

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ
СІКОРСЬКОГО»

ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ

КАФЕДРА СИСТЕМНОГО ПРОГРАМУВАННЯ І
СПЕЦІАЛІЗОВАНИХ КОМП'ЮТЕРНИХ СИСТЕМ

«На правах рукопису»
УДК 004.855

«До захисту допущено»
Завідувач кафедри СПСКС

_____ В.П.Тарасенко
(підпис) (ініціали, прізвище)
« ____ » _____ 2018р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності 123 Комп'ютерна інженерія

Системне програмування

на тему: **“Нейромережеві моделі та засоби біометричної ідентифікації користувачів компютерних мереж”**

Виконав: студент II курсу, групи КВ-з71мп
(шифр групи)

Войтко Максим Олегович

(прізвище, ім'я, по батькові)

(підпис)

Науковий керівник: доктор технічних наук, професор

Терейковський Ігор Анатолійович

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Рецензент

(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних посилань.

Студент _____
(підпис)

Київ – 2018 року

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ
СІКОРСЬКОГО»**

**ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ
КАФЕДРА СИСТЕМОГО ПРОГРАМУВАННЯ І
СПЕЦІАЛІЗОВАНИХ КОМП'ЮТЕРНИХ СИСТЕМ**

Рівень вищої освіти – другий (магістерський)

Спеціальність 123 Комп'ютерна інженерія

Комп'ютерні системи та компоненти

«Затверджую»

Завідувач кафедри СПСКС

_____ В.П.Тарасенко
(підпис) (ініціали, прізвище)

“ ___ ” _____ 2018р.

ЗАВДАННЯ

на магістерську дисертацію студента

Войтка Максима Олеговича

1. Тема дисертації: Нейромережеві моделі та методи біометричної ідентифікації користувачів комп'ютерних мереж,

науковий керівник дисертації: *д.т.н., проф. Терейковський І.А.,*

затверджені наказом по університету від «09» листопада 2018 р. № 4138-с.

2. Термін подання студентом дисертації: 7 грудня 2018 р.

3. Об'єкт дослідження: процес керування доступом до комп'ютерних мереж на основі біометричної ідентифікації користувачів за допомогою згорткових нейронних мереж.

4. Предмет дослідження: нейромережеві моделі та методи біометричної ідентифікації.

5. Перелік завдань, які потрібно розробити:

- Проаналізувати існуючі алгоритми, методи та моделі біометричної ідентифікації.
- Розробити нейромережеві моделі призначені для біометричної ідентифікації користувачів комп'ютерних мереж.
- Розробити методи нейромережевої ідентифікації по знімку вушної раковини.
- Провести експериментальні дослідження для верифікації результатів отриманих рішень.

6. Перелік ілюстративного матеріалу:

- Структурна схема архітектури програмної системи.

- Презентація.

7. Перелік публікацій: V Міжнародній науково-технічній Internet конференції в НУХТ на базі факультету АКС «Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами» (м. Київ, 22-23 листопада 2018р.), Міжнародна наукова інтернет-конференція "Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення"(випуск 34) (м.Тернопіль, 11 грудня 2018р.)

8. Дата видачі завдання 5 вересня 2018 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Ознайомлення з предметною галуззю	17.12.2017	
2	Визначення структури магістерської дисертації; вивчення літератури, пошук додаткової літератури, патентний пошук	05.02.2018	
3	Робота над першим розділом магістерської дисертації; проведення наукового дослідження	15.03.2018	
4	Проведення наукового дослідження; робота над другим розділом магістерської дисертації; розроблення програмного забезпечення	27.06.2018	
5	Проведення наукового дослідження; робота над статтею за результатами наукового дослідження	12.09.2018	
6	Проведення наукового дослідження; робота над третім розділом магістерської дисертації	03.10.2018	
7	Завершення роботи над основною частиною магістерської дисертації; підготовка ілюстративного матеріалу	16.10.2018	
8	Оформлення текстової і графічної частини магістерської дисертації	10.11.2018	
9	Попередній розгляд магістерської дисертації на кафедрі	26.11.2018	

Студент _____

Войтко М.О.

Науковий керівник дисертації _____

Терейковський І.А.

РЕФЕРАТ

Актуальність теми. Одним з основних способів керування та розмежування доступу до комп'ютерних мереж є аутентифікація. Проведені дослідження вказують на те, що класичні засоби аутентифікації з використанням статичних паролів не відповідають викликам сучасності та не забезпечують достатнього рівня захисту.

Сучасні системи біологічного розпізнавання є якісними та надійними засобами ідентифікації особистості. Біометричні інформаційні системи також володіють унікальними перевагами – вони не дозволяють відмовитись від досконалої транзакції і дають можливість визначити, коли індивідуум користується декількома акаунтами.

Проаналізувавши сучасні системи контролю доступу користувачів до комп'ютерних мереж можна зробити висновок, що відбувається активний розвиток біометричних методів завдяки їх зручності, надійності та достовірності.

Тобто на сьогодні актуальним є підвищення ефективності нейромережових моделей та засобів біометричної ідентифікації користувачів комп'ютерних мереж.

Об'єктом дослідження є процес керування доступом до комп'ютерних мереж на основі біометричної ідентифікації користувачів за допомогою згорткових нейронних мереж.

Предмет дослідження – нейромережові моделі та методи біометричної ідентифікації.

Мета роботи – підвищення ефективності нейромережових моделей для біометричної аутентифікації користувачів комп'ютерних мереж.

Завдання роботи:

1. Проаналізувати існуючі алгоритми, методи та моделі біометричної ідентифікації.
2. Розробити нейромережеві моделі призначені для біометричної ідентифікації користувачів комп'ютерних мереж.
3. Розробити методи нейромережевої ідентифікації по знімку вушної раковини.
4. Провести експериментальні дослідження для верифікації результатів отриманих рішень.

Методи дослідження:

- аналіз;
- аналогія;
- емпіричні: спостереження за похибкою при ідентифікації персоналу та експериментальна перевірка ефективності розробленого програмного забезпечення;
- індукція;
- дедукція;
- синтез.

Наукова новизна. Отримали подальший розвиток нейромережевої моделі біометричної ідентифікації користувачів комп'ютерних мереж на основі зображення вушної раковини.

Практична цінність. Результати даної роботи можуть бути використані для ідентифікації користувачів при розмежуванні та управлінні доступом до будь-якої комп'ютерної мережі, попередньо отримавши фотознімки вушних раковин.

Апробація роботи. Основні положення магістерської роботи доповідалися на V Міжнародній науково-технічній Internet конференції в

НУХТ на базі факультету АКС “Сучасні методи, інформаційне, програмне та технічне забезпечення систем керування організаційно-технічними та технологічними комплексами” (м. Київ, 22-23 листопада 2018р.), Міжнародна наукова інтернет-конференція "Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення" (випуск 34) (м.Тернопіль, 11 грудня 2018р.).

Структура та обсяг роботи. Магістерська дисертація складається із вступу, 4 розділів, висновків, списку використаних джерел та додатків.

У *вступі* подано загальну характеристику роботи, проаналізовано сучасне становище проблеми, визначено актуальність напрямку досліджень, поставлено мету і задачі досліджень, представлено наукову новизну отриманих результатів і практичну цінність роботи.

У *першому розділі* розглянуто та проаналізовано існуючі методи біометричної ідентифікації користувачів.

У *другому розділі* розроблено нейромережеві моделі призначені для ідентифікації по знімку вушної раковини та проаналізовано їх обґрунтування та особливості застосування.

У *третьому розділі* розроблено методи нейромережевої ідентифікації по знімку вушної раковини та проаналізовано їх обґрунтування.

У *четвертому розділі* проведені експериментальні дослідження спрямовані на верифікацію отриманих рішень.

У *висновках* проаналізовано отримані результати дослідження.

У *додатках* наведено уривок програми, презентацію, довідку про впровадження результатів дослідження та копії публікацій.

Магістерська дисертація представлена на 79 аркушах, містить чотири додатки та список використаних джерел із 42 найменувань. У роботі наведено 24 рисунки і 9 таблиць.

Ключові слова: аутентифікація, біометрична ідентифікація, нейронні мережі, згорткові нейронні мережі, машинне навчання, комп'ютерне бачення.

ABSTRACT

Theme urgency. One of the primary means to manage access to computer networks is authentication. Provided research indicates that classic authentication means with static passwords are not relevant anymore and didn't provide enough safety level.

Modern biological identification systems are high quality and reliable way of personal identification. Biological informational systems have unique advantages – they do not allow denying perfect transaction and allow multiple accounts usage detection.

Modern access control system analysis proves apparent march toward biometrical means, because of their convenience, reliability, and accuracy.

Thus increasing the efficiency of computer networks users biometrical identification neural networks models and means is exceptionally relevant.

The object of research is computer network access control process based on biometrical identification with convolutional neural networks.

The subject of research – neural network models and methods of biometrical identification.

The research objective is to improve computer networks users biometrical identification neural network model efficiency.

The research tasks:

1. Analyze existing biometrical identification algorithms, models and methods.
2. Develop neural network models for computer networks users biometrical identification.
3. Develop methods for ear recognition neural network biometrical identification method.

4. Conduct experiments to verify the developed solution.

The research methods:

- analysis;
- analogy;
- deduction;
- empirical: experimental verification of the proposed solution;
- induction;
- synthesis.

Results and scientific novelty. Further development of computer networks users ear recognition biometrical identification neural network model was reached.

Practical value. The thesis results might be used for computer network users identification, given user ears images.

Approbation. The main provisions and results of the work were presented and discussed at V International Scientific and Technical Conference “Modern methods, information, software and technical support of control systems for organizational, and technological complexes” (Kyiv, November 22-23, 2018), as well as International Scientific Internet Conference “Informational society: technological, economic and technical aspects” (Ternopil, December 11, 2018).

Thesis structure and content. The master thesis consists of the introduction, four sections, conclusions, references, and annexes.

The introduction describes the research, assess the scope of biometrical identification with neural network models problem, underlines the research relevance, finds the research purpose and objectives, presents the scientific novelty of the obtained results and the practical value of the work.

In *the first chapter* existing biometrical identification methods are described and analyzed.

In *the second chapter* ear recognition biometrical identification neural network models are developed, and their scientific basis and implementation details are analyzed.

The third chapter explores neural networks methods for ear recognition biometrical identification and their validation.

The fourth chapter describes the conducted experiments for solution verification.

The conclusions analyze the results of the conducted research.

In *the appendixes* include the following items: a presentation, program code abstract, the publications copy, and the implementation certificate.

The thesis is presented in 79 pages, and it contains 4 annexes and 42 references to the used sources. 24 figures and 9 tables are presented in the thesis.

Keywords: authentication, biometrical identification, neural networks, CNN, ConvNets, Machine Learning, Computer Vision.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	7
ВСТУП.....	9
1. АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ.....	11
1.1 Загальні відомості про біометричну ідентифікацію по знімку вушної раковини.....	11
1.2 Згорткові нейромережеві моделі розпізнавання знімків вушної раковини.....	18
2. НЕЙРОМЕРЕЖЕВІ МОДЕЛІ ДЛЯ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ ПО ЗНІМКУ ВУШНОЇ РАКОВИНИ.....	31
2.1 Архітектура та особливості застосування нейромережевої моделі VGGNet-19.....	32
2.2 Архітектура та особливості застосування нейромережевої моделі ResNet.....	39
3. МЕТОДИ НЕЙРОМЕРЕЖЕВОЇ ІДЕНТИФІКАЦІЇ ПО ЗНІМКУ ВУШНОЇ РАКОВИНИ ТА ЇХ ОБҐРУНТУВАННЯ.....	50
4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ВИКОРИСТАНОЇ НЕЙРОННОЇ МЕРЕЖІ.....	55
ВИСНОВКИ.....	60

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....	62
ДОДАТОК А. Текст програми.....	66
ДОДАТОК Б. Презентація.....	90
ДОДАТОК В. Довідка про впровадження.....	100
ДОДАТОК Г. Тези доповідей на конференціях	101

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

AI – Artificial Intelligence

BN – Batch Normalization

BSIF – Binarized Statistical Image Features

CNN – (ConvNet) – Convolutional Neural Network

CPU – Central processing unit

cuDNN – The NVIDIA CUDA[®] Deep Neural Network library

DNN – Deep Neural Network

FAR – False Acceptance Rate

FC – Fully connected layer

FEARID – Forensic Ear Identification Group

FRR – False Rejection Rate

GD – Gradient Descent

IAPR – International Association for Pattern Recognition

IEEE – Institute of Electrical and Electronics Engineers

ILSVRC – ImageNet Large Scale Visual Recognition Challenge

LBP – Local binary patterns

LPQ – Local Phase Quantization

LRN – Local Response Normalization

LSTM – Long Short Term Memory

ML – Machine Learning

NN – Neural Network

PCA – Principal Component Analysis – статистичний метод факторного аналізу, при якому відбувається ортогональне перетворення множини

спостережень з можливо пов'язаними змінними (сутностями з різними числовими значеннями) у множину головних компонент (змінних без лінійної кореляції).

RAM – Random-Access Memory

RC – Release Candidate

ReLU – Rectified Linear Units

ResNet – Residual Network

RGB – RGB color model – Red, Green, Blue (червоний, зелений, синій) – адитивна колірна модель для опису синтезу червоного, зеленого та синього кольорів.

SIFT – Scale-invariant feature transform

SGD – Stochastic Gradient Descent

SURF – Speeded Up Robust Features

UERC – Unconstrained Ear Recognition Challenge – перше змагання по розпізнаванню зображень вушних раковин без обмежень проведене на Міжнародній спільній конференції по біометриці 2017 IEEE/IAPR.

USTB – University of Science and Technology Beijing

VGG – Visual Geometry Group of Oxford University

ГЗНМ – Глибока згорткова нейронна мережа

ДКЧП – Довга короткочасна пам'ять

ЗНМ – Згорткова нейронна мережа

ЗНММ – Згорткова нейромережева модель

ЗШ – Зважені шари

ЗШН – Згортковий шар нейронів

КМ – Компютерна мережа

КС – Компютерна система

МН – Машинне навчання

НМ – Нейронна мережа

ОС – Операційна система

ПЗ – Програмне забезпечення

ПЗШН – Повністю з'єднаний шар нейронів

ПШН – Прихований шар нейронів

ШІ – Штучний інтелект

ШНМ – Штучна нейронна мережа

ВСТУП

Одним з основних способів керування та розмежування доступу до комп'ютерних мереж є аутентифікація. Проведені дослідження [36-38] вказують на те, що класичні засоби аутентифікації з використанням статичних паролів не відповідають викликам сучасності та не забезпечують достатнього рівня захисту.

Для забезпечення захисту інформації та контролю управління доступу до її ресурсів можуть бути використані біометричні технології, які дозволяють однозначно визначити суб'єкт доступу та його повноваження по відношенню до конкретного ресурсу. На відміну від традиційних методів аутентифікації (паролі, картки, різноманітні електронні ключі), біометричні ознаки людині дуже важко підробити та неможливо втратити, вкрати або передати в користування іншій особі. Біометричні зразки «знімаються» з об'єкта датчиками і надсилаються процесору, який витягує відмінні риси, відкидаючи всі інші компоненти. Оброблений біометричний зразок зберігається в базі даних як «шаблон», або порівнюється з певним шаблоном для визначення відповідності. Сучасні системи біологічного розпізнавання є якісними та надійними засобами ідентифікації особистості. Біометричні інформаційні системи також володіють унікальними перевагами – вони не дозволяють відмовитись від досконалої транзакції і дають можливість визначити, коли індивідуум користується декількома акаунтами.

Проаналізувавши сучасні системи контролю доступу користувачів до комп'ютерних мереж можна зробити висновок, що відбувається активний розвиток біометричних методів завдяки їх зручності, надійності та достовірності.

Тобто на сьогодні актуальним є підвищення ефективності нейромережових моделей та засобів біометричної аутентифікації користувачів комп'ютерних мереж.

Мета дослідження це підвищення ефективності систем біометричної ідентифікації користувачів комп'ютерних мереж.

Завдання дослідження:

1. Проаналізувати існуючі алгоритми, методи та моделі біометричної ідентифікації.
2. Розробити нейромережові моделі призначені для біометричної ідентифікації користувачів комп'ютерних мереж.
3. Розробити методи нейромережової ідентифікації по знімку вушної раковини.
4. Провести експериментальні дослідження для верифікації результатів отриманих рішень.

В сучасній науковій та технічній літературі значна увага приділяється проблемі дослідження, розробки та використання засобів біометричного контролю доступу. Серед робіт по даній тематиці слід відзначити роботи таких вчених: В.М. Колешко, Г.Є. Ляшенко, А.А. Астраханцев, С.Г. Удовенко, О.М. Яшина, А.К. Jain, А. Ross, S. Prabhakar, G. Kaur, Ch. K. Verma, J. Fierrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero, Hui Chen, B. Bhanu, S. Malki and L. Spaanenburg.

1. АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ

1.1 Загальні відомості про біометричну ідентифікацію по знімку вушної раковини

Проблематиці розпізнавання зображень вушних раковин було приділено значно менше уваги, ніж безлічі альтернативних біометричних засобів, включаючи розпізнавання знімків обличчя, відбитків пальців та розпізнавання зображень райдужної оболонки ока. Особливе зацікавлення викликає розпізнавання зображень вушних раковин, тому що це неінвазивний спосіб і тому, що на нього не впливають фактори навколишнього середовища, такі як настрій, здоров'я та одяг.

