

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ  
імені ІГОРЯ СІКОРСЬКОГО»  
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ  
Кафедра математичного моделювання та аналізу даних

До захисту допущено  
Завідувач кафедри

\_\_\_\_\_ Наталія КУССУЛЬ

(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 2022 р.

**Дипломна робота**  
на здобуття ступеня бакалавра  
за освітньо-професійною програмою «Математичні методи  
моделювання, розпізнавання образів та безпеки даних»  
спеціальності 113 «Прикладна математика»

на тему: Інформаційна система для знайомств та спілкування, заснована на  
технології розпізнавання голосу та обличь

Виконав: здобувач вищої освіти IV курсу, групи ФІ-82  
(шифр групи)

\_\_\_\_\_ Захарченко Леонід Максимович \_\_\_\_\_  
(прізвище, ім'я, по батькові) (підпис)

Керівник д.т.н., проф. Шелестов Андрій Юрійович \_\_\_\_\_  
(посада, науковий ступінь, вчене звання, прізвище, ім'я, по батькові) (підпис)

Рецензент \_\_\_\_\_  
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище, ім'я, по батькові) (підпис)

Засвідчую, що у цій дипломній роботі  
немає запозичень з праць інших авторів  
без відповідних посилань.

Здобувач вищої освіти \_\_\_\_\_  
(підпис)

Київ – 2022 року

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

**НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ  
Кафедра математичного моделювання та аналізу даних**

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 113 «Прикладна математика»

Освітня програма «Математичні методи моделювання, розпізнавання  
образів та безпеки даних»

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_ Наталія КУССУЛЬ

«\_\_» \_\_\_\_\_ 2022 р.

**ЗАВДАННЯ**

**на дипломну роботу**

Студент: Захарченко Леонід Максимович

1. Тема роботи «Інформаційна система для знайомств та спілкування, заснована на технології розпізнавання голосу та обличчя», керівник роботи д.т.н., професор Шелестов А.Ю., затверджені наказом по університету № від «\_\_» \_\_\_\_\_ 2022 р.

2. Термін подання студентом роботи «14» червня 2022 р.

3. Вихідні дані до роботи – статті з відкритих джерел на тему метод та алгоритмів розпізнавання образів, наукова література, дані соц.опитувань, відкриті статистичні дані тощо.

4. Зміст роботи: пошук та огляд базових відомостей з розпізнавання образів, огляд методологій, їх класифікація та вибір найоптимальніших з них, вдосконалення вже готової нейромережі за результатами огляду, огляд проекту програмного рішення на базі досліджуваної нейромережі, аналіз

отриманих статистичних даних, огляд конкурентів, прогнозування, створення готової моделі проєктованого додатку у середовищі Figma.

5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо): презентація доповіді, відеодемонстраційна модель додатку

6. Дата видачі завдання: 01.01.2021

#### Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів дипломної роботи	Примітка
1	Отримання завдання	01.01.2022	виконано
2	Огляд та опрацювання інформаційних джерел	15.01.2022 – 30.01.2022	виконано
3	Аналіз базових відомостей з розпізнавання образів та підбір методів	15.02.2022 – 01.03.2022	виконано
4	Аналіз відкритих статистичних даних з ринків дейтингу	06.03.2022 – 20.03.2022	виконано
5	Прогнозування моделі розвитку додатку, що реалізовує засоби нейромережі	21.03.2022 – 05.04.2022	виконано
6	Дослідження рішень конкурентів	06.04.2022 – 16.04.2022	виконано
7	Проведення опитування, збір та аналіз отриманих даних	17.04.2022 – 30.04.2022	виконано
8	Практична побудова моделі додатку у середовищі фігма, узагальнення інформації	01.05.2022 – 28.05.2022	виконано

Здобувач вищої освіти

\_\_\_\_\_

(підпис)

Леонід ЗАХАРЧЕНКО

(Власне ім'я, ПРІЗВИЩЕ)

Керівник роботи

\_\_\_\_\_

(підпис)

Андрій ШЕЛЕСТОВ

(Власне ім'я, ПРІЗВИЩЕ)

## РЕФЕРАТ

Обсяг роботи 31 сторінка, 8 ілюстрацій, 2 таблиці, 1 графік, 1 додаток, 5 джерел літератури.

Суб'єктом дослідження даної праці є актуальна нині проблематика додатків для знайомств, за якою криється більш глобальна проблема підвищеного загального рівня стресу та усамітнення серед людей, рішенням якої є створення додатку на базі нейромережі, що вміє розпізнавати голос та обличчя.

Метою дипломної роботи є розробка та вдосконалення нейромережі, що вмітиме розпізнавати зовнішність користувачів та аналізувати її для авторизації, голос, а саме голосові команди, психо-емоційний аналіз за рядом характеристик, та подальша інтеграція і синхронізація з додатком.

Завданням дипломної роботи є аналіз наявних методів розпізнавання оптичних образів, перевірка на ефективність; безпосередня інтеграція і вдосконалення вже створеної нейромережі за наявними методами та з застосування існуючих OpenCV бібліотек; дослідження і аналітика метрик та рішень вже існуючих конкурентних додатків, розробка плану розвитку та створення практичної демонстраційної моделі додатку, у який безпосередньо інтегрується нейромережа.

Ключові слова: дейтинг, нейромережа, розпізнавання образів, аналіз даних, комп'ютерний зір.

## **ABSTRACT**

The volume of work is 31 pages, 13 illustrations, 1 table, 2 graphics, 1 diagram, 1 appendices, 5 sources of literature.

The subject of this study is the current issue of dating apps, which is a more global problem of increased overall stress and loneliness among people, the solution of which is to create an application based on a neural network that can recognize voice and face.

The aim of the thesis is to develop and improve the neural network that will be able to recognize the appearance of users and analyze it for authorization, voice, voice commands, psycho-emotional analysis of some number of characteristics, and further integration and synchronization with the application.

