

## РОЗПІЗНАВАННЯ ВИРАЗУ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ ГЛИБОКОГО НАВЧАННЯ

Вінницький національний технічний університет

### Анотація

*Розглядаються технології глибокого навчання, а також способи їх застосування в рамках розпізнавання обличчя.*

**Ключові слова:** обличчя, інтелектуальний аналіз даних, штучний інтелект, теорія розпізнавання образів.

### Abstract

*Technologies of deep learning are considered, as well as ways of their application in the face recognition.*

**Keywords:** face, intelligent data analysis, artificial intelligence, theory of pattern recognition.

### Вступ

В сучасному науковому суспільстві в останні декілька років набули популярності алгоритми глибокого навчання. Нещодавні досягнення у підвищенні їх ефективності в області машинного навчання вказують на можливість створення інформаційної технології розпізнавання виразу обличчя з використанням глибокого навчання та підвищення точності розпізнавання виразу обличчя. Враховуючи новизну області можна стверджувати про можливість значного покращення точності розпізнавання обличчя при використанні глибокого навчання [1]

Метою розробки є процес перетворення вхідної фото, відео інформації в зображення з ім'ям людини.

Об'єктом дослідження постають технології глибокого навчання та аналізу даних.

Предметом дослідження є засоби реалізації технології глибокого навчання.

Головною задачею є розробка програмного продукту з використанням глибокого навчання та аналізу даних.

### Результати дослідження

Глибоке навчання — це галузь машинного навчання, основана на наборі алгоритмів, які намагаються моделювати високорівневі абстракції в даних. Часто використовують моделі нейронних мереж з великою кількістю прихованих шарів. Шари в такій моделі відповідають різним рівням концептуалізації, де концепції (ознаки, функції, представлення) більш високого рівня визначаються з концепцій більш низького рівня. Таким чином, одні й ті самі концепції низького рівня допомагають визначити багато високорівневих концепцій [2]. Глибоке навчання базується на розподілених інтерпретаціях, припускаючи, що спостережувані дані були створені комбінацією багатьох факторів, не всі з яких відомі спостерігачу, тому інформація про певний фактор, отримана з деяких конфігурацій інших факторів, часто може узагальнюватись до інших, невідомих конфігурацій факторів. Глибоке навчання додає до розподілених інтерпретацій припущення, що ці фактори мають багаторівневу організацію у відповідності до різних рівнів абстракції або складу об'єкту і високорівневі представлення отримуються шляхом перетворення та комбінування низькорівневих.

До моделей глибокого навчання можна віднести згорткові нейронні мережі, глибокі мережі довіри, звичайні багатошарові нейромережі прямого розповсюдження та ін. Приклад схеми багатошарової нейронної мережі прямого розповсюдження наведено на рисунку 1.

