Вінницький національний технічний університет

Факультет менеджменту та інформаційної безпеки

Кафедра менеджменту та безпеки інформаційних систем

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

"Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-team memory (LSTM) для поєднання статичних і динамічних ознак"

Виконав: ст. 2-го курсу, групи КІТС-23м
спеціальності 125– Кібербезпека
Освітня програма – Кібербезпека
інформаційних технологій та систем
(шифр і назва напряму підготовки, спеціальності)

Кухарук П.О.
(прізвище та ініціали)

Керівник: к.ф.м.н., доц. каф. МБІС
Шиян А.А.
(прізвище та ініціали)
«____» _____ 2024 р.

Опонент: к.т.н., професор, каф. ОТ,
Захарченко С.М.
(прізвище та ініціали)
«____» _____ 2024 р.

Допущено до захисту
Голова секції УБ кафедри МБІС

Юрій ЯРЕМЧУК
"09" грудня 2024 р

Вінниця ВНТУ - 2024 рік

Вінницький національний технічний університет
Факультет менеджменту та інформаційної безпеки
Кафедра менеджменту та безпеки інформаційних систем

Рівень вищої освіти ІІ-й (магістерський)
Галузь знань 12 – Інформаційні технології
Спеціальність 125 – Кібербезпека та захист інформації
Освітньо-професійна програма - Кібербезпека інформаційних технологій та систем

ЗАТВЕРДЖУЮ
Голова секції УБ/ кафедра МБІС

_____ **Юрій ЯРЕМЧУК**
"20" вересня 2024 р.

З А В Д А Н Н Я
на магістерську кваліфікаційну роботу студенту

Кухаруку П.О.
(прізвище, ім`я, по-батькові)

1. Тема роботи : " Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-team memory (LSTM) для поєднання статичних і динамічних ознак "

Керівник роботи_____ к.ф.м.н., доц. каф. МБІС _Шиян А.А.__
(прізвище, ім`я, по-батькові, науковий ступінь, вчене звання)
затверджені наказом вищого навчального закладу від "17" вересня 2024 року № 310

2. Строк подання студентом роботи_____

3. Вихідні дані до роботи: Електронні джерела, наукові статті, книги та документи, пов'язані з темою. Існуюче програмне забезпечення для аналізу біометричних даних, інструменти для машинного навчання, технічна документація, вимоги та обмеження до програмного забезпечення.

4. Зміст текстової частини: Для досягнення мети роботи були поставлені наступні задачі: дослідити актуальність загроз для систем біометричної аутентифікації мереж та визначити існуючі вразливості; проаналізувати методи, що використовуються для поєднання статичних і динамічних ознак у багатофакторній біометричній аутентифікації; дослідити можливості використання convolutional neural networks (CNN) та long short-term memory (LSTM) для покращення процесу біометричної аутентифікації; розробити вдосконалений алгоритм для поєднання статичних і динамічних ознак з використанням CNN та LSTM; реалізувати програмний засіб для аналізу біометричних даних та проведення аутентифікації з використанням Python та відповідних бібліотек для глибокого навчання.

5. Перелік ілюстративного матеріалу (з точним зазначенням обов`язкових креслень)

Зображення поетапної реалізації та налаштування веб-додатку
6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада Консультанта	Підпис, дата	
		завдання видав	завдання прийняв
I	к.ф-м.н., доц., кафедри МБІС Шиян А.А		
II	к.ф-м.н., доц., кафедри МБІС Шиян А.А		
III	к.ф-м.н., доц., кафедри МБІС Шиян А.А		
Економічна частина			
IV	Бурєннікова Н. В. д.е.н., професор каф. ЕПВМ		

7.Дата видачі завдання 20 вересня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів магістерської кваліфікаційної роботи	Строк виконання етапів роботи		Примітка
1.	Визначення напрямку магістерської роботи, формулювання теми	20.09.2024	31.09.2024	
2.	Аналіз предметної області обраної теми	01.10.2024	15.10.2024	
3.	Розробка роботи	16.10.2024	26.10.2024	
4.	Написання магістерської роботи на основі розробленої теми	27.10.2024	15.11.2024	
5.	Передзахист магістерської кваліфікаційної роботи	16.11.2024	24.11.2024	
6.	Виправлення, уточнення, корегування магістерської кваліфікаційної роботи	27.11.2024	04.12.2024	
7.	Захист магістерської кваліфікаційної роботи	10.12.2024	17.12.2024	

Студент _____ _____
(підпис)

Керівник роботи _____ _____

АНОТАЦІЯ

УДК 621.374.415

Кухарук П.О. Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-term memory (LSTM) для поєднання статичних і динамічних ознак. Магістерська кваліфікаційна робота зі спеціальності 125 – Кібербезпека, освітня програма – Кібербезпека інформаційних технологій та систем. Вінниця: ВНТУ, 2024. 95 с.