The task of the thesis is the analysis of existing methods of optical image recognition, testing for efficiency; direct integration and improvement of the already created neural network using existing methods and the use of existing OpenCV libraries; research and analysis of metrics and solutions of existing competitive applications, development of a development plan and creation of a practical demonstration model of the application, which directly integrates the neural network.

Keywords: dating, neural network, pattern recognition, data analysis, computer vision.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	7
ВСТУП .....	9
1 ОСНОВНІ ВІДОМОСТІ ПРО РОЗПІЗНАВАННЯ ОБРАЗІВ ТА ОГЛЯД МЕТОДОЛОГІЙ.....	11
2. Огляд методів та алгоритмів.....	13
2.1. Частина перша: Фільтрація .....	13
2.2. Частина друга: Логічна обробка результатів фільтрації.....	18
2.3. Частина третя: Дослідження .....	22
3. Практичне застосування.....	25
3.1. Аналітика .....	26
3.2. Ринок та монетизація.....	27
3.3. Конкуренти .....	27
3.4. Прогнози .....	28
Висновки .....	29
Перелік джерел посилань .....	30
Додаток.....	31

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

Нейронна мережа – математична модель, а також її програмне або апаратне втілення, побудована за принципом організації та функціонування біологічних нейронних мереж – мереж нервових клітин живого організму. Це поняття виникло при вивченні процесів, що протікають в мозку, і при спробі змоделювати ці процеси.

Дейтинг – онлайн-додаток або сайт для знайомств та спілкування. Відомі приклади: Tinder, Badoo, Mamba, Grindr.

Комп'ютерний зір — це область штучного інтелекту (ШІ), яка дає змогу комп'ютерам і системам отримувати значущу інформацію з цифрових зображень, відео та інших візуальних даних — і виконувати дії або давати рекомендації на основі цієї інформації. Якщо ШІ дозволяє комп'ютерам думати, то комп'ютерний зір дозволяє їм бачити, спостерігати і розуміти.

Розпізнавання образів – це метод аналізу даних, який використовує алгоритми машинного навчання для автоматичного розпізнавання образів і закономірностей у даних. Ці дані можуть бути будь-якими: від тексту та зображень до звуків чи інших визначених якостей. Системи розпізнавання образів можуть швидко і точно розпізнавати знайомі шаблони. Вони також можуть розпізнавати та класифікувати незнайомі об'єкти, розпізнавати форми та об'єкти під різними кутами, а також ідентифікувати об'єкти, навіть якщо вони частково затемнені.

Процес бінаризації – це переведення кольорового (або в градаціях сірого) зображення у двоколірне чорно-біле. Головним параметром такого перетворення є поріг  $t$  - значення, з яким порівнюється яскравість кожного пікселя. За результатами порівняння, пікселю надається значення 0 або 1.

OpenCV (Open Source Computer Vision Library) — бібліотека комп'ютерного бачення та машинного навчання з відкритим вихідним кодом. OpenCV створено для забезпечення загальної інфраструктури для додатків

комп'ютерного зору та для прискорення використання машинного навчання в комерційних продуктах.

Класифікація - система розподілу предметів, явищ або понять на класи, групи тощо за спільними ознаками, властивостями.



## ВСТУП

Суб'єктом дослідження даної праці є актуальна нині проблематика додатків для знайомств, за якою криється більш глобальна проблема високого рівня стресу та усамітнення серед людей.

Різноманіття подібних додатків не тільки підвищило шанси знайти потрібну людину, але й сформувало неправильне сприйняття анкети потенційного партнера. Сучасні додатки для знайомств спонукають користувачів оцінювати потенціальних партнерів тільки за зовнішністю, а алгоритми підбору анкет не відповідають запитам користувачів, як і більшість додатків для знайомств загалом, так як велика кількість з них оперують анкетами ботів, нав'язують зайві сервіси, та в цілому викликають у людей відразу та віднімають бажання користуватись такими продуктами.

Ще більш актуальною дана проблематика стала на тлі сьогоденних подій, внаслідок яких велика кількість людей втратила щонайменше дім над головою, щонайбільше – рідних, близьких та друзів. Майже третина українців потрапили в умови вимушеного переселення, велика доля з яких опинилась закордоном. Хтось, маючи більші навички спілкування знаходить контакти волонтерів, організовує побут, а інтроверти та люди з тяжкими психологічними проблемами здебільшого здаються на волю долі, так і не знайшовши підходящої підтримки.

Метою роботи є розробка та вдосконалення нейромережі, що вмітиме розпізнавати зовнішність користувачів та аналізувати її для авторизації, голос, а саме голосові команди, психо-емоційний аналіз за рядом характеристик, та подальша інтеграція і синхронізація з додатком.

Завданням дипломної роботи є аналіз наявних методів розпізнавання оптичних образів, перевірка на ефективність; безпосередня інтеграція і вдосконалення вже створеної нейромережі за наявними методами та з застосування існуючих OpenCV бібліотек; дослідження і аналітика метрик та

рішень вже існуючих конкурентних додатків, розробка плану розвитку та створення практичної демонстраційної моделі додатку, у який безпосередньо інтегрується нейромережа.

# 1 ОСНОВНІ ВІДОМОСТІ ПРО РОЗПІЗНАВАННЯ ОБРАЗІВ ТА ОГЛЯД МЕТОДОЛОГІЙ

Почнемо з базових принципів, які потрібно розуміти для ознайомлення та початку роботи з візуальним розпізнаванням.

- При вирішенні задачі завжди йти від найпростішого. Набагато простіше повісити на персону маркер оранжевого кольору, ніж слідкувати за людиною, виділяючи її каскадами. Набагато простіше взяти камеру з більшою роздільною здатністю, ніж розробляти алгоритм з надроздільною здатністю.

- Суворі постановки завдання у методах оптичного розпізнавання на порядки важливіше, ніж у задачах системного програмування: одне зайве слово у ТЗ може додати 50% роботи.