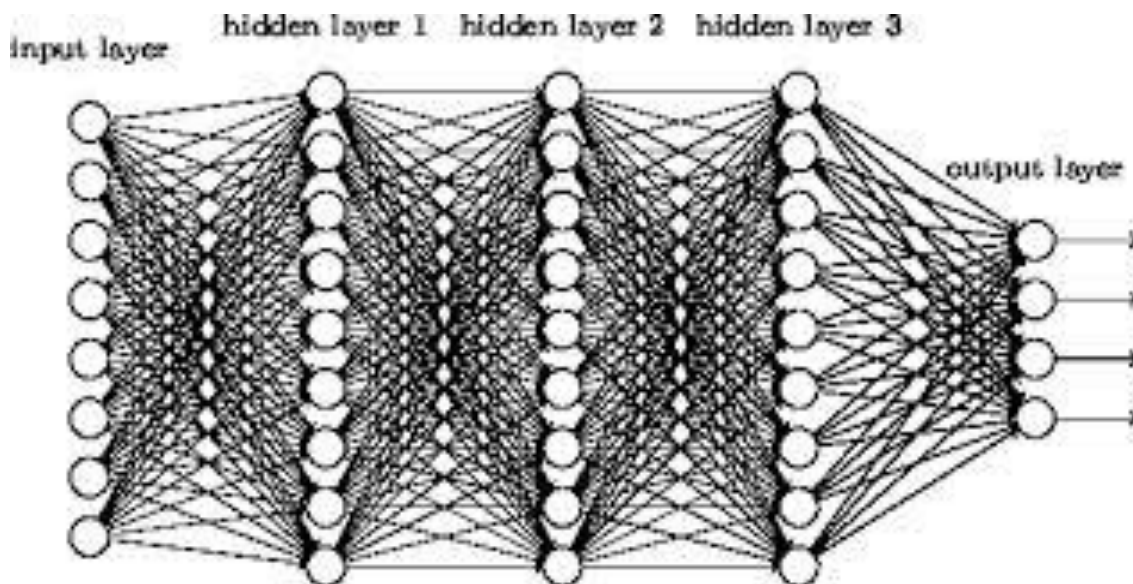


Рис. 1. Приклад схеми багат шарової нейронної мережі прямого розповсюдження

В 2008 році, Бітом Фейзелем була запропонована архітектура згорточної нейронної мережі що складається з 5-ти шарів, 2 з яких – шари підвибірки і 2 – згортчні шари [3]. Така архітектура дозволила значно знизити навантаження на модуль попередньої обробки і не використовувати особливих алгоритмів для попередньої обробки. Дана система досить ефективно розпізнає вирази обличчя людини і проводить розпізнавання з досить великою швидкістю. Тим не менше для підвищення швидкодії було прийнято рішення зменшити кількість нейронів і зв'язків між ними, зберігаючи архітектуру запропоновану Фейзелем. Шари згортки виконують двовимірну згортку зі зміщенням областей зображення, що перекриваються, в той час як шари підвибірки виділяють максимальне значення в областях зображення, що не перекриваються, і передають його до наступного шару згортки. Навчання ваг шарів згортки дозволяє виділення ознак, що залежать від наявних вхідних даних, тоді як шари підвибірки збільшують інваріантність об'єктів, що залежать від місця, що нас цікавить, та зменшують розмірність задачі. Розділення ваг (використання одного й того самого набору ваг для кожної оброблюваної області) дозволяє значно зменшити кількість незалежних параметрів, що в свою чергу покращує здатність до узагальнення.

Б.Фейзель дослідив декілька архітектур згорточної нейромережі і виділив вищевказану, за допомогою якої було досягнуто точності розпізнавання у 82.9% на базі даних JAFFE (Japanese Female Facial Expression Database). Основним недоліком такої архітектури є об'ємність обчислень, необхідних при розпізнаванні: хоча кількість параметрів мережі є малою, на кожному кроці згортки необхідно обробляти велику кількість областей зображення, що перекриваються. Слід зазначити, що цей недолік може бути усунений з використанням паралельних обчислень, оскільки кожна така область може бути обробленою одночасно з іншими з використанням однакового набору параметрів.

У 2010 році Томом Мак-Лафінім та іншими було використано підхід, що базується на глибоких нейромережах прямого розповсюдження із попереднім тренуванням за допомогою обмежених машин Больцмана [4] — типу нейронних мереж, який отримав значне розповсюдження останнім часом у зв'язку із порівняною легкістю використання, відсутністю потреби у великих об'ємах помічених даних (порівняно з іншими типами нейромереж, такими як згортчні) та нещодавно розробленими ефективними алгоритмами тренування.

## Висновки

Приведено актуальність дослідження розпізнавання обличчя. Проаналізовано існуючі досягнення в розпізнаванні виразу обличчя за допомогою нейронних мереж. Проаналізовано деякі вдалі підходи до вирішення задачі розпізнавання обличчя за допомогою глибокого навчання. Здійснений аналіз свідчить, що глибоке навчання є перспективним підходом для вирішення поставленої задачі через свою відносну новизну, можливість великої кількості варіацій, а також через підвищення швидкодії комп'ютерів згідно закону Мура, розповсюдження обчислень на нових, потужних графічних модулях, які дозволяють проводити обчислення з високим ступенем паралелізації [5], більшій кількості даних для тренування та тестування розпізнавання виразу обличчя людини

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. P. Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage. New York: W. W. Norton & Company, 2009. 416 p.
2. Y. Bengio, Learning Deep Architectures for AI. Montreal: Universit? de Montr?al, 2009. 130 p
3. B. Fasel, Multiscale Facial Expression Recognition using Convolutional Neural Networks, IDIAP 2009. 2009. No. 13. P. 121–137
4. T. McLaughlin, M. Le, N. Bayanbat, Emotion Recognition with Deep-Belief Networks, Stanford CS 229 Machine Learning Final Projects. 2010. No. 14. P. 62–66
5. Яровий А. А. Аналіз обчислювальної складності GPU-орієнтованих паралельно-ієрархічних обчислювальних систем та оцінювання продуктивності їх апаратного забезпечення / А. А. Яровий, Н. І. Кокряцька, С. В. Наконечна, М. С. Матейчук, Т. Д. Польгуль // Оптико-електронні інформаційно-енергетичні технології. – 2014. – № 1. – С. 18–25

**Стеблюк Дмитро Євгенійович**, студент групи ІПі-14б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: steblyukdm@gmail.com

Науковий керівник: **Коваленко Олена Олексіївна**, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, e-mail: kovalenka88@gmail.com

**Dmytro Steblyuk**, student of group ІPI-14b, Faculty for Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: steblyukdm@gmail.com

Supervisor: **Olena Kovalenko**, Associate Professor of Software Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: kovalenka88@gmail.com