На укр. мові. Бібліогр.: 35 назв; рис.: 23; табл. 2.

У магістерській кваліфікаційній роботі розроблено програму для удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-term memory (LSTM) для поєднання статичних і динамічних ознак. У загальній частині роботи розглянуто сучасні підходи до здійснення біометричної аутентифікації, проведено аналіз методів біометричної аутентифікації, приведено аналіз існуючих реалізацій методів біометричної ідентифікації, приведені аргументи для обгрунтування необхідності удосконалення багатофакторної біометричної аутентифікації, представлено алгоритм розробки програмного забезпечення для удосконалення моделі біометричної аутентифікації. У технологічній частині розроблений програмний засіб для підвищення удосконалення моделі біометричної аутентифікації на основі використання CNN та LSTM.

В економічному розділі обґрунтовано доцільність створення даного програмного засобу.

Ключові слова: біометрична аутентифікація, захист персональних даних, CNN, LSTM, Python.

ABSTRACT

Kukharuk P. O. improvement of multi-factor biometric authentication based on convolution neural networks (CNN) and long short-term memory (LSTM) for combining static and dynamic features. Bachelor's thesis in specialty 125 - Cybersecurity, educational program-cybersecurity of information technologies and systems. Vinnitsa: VNTU, 2024. – 95 p.

In Ukrainian language. Bibliographer: 35 titles; fig.: 23; tabl. 2.

In the master's qualification work, a program was developed to improve multi-factor biometric authentication based on convolution neural networks (CNN) and long short-term memory (LSTM) for combining static and dynamic features. In the general part of the work, modern approaches to the implementation of biometric authentication are considered, biometric authentication methods are analyzed, existing implementations of biometric identification methods are analyzed, arguments are given to justify the need to improve multi-factor biometric authentication, and a software development algorithm for improving the biometric authentication model is presented. In the technological part, a software tool has been developed to improve the biometric authentication model based on the use of CNN and LSTM.

In the economic section, the expediency of creating this software tool is justi

Keywords: biometric authentication, personal data protection, CNN, LSTM, Python.

ЗМІСТ

Вступ 10
1. АНАЛІЗ ПРОБЛЕМИ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ 12
1.1 Сучасні підходи до реалізації біометричної аутентифікації 12
1.2 Аналіз методів біометричної аутентифікації 19
1.3 Аналіз існуючих реалізацій методів біометричної ідентифікації 23
1.4 Висновки та постановка задачі 28
2 УДОСКОНАЛЕННЯ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІ З ВИКОРИСТАННЯМ CNN ТА LSTM 29
2.1 Обґрунтування необхідності удосконалення багатофакторної біометричної аутентифікації 29
2.2 Особливості удосконалення багатофакторної біометричної аутентифікації35
2.3 Розробка алгоритму удосконалення багатофакторної біометричної аутентифікації 36
2.4 Висновки до розділу 39
3. ПРОГРАМНА РЕАЛІЗАЦІЯ УДОСКОНАЛЕННЯ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ 40
3.1 Обґрунтування вибору мови програмування 40
3.2 Опис фрагментів лістингу програми 44
3.3 Тестування програмного засобу 48
3.4 Висновки до розділу 53
4. ЕКОНОМІЧНА ЧАСТИНА 54
4.1 Оцінювання комерційного потенціалу розробки 54
4.2 Прогнозування витрат на виконання науково-дослідної роботи 58
4.3 Розрахунок економічної ефективності науково-технічної розробки 63
4.4 Розрахунок ефективності вкладених інвестицій та періоду їх окупності 64
4.5 Висновки до розділу 67
ВИСНОВКИ 68
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ 69

ДОДАТКИ..... 73

9. Стадії та етани розробки

№ з/п	Назва етапів магістерської кваліфікаційної роботи	Початок	Закінчення
1	Визначення напрямку магістерської роботи, формулювання теми		
2	Аналіз предметної області обраної теми		
3	Апробація отриманих результатів		
4	Розробка алгоритму роботи		
5	Написання магістерської роботи на основі розробленої теми		
6	Розробка економічної частини		
7	Передзахист магістерської кваліфікаційної роботи		
8	Виправлення, уточнення, корегування магістерської кваліфікаційної роботи		
9	Захист магістерської кваліфікаційної Роботи		

10. Порядок контролю та прийому

10.1 До приймання магістерської кваліфікаційної роботи надається:

– ПЗ до магістерської кваліфікаційної роботи;

– програмний додаток;

– презентація;

– відзив керівника роботи;

– відзив опонента

Технічне завдання до виконання прийняв _____ Кухарук П.О.