- У завдання розпізнавання немає універсальних рішень. Практично неможливо зробити алгоритм, який просто «розпізнаватиме будь-який напис». Табличка на вулиці та аркуш з текстом – це принципово різні об'єкти. Можна створити загальний алгоритм, але це вимагатиме колосальної праці великої команди і матиме структуру з десятків різних підпрограм, один з найдолучніших прикладів - StreetView від Google, алгоритми якого навчилися розпізнавати номери будинків.

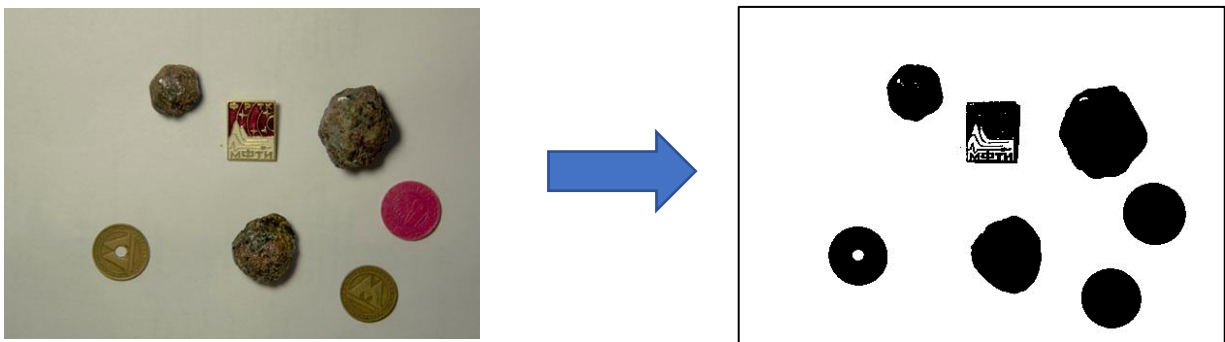
Дуже складно створити якусь структуру, або обрати готове рішення, навколо якого можна будувати рішення довільних завдань комп'ютерного зору. Умовний OpenCV - це біблія розпізнавання, в якій є безліч методів, і за допомогою якої можна вирішити 50% від обсягу майже будь-якого завдання, але OpenCV це лише мала частина того, що в реальності можна зробити. Не варто зациклюватися на OpenCV-шаблонах, а має сенс не лінуватися і тверезо оцінювати поточне завдання щоразу з нуля. У своїй роботі я постараюсь оглянути найбільш ефективні існуючі методи розпізнавання, розбити по групах за класифікацією, знайти та застосувати найоптимальніший з них.

Перша група - попередня фільтрація та підготовка зображення. Друга група – логічна обробка результатів фільтрації. Третя група – алгоритми прийняття рішень на основі логічної обробки. Рамки між групами дуже умовні, тобто, на вирішення завдання не завжди потрібно застосовувати методи з усіх груп, буває досить двох, а інколи навіть одного.

## 2. Огляд методів та алгоритмів

### 2.1. Частина перша: Фільтрація

У цій групі я розглянув кілька методів, які дозволяють виділити на зображеннях області, що цікавлять, без їх аналізу. На рівні фільтрації аналіз зображення немає, але точки, які проходять фільтрацію, можна як області з особливими характеристиками. Найпростіше перетворення — це бінаризація зображення на порозі. Для RGB зображення та зображення у градаціях сірого порогом є значення кольору. Трапляються ідеальні завдання, в яких такого перетворення достатньо. Припустимо, потрібно автоматично виділити предмети на білому папері (рис. 1.1):



**Рисунок 1.1. Виділення предметів на білому тлі шляхом бінаризації**

Вибір порога, за яким відбувається бінаризація, багато в чому визначає процес самої бінаризації. В даному випадку зображення було бінаризоване за середнім кольором. Зазвичай бінаризація здійснюється за допомогою алгоритму, що адаптивно вибирає поріг, і таким алгоритмом легко може стати вибір медіани, або математичного сподівання, наприклад.

Класичні методи фільтрації з радіолокації та обробки сигналів можна з успіхом застосовувати у безлічі завдань Pattern Recognition. Традиційним методом у радіолокації, який майже не використовується у зображеннях у чистому вигляді, є перетворення Фур'є. Одне з небагатьох винятків, у яких використовується одномірне перетворення Фур'є — компресія зображень. Для

аналізу зображень одновимірного перетворення зазвичай не вистачає, потрібно використовувати куди більш ресурсомістке двовимірне перетворення.

$$G_{uvw} = \frac{1}{NM} \sum_{n=1}^{N-1} \sum_{m=1}^{M-1} x_{mn} e^{-2\pi j \left[ \frac{un}{M} + \frac{vm}{N} \right]}$$

Мало хто його насправді розраховує, зазвичай, куди швидше і простіше використовувати згортку області, що цікавить, з вже готовим фільтром, заточеним на високі (ФВЧ) або низькі (ФНЧ) частоти. Найпростіші приклади фільтрів, що реалізують підкреслення низьких частот (фільтр Гаусса) та високих частот (Фільтр Габора). Для кожної точки зображення вибирається вікно та перемножується з фільтром того самого розміру. Результатом такої згортки є нове значення точки. При реалізації ФНЧ та ФВЧ виходять зображення такого типу:

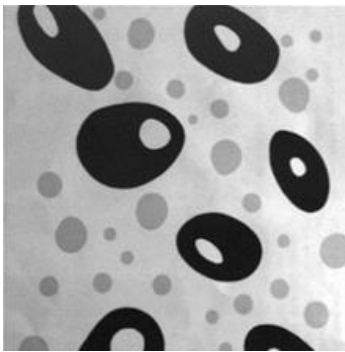


Рис.1.2. Вихідне зображення

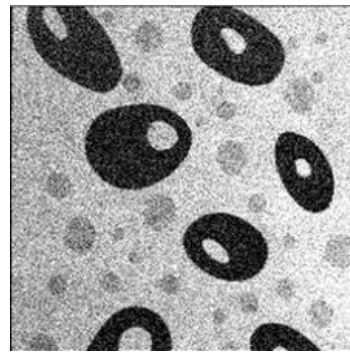


Рис.1.3. Слабко зашумлене зображення

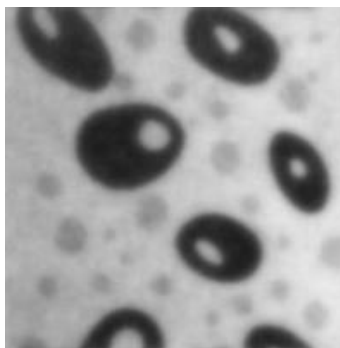


Рис.1.4. Низькочастотна фільтрація

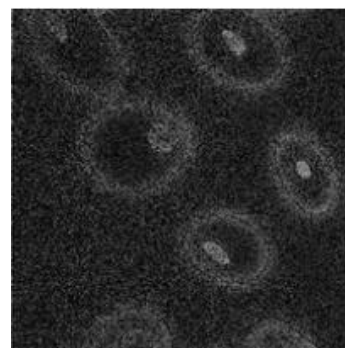


Рис.1.5. Високочастотна фільтрація

Також важливо згадати, що є випадки, коли виникає необхідність використовувати для згортки з сигналом якусь довільну характеристичну функцію. В такому разі, це називатиметься "Вейвлет-перетворення". Це визначення вейвлетів перестало бути коректним, але традиційно склалося, що у багатьох командах вейвлет-аналізом називається пошук довільного паттерна на зображенні з допомогою згортки з моделлю цього паттерна. Після такого досить вільного трактування вейвлетів, варто згадати власне кореляцію, яка лежить в їх основі. При фільтрації зображень це незамінний інструмент. Класичне застосування — кореляція відеопотоку знаходження зрушень чи оптичних потоків. Найпростіший детектор зсуву - теж у певному сенсі різницевий корелятор, тобто, там, де зображення не корелюють — був рух. Існує набір класичних функцій, що використовуються у вейвлет-аналізі. До них відносяться такі, як вейвлет Хаара (рис.1.6), вейвлет Морле (рис.1.7), вейвлет «мексиканський капелюх» (рис.1.8), та деякі інші.

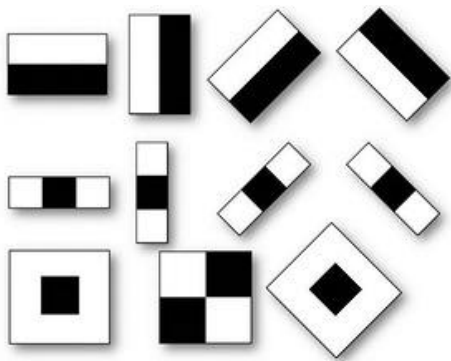


Рис.1.6. Вейвлет Хаара

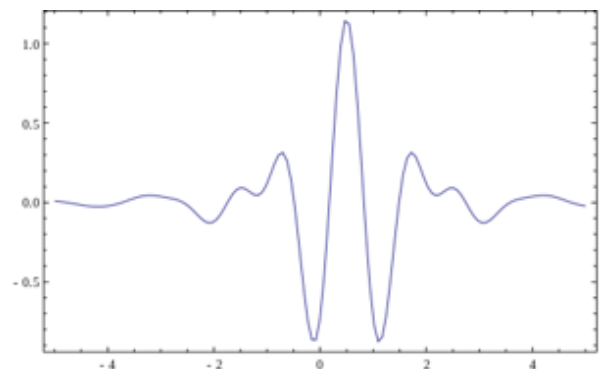


Рис.1.6. Вейвлет Морле

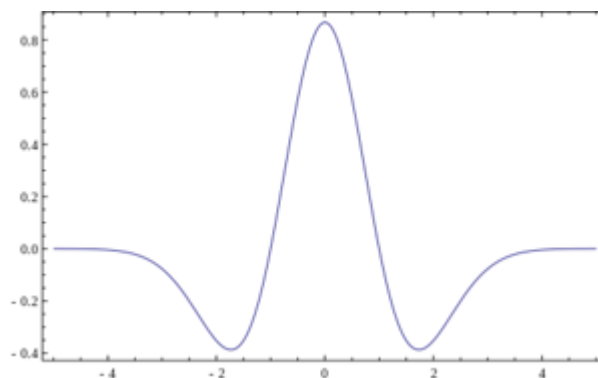


Рис.1.8. Вейвлет «мексиканський капелюх»

Хорошим прикладом використання вейвлетів є завдання пошуку відблиску в оці, на яку вейвлетом є сам відблиск (рис.1.9). Класичні вейвлети зазвичай використовуються для стиснення зображень, або їх класифікації.

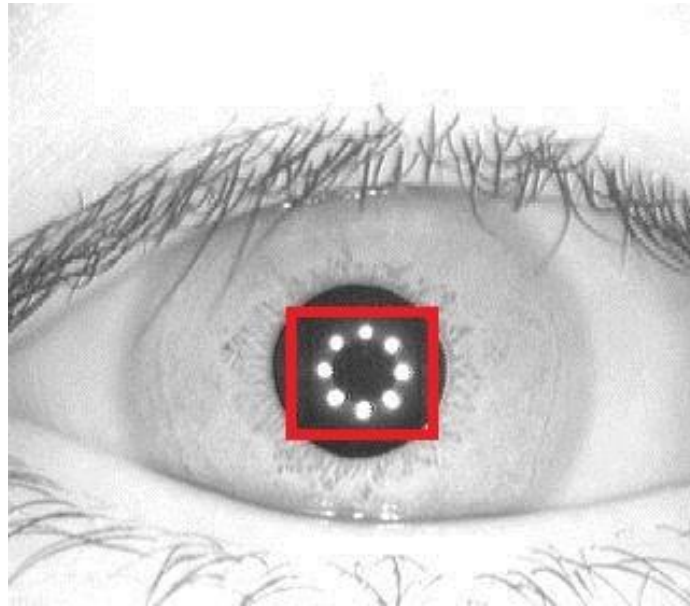


Рис.1.9. Пошук відблиску в оці

Окремий клас фільтрів - фільтрація меж та контурів. Контури корисні, коли ми хочемо перейти від роботи із зображенням до роботи з об'єктами на цьому зображенні. Коли об'єкт досить складний, але добре виділяється, часто єдиним способом роботи з ним є виділення його контурів. Існує ціла низка алгоритмів, що вирішують завдання фільтрації контурів, такі як Оператор Кенні, Оператор Собеля, Оператор Лапласа і т.д.

