**Міністерство освіти і науки України**

**Національний технічний університет України «КПІ» імені Ігоря Сікорського**

**Кафедра обчислювальної техніки ФІОТ**

**ЗВІТ**

**з лабораторної роботи №3**

**з навчальної дисципліни «Технології Computer Vision»**

**Тема:**

**ДОСЛІДЖЕННЯ ТЕХНОЛОГІЙ ІДЕНТИФІКАЦІЇ ОБ’ЄКТІВ НА**

**ЦИФРОВИХ ЗОБРАЖЕННЯХ ДЛЯ ЗАДАЧ COMPUTER VISION**

**Виконав:**

Студент 3 курсу кафедри ІПІ ФІОТ,

Навчальної групи ІП-11

Лошак В.І.

**Перевірив:**

Професор кафедри ОТ ФІОТ

Писарчук О.О.

**Київ 2024**

**І. Мета:**

дослідити принципи та особливості підготовки даних, синтезу, навчання та застосування штучних нейронних мереж (Artificial Neural Networks) для практичних задач ідентифікації в технологіях Computer Vision

**ІІ. Завдання:**

Розробити програмний скрипт мовою Python що реалізує обчислювальний алгоритм ідентифікації об’єктів на цифрових зображеннях за технологіями штучних нейронних мереж (Artificial Neural Networks): підготовка даних; конструювання нейромережі; навчання штучної нейронної мережі; застосування нейромережі:

**І рівень складності – максимально 8 балів**.

*Варіант 4*

*Технічні умови реалізації завдання:*

A close up of text

Description automatically generated

**ІІІ. Результати виконання лабораторної роботи.**

**Завдання І.**

1. ***Підготовка даних***

Джерелом даних для розпізнавання обрано набір зображень букв англійської мови написаних від руки: <https://www.kaggle.com/datasets/dhruvildave/english-handwritten-characters-dataset>

Перед початком використання дані було піддано наступним етапам обробки:

* Виокремлення зображень 5ти букв що будуть використовуватись в класифікації
* Інвертування кольорової гами зображення
* Перетворення зображень в одновимірну структуру даних
* Масштабування значень пікселів
* Використання One hot encoding для кодування labels.
* Зменшення роздільної здатності зображення

Кінцевим результатом обробки вводу є дані представлені в наступній формі:

A black and white photo of letters

Description automatically generated

Рис. 1— дані після обробки.

1. ***Тренування мережі***

Для класифікації даних було обрано архітектуру MLP глибокої нейронної мережі. Вибір впав на MLP через порівняну простоту архітектури та ефективність для простих задачі по типу MNIST, EMNIST класифікації.

*Архітектура MLP:*

* Вхідний шар (Input Layer): Приймає вхідні дані.
* Приховані шари (Hidden Layers): Виконують обчислення за допомогою ваг та активаційних функцій.
* Вихідний шар (Output Layer): Видає результат, наприклад, класифікацію чи регресійне значення.

Для правильної роботи MLP який по суті є імлементацією Universal approximation theorem, нам необхідно додати нелінійність до стандартного персептрона.

Для цього використовується функція активації:

**Функція активації (σ)**: Забезпечує нелінійність нейронної мережі. Популярні активаційні функції включають ReLU, сигмоїд та тангенс гіперболічний.

Релу:

Сигмоїд:

Формула для отримання виводу одного нейрону:

Обгрунтування вибору архітектури мережі

* *Кількість нейронів в шарах:* Перший прихований шар має 128 нейронів, а другий - 64. Це дозволяє нейромережі добре адаптуватися до різноманітності форм та стилів написання літер. Нейрони у прихованих шарах забезпечують достатньою потужністю для виявлення важливих особливостей зображень.
* *Функція активації ReLU:* Використовується через її ефективність в нейромережах та здатність зменшувати проблему зникнення градієнту. ReLU пропускає лише додатні значення, що робить нейромережу менш схильною до затухання градієнтів при навчанні.
* *Softmax у вихідному шарі:* Це дозволяє отримати ймовірності приналежності зображення до кожного з класів, що є зручним для класифікації.
* *Розмір пакету (batch):* 64 - це збалансований вибір, який дозволяє ефективно обробляти дані, забезпечуючи при цьому достатньо інформації для оновлення ваг в кожній ітерації.
* *Кількість епох 50:* Це дозволяє моделі проходити через дані достатньо разів для адекватного навчання, що підтверджується графіком зниження втрат.

A graph with a line

Description automatically generated

Рис. 2— Графік помилки по відношення до кількості епох.

На графіку видно, що втрати значно знижуються на початку навчання і стабілізуються ближче до кінця навчання, що свідчить про ефективність обраної архітектури та параметрів навчання. Коливання втрат можуть бути пов'язані з різноманітністю даних у пакетах, що вказує на потребу подальшої оптимізації або використання регуляризації.

В результаті виконання програми було досягнуто точності розпізнавання в 54%. Таку точність можна пояснити низькою кількістю даних для тренування не достатньо. Як видно з графіку, помилка на тренувальних даних сягає нуля, в той час як мережа ще явно недотренована. Всього було використано 290 зображень літер, з них лише 220 безпосередньо для тренування.   
В будь якому випадку ми показали хороший результат враховуючи співвідношення – помилка/розмір даних. Ми довели результативність використання багатошарового персептрону для розпізнавання зображень.

Код лабораторної доступний для перегляду в файлі.

A screenshot of a computer program

Description automatically generated

Рис 3. — Приклад класифікації зображення використовучи нейронну мережу

**IV. Висновки.**

В ході цієї лабораторної роботи я поглиблено вивчив основи структури та функціонування багатошарового персептрону (MLP), який є основою для багатьох сучасних алгоритмів computer vision. Основну увагу було приділено механізму прямого поширення сигналу (forward pass) та зворотнього поширення помилки (backpropagation), які є ключовими для здійснення тренування нейронних мереж. Крім того, я ознайомився з практичним застосуванням навченої моделі для класифікації рукописних літер, що є чудовим прикладом реального застосування глибинних нейронних мереж.

Виконав: студент ФІОТ Лошак В.І. ІП-11