Крім того, на зовнішність вушної раковини (зовнішнього вуха) відносно не впливає старіння, що робить розпізнавання за знімками вушної раковини можливим для довгострокового використання для ідентифікації.

Зображення вушних раковин можна легко отримати з відстані без відома відповідної особи. Тому ідентифікації на основі розпізнавання зображень вушних раковин підходить для додатків розвідки, підтримання безпеки, контролю доступу та моніторингу. Відбитки вушних раковин, знайдені на місці злочину, були використані як доказ у більш ніж кількох сотнях випадків в Нідерландах та США.

Дослідження, виконані Яннареллі, дають достатньо доказів, що вушні раковини є унікальними біометричними рисами.

У першому дослідженні було зроблене порівняння понад 10 000 зображень вушних раковин, випадковим чином вибраних із набору даних

зібраних в Каліфорнії, і у другому дослідженні було перевірено зображення вух двояйцевих та однояйцевих близнюків, в яких фізіологічні особливості як відомо, подібні. Результати цих досліджень підтверджують гіпотезу, що вушна раковина має унікальні фізіологічні особливості, тому що в обох дослідженнях всі розглянуті вушні раковини були визнані унікальними, хоча були однояйцеві близнюки мають схожі, але не ідентичні, вушні структури, особливо у зоні раковини та долини.

У медичній літературі вказується, що збільшення вуха після перших чотири місяців життя людини пропорційна. Виявляється, що навіть при тому, що зростання вуха пропорційне, гравітація може спричинити розтягнення вуха у вертикальному напрямку. Ефект від цього розтягнення найбільш виражений в зоні вушної раковини, і виміри показують, що зміна нелінійна. Швидкість розтягування приблизно в п'ять разів перевищує норму в період від чотирьох місяців до восьми років, після чого швидкість розтягнення залишається сталою до 70 років, і тоді знову зростає.

Основним недоліком біометрії вушних раковин є те, її неможливо застосовувати, коли вухо суб'єкта прикрите. У випадку активних систем ідентифікації, це не є недоліком, оскільки суб'єкт може відтягнути волосся назад і продовжити процес аутентифікації. Проблема виникає під час пасивної ідентифікації, оскільки в даному випадку допомога з боку суб'єкта не передбачається. У випадку коли вухо лише частково закрито волоссям, можливо розпізнати волосся і виділити його від зображення [27].

Нейронні мережі (НМ) показали свою ефективність у вирішенні задач розпізнавання зображень, таких як: сегментація зображень (image segmentation) [1], визначення та розпізнавання об'єктів (object detection and recognition) [21], збільшення зображення (image superresolution) [7] та інші. Один з ключових факторів, що зробили це можливим, є наявність значного корпусу даних для тренування, що дозволяє спроектувати і натренувати

глибокі нейромережеві моделі (deep CNNs). Зазначимо, що у загальному випадку математичне забезпечення системи нейромережевого розпізнавання зображень вушної раковини у першому наближенні можна описати за допомогою виразів:

$$W^{[l]} \in R^{n_h^{l+1} \times n_h^{l-1}}$$

$$b^{[l]} \in R^{l+1}$$

$$x^{(i)} \in R^{n_x}$$

$$y^{(\hat{i})} \in R^{n_y}$$

$$a = g^{[l]}(W_x x^{(i)} + b_1) = g^{[l]}(z_1)$$

$$y^{(\hat{i})} = \text{softmax}(W_h h + b_2)$$

де i – номер тренувального екземпляру;

l – номер шар НМ;

$g^{[l]}$ – функція активації l -го шару НМ;

$W^{[l]}$ – матриця вагів l -го шару;

n_h^l – кількість нейронів у наступному шарі;

n_h^{l-1} – кількість нейронів у попередньому шарі;

$b^{[l]}$ – вектор похибки (bias vector) l -го шару;

$y^{(\hat{i})}$ – прогнозований вихідний вектор i -го екземпляру;

n_x – розмір вхідного вектору;

n_y – розмір вихідного вектору (кількість вихідних класів);

$x^{(i)}$ – i -тий екземпляр представлений у вигляді вектора.

Загальна формула визначення активації нейронів:

$$a_j^{[l]} = g^{[l]}(\sum w_{jk}^{[l]} a_k^{[l-1]} + b_j^{[l]}) = g^{[l]}(z_j^{[l]})$$

Для проблем, де вдосталь даних для тренування таких нейронних мереж, є найкращим рішенням. Використання нейронних мереж (НМ) в основному вимагає значної кількості даних для тренування моделей. Як правило кількість зразків в наборах даних для розпізнавання зображень вушних раковин доволі обмежена [11], [18]. В той час як в інших випадках, де недостатньо даних вплив глибокого навчання обмежений. Одним з таких випадків є автоматичне розпізнавання на основі зображень вушних раковин, яке має значний потенціал у сфері безпеки, але при цьому недостатньо даних для повноцінного використання.

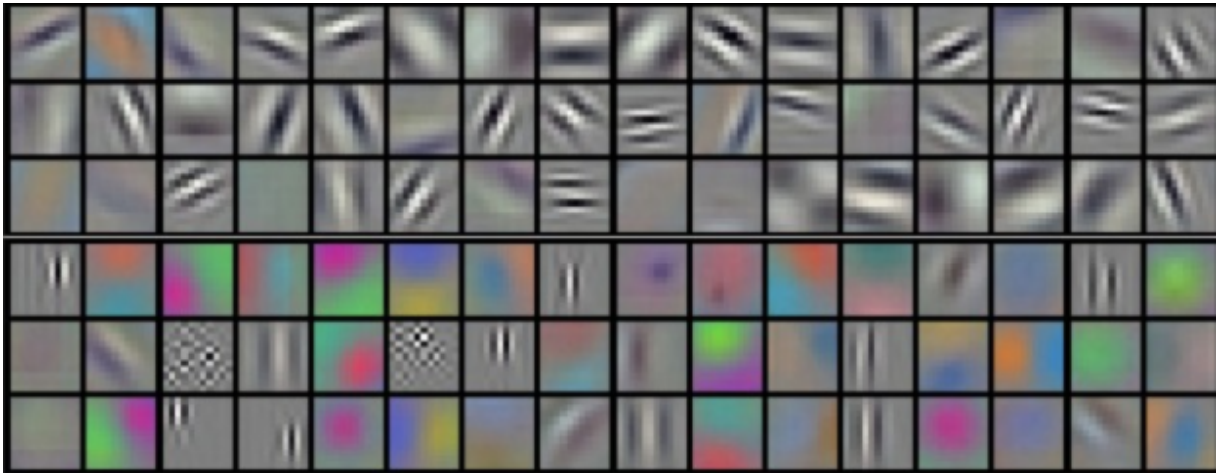


Рисунок 1.1 – Приклади фільтрів отриманих Крижевським та іншими у своєму дослідженні

Через це обмеження при використанні згорткових мереж для даної задачі в основному використовують уже попередньо треновану модель класифікації об'єктів, так звану модель претренованої глибокої згорткової нейронної мережі (pretrained deep CNN) з однією із добре відомих високопродуктивних архітектур згорткових нейронних мереж [14], [29]. Ці моделі доцільно попередньо натренувати на наборі даних ImageNet [5] для загальних цілей класифікації об'єктів.

Згорткові нейронні мережі (CNN) показали свою ефективність у сфері розпізнавання зображень. В попередніх дослідженнях розпізнавання зображень вушних раковин найбільш популярними методами визначення вхідних характеристик були SIFT [11], SURF [15] та LBP [26].

Згідно з Емершічем [7] існуючі техніки ідентифікації на основі розпізнавання зображень вушних раковин можна поділити на такі види (Рис.1):

1. Геометричні.
2. Холістичні (цілісні).
3. Локальні.
4. Гібридні.

Геометричні техніки описують геометричні властивості вушних раковин та витягують геометричні дані, що можуть бути використані для розпізнавання. Оскільки використовуються лише геометричні дані, очевидно, що можна розробити методи, що не будуть залежати від геометричних викривлень, таких як: поворот, масштабування та навіть невеликі зміни перспективи [7], [18].

Холістичні техніки описують загальний вигляд вушної раковини та використовують закодоване представлення вуха як цілого. Так як вигляд вушної раковини може відрізнитися залежно від пози та освітлення, необхідно підготувати та нормалізувати зображення перед витяганням характеристик [29].

Локальні техніки витягування характеристик з просторово обмежених зон зображення без вивчення глобальної інформації про загальну структуру вушної раковини. Отримані характеристики не обов'язково співставні з анатомічними частинами вуха, але можуть в загальному представити будь-яку точку зображення. Можна виділити два типи локальних технік: техніки, що спочатку визначають розташування

ключових точок зображення і тоді обчислюються дескриптори кожної з визначених ключових точок [16] та техніки, коли локальні дескриптори обчислюються густо для всього зображення без прив'язки до структурних характеристик зображення [7].

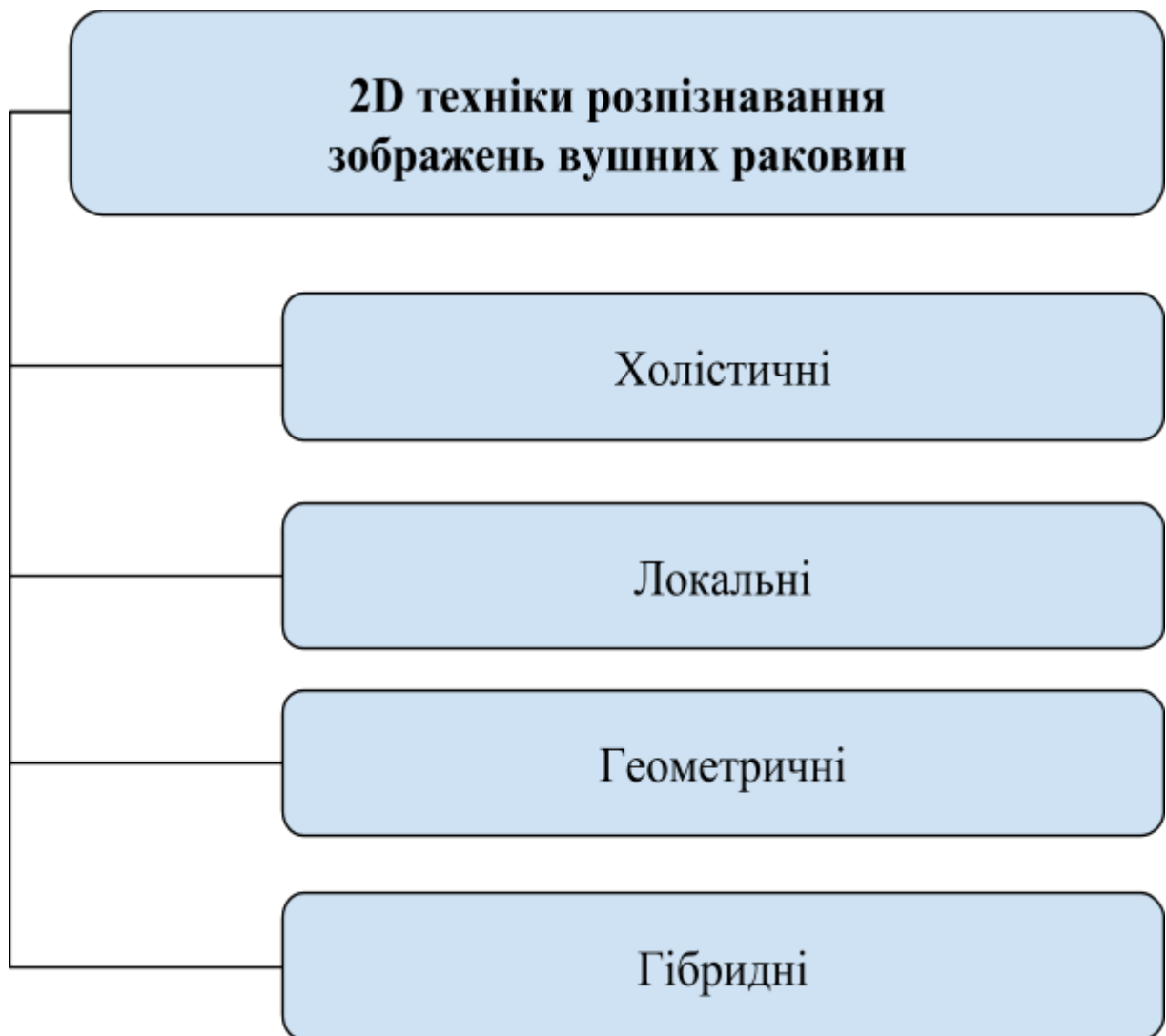


Рисунок 1.2 – Класифікація технік розпізнавання зображень вушних раковин

Гібридні техніки, які поєднують елементи різних категорій або використовують різні представлення для збільшення точності розпізнавання зображень вушних раковин. Техніки цієї групи показують найкращий результат порівняно з іншими групами, але часто за рахунок більшої складності розпізнавання зображень вушних раковин, гібридні

техніки різом з методами, що базуються на локальних дескрипторами на даний момент є найбільш досконалыми [7].

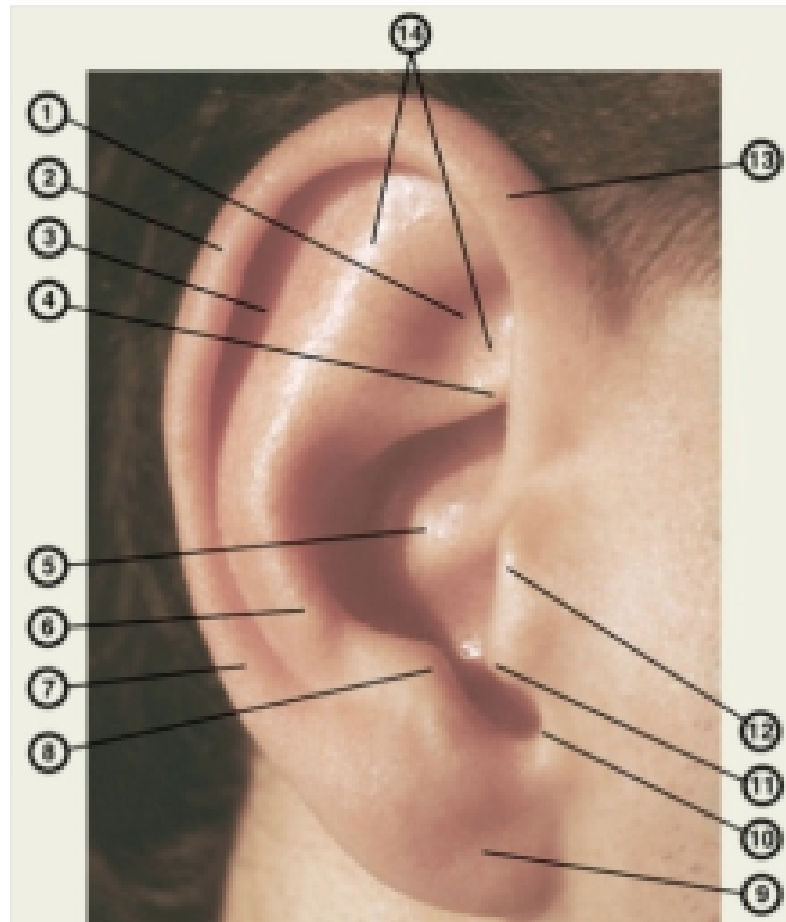


Рисунок 1.3 – Анатомічні компоненти вушної раковини: 1–Трикутна ямка. 2 – Горбок вушної раковини. 3 – Човноподібна ямка. 4 – Ніжка завитка. 5 – Раковина вуха. 6 – Протизавиток. 7 – Завиток. 8–Протикозелок. 9 – Вушна сережка. 10 – Міжкозелкова вирізка. 11– Козелок. 12 – Надкозелковий горбок. 13 – Завитково-козелкова борозда. 14–Ніжки протизавитка.

Ефективність роботи алгоритму стосовно заданого біометричного параметру зазвичай оцінюють за двома критеріями [31]:

FAR (False Acceptance Rate) – коефіцієнт помилкового доступу, процентний показник випадків, при яких перевірка особи виявилася помилково успішною.

FRR (False Rejection Rate) – коефіцієнт помилкової відмови в доступі, процентний показник випадків, при яких перевірка особи помилково завершилася невдачею.

У теорії статистики FRR і FAR прийнято називати в помилками 1-го та 2-го роду. Чим менше значення показників FRR і FAR, тим краща система. Проте залежно від конкретної області застосування одна з метрик має більшу вагу, ніж інша. Для системи контролю логічного або фізичного доступу критичним пріоритетом є заборона доступу неуповноважених осіб за будь-яких обставин. Тобто більшу вагу має метрика FRR.

Ефективність розпізнавання зображень залежить від стійкості біометричного шаблону до різного роду перешкод, спотворень у вихідному фото- або відеозображенні [39, 40].