Додаток Б. Лістинг програми 78
Додаток В. Ілюстративний матеріал 83

Вступ

Актуальність. Сучасний розвиток технологій кібербезпеки вимагає впровадження ефективних методів захисту персональних даних, особливо з урахуванням зростання обсягів цифрової інформації та загроз, пов'язаних із компрометацією систем. Біометрична аутентифікація забезпечує високий рівень захисту, використовуючи унікальні фізичні або поведінкові характеристики для підтвердження особи. Цей підхід знаходить застосування як у фізичному, так і в цифровому контекстах, наприклад, для доступу до приміщень, пристроїв чи конфіденційної інформації.

Попри високий рівень надійності, існуючі методи біометричної аутентифікації мають недоліки, такі як можливість хибних спрацьовувань, проблеми конфіденційності, уразливість до атак типу "спуфінг" і ризики втрати або компрометації даних. Водночас прогрес у машинному навчанні, зокрема використання convolutional neural networks (CNN) та long short-term memory (LSTM), відкриває нові можливості для вдосконалення багатофакторної біометричної аутентифікації, що дозволяє комбінувати статичні та динамічні ознаки для підвищення точності та надійності систем.

Впровадження таких інновацій сприяє зменшенню ризиків шахрайства та підвищенню рівня довіри до біометричних систем. Саме тому розробка вдосконалених методів багатофакторної біометричної аутентифікації є одним із актуальних напрямків у сфері захисту персональних даних.

Мета та завдання дослідження.

Метою кваліфікаційної роботи є розробка вдосконаленого методу багатофакторної біометричної аутентифікації з використанням CNN та LSTM, що забезпечує високу точність і надійність за рахунок поєднання статичних і динамічних ознак.

Для досягнення цієї мети поставлено такі завдання:

– Розглянути основні концепції біометричної аутентифікації та підходи до її реалізації.

– Проаналізувати сучасні методи багатофакторної біометричної аутентифікації.

– Оцінити існуючі рішення у сфері біометричної ідентифікації.

– Обґрунтувати необхідність удосконалення існуючих підходів до біометричної аутентифікації.

– Розробити алгоритм багатофакторної аутентифікації на основі CNN та LSTM.

– Провести тестування запропонованого алгоритму та представити результати.

– Виконати оцінку економічної доцільності впровадження запропонованого рішення.

Об'єкт і предмет дослідження.

Об'єктом дослідження є методи біометричної багатофакторної аутентифікації.

Предметом дослідження є застосування CNN і LSTM для вдосконалення багатофакторної біометричної аутентифікації.

Наукова новизна кваліфікаційної роботи полягає у розробці вдосконаленого методу багатофакторної біометричної аутентифікації, що поєднує статичні та динамічні ознаки за допомогою CNN та LSTM. Запропонований підхід дозволяє суттєво підвищити точність і надійність аутентифікації, знижуючи ризики шахрайства та компрометації даних.

Практична цінність.

Розроблена система має широкий спектр застосувань, включаючи доступ до інформаційних ресурсів, фізичних об'єктів, сертифікаційних платформ та інших критично важливих систем. Впровадження таких технологій сприяє підвищенню рівня кібербезпеки та довіри до цифрових платформ, що є важливим у сучасному світі.

1. АНАЛІЗ ПРОБЛЕМИ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ

В даному підрозділі буде здійснено аналіз проблеми багатофакторної біометричної аутентифікації та розглянуто сучасні підходи до реалізації біометричної аутентифікації.

1.1 Сучасні підходи до реалізації біометричної аутентифікації

Біометрія – це технічний термін для позначення фізичних або поведінкових рис людини. Біометрична аутентифікація – це поняття в безпеці даних. Рішення для біометричної автентифікації створюють модель, згенеровану даними, яка представляє людину. За допомогою цієї моделі та біометричної інформації системи безпеки можуть автентифікувати доступ до додатків та інших мережевих ресурсів. Біометрична автентифікація швидко стає популярним компонентом стратегій багатофакторної автентифікації, оскільки вона поєднує в собі серйозну проблему автентифікації з мінімальним рівнем взаємодії з користувачем.