Найчастіше використовується саме Кенні (рис.1.10), який добре працює і реалізація якого є в OpenCV (Собіль там теж є, але він гірше шукає контури).

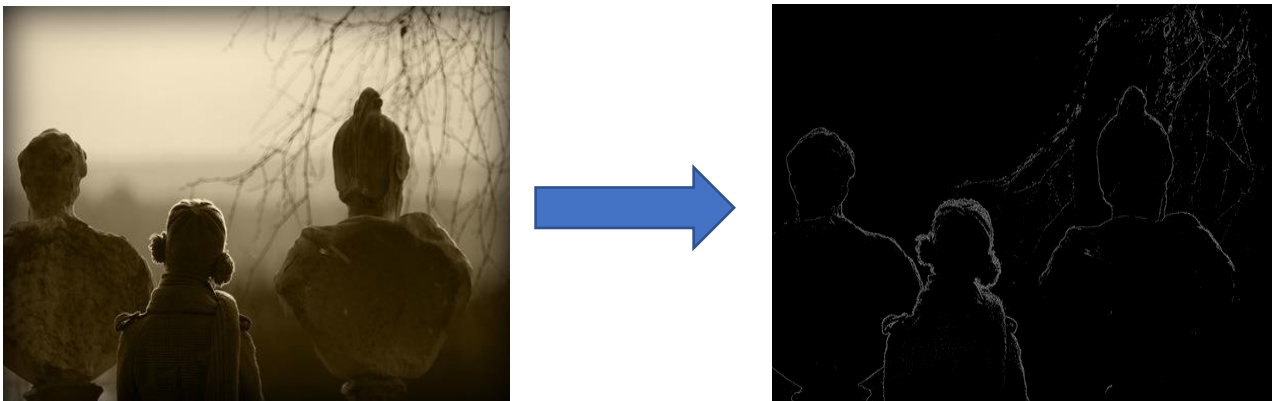


Рис.1.10. Використання оператора Кенні



## Висновки до розділу 1

В першому розділі був зроблений огляд основних відомостей про засоби та методи фільтрації вихідних зображень, їх використання і процес вибору найоптимальнішого з них.

Після проведеного огляду було зроблено висновок, що базові методи фільтрації, такі як бінаризація, або радіолокація вже застаріли, тому було прийнято рішення використовувати для фільтрації методи фільтрації по вейвлетам, так як вони більше підходять для детального аналізу маленьких деталей, яких на обличчі вдосталь, та фільтрація меж і контурів.

Наступний розділ буде присвячено опису методів, які допоможуть в обробці результатів фільтрації, як першого етапу.

## 2.2. Частина друга: Логічна обробка результатів фільтрації

Фільтрування дає набір придатних для обробки даних, але часто не можна просто взяти та використовувати ці дані без їх обробки. У цьому розділі буде кілька класичних методів, що дозволяють перейти від зображення до властивостей об'єктів, або самих об'єктів.

Переходом від фільтрації до логіки є методи математичної морфології. По суті, це найпростіші операції нарощування та ерозії бінарних зображень. Ці методи дозволяють прибрати шуми з бінарного зображення, збільшивши або зменшивши наявні елементи. На основі математичної морфології існують алгоритми оконтурювання, але зазвичай користуються якимись гібридними алгоритмами чи алгоритмами зв'язки.

У розділі фільтрації згадувалися алгоритми отримання кордонів. Отримані межі досить легко перетворюються на контури. Для алгоритму Кенні це відбувається автоматично, для інших алгоритмів потрібна додаткова бінаризація. Контур є унікальною характеристикою об'єкта. Часто це дозволяє ідентифікувати об'єкт за контуром. Існує потужний математичний апарат, що дозволяє це зробити, і називається він контурним аналізом. Отримати контур для бінарного алгоритму можна, наприклад, алгоритмом жука (рис.1.11).

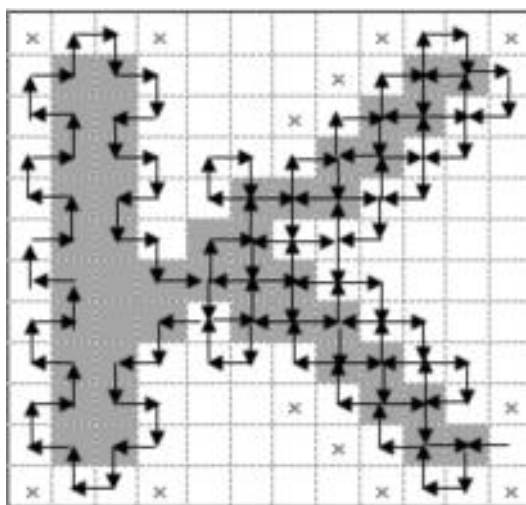


Рис.1.11. Демонстрація алгоритму Жука

"Жук" починає рух з білої області у напрямку до чорної. Як тільки він потрапляє на чорний елемент, він повертає ліворуч і переходить до наступного елемента. Якщо цей елемент білий, то жук повертається праворуч, інакше - ліворуч. Процедура повторюється до того часу, поки жук повернеться у вихідну точку.

Контурний аналіз чудовий варіант у разі, якщо потрібно щось розпізнавати в ідеальних умовах, в інших випадках зачасту або межа не знайдеться, або шумів занадто багато для роботи з зображенням.