1.2 Аналіз існуючих рішень

Розпізнавання ідентичності на основі зображень вушних раковин є активною сферою досліджень у біометричній спільноті. Жіга Емершч, Деян Штепец, Вітомір Штруц та Пітер Пір у своїй статті “Тренування згорткових нейронних мереж на малому наборі даних для розпізнавання зображень вушних раковин у природних умовах” [8] відзначають, що на відміну від інших біометричних модальностей, коли великі набори даних, отримані в неконтрольованих умовах, є легко доступними, набори даних зображень вушних раковин залишаються обмеженими і в основному доступні зображення лабораторної якості. Внаслідок цього технологія розпізнавання зображень вушних раковин ще не була вдосконалена згідно з останніми досягненнями глибокого навчання та згорткових нейронних

мереж (CNN) і до цього часу відстає інших модальностей, які розвинулися за рахунок глибоким технологіям розпізнавання зображень. У цій статті автори розглянули цю проблему і побудували модель розпізнавання зображень вушних раковин використовуючи згорткові нейронні мережі (CNN). Було вивчено різні стратегії щодо навчання нейромережевої моделі з обмеженою кількістю навчальних даних і показано, що, вибравши відповідну архітектуру моделі, використовуючи агресивне аугментацію даних (data augmentation) та вибіркоче навчання (selective learning) на існуючих попередньо навчених моделях (pretrained models), можливо ефективну згорткову нейромережеву модель (CNN), використовуючи трохи більше 1300 зображень для навчання.

Результатом роботи став перший згортковий нейромережевий підхід (CNN-based approach), до розпізнавання зображень вушних раковин, який також став відкритим для наукового співтовариства. Запропонована модель покращила показник розпізнавання попередньою найкращою моделі більш ніж на 25% на досліджуваному наборі даних зібраному в Інтернеті (у природніх умовах).

Галдамез та інші у своїй праці [11] визначили, що незважаючи на значну увагу до проблематики розпізнавання людей по зображенню вушних раковин виникають труднощі у зв'язку із різноманіттям форм, умов освітленості та зміні профілю форми, що є плоским відображенням складного об'єкту.

Авторами була запропонована для ідентифікації особи система розпізнавання, що включає згорткові нейронні мережі, якій на вхід подається зображення. Запропонований метод досягав результатів розпізнавання інших традиційних підходів при аналізі чистих фотографій. Однак F1 метрика результатів показала покращення при використанні запропонованої системи.

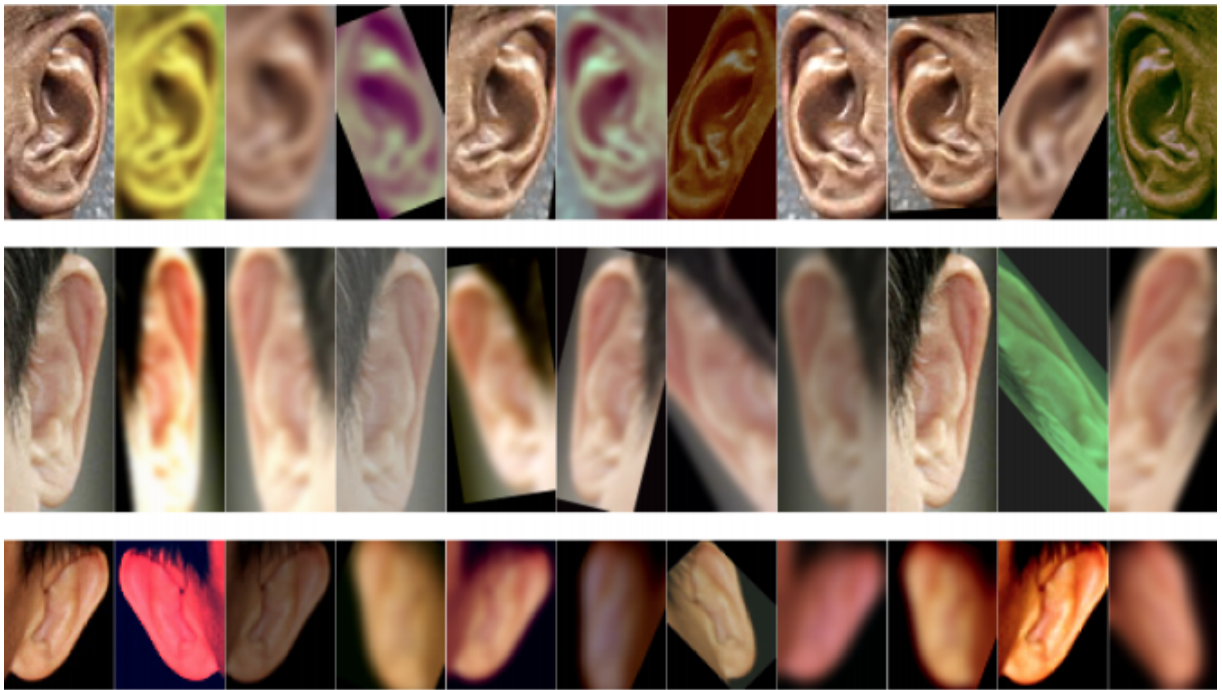


Рисунок 1.4 – Приклади аугментації зображень вушних раковин: крайнє зліва зображення кожного ряду є оригіналом з набору даних, наступні зображення результати аугментації

Дослідники також презентували техніку покращення швидкості тренування згорткової нейронної мережі застосованої до великих вхідних зображень завдяки оптимізації підходу ковзаючого вікна (sliding window). Цей алгоритм визначається двома параметрами: розмір вікна S (window size), який зазвичай встановлюється так, щоб відповідати розміру вхідного вектора НМ; та кроком вікна (window stride) T , який визначає відстань, якою розділені послідовні вікна. Крок визначає загальну кількість вікон W , яка будуть застосовуватись до даного вхідного зображення:

$$W = \left(\frac{I_w - S}{T} + 1 \right) \left(\frac{I_h - S}{T} + 1 \right)$$

$$W = \frac{I_w \cdot I_h}{T^2} ,$$

де I_w – ширина зображення;

I_h – висота зображення;

T – крок ковзаючого вікна;

S – розмір ковзаючого вікна.

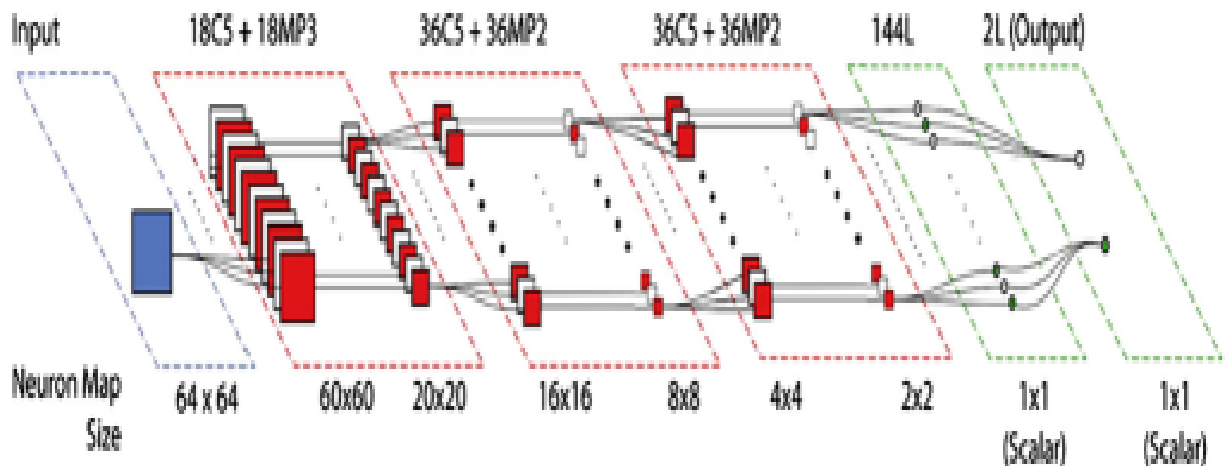


Рисунок 1.5 – Архітектура ГЗНН запропонованої Галдамезом, Ревеаном та Арріеттою у роботі “Короткий огляд процесу розпізнавання зображень вушних раковин використовуючи глибокі нейронні мережі”

У своїй статті [42] Сабанов А.Г. та Смолина С.Г. провели порівняльний аналіз методів біометричної ідентифікації особи, проаналізували достовірність ідентифікації виконаної з використанням різноманітних технологій, в тому числі із застосуванням біометрії. Виконаний огляд найбільш відомих способів і технічних рішень біометричної ідентифікації. Показано, що достовірність біометричної ідентифікації особистості із застосуванням технологій, що найбільш широко використовуються можна порівняти із достовірністю традиційних ідентифікаторів. Напрацьовані рекомендації із застосування певних технологій біометричної ідентифікації для невеликих систем.

Емершіч, Штруц, та Пір у своїй статті “Розпізнавання зображень вушних раковин: більше ніж опитування” [7] зазначають, що у останні роки було зроблено багато досліджень у розпізнаванні зображень вушних раковин, але все ще залишаються відкритими питання для комерційного застосування технологій. Було виконано огляд сфери автоматичного розпізнавання зображень вушних раковин (2D зображень). У даній сфері було детально розглянуто останні методи, що базуються на дескрипторах. Відкриті змагання були вивчені та розглянуті потенційні напрями подальших досліджень з метою визначення. Також було запропоновано новий набір даних зображень вушних раковин без обмежень зібраних в Інтернеті та набір технік для розпізнавання зображень вушних раковин.

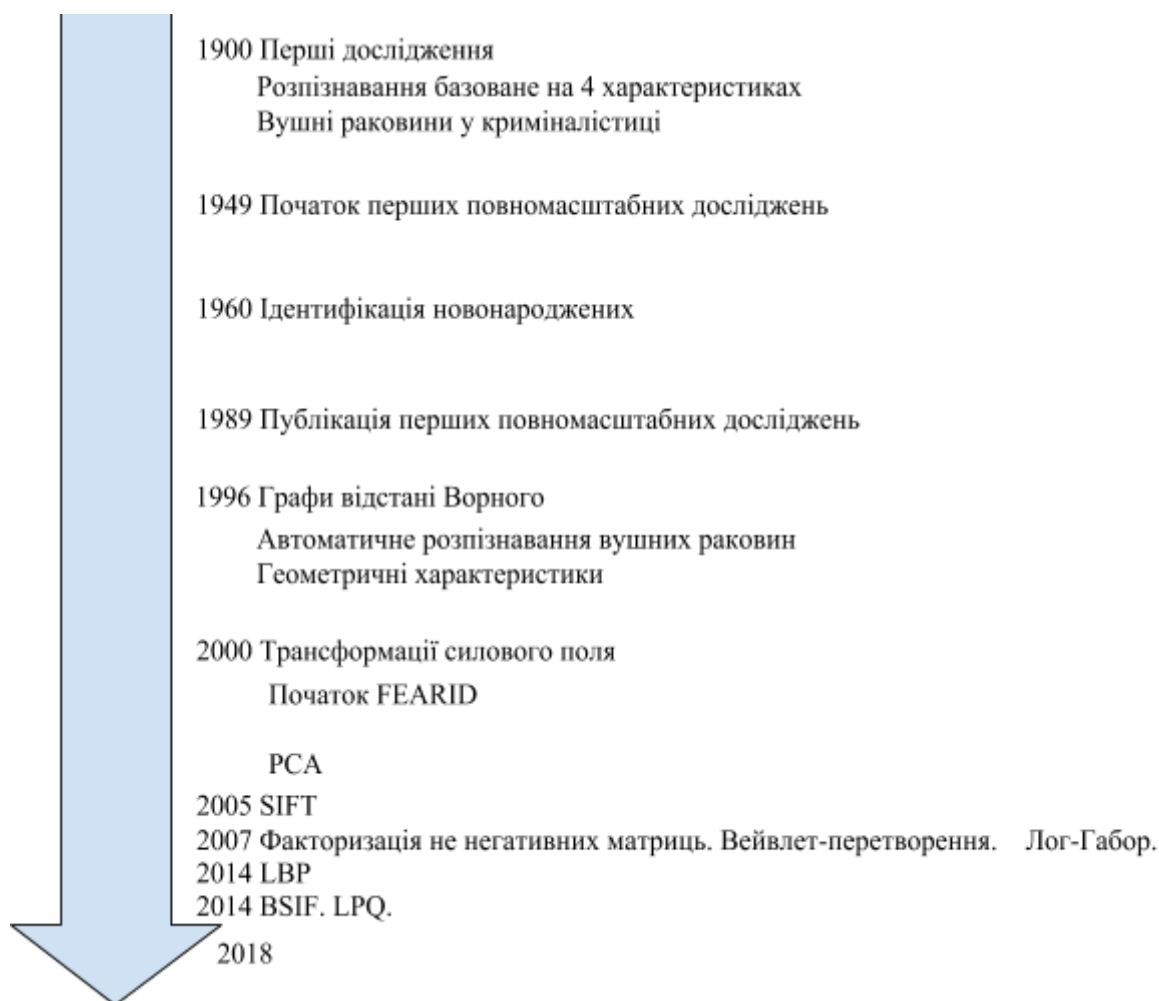


Рисунок 1.6 – Хронологія розвитку досліджень розпізнавання зображень вушних раковин

Вчені Ірем Февзіє Ейокур, Яман Догукан та Хазім Кемал Укенел із Стамбульського Технічного Університету у своїй статті “Адаптація до області застосування для розпізнавання зображень вушної раковини використовуючи глибокі згорткові нейронні мережі” [9] глибоко дослідили проблематику розпізнавання зображень вушної раковини без додаткових обмежень. Вони вперше показали необхідність адаптації до області застосування, коли глибокі згорткові нейромережеві моделі використовуються для розпізнавання зображень вушної раковини.

Для адаптації дослідники зібрали новий набір даних, використовуючи Multi-PIE набору даних облич, який був названий Multi-PIE набір даних зображень вушних раковин. Щоб покращити результати роботи, було поєднано різні глибокі згорткові нейромережеві моделі.

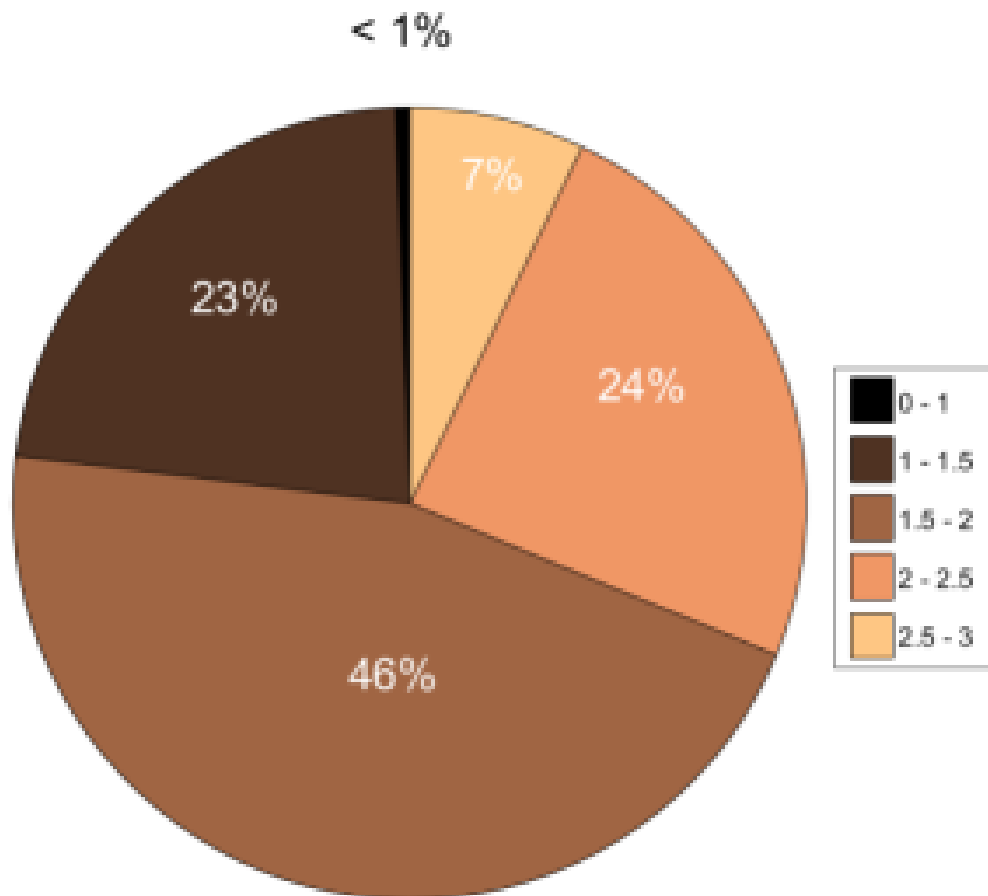
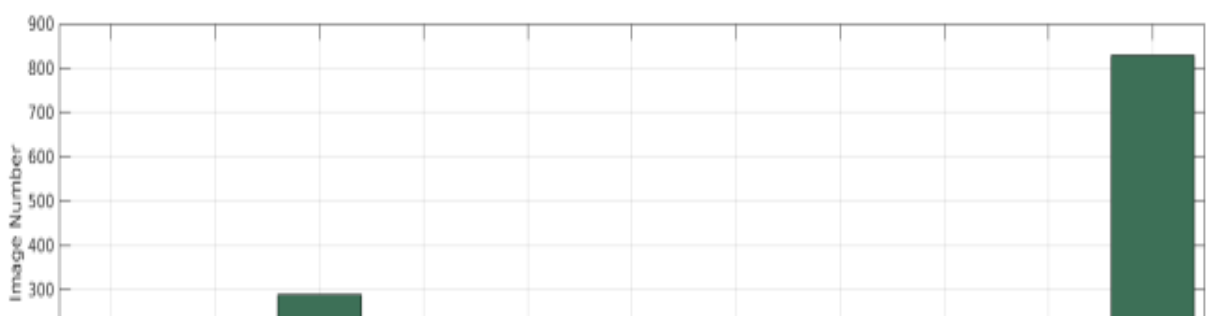


Рисунок 1.7 – Розподіл кількості зображень відповідно до співвідношення сторін наборі даних UERC

У данній статті було глибоко проаналізовано ефекти якості зображень вушних раковин, таких як: освітлення і співвідношення сторін, на результати роботи моделі. Було розглянуто проблему перекосу (bias) в наборі даних у сфері розпізнавання зображень вушних раковин. Проведені експерименти на UERC наборі даних показали, що адаптація до області застосування дає серйозне покращення результатів. Для прикладу, при використанні моделі VGG-16 із адаптацією до області застосування було досягнуто покращення результату на 10%. Поєднання різних глибоких згорткових нейромережевих моделей дало авторам



підвищення точності ще на 4%. Також було виявлено, що якість зображень має вплив на результат.