Імена користувачів і паролі були основним заходом безпеки протягом десятиліть, але не більше. Численні резонансні злами у великих фінансових і бізнес-установах призвели до того, що мільйони комбінацій імені користувача та пароля були викрадені та виставлені на продаж у Dark Web. Додати до цього тенденцію до повторення паролів для кількох облікових записів, і масштаб вразливості стане більш очевидним.

Системи біометричної аутентифікації менш схильні до цієї вразливості, оскільки біометричні дані користувача унікальні. Зловмиснику дуже складно шахрайським шляхом відтворити сканування відбитків пальців або обличчя особи, якщо вони використовуються надійними рішеннями з сильним виявленням активності/підробки, і все ж потрібен лише мить для автентифікації відповідного користувача. Через це біометрія вважається зручнішою за паролі та безпечнішою.

Розпізнавання обличь є дуже відомою формою біометричної аутентифікації, популяризованою в багатьох шпигунських драмах і науково-фантастичних казках

та у популярних засобах масової інформації. Щодня використовується розпізнавання обличь, щоб ідентифікувати друзів і родичів і відрізняти незнайомців. При аутентифікації принципи цього процесу оцифровуються, щоб смартфон або мобільний пристрій міг розпізнавати обличчя майже таким же чином.

Програмне забезпечення для розпізнавання обличь аналізує геометрію обличчя, включаючи відстань між очима, відстань між підборіддям і носом тощо, щоб створити зашифровану цифрову модель для даних про обличчя. Під час аутентифікації інструмент розпізнавання обличь скануватиме обличчя в режимі реального часу та порівнюватиме модель із тією, що зберігається в системі. Незважаючи на те, що розпізнавання обличь розвивається, все ще існують деякі ризики.

Плюси:

– Мобільні пристрої широко поширені, і більшість, якщо не всі, з них мають камеру.

– Дуже мало налаштувань. У більшості сучасних мобільних пристроїв ці можливості входять до стандартної комплектації.

– Розпізнавання обличь є одним із найзручніших способів біометричної автентифікації. Погляд у камеру пристрою пов'язаний з меншим тертям, ніж сканування відбитків пальців або код автентифікації.

Мінуси:

– Не всі системи розпізнавання обличь створені однаково. Деякі з них легше підробити, ніж інші.

– Нативні рішення не такі ефективні, як сторонні або фірмові рішення.

– Системи розпізнавання обличь з «активним визначенням жвавості» вимагають від користувача поворухнути головою, моргнути або виконати інші дії в момент для перевірки запиту. Цей процес може бути простішим для аналізу та обходу зловмисником і може створити незручний досвід користувача, тоді як «пасивне виявлення живості» відбувається за лаштунками, тому воно не заважає користувачеві та не важче ідентифікувати та зрозуміти зловмиснику.

Розпізнавання відбитків пальців

Правоохоронці роками використовують відбитки пальців як форму ідентифікації. Сканер відбитків пальців працює за тими ж принципами, але весь процес оцифрований. Відбитки пальців у кожної людини унікальні. Таким чином, аналізуючи виступи та малюнок відбитка, сканери відбитків пальців створюють цифрову модель, яку порівнюють із майбутніми спробами аутентифікації.

Плюси:

- Використовується в багатьох галузях промисловості
- Серед найбільш поширених модальностей

Мінуси:

- Продуктивність може постраждати через якість відбитка пальця або поточні умови, такі як мокрі або брудні пальці.

Всупереч поширеній думці, насправді існує два методи сканування ока з метою аутентифікації. Сканування використовує розпізнавання райдужної оболонки ока або сітківки ока для ідентифікації користувачів.

Під час сканування сітківки ока автентифікатор короткочасно світить світлом, щоб висвітлити унікальний малюнок кровоносних судин в оці. Зіставляючи цей шаблон, інструмент розпізнавання очей може порівнювати очі користувача з оригіналом. Сканування райдужної оболонки ока працює аналогічно, але воно аналізує кольорові кільця, знайдені в райдужній оболонці.

Плюси:

- У деяких реалізаціях розпізнавання очей може бути таким же швидким і точним, як і розпізнавання обличчя (хоча і менш зручним для користувача).

Мінуси:

- Може бути важко дістати зразок для порівняння, перебуваючи на сонячному світлі (зіниці звужуються).
- Залежно від реалізації, для цього може знадобитися спеціалізоване обладнання.

Розпізнавання голосу аналізує звучання голосу користувача. Унікальний голос кожної людини визначається довжиною голосового тракту та формою носа,

рота та гортані. Всі ці фактори роблять аналіз голосу користувача сильним методом аутентифікації.