Ще один варіант - оброблення продуктів фільтрації за особливими точками. Особливі точки - це деякі унікальні характеристики об'єкта, які дозволяють співставляти об'єкт з самим собою, або зі схожими класами об'єктів. Існує кілька десятків способів, що дозволяють виділити такі точки. Деякі способи виділяють особливі точки в сусідніх кадрах, деякі через великий проміжок часу та при зміні освітлення, деякі дозволяють знайти особливі точки, які залишаються такими навіть при поворотах об'єкта. Почнемо з методів, що дозволяють знайти особливі точки, які не такі стабільні, проте швидко розраховуються, а потім підемо за зростанням складності:

Перший клас - це особливі точки, які є стабільними протягом секунд. Такі точки служать для того, щоб вести об'єкт між сусідніми кадрами відео, або для зведення зображення з сусідніх камер. До таких точок можна віднести локальні максимуми зображення, кути на зображенні (кращий детектор, мабуть, детектор Харіса), точки в яких досягається максимуми дисперсії, певні градієнти і.т.д.

Другий клас. Особливі точки, що є стабільними при зміні освітлення та невеликих рухах об'єкта. Такі точки служать насамперед на навчання і наступної класифікації типів об'єктів. Наприклад, класифікатор пішохода чи класифікатор особи — це продукт системи, побудованої саме у таких точках. Деякі з згаданих вейвлетів можуть бути базою для таких точок. Наприклад, примітиви Хаара, пошук відблисків, пошук інших специфічних функцій. До

таких точок відносяться точки, знайдені методом гістограм спрямованих градієнтів (HOG).

Третій клас - стабільні точки. Мені вдалося виділити лише два методи, які дають повну стабільність та їх модифікації. Це SURF та SIFT. Вони дозволяють знаходити спеціальні точки навіть при повороті зображення. Розрахунок таких точок здійснюється довше, ніж іншими методами, але все ж, досить обмежений час. Проте, на жаль, ці методи запатентовані, тож їх використання потребує отримання особливих дозволів, тощо.

## Висновки до розділу 2

В другому розділі був зроблений огляд основних відомостей про засоби та методи обробки вихідних зображень після фільтрації, їх використання і процес вибору найоптимальнішого з них.

Після проведеного огляду було зроблено висновок, що найбільш ефективним серед розглянутих засобів є класифікація за стабільними точками, але зважаючи на авторське право, патенти і т.д. використати його змоги немає, тому, певне, найдоцільнішим варіантом у даному випадку стане алгоритм контурування, так як у розпізнаванні обличь контур дійсно грає одну з вирішальних ролей.

Наступний, останній теоретичний розділ буде присвячено опису методів навчання, які допоможуть в постановці задачі створення моделі та прийнятті рішень.

### 2.3. Частина третя: Дослідження

У 80% ситуацій суть навчання у задачі розпізнавання в наступному: Є тестова вибірка, де є кілька класів об'єктів. У випадку програми, що розробляється, розглянемо приклад де буде досліджуватися наявність/відсутність особи людини на фотографії. Для кожного зображення є набір ознак, які були виділені якоюсь ознакою. Алгоритм навчання повинен побудувати таку модель, за якою він зможе проаналізувати нове зображення і прийняти рішення, який об'єкт є на зображенні.

Як це робиться? Кожне з тестових зображень це точка в просторі ознак. Її координати - це вага кожної з ознак на зображенні. Нехай нашими ознаками будуть: «Наявність очей», «Наявність носа», «Наявність вух», і т.д... Всі ці ознаки ми виділимо детекторами, що існують у нас, які навчені на частини тіла, схожі на людські. Для людини у такому просторі буде коректною точка [1;1;1;1;...]. Для мавпи точка [1; 0; 1; 0 ...] для коня [1; 0; 0; 0 ...]. Класифікатор навчається з вибірки прикладів. Але на всіх фотографіях чітко виділилося обличчя, інших немає очей, але в третьому, наприклад, у мавпи через помилки класифікатора з'явився людський ніс. Класифікатор людини, що навчається, автоматично розбиває простір ознак таким чином, щоб сказати: якщо перша ознака лежить в діапазоні  $0.5 < x < 1$ , другий  $0.7 < y < 1$ , і т.д., тоді це людина.

Розберемо на прикладі найпростіший випадок класифікації, коли простір ознак одновимірний, а нам потрібно розділити 2 класи. Ситуація зустрічається частіше, ніж може бути: наприклад, коли потрібно відрізнити два сигнали, або порівняти патерн із зразком. Нехай у нас є навчальна вибірка. При цьому виходить зображення, де по осі X буде міра схожості, а по осі Y - кількість подій з таким заходом. Коли об'єкт, що шукається, схожий на себе — виходить ліва гауссіана. Коли не схожий, права. Значення  $X=0.4$  розділяє вибірки так, що помилкове рішення мінімізує можливість прийняття будь-якого неправильного рішення. Саме пошуком такого роздільника є завдання класифікації.

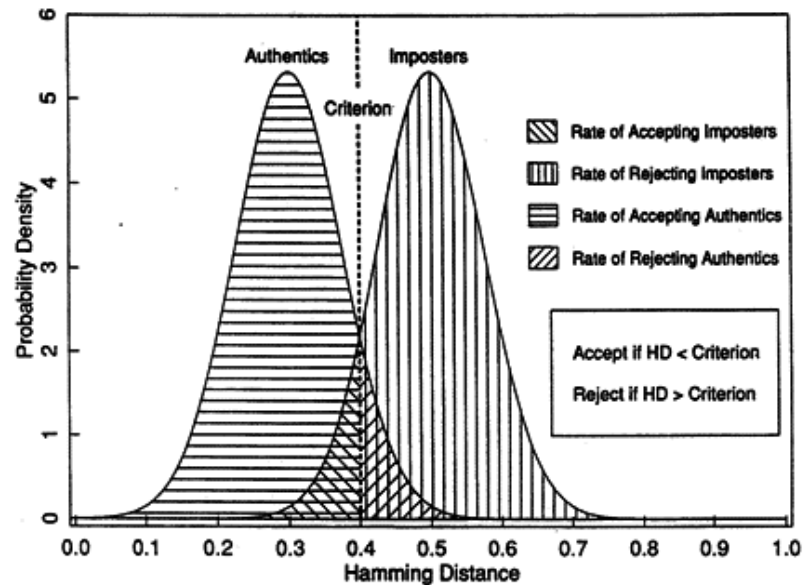


Рис.1.12. Демонстрація порівняння патернів

Також, розглянемо випадки, в яких простір є багатомірним. Для таких випадків є величезна купа алгоритмів, але ознайомимось із ними на прикладі декількох основних.