Рисунок 1.8 – Розподіл кількості зображень відповідно до роздільної здатності у (a) тренувальному та (b) тестовому наборі даних UERC

При проведенні експериментів спрямованих на дослідження перекосу набору даних, було класифіковано набір даних із точністю 99,71%, що вказує на значну похибку серед наборів даних для розпізнавання зображень вушних раковин.

Ліанг Тьян та Жічун Му у своїй праці “Розпізнавання зображень, що базується на глибоких згорткових нейронних мережах” [38] запропонували принципово новий алгоритм для розпізнавання зображень вушних раковин використовуючи глибокі згорткові нейронні мережі та підтвердили показали візуалізацію навченої мережі.

У таблиці 1 наведені параметри згорткової нейронної мережі запропонованої авторами. Глибина зростає зліва (від А) направо (до Е) (додані шари виділені жирним). Функція активації ReLU не наведена для забезпечення лаконічності. Параметри згорткового шару вказані як conv {розмір рецепторного поля} - {к-сть каналів}

Таблиця 1.1 – Параметри ЗНН запропонованої Ліанг Тьяном та Жічун Му

№	А	А-LRN	В	С	Д	Е
1	11 ЗШ	11 ЗШ	13 ЗШ	16 ЗШ	16 ЗШ	19 ЗШ
2	Вхідний шар (224 × 224 RGB зображення)					
3	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64
4		LRN	conv3-64	conv3-64	conv3-64	conv3-64

5	maxpool					
6	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
7			conv3-128	conv3-128	conv3-128	conv3-128
8	maxpool					
9	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
10	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
11				conv3-256	conv3-256	conv3-256
12						conv3-256
13	maxpool					
14	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
15	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
16				conv3-512	conv3-512	conv3-512
17						conv3-512
18	maxpool					
19	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
21	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
22				conv3-512	conv3-512	conv3-512
23						conv3-512
24	maxpool					

25	FC-4096
26	FC-4096
27	FC-1000
28	softmax

Авторами було розроблено згорткову нейронну мережу з 3 згортковими шарами нейронів, повністю з'єднаним шаром (fully-connected layer) та softmax класифікатором. Результати проведених дослідниками експериментів на USTB наборі даних знімків вушних раковин показали, що запропонований метод дозволяє меншими зусиллями добитися вищої точності та показує результати кращі зі традиційні методи боротьби із частковим загородженням.

Сучасні глибокі згорткові нейронні мережі показують надзвичайні результати у вирішенні ряду задач комп'ютерного бачення, таких як: розпізнавання зображень, класифікація та визначення об'єктів. Джаміль Н. та інші провели ґрунтовне дослідження впливу освітленості зображень при розпізнаванні зображень вушних раковин при використанні глибоких згорткових нейронних мереж [14].

Автори дослідили можливість згорткових нейронних мереж показувати однаково хороші результати при недовитриманих або перевитриманих вхідних зображеннях відомих як зображення з нерівномірно розподіленим освітленням (uniform illumination invariant).

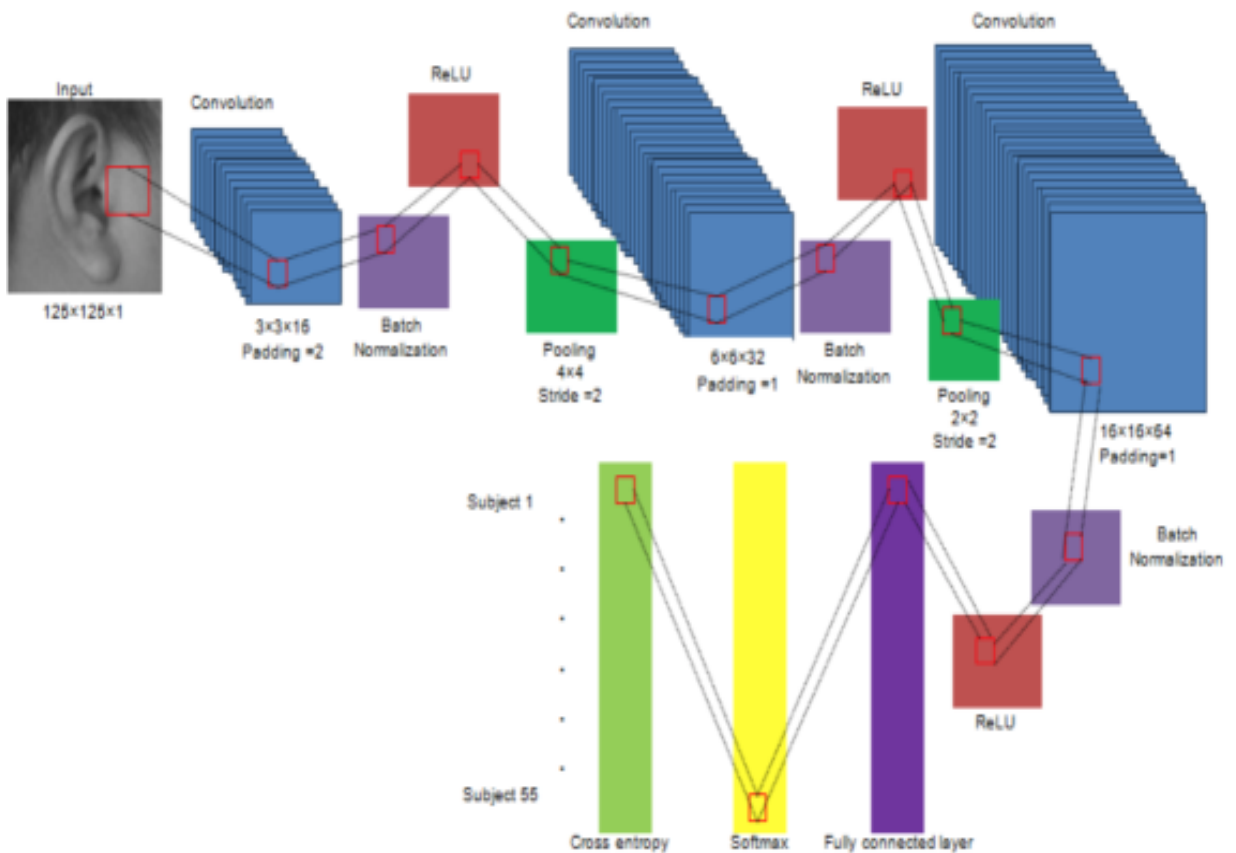


Рисунок 1.9 – Архітектура згорткової нейронної мережі (ЗНМ) запропонована [13]

В своїй роботі вчені зібрали зображення вушних раковин при різних умовах розподілення освітленості з варіацією люменів або люксів від 2 люксів до 10700 люксів. Всього 1100 зображень лівих і правих вушних раковин було зроблено при природному освітленні. Оскільки згорткові нейронні вимагають достатньо великої кількості даних, до зібраних зображень був застосований поворот на кожні 5 градусів, щоб згенерувати 25300 зображень. Для кожного перетворення було використано 50 зображень в якості валідаційного/тестового набору даних, решта зображень були використані для тренування. Запропонована згорткова нейромережева модель була натренована, валідаційні та тестові результати показали точність розпізнавання 97%. Результати показали досягнення точності у 100% була досягнута для зображень із освітленістю від більше 30, але є проблеми із зображеннями із освітленістю менше 10 люксів.

Таблиця 1.2 – Характеристики першого згорткового шару ЗНМ

<i>№</i>	<i>Тип шару</i>	<i>Деталі шару</i>
1	Вхідний шар	(125, 125, 1)
2	Фільтри	16
3	Розмір фільтрів	3 x 3
4	Кількість вагів на фільтр	$3*3*1 = 9$
5	Кількість параметрів на шар	$(9 + 1) * 16 = 160$
6	Крок (stride)	1
7	Кількість нейронів кожної карти характеристик (feature map)	$(125 - 3 + 2*1) / 1 + 1 = 125$
8	Загальна кількість нейронів шару	$125*125*16 = 250000$

З метою запобігання перенавчання НМ у повністю з'єднаному шарі (FC) дослідники використовували техніку регуляризації під назвою дропаут (dropout).

Таблиця 1.3 – Характеристики другого згорткового шару ЗНМ

<i>№</i>	<i>Тип шару</i>	<i>Деталі шару</i>
1	Фільтри	32
2	Розмір фільтрів	6 x 6
3	Кількість вагів на фільтр	$6*6*1 = 36$
4	Кількість параметрів на шар	$(36 + 1) * 32 = 1184$
5	Крок (stride)	1

6	Кількість нейронів кожної карти характеристик (feature map)	$(125 - 32 + 2*1) / 1 + 1 = 96$
7	Загальна кількість нейронів шару	$96*96*16 = 294912$

Функція помилок (втрат) визначалася через функцію крос-ентропії:

$$E(\theta) = - \sum_{i=1}^n \sum_{j=1}^k t_{ji} \cdot \ln y_j(x_i, \theta)$$

$$y_j(x_i, \theta) \sim P(t_j = 1 | x_i),$$

де θ – вектор параметрів,

t_{ji} – індикатор, що i -тий екземпляр належить до j -го класу,

x_i – вхідний вектор i -го екземпляру,

$y_j(x_i, \theta)$ – вихідний результат i -го екземпляру, який можна вважати

ймовірністю, що мережа асоціює i -тий зразок із j -тим класом.

Таблиця 1.4 – Характеристики третього згорткового шару ЗНМ

<i>№</i>	<i>Тип шару</i>	<i>Деталі шару</i>
1	Фільтри	64
2	Розмір фільтрів	16 x 16
3	Кількість вагів на фільтр	$16*16*1 = 256$
4	Кількість параметрів на шар	$(256 + 1) * 64 = 16448$
5	Крок (stride)	1
6	Кількість нейронів кожної карти характеристик (feature map)	$(16448 - 64 + 2*1) / 1 + 1 = 64$

7	Загальна кількість нейронів шару	$64*64*64 = 262144$
---	----------------------------------	---------------------

2. НЕЙРОМЕРЕЖЕВІ МОДЕЛІ ДЛЯ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ ПО ЗНІМКУ ВУШНОЇ РАКОВИНИ

За останні роки були розроблені та представлені різноманітні архітектури згорткових нейронних мереж (CNN). Однією з найбільш відомих проблем, що підкреслила потенціал та потужність підходів, що базуються на згорткових нейронних мережах, була класифікація об'єктів через змагання *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) [5].

Різні ЗНМ моделі мають різну кількість параметрів моделі, які необхідно визначати протягом тренування. Внаслідок цього виникають певні вимоги кількості тренувальних даних, що необхідні для вивчення параметрів. В той же час раціонально допустити, що легші архітектури (архітектури з меншою кількістю параметрів) вимагають меншої кількості даних для проектування, конвергенцію процедури (навчання) зворотного поширення помилки також необхідно розглянути, оскільки певні архітектури можуть піддаватися більш швидкому навчанню.

Архітектура AlexNet запропонована Александром Крижевським та іншими [14] досягла безпрецедентних результатів на наборі даних ImageNet та дала поштовх використанню та популярності згорткових нейронних мереж у розпізнаванні зображень.

Наступним кроком у покращенні результатів ILSVRC стала розробка 16 шарової VGGNet-16 та 19 шарової VGGNet-19 [26] .

Наведені архітектури використовуються до сьогоднішнього дня. Більш недавньою розробкою є архітектура ResNet представлена Хе,

Жангом та іншими, яка запровадила скорочені шляхи до згорткових нейронних мереж. Це дало можливість надійно тренувати глибокі мережі з кількома сотнями і навіть тисячами шарів [32].

2.1 Архітектура та особливості застосування нейромережевої моделі VGGNet-19

Вхідними даними для ЗНМ у загальному випадку є RGB зображення фіксованого розміру 224×224 пікселів. Єдиним передпроцесинговим кроком, який необхідно виконати, є нормалізація зображення шляхом віднімання середнього значення RGB обчисленого на тренувальному наборі даних від кожного пікселя.

Зображення пропускається через набір згорткових шарів, де застосовуються фільтри із малим рецептивним полем 3×3 (найменший розмір для отримання поняття про верх / низ, ліво / право, центр).

Згортковий шар являє собою набір карт (карти характеристик), у кожної карти є синаптичне ядро (скануюче ядро або фільтр).

Кількість карт визначається вимогами до задачі та архітектурою нейромережевої моделі. При збільшенні кількості карт збільшиться якість розпізнавання, але при цьому збільшується і обчислювальна складність.

Ми пропонуємо пропустити зображення через набір згорткових шарів, де застосовуються фільтри із малим рецептивним полем 3×3 (найменший розмір для отримання поняття про верх / низ, ліво / право, центр). Крок конволюції зафіксовано у 1 піксель; просторовий паддінг (padding) вхідних даних є таким, що просторова роздільна здатність зберігається після згортання, наприклад паддінг складає 1 піксель для згорткових шарів розміром 3×3 .

Розмір карт можна визначити за формулою 2.1:

$$(w, h) = (mW - kW + 1, mH - kH + 1), \quad (2.1)$$

де (w, h) – розмір згорткової карти, який обраховується;

mW – ширина попередньої карти;

mH – висота попередньої карти;

kW – ширина ядра (фільтру);

kH – висота ядра (фільтру).

Ядро являє собою фільтр або вікно, яке ковзає по всій області попередньої карти і знаходить певні характеристики об'єкта. Якщо розміри ядра малі, то неможливо буде визначити характеристики. Якщо розміри ядра занадто великі, то збільшується кількість зв'язків між нейронами і відповідно складність обчислень. Також розмір ядра вибирається таким, щоб розмір карт згорткового шару був парним, що дозволяє не втрачати інформацію при зменшенні розмірності у шарі пулінгу.

Ядро є системою розділяємих вагів або синапсів. У звичайній багат шаровій нейромережі дуже багато зв'язків між нейронами, тобто синапсів, що уповільнює процес детектування. У згортковій мережі – навпаки, загальні ваги дозволяють зменшити кількість зв'язків і дозволити знаходити одну й ту ж характеристику по всій області зображення.

Початково значення кожної карти згорткового шару рівні 0. Значення вагів ядер задаються випадково в області від -0,5 до 0,5. Ядро ковзає по попередній карті і виконує операцію згорткування, яка часто використовується для обробки зображень. Математично дану операцію можна описати наступною формулою 2.2:

$$(f * g)[m, n] = \sum_{k,l} f[m - k, n - l] * g[k, l], \quad (2.2)$$

де f – початкова матриця зображення;

g – фільтри згортки.

При цьому залежно від методу обробки країв початкової матриці результат може бути менше початкового зображення (valid), такого ж розміру (same) або більшого розміру (full) як показано на рис.2.1.

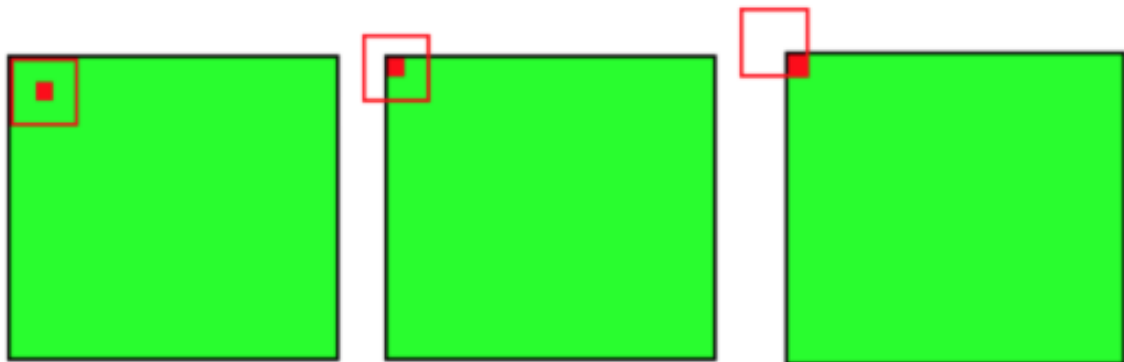


Рисунок 2.1 – Три методи обробки країв початкового зображення

В спрощеному вигляді цей шар можна описати формулою 2.3:

$$x^l = f(x^{l-1} * k^l + b^l), \quad (2.3)$$

де x^l – вихід шару l ;

$f(x)$ – функція активації;

b^l – коефіцієнт зсуву шару l ;

$*$ – операція згортки входу x із ядром k .

При цьому за рахунок крайових ефектів розмір початкових матриць зменшується:

$$x_j^l = f \left(\sum_i x_i^{l-1} * k_j^l + b_j^l \right), \quad (2.4)$$

де x_i^l – вихід шару l – карта j ;

$f(x)$ – функція активації;

b_j^l – коефіцієнт зсуву карти j шару l ;

* – операція згортки входу x із ядром k .