Плюси:

- Пропонує зручний досвід автентифікації
- У деяких програмах передбачена фраза для користувача

Мінуси:

- Фоновий шум може спотворювати записи.
- Застуда, бронхіт або інші поширені захворювання можуть спотворити голос і порушити аутентифікацію.
- У громадських ситуаціях людина може почуватися незручно, коли говорить вголос (наприклад, у поїзді чи автобусі).

Біометрична аутентифікація використовується в широкому спектрі додатків у багатьох галузях. Ось лише кілька прикладів того, як ці галузі використовують використання біометрії для підвищення безпеки та ефективності існуючих процесів.

Деякі авіакомпанії та аеропорти пропонують своїм пасажирам можливість зареєструватися на свій рейс за допомогою розпізнавання обличчя. Подібним чином готелі та готельні компанії починають включати самостійну реєстрацію за допомогою біометричної автентифікації.

Безпека та автентифікація мають важливе значення в багатьох галузях, але особливо в мобільному банкінгу. Фінансові установи використовують біометричну автентифікацію як частину стратегії двофакторної автентифікації або багатофакторної автентифікації для захисту банку та його клієнтів від атак на захоплення рахунків.

Біометрична автентифікація у вигляді сканерів відбитків пальців, сканерів райдужної оболонки ока та розпізнавання обличчя може допомогти лікарням підтвердити особу пацієнта, забезпечити доступ осіб, які доглядають за пацієнтами, до потрібної медичної інформації тощо.

Хоча системи біометричної аутентифікації набирають популярності у сфері безпеки, все ще існує кілька поширених міфів навколо біометрії, які уповільнюють

впровадження. Ось чотири найбільш значущі помилки щодо біометричної аутентифікації:

Міф – Біометричні технології – це вторгнення в особисте життя: Існує значна різниця між варіантами біометричної аутентифікації, які вимагають від користувача згоди, і технологією розпізнавання обличчя, розгорнутою в громадських місцях. Рішення для біометричної автентифікації за своєю природою вимагають згоди користувача, оскільки користувач спочатку повинен зареєструвати свої біометричні дані. Крім того, фотографічні зображення обличь не зберігаються в базі даних. Натомість, математична модель обличчя шифрується і зберігається у файлі для порівняння. Вона по суті марна для зловмисника, навіть якщо її вкрали.

Міф – біометричну ідентифікацію можна обдурити статичними зображеннями та фотографіями: Можливо, це було правдою в старих або менш складних ітераціях технології аутентифікації. Однак сучасні рішення для біометричної автентифікації включають можливості виявлення живості, які можуть визначити, чи є представлена біометрична ознака автентичною маскою, моделлю, зображенням або навіть відео.

Для аутентифікації користувача може бути запропоновано моргнути або повернути голову, але інші можливості виявлення жвавості працюють повністю у фоновому режимі.

Міф – Термін придатності біометричних моделей закінчується в міру старіння користувача або зміни функцій: Занепокоєння полягає в тому, що з віком обличчя користувача повільно змінюватиметься з часом, доки воно не перестане реєструватися як збіг. У програмах біометричної автентифікації користувач зазвичай проходить аутентифікацію достатньо регулярно, щоб ці невеликі зміни в зовнішньому вигляді не були достатньо великими, щоб визнати збіг недійсним. Замість цього математична модель буде оновлюватися в міру того, як вона розпізнає зміни в зовнішньому вигляді.

Біометрична ідентифікація застосовується лише в тому випадку, якщо користувач вже відомий: Проаналізувавши способи, за допомогою яких окремий

користувач взаємодіє зі своїм пристроєм, те, як вони тримають телефон, проводять пальцем, друкують на клавіатурі тощо, може бути використано для розробки профілю, за допомогою якого можна автентифікувати користувача або визначити відносний ризик транзакції. Наприклад, у новому сценарії відкриття облікового запису поведінкова біометрія може порівнювати поведінку заявника з репрезентативним пулом користувачів, щоб визначити, чи виглядає новий заявник справжнім, законним користувачем або ботом чи зловмисником

Біометрична автентифікація за геометрією обличчя

Розпізнавання обличчя має перевагу бути пасивною, ненав'язливою системою, яка перевіряє особистість. Загалом, біометричні пристрої можна визначити за допомогою трьох етапної процедури:

1. Коли датчик проводить спостереження. Тип датчика і його спостереження залежать від типу біометричного використовувані пристрої. Це спостереження дає біометричну сигнатуру фізичної особи [10].