- k-means - один із найпростіших алгоритмів навчання. Звичайно, він використовується переважно для задач кластеризації, але й навчити через нього теж можна. Працює у ситуації, коли групи об'єктів мають непогано рознесений центр мас і мають великого перетину.
- АдаБуст - один із найпоширеніших класифікаторів. Наприклад, каскад Хаара побудований саме на ньому. Зазвичай використовують коли потрібна бінарна класифікація, але нічого не заважає навчити більшу кількість класів.
- SVM - один із найпотужніших класифікаторів, що має безліч реалізацій. Працює майже аналогічно Адабусти, вважається досить швидким, але його навчання складніше, ніж у Адабусти, а також він потребує вибору правильного ядра.

### **Висновки до розділу 3**

У третьому розділі був зроблений огляд основних методів навчання, які допомагають в постановці задачі створення моделі та прийнятті рішень, виділили основні із них, ознайомились з принципом дії, та от до яких висновків прийшли.

Так як аналіз обличь відбувається у двох- або трьохмірних просторах, для реалізації у нейромережі було обрано для використання алгоритми k-means та Адабуст як найбільш доцільні у заданому випадку.



### 3. Практичне застосування

**Проект Veetok** – анонімний додаток для знайомств та спілкування, заснований на досліджуваній мережі, що здатна до розпізнавання голосу, та психо-емоційного аналізу спілкування користувачів. Додаток переосмислює аналіз спілкування та підбору анкет, а комбінація аватару, нейронної мережі та голосу дозволить сконцентруватись на особистості людини, виносячи зовнішність взагалі на інший план.

Серед попередніх планів щодо функціоналу – додаток передбачатиме окреме існування дейтингової складової, голосових та ігрових кімнат, графічного редактору профілю та аватарів, внутрішнього магазину, служби підтримки тощо.

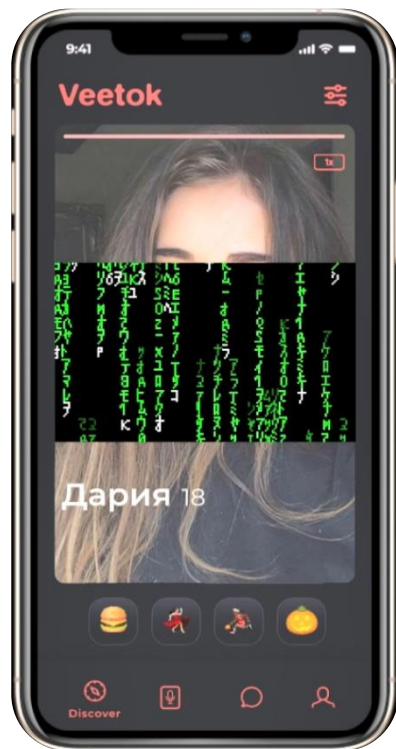


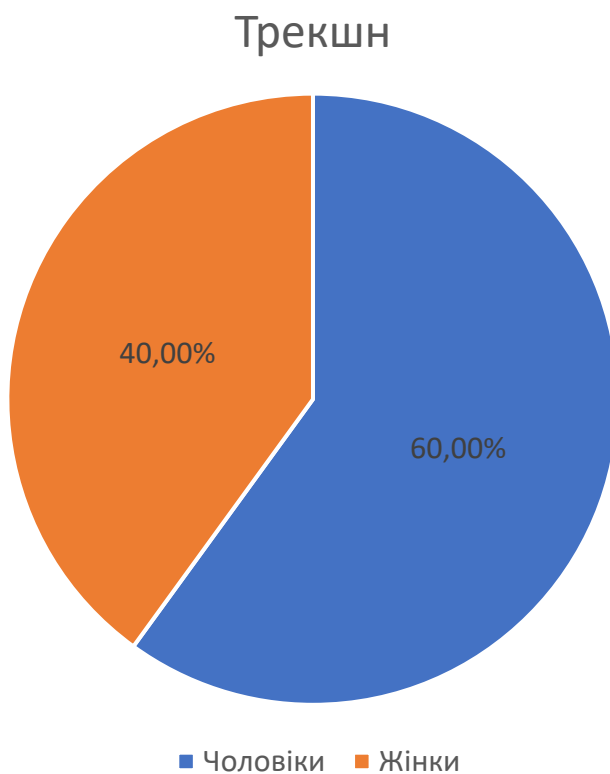
Рис.1.13. Демонстрація головної сторінки на моделі додатку



Рис.1.14. Демонстрація сторінки діалогу на моделі додатку

### 3.1. Аналітика

Було проведено опитування серед ряду середньостатичних користувачів подібних додатків, та встановлено загальну картину цільової аудиторії. Здебільшого, це чоловіки та жінки віком від 18 до 30 років, студенти або офісні співробітники. Також, це люди, що ведуть активний образ життя, та хоча б раз користувались онлайн-дейтингом. Серед 200 опитаних 60% - чоловіки, 40% - жінки, 178 з яких були внесені в список очікування. Зокрема, дослідження показало, що такі люди люблять ігри та слухають подкасти.



Діаграма 1. Демонстрація  
трекшну

### 3.2. Ринок та монетизація

За останніми даними, об'єм всесвітнього ринку мобільних дейтингів за 21-й рік становить 3.08 млрд. доларів, об'єм російськомовного ринку – трохи більше 200-та мільйонів доларів. Потенційно, додаток зможе претендувати на 1-2%. Монетизацію планується реалізувати за рахунок нативної аудіореклами, декількох планів підписки з додатковим функціоналом, та внутрішнього магазину подарунків. Приблизну Roadmapу буде додано за потреби.