Мета пулінгового шару – зменшення розмірності карт попереднього шару. Якщо на попередній операціях згортки уже виявлені деякі характеристики, то для подальшої обробки настільки детальне відображення уже не потрібне, і воно ущільнюється до менш детального. Фільтрація вже непотрібних деталей допомагає запобігти перенавчанню.

В процесі сканування фільтрами шару пулінгу карти попереднього шару, скануючий фільтр не перетинається на відміну від згорткового шару.

Просторовий пулінг (pooling) виконується п'ятьма шарами макс-пулінгу (max-pooling), які слідуєть за деякими згортковими шарами (не за всіма згортковими шарами слідуєть шари макс-пулінгу). Макс-пулінг виконуєтья з вікном розміром 2×2 пікселя та кроком 2 пікселя. Шар пулінгу зменшує розмірність простору незалежно в кожному зрізі по глибині вхідних даних. Зліва на рис.2.1 зображено вхідні дані розміру $[224 \times 224 \times 64]$ агреговані із фільтром розміру 2, відступом 2 до вхідних даних розміром $[112 \times 112 \times 64]$. При цьому глибина масиву зберігаєтья. Зправа на рис. Зображений найбільш популярний спосіб агрегації макс-пулінг із відступом 2. Це означає, що із чотирьох чисел (квадрату 2×2) вибираєтья одне максимальне.

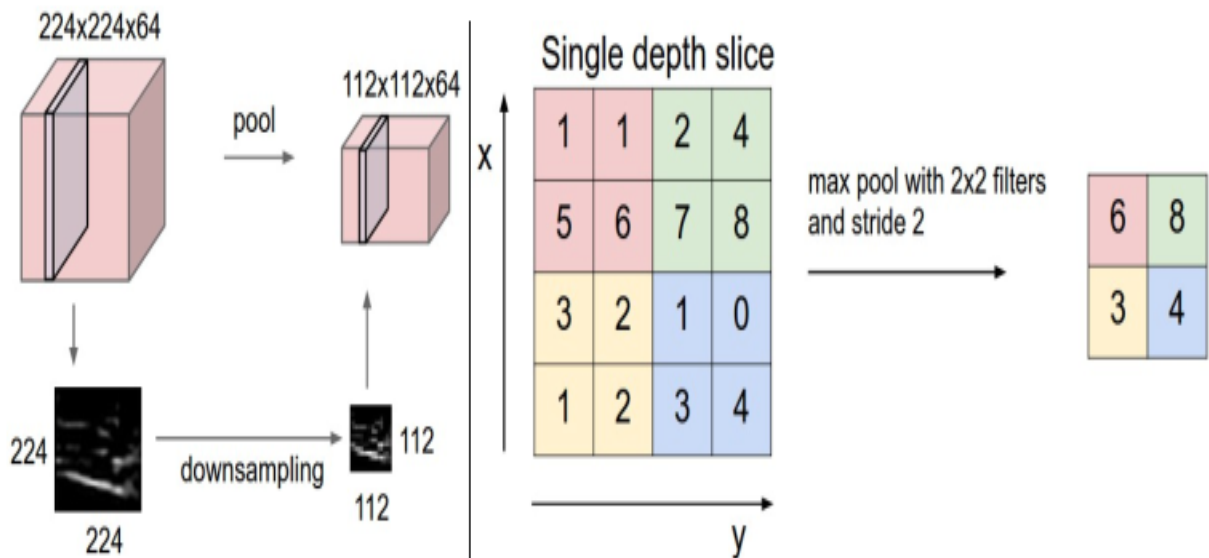


Рисунок 2.2 – Зліва: Пулінг із зменшенням розмірності з 224 x 224 до 112 x 112. Зправа макс-пулінг із відступом 2.

За блоком згорткових шарів (які мають різну глибину та архітектури) слідує три повністю з'єднаних (fully-connected) шари: перших два мають 4096 канали кожен, третій є відповідальним за класифікацію і має 1000 каналів (по одному каналу на кожен клас). Останнім шаром є шар софт-макс (soft-max) класифікатору. В першому наближенні математичне забезпечення софт-макс шару можна виразити наступною формулою 2.5:

$$S(x_i) = \frac{e^{x_i}}{\sum_{i=1}^K e^{x_i}}, \quad (2.5)$$

де $e=2,71828$ – стала Ейлера;

x_i – вхідний вектор i -го екземпляру,

i – номер тренувального екземпляру;

K – кількість класів.

Всі приховані шари активуються нелінійною трансформацією ReLU, яка може бути визначена за наступною формулою 2.6:

$$A(x) = \max(0, x), \quad (2.6)$$

де x_i – вхідний вектор i -го екземпляру,

$\max()$ – функція вибору максимального із двох чисел.

В даній мережі не використовується шари локальної нормалізації (LRN): оскільки така нормалізація не покращує продуктивність роботи згорткової нейромережевої моделі на виконанні наших завдань, але призводить до збільшення об'єму пам'яті, яка використовується та часу обчислень.

Нейрони у повністю з'єднаному шарі мають повне з'єднання з усіма активаціями попереднього шару як у звичайній нейронній мережі. Їх активації можуть бути обчислені за допомогою перемноження матриць разом із додаванням похибки (bias).

Тобто нейрони кожної карти попереднього шару з'єднані із одним нейроном прихованого шару. Таким чином число нейронів прихованого шару дорівнює числу карт попереднього шару, але зв'язки можуть бути необов'язково такими, наприклад, що тільки частина нейронів якоїсь із карт попереднього шару пов'язана із першим нейроном прихованого шару, а частина, що залишилася із другим, або всі нейрони першої карти пов'язані із нейронами першого та другого прихованого шару. Обчислення значення нейрона можна описати наступною формулою 2.7:

$$x_j^l = f \left(\sum_i x_i^{l-1} * w_{ij}^{l-1} + b_j^{l-1} \right), \quad (2.7)$$

де x_j^l – карта признаков j (вихід шару l);

$f()$ – функція активації;

b_j^{l-1} – коефіцієнт зсуву шару l ;

w_{ij}^{l-1} – матриця вагових коефіцієнтів шару l .

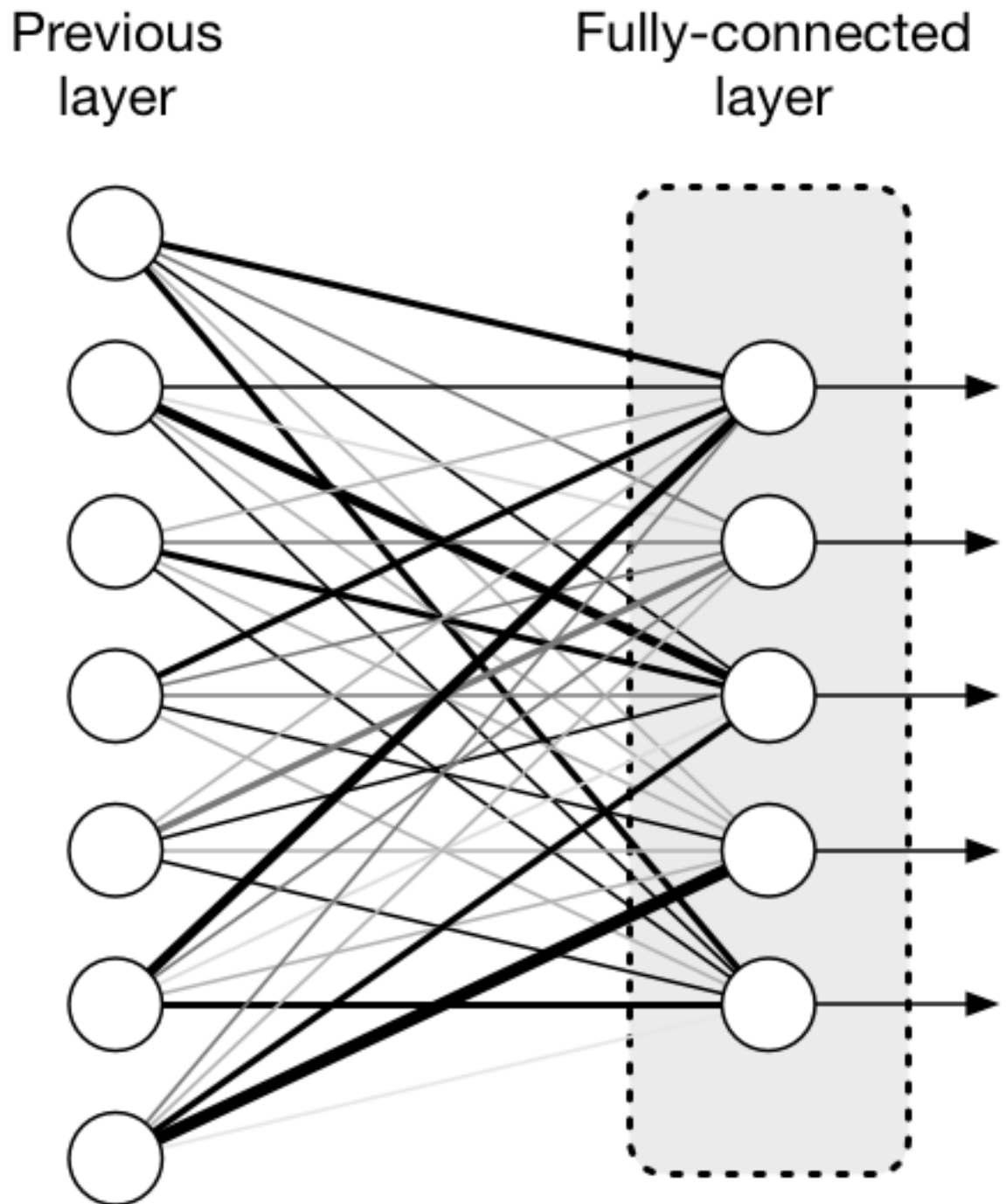


Рисунок 2.3 – Повністю з'єднаний шар нейронів згорткової нейромережевої моделі VGG-19

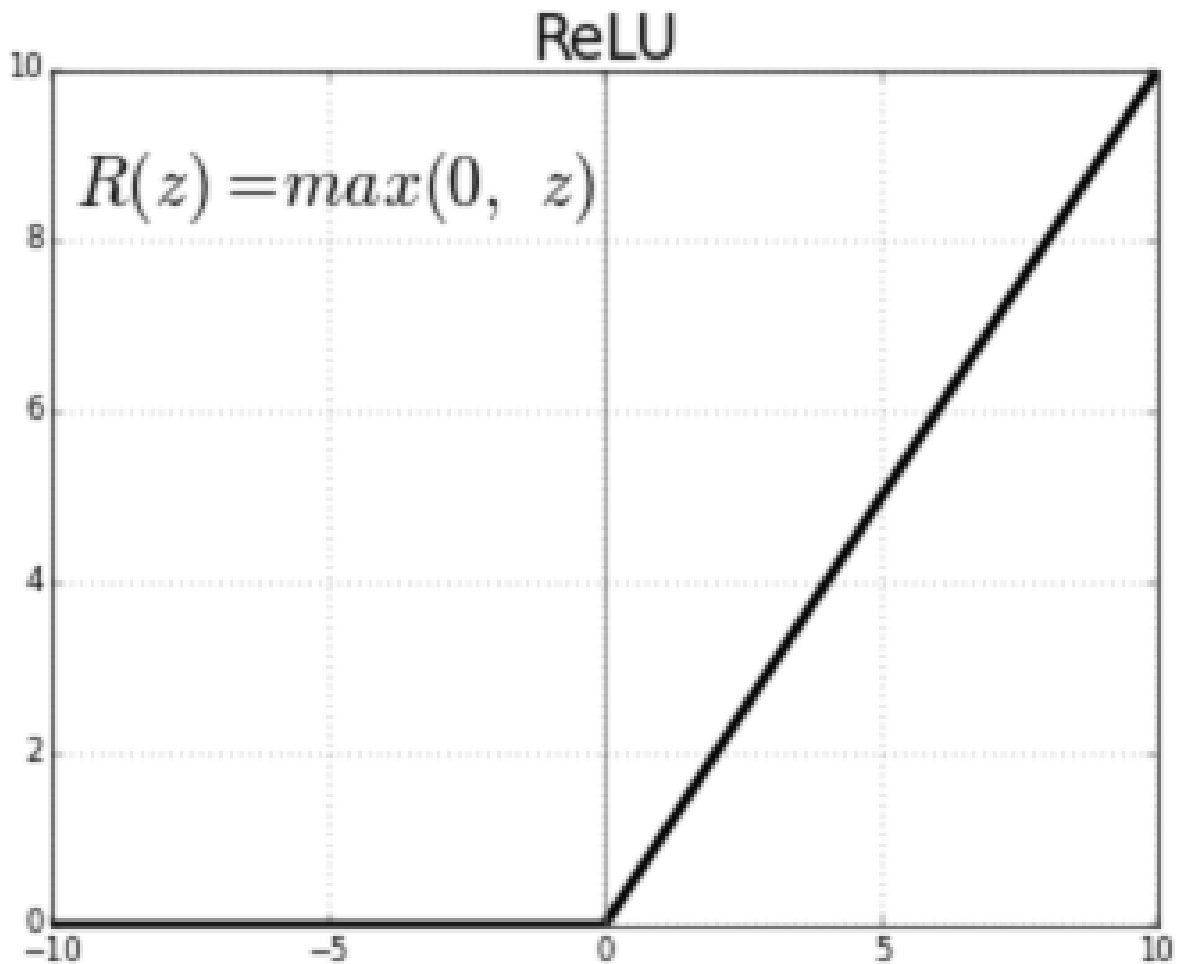


Рисунок 2.4 – Функція нелінійної трансформації ReLU.

В таблиці 2.1 наведемо кількість параметрів кожного шару ЗНМ. Незважаючи на значну глибину, кількість вагів нашої мережі не перевищує кількість вагів у більш компактних мережах із більшою глибиною згорткового шару та більшими рецептивними полями (144 000 000 вагів у [25]).

Зобразимо архітектуру згорткової нейромережевої моделі, що базується на архітектурі VGG-19 на рис.2.5.

Таблиця 2.1 – Кількість параметрів мережі VGG-19

№	Характеристика шару	Обчислення	Кількість п-рів	Кількість ваг
1	Вхідний: [224x224x3]	224*224*3	150 000	0
2	CONV3-64 [224x224x64]	224*224*64	3 211 264	$(3*3*3)*64 = 1\,728$
3	CONV3-64 [224x224x64]	224*224*64	3 211 264	$(3*3*3)*64 = 1\,728$
4	POOL2 [112x112x64]	112*112*64	802 816	0
5	CONV3-128: [112x112x128]	112*112*128	1 605 632	$(3*3*64)*128 = 73\,728$
6	CONV3-128: [112x112x128]	112*112*128	1 605 632	$(3*3*128)*128 = 147\,456$
7	POOL2: [56x56x128]	56*56*128	401 408	0
8	CONV3-256: [56x56x256]	56*56*256	802 816	$(3*3*128)*256 = 294\,912$
9	CONV3-256: [56x56x256]	56*56*256	802 816	$(3*3*256)*256 = 589\,824$
10	CONV3-256: [56x56x256]	56*56*256	802 816	$(3*3*256)*256 = 589\,824$
11	POOL2: [28x28x256]	28*28*256	200 704	0
12	CONV3-512: [28x28x512]	28*28*512	401 408	$(3*3*256)*512 = 1\,179\,648$
13	CONV3-512: [28x28x512]	28*28*512	401 408	$(3*3*512)*512 = 2\,359\,296$
14	CONV3-512: [28x28x512]	28*28*512	401 408	$(3*3*512)*512 = 2\,359\,296$
15	POOL2: [14*14*512]	14*14*512	100 352	0

	[14x14x512]			
16	CONV3-512: [14x14x512]	14*14*512	100 352	$(3*3*512)*512=2\ 359\ 296$
17	CONV3-512: [14x14x512]	14*14*512	100 352	$(3*3*512)*512=2\ 359\ 296$
18	CONV3-512: [14x14x512]	14*14*512	100 352	$(3*3*512)*512=2\ 359\ 296$
19	POOL2: [7x7x512]	7*7*512	25 088	0
20	FC: [1x1x4096]	1*1*4096	4096	$7*7*512*4096=102\ 760\ 448$
21	FC: [1x1x4096]	1*1*4096	4096	$4096*4096=16\ 777\ 216$
22	FC: [1x1x1000]	1*1*1000	1000	$4096*1000=4,096,000$
Всього			15 241 080	135 949 696

Загальна кількість параметрів даної моделі дорівнює 151 190 776.

Можна підсумувати, що головною ідеєю нейромережевої моделі VGG-19 є зменшення кількості параметрів нейронної мережі, які необхідно визначати в ході тренування. Наприклад, візьмемо згортку розміром 5 x 5 пікселів. Вона містить 25 параметрів. Якщо її замінити стеком з двох шарів із згортками 3 x 3 пікселі, то ми отримаємо таке ж відображення, але 25% менше. Якщо ж замінити згортку розміром 11 x 11 пікселів на стек із чотирьох згорток розміром 3 x 3 пікселі, то різниця буде набагато більш драматичною: $11 * 11 = 121$ проти $3*3 + 3*3 + 3*3 + 3*3 + 3*3 = 36$, а це на 70% менше.