2. Комп'ютерний алгоритм «нормалізує» біометричну сигнатуру, щоб він був у тому самому форматі (розмір, роздільна здатність, вид тощо) як сигнатуру в базі даних системи. Нормалізація біометричних даних сигнатури дає нам «нормалізовану сигнатуру» особи [10].

3. Система порівнює нормалізовану сигнатуру з набором нормалізованих підписів у базі даних системи і надає «оцінку подібності», яка порівнює індивідуальну нормовану сигнатуру [10].

Розпізнавання обличчя починається з виявлення візерунків обличчя іноді захаращені сцени, протікає шляхом нормалізації обличчя зображення для врахування геометричних змін і зміни освітлення, можливо, використовуючи інформацію про місце розташування та зовнішній вигляд орієнтирів обличчя, ідентифікує обличчя за допомогою відповідних алгоритмів класифікації та постобробки результатів за допомогою схеми на основі моделі та логістичний зворотний зв'язок [3]. Застосування техніки розпізнавання обличчя може бути поділяються на дві основні частини: застосування правоохоронних органів і комерційне застосування. На рисунку 1.1 зображена схема розпізнавання обличчя.

Рис. 1.1 - Схема розпізнавання зображень обличчя

Усі алгоритми розпізнавання обличь узгоджуються з двома основними частини: (1) виявлення та нормалізація обличчя та (2) ідентифікація обличчя. Алгоритми, які складаються з обох частин називають повністю автоматичними алгоритмами і тими, які складаються з лише другої частини називаються частково автоматичними алгоритми.

Загальна конфігурація може бути описана вектором, що представляє положення та розмір основних рис обличчя, такі як очі та брови, ніс, рот, і форму контуру обличчя (рис.1.2)

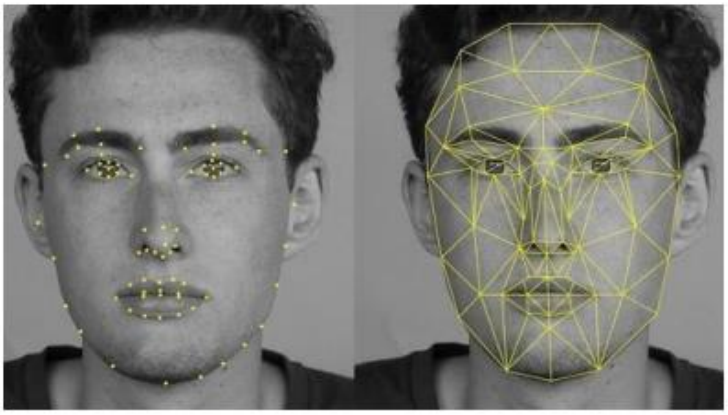


Рис. 1.2 - Зображення геометрії обличчя

Зіставлення графіків. Це ще один підхід до розпізнавання обличчя. При дослідженнях [12] представлена динамічна структура зв'язку для спотворення інваріантного розпізнавання об'єкта, який використовується еластичним зіставлення графіка, щоб знайти найближчий збережений графік. Динамічне посилання архітектури є розширенням класичної штучної нейронної системи мереж.

Підхід (Line Edge Map) LEM не тільки володіє перевагами підходів на основі функцій, таких як незмінність до освітлення та низькі вимоги до пам'яті, але також має перевагу високої продуктивності розпізнавання відповідність шаблону. Line Edge Map інтегрує структурну інформацію з просторова інформація зображення обличчя шляхом групування пікселів обличчя відображення країв у відрізки (рис. 1.3)

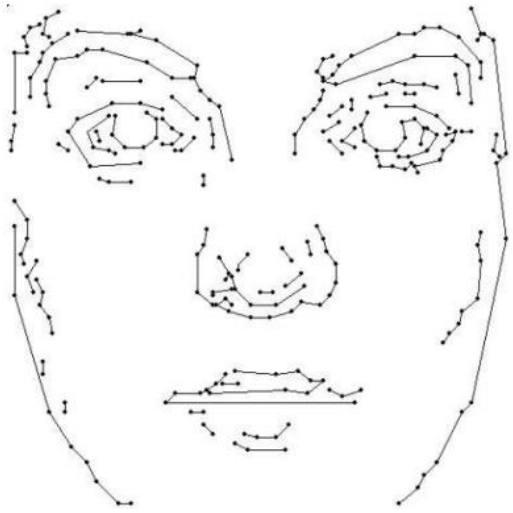


Рис. 1.3 - Підхід Line Edge Map

Після стоншення карти контуру процес підгонки полігональної лінії [10] використовується для створення LEM обличчя. Прикладом людського лобового обличчя LEM є показано на рисунку. 1.3. Подання LEM зменшує вимоги до зберігання, оскільки він записує лише кінцеві точки лінії відрізків на кривих.

