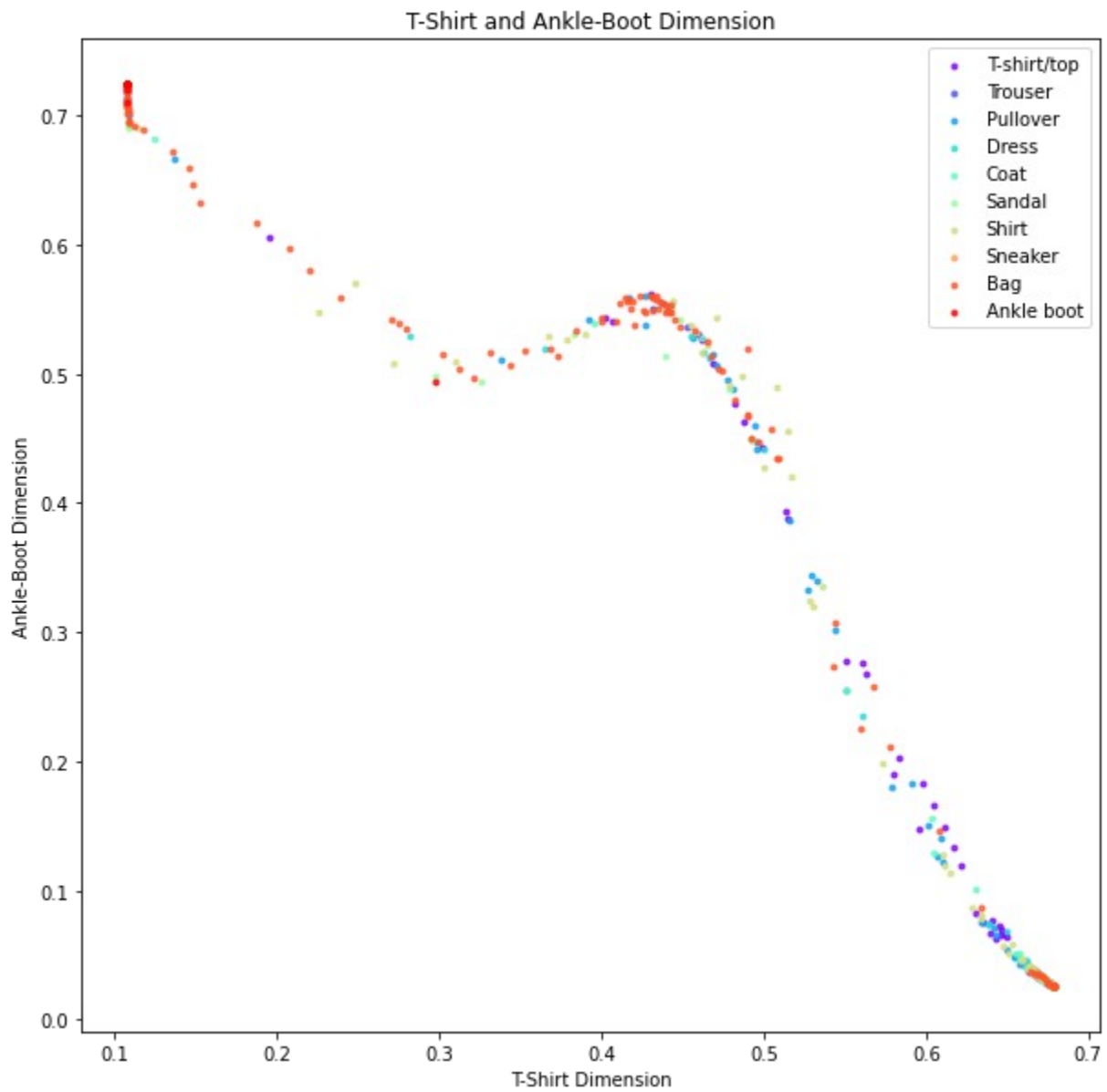
0.67854476





t-SNE





Висновок: набула практичних навиків у розв'язанні задачі пошуку подібних зображень на прикладі організації CNN класифікації.