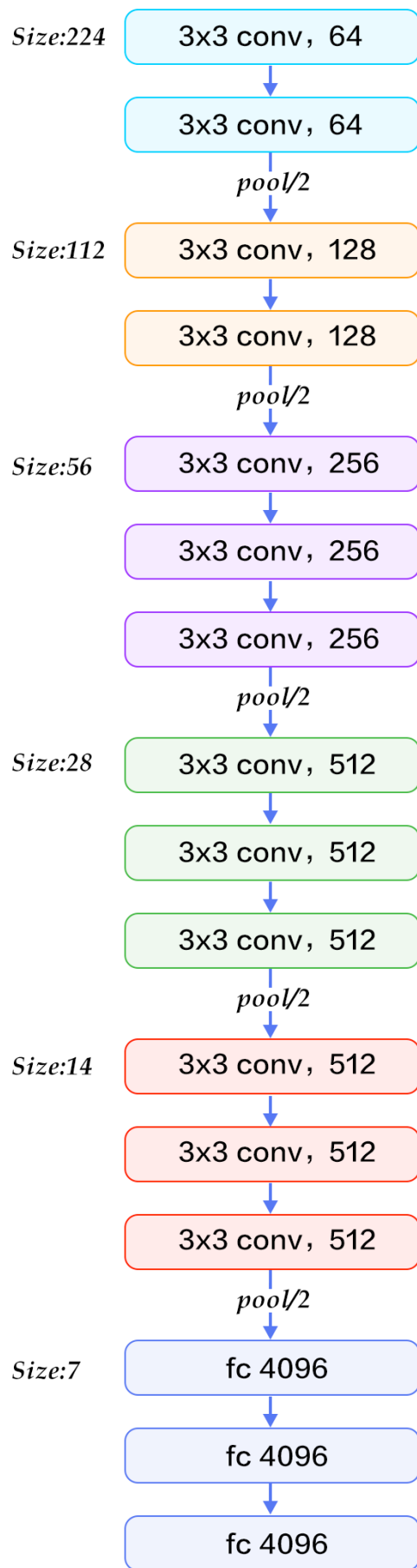


Рисунок 2.5 – Архітектура мережі на основі ЗНМ VGG-19

2.2 Архітектура та особливості застосування нейромережевої моделі на основі архітектури ResNet

Згідно з універсальною теоремою апроксимації: маючи достатньо потужностей, прямої нейронної мережі із одним шаром достатньо для того, щоб відобразити будь-яку функцію. При цьому шар має бути масивним і нейронна мережа матиме тенденцію до перенавчання на наборі даних. Тому склалася загальна тенденція у дослідницькому середовищі до того, щоб робити нейромережеві архітектури глибшими.

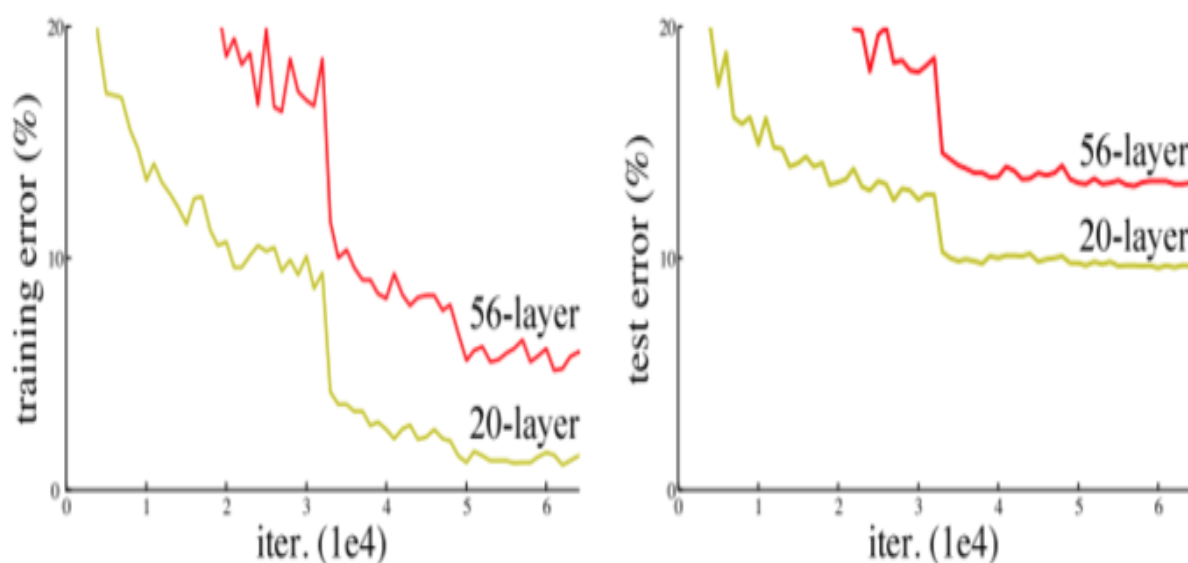


Рисунок 2.6 – Приклад зникання градієнту (vanishing gradient) при поглибленні нейронних мереж

Однак збільшення глибини мережі не дає бажаного покращення результату при простому розміщенні шарів один за одним. Глибокі мережі важко тренувати через серйозну проблему зникання градієнту – при зворотньому поширенні градієнту до ранніх шарів, повторюване перемноження може спричинити градієнт стати нескінченно

малим. В результаті із поглибленням нейронної мережі її продуктивність затухає або навіть починає стрімко деградувати.

Перед ResNet було декілька способів справлятися із проблемою зникання градієнту, наприклад, Шегеди, Ліу, Джіа, Серманет, Рід, Ангуелов, Ерхан, Вануке та Рабінович у своїй праці “Заглиблення з конволюціями” [30] додають додаткові втрати у середньому шарі як додатковий засіб, але жоден спосіб не міг справитися із проблемою повністю.

На відміну від традиційних послідовних нейромережевих архітектур таких як: AlexNet, OverFeat, and VGG, ResNet натомість є формою “екзотичних” архітектури, що базується на мікроархітектурних модулях, так званих, архітектур “мережа-в-мережі” (“network-in-network architecture”).

Термін мікроархітектура застосовують до набору “будівельних блоків”, які використовують для конструювання згорткової мережі. Набір мікроархітектурних будівельних блоків разом із стандартними згортковими, пуллінговими та іншими шарами утворює макроархітектуру (кінцеву мережу).

Ключова ідея ResNet у запровадженні, так званого, “одиночного з'єднання короткого шляху” (identity shortcut connection), що пропускає один або кілька шарів як показано на рис.2.7.

Автори [34] заперечують, що послідовне розміщення шарів один за одним не призведе до зниження продуктивності мережі, тому що можна просто розмістити одиничні маппінги (шари, що не роблять нічого) у нейронній мережі, і в результаті архітектура ЗНМ буде показувати той самий результат. Це означає, що глибші згортково нейромережеві моделі не будуть показувати тренувальну помилку вище ніж їх компактніші аналоги. Вчені висувають гіпотезу, що прилаштувати послідовні блоки (residual blocks) до

остаточного маппінгу (residual mapping) легше, ніж підлаштувати послідовні шари безпосередньо до бажаних нижчележачого маппінгу. І остаточний блок (residual block) чітко дозволяє зробити саме це.

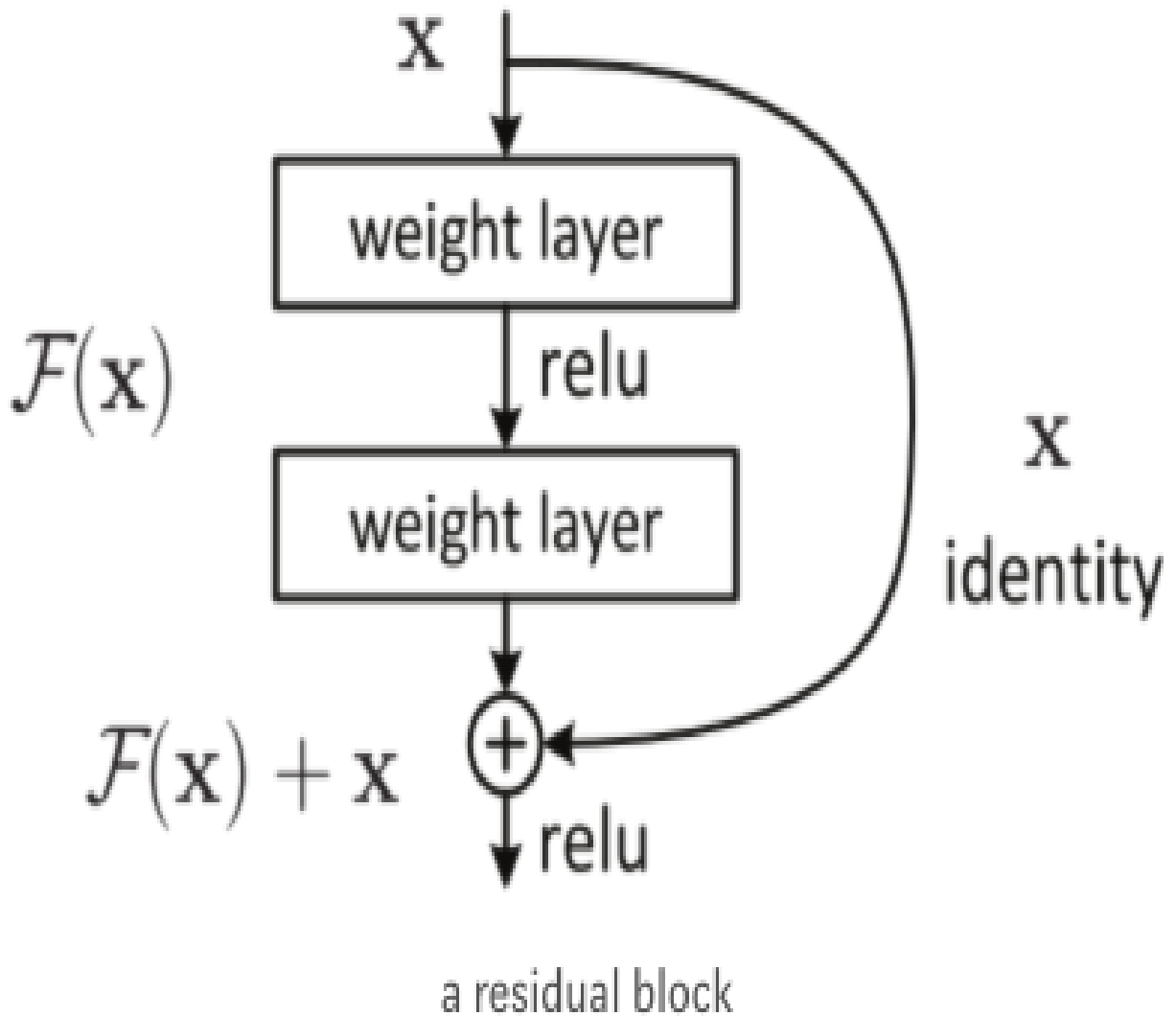


Рисунок 2.7 – Залишковий блок (residual block) згорткової нейронної мережі архітектури ResNet

Як згадувалось раніше, ResNet була не першою нейронною мережею, яка використала з'єднання короткого шляху (shortcut connections), Highway Network [28] запропонували огорожені з'єднання короткого шляху (gated shortcut connections). Ці параметризовані огороження контролюють об'єм інформації, що проходить через з'єднання короткого шляху. Схожу ідею можна побачити у комірках Довгої короткочасної пам'яті (ДКЧП, LSTM) [24], у яких є параметризоване огороження

відповідальне за забування, що контролює наскільки багато інформації потрапить до наступного кроку. Таким чином, ResNetf можна розглядати як частковий випадок нейронної мережі Highway Network.

Математичне забезпечення для остаточних блоків у першому наближенні можна представити наступними формулами:

$$F(x) = \sum_{i=1}^C \tau_i(x), \quad (2.8)$$

де C – розмір набору трансформації, які будуть агрегуватися,

$\tau_i(x)$ – довільна функція, яка проектує x на (ймовірно маловимірні) ембедінги (embedding) і тоді трансформ їх,

x – вхідний вектор i -го екземпляру,

i – номер тренувального екземпляру.

У формулі 2.8 Кантор у своїй праці [3] визначає C як кардинальне число (cardinality), яке контролює кількість більш складних трансформацій.

Водночас експерименти вказують на те, що нейронна мережа Highway Network показує результати гірші за ResNet. Даний факт є доволі дивним, оскільки простір рішень Highway Network містить ResNet, і таким чином вона має показувати щонайменше результати настільки ж хороші як і ResNet. При цьому допускається, що важливіше не стримувати прагнення цих “швидкісних градієнтів” (“gradient highways”) до більшого простору рішень.

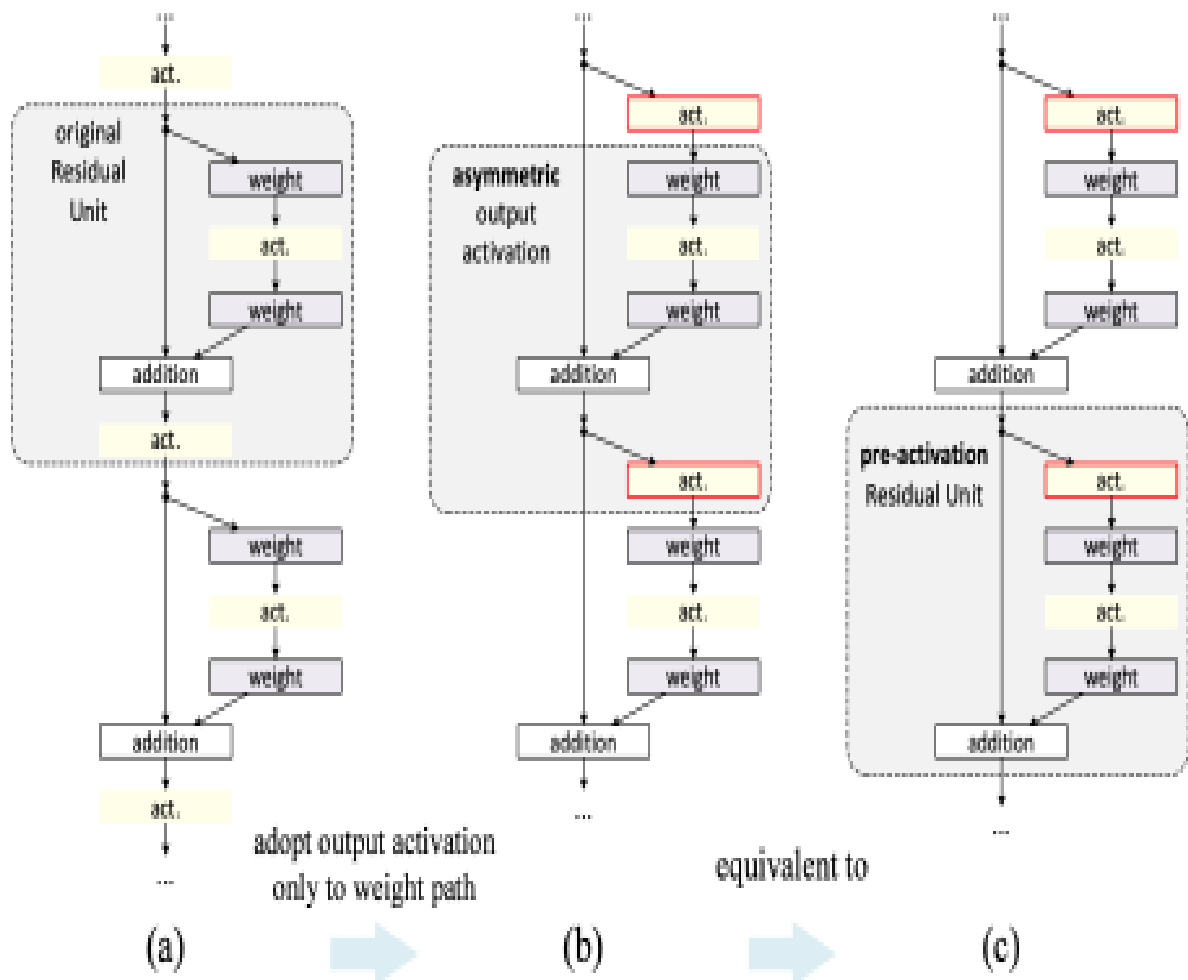


Рисунок 2.9 – Використання асиметричної активації післядодавання (asymmetric after-addition activation) рівносильне конструюванню преактиваційного залишкового блоку (pre-activation Residual Unit)

Розвиваючи цю гіпотезу, автори [35] звільнилися від остаточного блоку (residual block) і запропонували натомість перед-активаційний варіант остаточного блоку (pre-activation variant of residual block) [36], у

якому градієнти можуть перейти через з'єднання короткого шляху (shortcut connections) до будь-якого попереднього шару без перешкод. Фактично, використання оригінального остаточного блоку (residual block) у [35] при тренуванні 1202-шарової ЗНММ ResNet призвело до гіршого результату за її 110-шаровий аналог.