1.2 Аналіз методів біометричної аутентифікації

Метод аутентифікації за райдужкою ока. Процес розпізнавання складається з послідовного виконання наступних етапів: пошук, нормалізація, вибір ключових точок, порівняння цих точок.

Локалізація райдужної оболонки. Межа між райдужною оболонкою і склерою є плавним переходом за кольором, у зв'язку з цим завдання виявлення межі переходу стає більш складним, але межа зіниці і райдужки досить чітка.

Його пошук набагато простіше. З цієї причини пошук райдужної оболонки починається з пошуку зіниці. Процес визначення місця розташування райдужної

оболонки складається з двох кроків: знаходження зіниці та знаходження райдужної оболонки біля зіниці. Детектор використовує інформацію про контури, що отримує в результаті обчислення градієнта зображення, для уточнення цих контурів за допомогою подвійної порогової фільтрації та відстеження.

Нормалізація райдужної оболонки. Під час запису зображення в різних умовах око може мати різну форму, наприклад, при збільшенні або звуженні зіниці змінюється кількість пікселів, призначених для зображення райдужної оболонки. Розмір очей на знімку також може бути різним через фізичний стан людини по відношенню до камери. Для врахування таких факторів необхідно привести кільцеподібний малюнок райдужної оболонки до стандартизованої форми, інакше обчислення подібності двох райдужок буде неможливим або близьким до нього.

Виділення характерних ознак. Для того щоб виділити відмінні риси райдужної оболонки, необхідно підвищити контрастність зображення, що зробить текстуру райдужної оболонки більш чіткою. Це означає, що дискретне перетворення зможе отримувати детальну інформацію набагато ефективніше. Посилення контрасту досягається шляхом застосування операції вирівнювання гістограми до нормованого зображення райдужної оболонки [7].

Метод аутентифікації за геометрією обличчя. Процеси розпізнавання обличчя на основі відео використовують зображення обличчя або серію зображень, знятих відеокамерою. Точність розташування та освітлення об’єкта може вплинути на роботу системи. Зазвичай повністю фіксується зображення обличчя, на яке можна прикріпити ключові точки обличчя людини. Наприклад, розташування очей, рота і ніздрів може бути таким, що створюється унікальний візерунок. Тривимірні моделі обличчя можна створювати різними способами, наприклад, шляхом проектування ІЧ-сітки («структурованого світла»), об’єднання кількох зображень.

Тепловізорне зображення обличчя показує кількість тепла, що створюється припливом крові в обличчя. Тепловізорна камера фіксує невидимий тепловий малюнок кровоносних судин під шкірою. Оскільки під час зйомки обличчя за допомогою ІЧ-камери освітлення не потрібно, системи можуть робити зображення в темряві. Однак ІЧ-камери дорожчі за інші типи відеокамер.

За допомогою спеціальних алгоритмів або нейтронної мережі в ядрі розпізнавання біометричної системи зображення обличчя перетворюється в шаблон, а потім в унікальний математичний код. Цей код зберігається як шаблон для певної особи.

Голосовий метод аутентифікації. Обробка мовного сигналу розділяється на кілька основних кроків: попередня обробка сигналу, вибір критеріїв, розпізнавання мовця. Основні характеристики голосу формуються трьома основними властивостями: механікою коливань голосових складок, анатомією мовного тракту та системою управління артикуляцією.

Перевагами є низька вартість цього методу, оскільки потрібні лише мікрофон і звукова карта, які зараз є в кожному комп'ютері. Під час розпізнавання голосу аналізуються висота, модуляція, інтонація тощо.

Основні ознаки, за якими приймається рішення про особу, голос якої перевіряється, формуються з урахуванням усіх факторів процесу формування мовлення: джерела голосу, резонансних частот мовленнєвого тракту та їх загасання, а також динаміки артикуляції. контроль. Крім того, враховується тріада слова, ритм (розподіл наголосу), рівень сигналу, частота та тривалість перерв. Однак надійність і точність цього методу невисокі, оскільки голос може залежати від здоров'я та поведінкових факторів [7].

Метод аутентифікації підпису. Пристрої ідентифікації динаміки підпису використовують геометричні або динамічні функції почерку підпису в реальному часі. Підпис робиться користувачем на спеціальній сенсорній панелі, за допомогою якої змінюється тиск ручки (швидкість, прискорення) перетворюється в електричний аналоговий сигнал. Електронна схема перетворює цей сигнал у цифрову форму, придатну для машинної обробки.