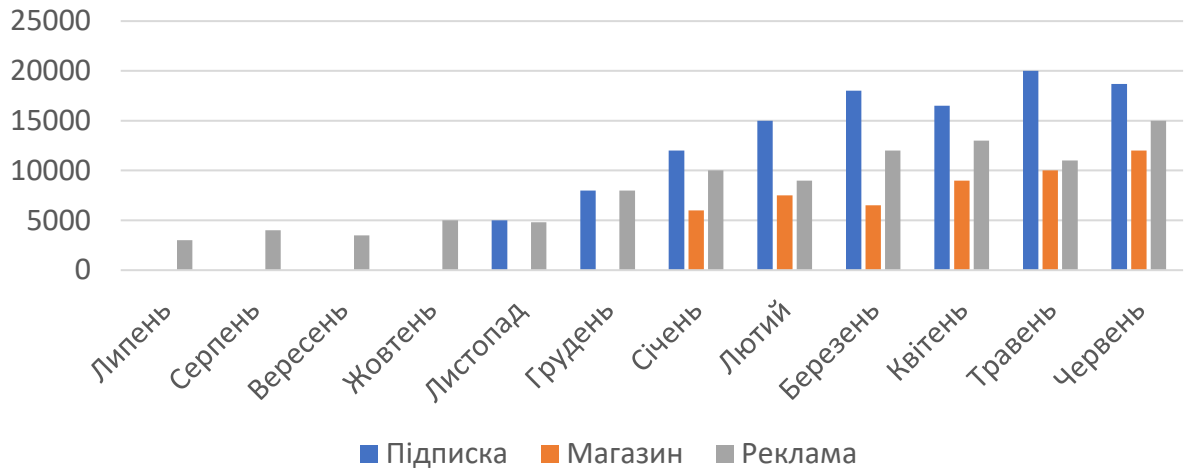
### 3.3. Конкуренти

Окрім отриманої загальної картинки ринку, також були враховані конкуренти з російсько- та україномовного сегменту, а саме:

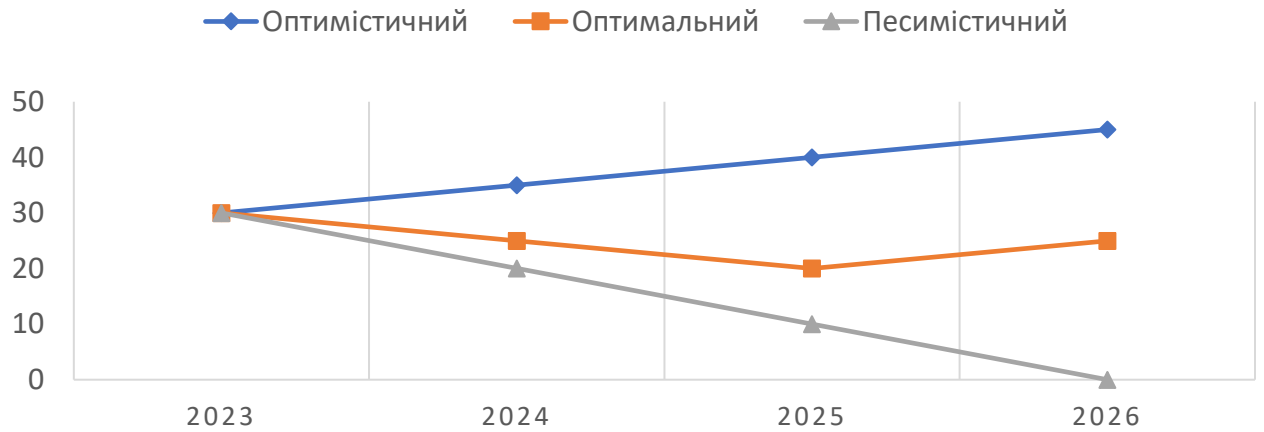
Назва	Рік запуску	Інвестиції	Функціонал		
Vox	2020	3000000\$	Голосові анкети з заблюреною фотографією	Текстові чати та дзвінки. Відсутність голосових чатів	Елементи соц.мережі винесені до окремого додатку
TalkNow	2020	100000\$	Голосові анкети з фотографією	Голосові чати. Відсутність текстових чатів та дзвінків	Групові голосові чати

Табл.1. Порівняльна таблиця потенційних конкурентів

### 3.4. Прогнози



Графік 1. Прибуток після релізу 2022-2023(в грн)



Графік 2. Сценарій фінансового розвитку

## Висновки

Метою даної роботи було оглянути та застосувати ефективні алгоритми та методи для вдосконалення нейромережі, що вміє розпізнавати голос та зовнішність, а також розробка та подальше втілення у життя практичної реалізації, в яку буде інтегрована дана мережа.

Проведений аналіз методів виявив фаворитів та показав, з якими недоліками можливо зіштовхнутись при обранні того чи іншого, недоцільного варіанту. Також, за висновками проведеного аналізу ринку та метрик ринку додатків для знайомств було встановлено, що на даний момент аналоги з близьким функціоналом нехай і існують, але мають за базис не дуже вдалий концепт, або обмежені у своїх можливостях, таким чином даючи змогу розвинути та зайняти свою нішу проектованому додатку.

Також, на практиці було продемонстровано створену власноруч модель готового додатку, що дало можливість розвинути свої навички візуального програмування та показати потенційні можливості додатку.

Розглянута у дипломі тематика є дуже актуальною, та потребує подальшого дослідження, наприклад, на предмет появи нових, більш ефективних методів розпізнавання, виявлення подальших пріоритетів, аналітики ринку в нових умовах тощо.

## Перелік джерел посилань

1. <https://habr.com/ru/post/208090/>
2. <https://habr.com/ru/company/droider/blog/568764/>
3. <https://www.ibm.com/topics/computer-vision>
4. <https://www.arm.com/glossary/pattern-recognition>
5. <https://robocraft.ru/computervision/1061>

## Додаток

Демонстративна анімація того, як виглядатиме та працюватиме даний продукт. (для перегляду клікнути двічі)



veetok\_tutor.mp4