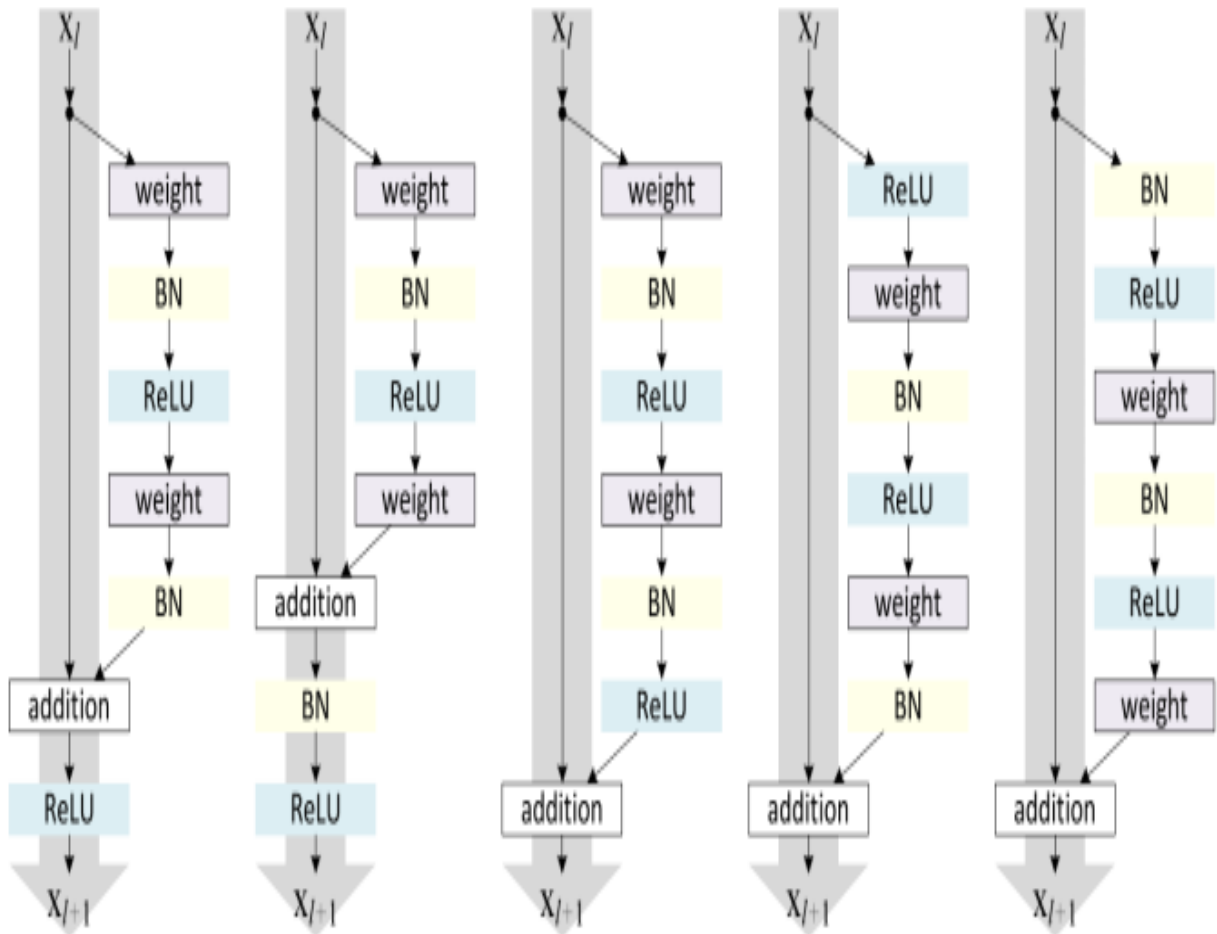


Рисунок 2.10 – Варіанти залишкових блоків. Зліва направо: оригінальний; Batch Normalization після додавання; ReLU після додавання; ReLU переддодавання; ReLU тільки як передактивація; повна преактивація.

В результаті автори [35] продемонстрували, що можливо натренувати 1001-шарову глибоку згорткову нейромережеву модель ResNet, яка буде показувати кращий результат за компактніші аналоги. Завдяки своїм інтригуючим результатам нейромережева модель ResNet швидко завоювала популярність серед найбільш продуктивних моделей у задачах комп'ютерного бачення.

При тренуванні ЗНМ ResNet для вирішення реальних проблем Хе та інші у своїй праці “Глибоке остаточне навчання для розпізнавання зображень” [35] радять послуговуватись наступними методами та техніками:

- Оптимізувати пакетами (batch normalization) після кожного згорткового шару.
- Використовувати ініціалізації Kсаве/2 (Xavier/2 initialization).
- Використовувати SGD разом із моментумом (momentum).
- За коефіцієнт швидкості навчання (learning rate) взяти 0.1 та зменшити його у 10 разів при досягненні валідаційною помилкою.
- Визначити розмір міні-пакету (mini-batch) 256.
- Встановити зниження вагів (weight decay) 10^{-5} .
- Не використовувати техніку регуляризації дропаут (dropout).

2.3 Архітектура та особливості застосування нейромережевої моделі на основі архітектури Сіамських нейронних мережах (Siamese neural networks)

Навчання “з одного погляду” – це техніка навчання представлення з єдиного зразку. Враховуючи, що ми розробляємо систему ідентифікації користувачів комп'ютерних мереж, ми не можемо використати і натренувати типову згорткову нейромережеву модель, тому що вона не зможе вивчити характеристики із такою малою кількістю даних. Відповідно перед нами стоїть завдання навчання “з одного погляду”, при якому ми визначаємо функцію подібності, яка порівнює два зображення вушних раковин і повідомляє чи вони співпадають.

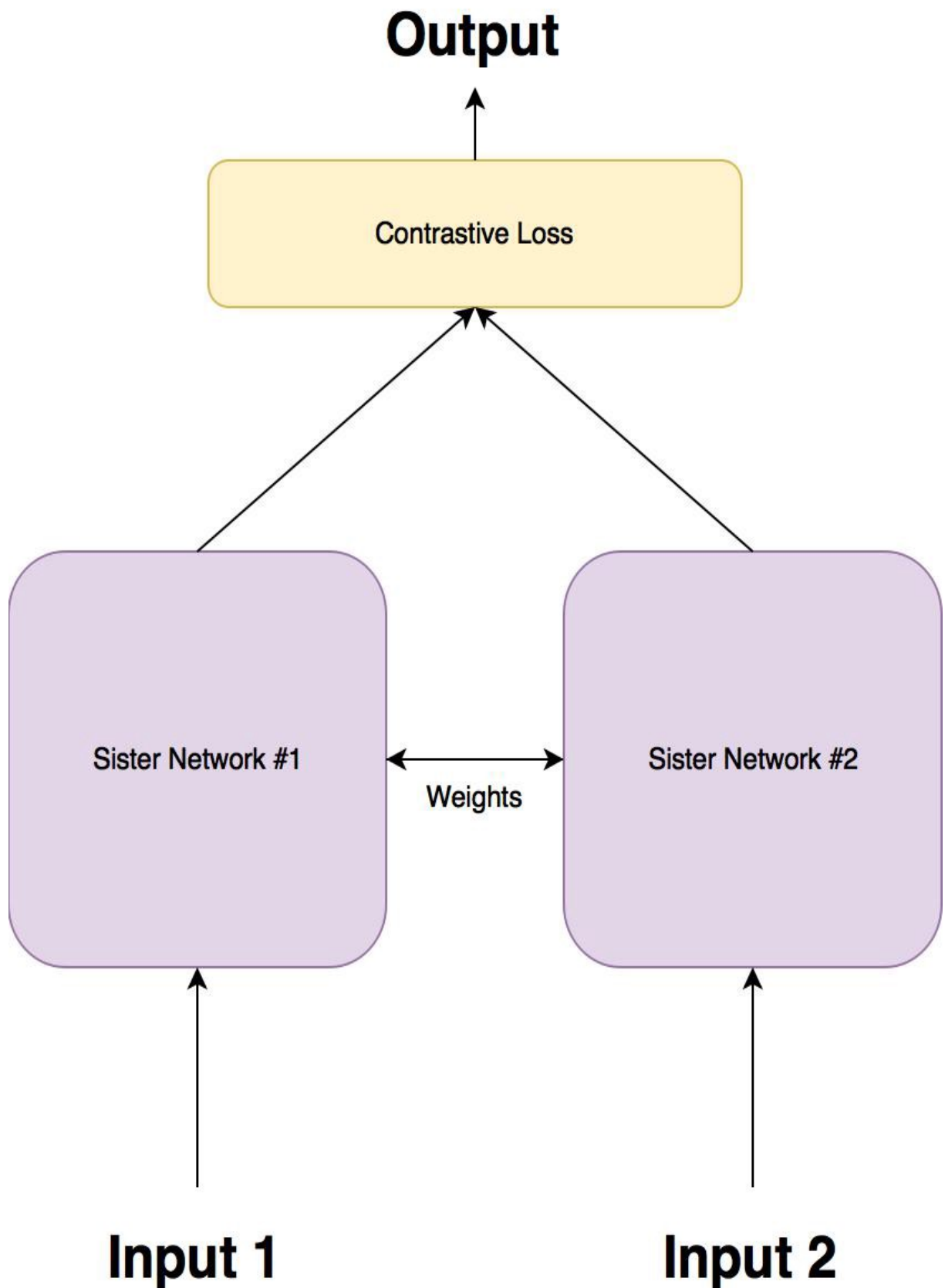


Рисунок 2.11 – Укрупнена схема архітектури нейромережевої моделі, що базується на Сіамських нейронних мережах (Siamese neural networks)

Функція подібності буде повертати значення ймовірності співпадіння і якщо значення більше певної рівня, можна сказати, що зображення подібні.

Функцію втрат для такої мережі можна представити наступною формулою:

$$L(x_1^{(i)}, x_2^{(i)}) = y(x_1^{(i)}, x_2^{(i)}) \cdot \log p(x_1^{(i)}, x_2^{(i)}) + (1 - p(x_1^{(i)}, x_2^{(i)})) \cdot \log(1 - p(x_1^{(i)}, x_2^{(i)})) + \lambda^T |w|^2, \quad (2.9)$$

де i – ідекс the i -го мініпакету;

$(x_1^{(i)}, x_2^{(i)})$ – значення обох зображень мініпакету;

$y(x_1^{(i)}, x_2^{(i)})$ – вектор маркування мініпакету;

λ^T – регуляризаційні ваги.

Сіамські нейронні мережі були запропоновані на початку 1990-х років Бромлі та ЛеКуном для вирішення задачі верифікації підпису як проблеми подібності зображень [2]. Сіамська нейронна мережа складається із мереж близнюків, які приймають різні вхідні дані, але об'єднані однією функцією енергії.

Ця функція обчислює деяку метрику між представленнями характеристик найвищого рівня з кожного боку. Параметри між мережами близнятами пов'язані. Привязка вагів гарантує, що два екстремально схожі зображення не можуть бути спроектовані відповідними мережами у дуже різні частини простору характеристик, оскільки обчислю спільну функцію. Також нейронні мережі симетричні, а отже, якщо ми надамо 2 різні зображення нейромережам близнюкам, верхній спільний шар обчислить одну й ту ж метрику так неначе ми надали такі ж два зображення, але протилежним близнюкам.

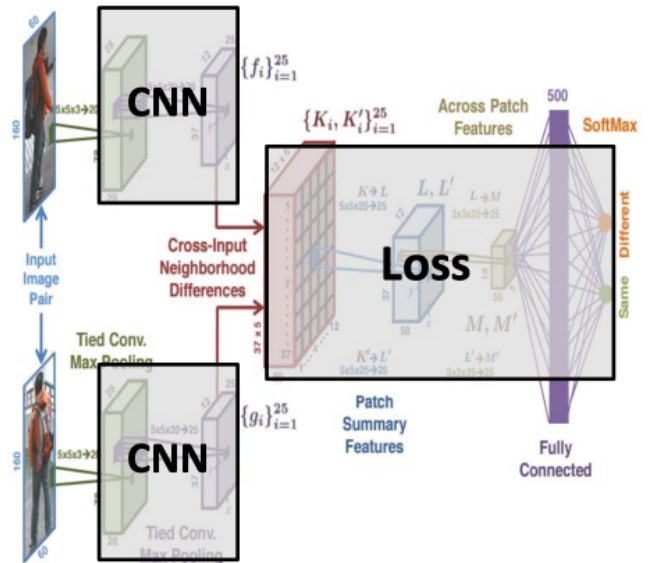
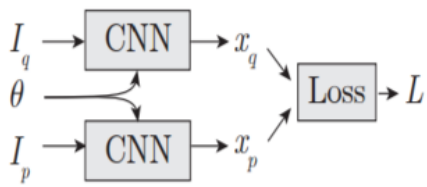


Рисунок 2.12 – Деталізація блоків нейромережевої моделі, що базується на Сіамських нейронних мережах (Siamese neural networks)

3. МЕТОДИ НЕЙРОМЕРЕЖЕВОЇ ІДЕНТИФІКАЦІЇ ПО ЗНІМКУ ВУШНОЇ РАКОВИНИ ТА ЇХ ОБҐРУНТУВАННЯ

Етап 1. Визначення задачі згорткової нейронної мережі.

Згідно із метою і завданнями даної магістерської дисертації очевидно, що задачею згорткової нейронної мережі є задача визначення схожості (similarity) та верифікації вушної раковини.

Етап 2. Визначення обмежень задачі.

Оскільки наша робота є одними з перших кроків у дослідженні данного напрямку, у першому наближенні ми не ставимо жодних обмежень щодо швидкості та точності процесу ідентифікації на основі розпізнавання зображень вушних раковин.

Для отримання вхідних даних зробимо фото вушних раковин користувачів нашої мережі за допомогою веб-камер із роздільною здатністю Full HD, тобто 1920 x 1080 пікселів

У дослідженні [31] визначення кількості вхідних і вихідних даних є відправною точкою у розробці моделі нейронної мережі, що і визначає наступний етап.

Етап 3. Визначення кількості вхідних параметрів.

В більшості випадків кількість вихідних нейронів приймається такою ж як кількість користувачів, яких слід розпізнати, а кількість вхідних нейронів такою ж як кількість пікселів зображення для розпізнавання (рисунок вушної раковини). У зв'язку із тим, що специфікою задачі нашої нейронної мережі є визначення подібності вхідного зображення вушної раковини користувача із зображеннями користувачів, яким наданий доступ

до мережі, то вхідними даними для наших згорткових нейромережових моделей є два зображення вушних раковин та відповідне бінарне цільове значення, яке вказує на те чи це зображення однієї людини.

При використанні таких сучасних веб-камер для виконання знімків повний розмір зображення складатиме 1920 x 1080 пікселів. Початково можна вважати, що зроблене зображення відповідає профілю голови користувача КМ. При проведенні експериментів в натурі встановлено, що лінійні розміри проекції вушної раковини складають 25% від лінійних розмірів профіля голови користувача. Шляхом простих арифметичних обчислень можна встановити розміри квадрата, що описує зовнішній контур вушної раковини складають 480 x 270 пікселів. Оскільки вхідне поле класичної ЗНМ (CNN) має співвідношення сторін 1:1 і відповідно розмір сторони може складати від 270 до 480 пікселів. Для спрощення обчислень приймемо стандартний розмір вхідного зображення 270 x 270 пікселів.

Оскільки кожне зображення розбивається на три канали: червоний, синій, Зелений, то ми отримаємо три зображення розміром 270 x 270 пікселів.

Вхідний шар враховує двомірну топологію зображення і складається із трьох карт (матриць), що відповідає кількості каналів кольорового зображення. Кожна карта відповідає зображенню з конкретним каналом (червоним, синім, зеленим).

Вхідні дані кожного конкретного пікселя необхідно нормалізувати в діапазон від 0 до 1. Таке перетворення можна здійснити за наступною формулою 3.1:

$$f(p, min, max) = \frac{p - min}{max - min} = \frac{p - 0}{255 - 0} = \frac{p}{255}, \quad (3.1)$$

де f – функція нормалізації;

p – значення конкретного кольору пікселя від 0 до 255;

min – мінімальне значення пікселя – 0;

max – максимальне значення пікселя – 255.

Етап 4. Задаємо конфігурацію мережі згідно з обраними базовими архітектурами.

Метою тренування нейромережевої моделі є максимізація ймовірності передбачення правильного класу. Ми можемо досягнути її мінімізуючи крос-ентропію для кожного тренувального зразка. Математичне забезпечення для функції крос-ентропії можна представити наступною формулою 3.2:

$$L = -\log p_k, \quad (3.2)$$

де L – функція крос-ентропії;

p – ймовірності передбачення класу k ;

k – правильний клас вхідного зображення.

Мінімізуємо функція втрат, обчисливши градієнт L і взявши до уваги параметри та оновлюючи значення параметрів використовуючи ітеративний метод оптимізації градієнтного спуску – стохастичний градієнтний спуск (SGD). Градієнти обчислимо за допомогою стандартного методу зворотнього поширення помилки (backpropagation) [23].

У нашому випадку ми обрали за основу архітектури VGG-19 та ResNet для кодування зображень.

Подальша розробка методу базувалася на принципах адаптації мережі до специфіки завдання (domain adaptation).

Етап 5. Навчання нейромережевої моделі.

Ініціалізація вагів. Ініціалізуватимемо всі ваги нейронної мережі згорткових шарів з нормального розподілу із середнім значенням 0 і стандартним відхиленням 10^{-2} . Зсуви (biases) доцільно ініціалізовувати так само як і згорткові шари із нормального розподілу, але із середнім значенням 0,5 та стандартним відхиленням 10^{-2} . У повноз'єднаних шарах, зсуви (biases) необхідно ініціалізовувати так як і у згорткових шарах, але ваги брати із значно ширшого нормального розподілу з середнім значенням 0 та стандартним відхиленням $2 * 10^{-1}$.

Очевидно, що у зв'язку з малою кількістю даних для тренування необхідно застосувати переніс навчання (transfer learning), адаптація до предметної області (domain adaptation), аугментацію даних (data augmentation) та точне налаштування (fine tuning).

Оптимізація гіперпараметрів. Вважаємо оптимальним використати реалізацію оптимізації Ветаба та Баєсівську оптимізацію із програмної бібліотеки для підбору гіперпараметрів нейромережевої моделі.

Для навчання і регуляризації, темп навчання задавався $\eta_j \in [10^{-4}, 10^{-1}]$, моментум $\mu_j \in [0, 1]$ і параметр L2 регуляризації $\lambda_j \in [0, 0.1]$. Повноз'єднані шари мали кількість нейронів від 128 до 4096 (такі що ділилися на 16). Оптимізатор задається так, щоб максимізувати точність валідаційного набору “з одного погляду”.

Переніс навчання (transfer learning) застосовувався в основному двома різними шляхами у згорткових нейромережевих моделях і залежить від розміру і подібності між передтренуваним (pretrained) і цільовим (target) наборами даних.

Перший підхід використовувати передтреновану глибоку ЗНММ безпосередньо для отримання характеристик з вхідного зображення. Отримані характеристики передаються, наприклад, класифікатору на основі методу опорних векторів (SVM classifier), щоб навчитися відрізняти різні класи один від одного. Цей підхід застосовується, коли у цільовий набір даних містить незначну кількість зразків зображень.

Другий підхід це точне налаштування передтренованої глибокої нейромережевої моделі на цільовому наборі даних. Відповідно необхідно ініціалізувати ваги передтренованої нейронної мережі і далі тренувати і проводити тонке налаштування вагів на цільовому наборі даних. Цей метод корисний, коли цільовий набір даних має достатньо тренувальних зразків, оскільки виконання точного налаштування на цільовому наборі даних (target dataset) із обмеженою кількістю зразків може призвести до перетренування (overfitting) [33]. Залежно від ступеню схожості між наборами даних і кількістю зразків доступних для тренування з цільового набору даних, необхідно вибрати один з двох підходів.

Етап 5.1. Вибір набору даних для початкового тренування.

У нашій роботі ми використаємо передтреновані нейромережеві моделі, що базуються на архітектурах VGG-19 та ResNet, що були натреновані на наборі даних ImageNet, тоді ми провели точне налаштування на наборі даних зображень вушних раковин.

Розглянемо набір даних для розпізнавання зображень вушних раковин, який містить обмежену кількість зразків для тренування (до 1000 зображень). Цієї кількості зразків достатньо для точного налаштування, але її б не вистачило для тренування глибокої згорткової нейромережевої моделі від початку.

Етап 5.2. Точне налаштування передтренованої глибокої згорткової нейромережевої моделі

Пропонуємо провести точне налаштування передтренованої глибокої згорткової нейромережевої моделі для розпізнавання зображень у два етапи. Для цього підходу спочатку сформуємо наш набір даних зображень вушних раковин з Multi-PIE набору даних зображень облич. Тоді проведемо точне налаштування передтренованої глибокої згорткової нейромережевої моделі на цьому наборі даних. Таким чином ми спочатку здійсимо адаптацію до предметної області (domain adaptation) для глибокої згорткової нейромережевої моделі. На другій фазі ми проведемо операцію з точного налаштування, використовуючи цільовий набір даних, яким стане UERC набір даних. Останній етап точного налаштування дозволяє більш направлену адаптацію до предметної області та завдання.

В нашому випадку проведемо адаптацію для природніх неконтрольованих умов. Цей крок насправді дуже важливий, оскільки як показано у дослідженні [28] є похибка (bias) серед наборів даних зображень вушних раковин.

При виконанні точного налаштування, параметри ініціалізувалися значеннями від передтренованої нейромережевої моделі. Значення темпу навчання (learning rate) останнього повноз'єданого шару було збільшено у 10 разів. Це загальноприйнята практика точного налаштування, оскільки ранні шари в основному фокусуються на отриманні низькорівневих характеристик (low-level features), а останні шари в основному відповідальні за класифікацію. Глобальний темп навчання обраний 0,001 і для згорткової нейромережевої моделі на основі архітектури VGG-19 і для конкурентної моделі на основі архітектури ResNet. Темп навчання зменшується у 10 разів кожні 20 тисяч ітерацій.

Відобразимо запропоновані методи неймережевої ідентифікації користувачів комп'ютерних мереж на рис.3.1.

4. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ВИКОРИСТАНОЇ НЕЙРОННОЇ МЕРЕЖІ

Було розроблено сіамську нейронну мережу для знайдення подібності між парами зображень вушних раковин для біометричної ідентифікації користувачів КМ. Дана модель особливо корисна якщо кількість тренувальних зразків обмежена зображеннями вушних раковин конкретної групи людей.

Розроблена згортова нейромережева модель (CNN) стала основою для розробки програмного забезпечення – DeepEar – призначеного для ідентифікації користувачів комп'ютерних мереж на основі зображення вушної раковини. Програма написана на мові програмування Python версії 3.7 із використанням бібліотеки PyTorch.

Система ідентифікації навчалася та тестувалася на вибірці зібраній автором. В експериментах розраховувалася точність розпізнавання тестових зразків при різних кількостях епох навчання (epochs) та карт характеристик.

У табл.4.1 та табл.4.2 наведено порівняння помилок та швидкості роботи ЗНММ архітектур, що базуютьс на основі нейронних мереж VGG-19 та ResNet-34.

Під швидкістю ми розуміємо загальний час виконання прямого та зворотнього проходу. Тренування відбувалося за допомогою відеокарти Pascal Titan X.

Таблиця 4.1 – Порівняння ефективності нейромережових моделей досліджуваних архітектур

№	Модель, на якій базується згортова нейронна мережа	Помилка, %	Швидкість, мс
1	VGG-19	27,30	147,32
2	ResNet-34	26,73	51,59

Відеокарта Pascal Titan X була встановлена на машині з процесором Intel Core i5-6500 (CPU) та оперативною пам'яттю 16GB RAM під управлінням ОС Ubuntu 16.04 із CUDA 8.0 RC. Детальні характеристики відеокарти наведені у табл.4.3.

Таблиця 4.2 – Порівняння швидкості роботи нейромережових моделей

№	ЗНММ	Швидкість зворотнього проходження, мс	Швидкість прямого проходження, мс	Загальна швидкість, мс
1	VGG-19	99,23	48,09	147,32
2	ResNet-34	34,58	17,01	51,59

При тренуванні використовувалася версія 5.1.05 cuDNN.

Таблиця 4.3 – Характеристики відеокарти, що використовувалася для експериментів

№	Відеокарта	Пам'ять	Архітектура	Ядра CUDA	FP32 TFLOPS	Дата релізу
1	Pascal Titan X	12GB GDDR5X	Pascal	3584	10,16	Серпень 2016р.

Результати експериментів частково представлені у табл.4.4. Максимальна точність розпізнавання становить близько 84,51%.

Таблиця 4.4 – Порівняння точності роботи сіамських нейронних мереж із використанням обраних базових архітектур

№	Базова архітектура сіамської нейронної мережі	Точність ідентифікації, %
1	VGG-19	82,1
2	ResNet-34	83,6

Досягнені результати дозволяють стверджувати, що використання нейромережева біометрична ідентифікація користувачів на основі розпізнавання зображень вушної раковини має подальшу перспективу.

ВИСНОВКИ

В даній дисертації було отримано подальший розвиток у вирішенні актуальної науково-технічної проблеми біометричної ідентифікації користувачів комп'ютерних мереж.

Зображення вушних раковин можна легко отримати з відстані без відома відповідної особи. Тому ідентифікації на основі розпізнавання зображень вушних раковин підходить для додатків розвідки, підтримання безпеки, контролю доступу та моніторингу. Відбитки вушних раковин, знайдені на місці злочину, були використані як доказ у більш ніж кількох сотнях випадків в Нідерландах та США.

Дослідження, виконані Яннареллі, дають достатньо доказів, що вушні раковини є унікальними біометричними рисами.

Для таких проблем комп'ютерного бачення згорткові нейронні мереж, є найкращим рішенням.

Вважаємо, що було досягнуто мету даної роботи, а саме підвищення ефективності систем біометричної ідентифікації користувачів комп'ютерних мереж.

Основні результати даного дослідження:

1. Детально проаналізовано існуючі алгоритми, методи та моделі біометричної ідентифікації.
2. Розроблено та реалізовано два варіанти самських нейромережових моделей призначених для біометричної ідентифікації користувачів комп'ютерних мереж на основі архітектур VGG-19 та ResNet.
3. Розроблено метод нейромережової ідентифікації по знімку вушної раковини.

4. Проведено експериментальні дослідження, які підтверджують ефективність та доцільність запропонованих рішень, оскільки запропоновані моделі показують тоність більше 82%.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Bhattacharyya D. Biometric Authentication: A Review / Bhattacharyya Debnath, Ranjan Rahul, Alisherov Farkhod, Minkyu Choi // International Journal of u- and e-Service, Science and Technology. – № 2. – 2009. – С.335–341.
2. Bromley J. Signature verification using a siamese time delay neural network / Jane Bromley, James W Bentz, Leon Bottou, Isabelle Guyon, Yann LeCun, Cliff Moore, Eduard Sackinger, Roopak Shah // International Journal of Pattern Recognition and Artificial Intelligence. – №7(04). – 1993. – С.669–688.
3. Cantor G. Über unendliche, lineare Punktmannichfaltigkeiten, arbeiten zur mengenlehre aus den jahren 1872–1884. – Leipzig: BSB B.G. Teubner Verlagsgesellschaft, 1984. – 262 с.
4. Chen L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs [Електронний ресурс] / Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille // CoRR. – 2014. – Режим доступу до ресурсу: <http://arxiv.org/abs/1412.7062>.
5. Deng J. ImageNet: A large-scale hierarchical image database / J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2009. – С. 248–255.
6. Dong C. Image super-resolution using deep convolutional networks / C.Dong, C.C. Loy, K. He, X. Tang // IEEE transactions on pattern analysis and machine intelligence. – Vol. 38, № 2. – 2016. – С. 295–307.
7. Emeršič Ž. Ear Recognition: More Than a Survey / Z. Emersic, V. Struc, P. Peer // Neurocomputing. – 2017.
8. Emeršič Ž. Training Convolutional Neural Networks with Limited Training Data for Ear Recognition in the Wild / Žiga Emeršič, Dejan Štepec,

Vitomir Štruc, Peter Peer // 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). – 2017

9. Eyiokur F.I. Domain Adaptation for Ear Recognition Using Deep Convolutional Neural Networks / F.I. Eyiokur, D. Yaman, H.K. Ekenel // IET Biometrics. – Vol. 7, No. 3. – 2018. – C. 199-206.

10. Frejlichowski D. The west pomeranian university of technology ear database – a tool for testing biometric algorithms / D. Frejlichowski, N. Tyszkiewicz // Image Analysis and Recognition. – 2010. – C. 227–234.

11. Galdámez P. L. A brief review of the ear recognition process using deep neural networks / P. L. Galdámez, W. Raveane, and A. G. Arrieta // Journal of Applied Logic. – 2016.

12. Hui Chen. Shape Model-Based 3D Ear Detection from Side Face Range Images / Hui Chen, B. Bhanu // Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2005. – C. 122.

13. Hurley D.J. Ear biometrics by force field convergence / D.J. Hurley, M.S. Nixon, J.N. Carter // International Conference on Audio and Video-Based Biometric Person Authentication (AVBPA). – 2005. – C. 386–394.

14. Jamil N. Can Convolution Neural Network (CNN) Triumph in Ear Recognition of Uniform Illumination Invariant? / Nursuriati Jamil, Ali Abd Almisreb, Syed Mohd Zahid Syed Zainal Ariffin, N. Md Din, Raseeda Hamzah // Indonesian Journal of Electrical Engineering and Computer Science. – Vol. 11, No. 2. – August 2018. – C. 558–566

15. Krizhevsky A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in Neural Information Processing Systems (NIPS). – 2012. – C. 1097–1105.

16. Kumar A. Automated human identification using ear imaging / A. Kumar, C. Wu. // Pattern Recognition. – №45 (3). – 2012. – C. 956–968.

17. Lowe D.G. Object recognition from local scale-invariant features / D.G.Lowe // Computer vision. The proceedings of the seventh IEEE international conference. – Vol. 2. – 1999. – C. 1150–1157.
18. Morales A. Earprint recognition based on an ensemble of global and local features / A. Morales, M. Diaz, G. Llinas-Sanchez, M.A. Ferrer // Security Technology (ICCST). International Carnahan Conference on IEEE. – 2015. – C.253–258.
19. Pflug A. Ear biometrics: A Survey of Detection, Feature Extraction, and Recognition Methods / A. Pflug, C. Busch // IET biometrics. – Volume 1, Issue 1. – 2012. – C.114–129.
20. Prabhakar S. An introduction to biometric recognition / A.K. Jain, A. Ross, and S. Prabhakar // IEEE Trans. Circuits Syst. Video Technology, Special Issue Image- and Video-Based Biomet. – Volume 14, Issue 1. – 2004. – C. 4–20.
21. Prakash S. An efficient ear recognition technique invariant to illumination and pose / S. Prakash, P. Gupt // Telecommunication Systems. – №52 (3). – 2013. – C.1435–1448.
22. Redmon J. You only look once: Unified, real-time object detection. / J. Redmon, S. Divvala, R. Girshick, A. Farhadi // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – C. 779–788.
23. Rumelhart D. Learning representations by back-propagating errors / D. Rumelhart, G. Hinton, R. Williams // Nature. – 1986.
24. Schmidhuber J. Long Short-term Memory / S. Hochreiter, J. Schmidhuber // Neural Computation. – №9 (8). – 2015. – C.1735–1780.
25. Sermanet P. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks / P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun // Proceedings of the International Conference on Learning Representations ICLR. – 2014.

26. Simonyan, K. Very deep convolutional networks for large-scale image recognition / K. Simonyan, A. Zisserman // Proceedings of the International Conference on Learning Representations (ICLR). – 2015.

27. Spaanenburg L. CBAS: A CNN-based biometrics authentication system / S. Malki and L. Spaanenburg // 2010 12th International Workshop on Cellular Nanoscale Networks and their Applications. – CNNA, Berkeley, CA. – 2010. – C.1–6.

28. Srivastava R. Training Very Deep Networks / R. Srivastava, K. Greff, J. Schmidhuber // NIPS Proceedings. – 2015. – Режим доступа до пецыцы: <https://papers.nips.cc/paper/5850-training-very-deep-networks.pdf>.

29. Subramanian R.K. Robust Multi-biometric Recognition Using Face and Ear Images / Nazmeen Bibi Boodoo*, R.K. Subramanian // (IJCSIS) International Journal of Computer Science and Information Security . – Vol. 6, No. 2 . – 2009.

30. Szegedy C. Going deeper with convolutions / C. Szegedy, W. Liu., Jia Y., P. Sermanet., S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – C.1–9.

31. Tereykovsky I. Perspectives for using classical neural network models and methods of counteracting attacks on network resources of informational systems / B. Aichanov, A. Korchenko, I. Tereykovskyi, I. Bapiev // News of the national academy of sciences of the republic of Kazakhstan. Series of geology and technical sciences. – Volume 5, Number 425. – 2017. – C.202–212.

32. Torralba A. Unbiased look at dataset bias / A. Torralba, A.A. Efros // IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2011. – 2011. – C.1521–1528.

33. Wang, Z.Q. Multi-scale feature extraction algorithm of ear image. / Z.Q. Wang, X.D. Yan // IEEE International Conference on Electric Information and Control Engineering (ICEICE). – 2011. – С.528–531.
34. Yosinski J. How transferable are features in deep neural networks? / J. Yosinski, J. Clune, Y. Bengio, H. Lipson // Advances in Neural Information Processing Systems (NIPS). – 2014. – С. 3320–3328.
35. Yuan L. Ear recognition using improved non-negative matrix factorization / L. Yuan, Z.-c. Mu, Y. Zhang, and K. Liu // Pattern Recognition. ICPR. 18th International Conference. – Vol. 4. – 2006. – С. 501–504.
36. Zhang X. Deep residual learning for image recognition / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2016. – С. 770–778.
37. Zhang X. Identity Mappings in Deep Residual Networks / Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun // Microsoft Research. – 2016. – Режим доступу до ресурсу: <https://arxiv.org/pdf/1603.05027v3.pdf>
38. Zhichun Mu. Ear recognition based on deep convolutional network / Mu Zhichun, Tian Liang // 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). – Datong. – 2016. – С. 437–441.
39. Корпань Я.В. Аналіз методів та алгоритмів компресії-декомпресії цифрових відеоданих [Текст] / Я.В. Корпань // Вісник Хмельницького національного університету. – 2015. – №3. – С. 175–179.
40. Корпань Я.В. Методи та алгоритми компактного представлення графічної інформації в комп'ютерних системах [Текст] / Я.В. Корпань // Технологічний аудит та резерви виробництва. – 2015. – Т. 3, №2 (23). – С.124–134.

41. Лысак А.Б. Идентификация и аутентификация личности: обзор основных биометрических методов проверки подлинности пользователя компьютерных систем [Текст] / А.Б. Лысак // Математические структуры и моделирование. – 2012. – №2 (26). – С. 124–126.

42. Сабанов А.Г. Сравнительный анализ методов биометрической идентификации личности / А.Г. Сабанов, С. Г. Смолина // Труды института системного анализа РАН. – Том 66. – 2016. – С. 11–20.