Метод аутентифікації за відбитком пальця.

Є два типи ємнісних сенсорів: пасивні (кожен осередок сенсора має лише одну з пластин конденсатора) і активні (комірка сенсора містить обидві пластини конденсатора).

Це найбільш широко поширений тип напівпровідникових сканерів, в яких для отримання зображення відбитка пальця використовується ефект зміни ємності pn-переходу напівпровідникового приладу при зіткненні гребеня папілярного візерунка з елементом напівпровідникової матриці. Існують модифікації описаного сканера, в яких кожен напівпровідниковий елемент в матриці сканера виступає в ролі однієї пластини конденсатора, а палець - в ролі іншої.

При додатку пальця до сенсора між кожним чутливим елементом і виступом або впадиною папілярного візерунка утворюється якась ємність, величина якої визначається відстанню між поверхнею пальця і елементом. Матриця цих ємностей перетворюється в зображення відбитка пальця.

Ємнісні сенсори неможливо обдурити, просто надрукувавши зображення папіломного малюнка на папері. Більш значуща перевага ємнісних сканерів полягає в тому, що вони компактніші і тому легко інтегруються в портативні пристрої. Саме за рахунок цієї їх особливості вони й набули на даний момент найширшого поширення у смартфонах.

Порівняння вище розглянутих методів можна побачити в таблиці 1.2. Дані методи були порівняні за характеристиками універсальності застосування, стійкості до муляжів, унікальності, та продуктивності сканерів.

Таблиця 1.2 Порівняння біометричних методів аутентифікації

Метод	Універсальність	Унікальність	Постійність	Продуктивність	Стійкість
Обличчя	Висока	Середня	Середня	Висока	Висока
Палець	Середня	Середня	Висока	Висока	Середня
Геометрія руки	Середня	Середня	Середня	Середня	Середня
Райдужка	Висока	Висока	Висока	Висока	Низька
Сітківка	Висока	Висока	Середня	Висока	Низька
Підпис	Низька	Низька	Низька	Низька	Висока
Голос	Середня	Низька	Низька	Низька	Висока

Порівняння біометричних методів аутентифікації демонструє, що кожен з них має свої сильні та слабкі сторони. Методи на основі обличчя та райдужної оболонки ока вирізняються високими показниками універсальності, продуктивності та унікальності, проте райдужка та сітківка ока мають низьку стійкість до зовнішніх факторів. Методи на основі відбитків пальців та геометрії руки забезпечують середній баланс показників, тоді як голос і підпис поступаються за стабільністю та точністю. Найбільш ефективними у контексті багатофакторної аутентифікації є методи, які поєднують високу унікальність та продуктивність.

1.3 Аналіз існуючих реалізацій методів біометричної ідентифікації

Засоби реалізації біометричної аутентифікації надають або відмовляють користувачам у доступі до корпоративних систем і даних залежно від того, чи може користувач довести, що він є тим, за кого себе видає. Щоб перевірити особистість користувачів, засоби біометричної аутентифікації аналізують живі біологічні характеристики користувачів, такі як відбитки пальців, структура обличчя або шаблони друку, і порівнюють їх із записом, який постачальник аутентифікації зберіг у захищеній базі даних. Якщо біометричні дані користувача в реальному часі збігаються з записаними даними, користувачеві надається доступ [5].

Біометрична аутентифікація, можливо, є одним із найбезпечніших методів аутентифікації користувача через той простий факт, що викрасти відбиток пальця користувача, наприклад, набагато складніше, ніж викрасти або зламати його пароль. Таким чином, засоби біометричної аутентифікації стають все більш популярними як основний метод аутентифікації, так і як частина процесу багатофакторної аутентифікації (MFA).

Існує два основних типи засобів реалізації біометричної аутентифікації: ті, що базуються на фізіологічній біометрії, і ті, що базуються на поведінковій біометрії. Фізіологічна біометрія ґрунтується на фізичних особливостях користувача, таких як структура обличчя, малюнок райдужної оболонки ока та сітківки ока або відбитки пальців. Поведінкова біометрія зосереджена на

поведінкових характеристиках користувача, таких як його характер друку, хода або тяга миші.

Приклади засобів реалізації біометричної аутентифікації.

1. Prove

Prove є провідним постачальником інноваційних рішень для перевірки особи та підтвердження, які перевіряють особистість користувачів лише за допомогою їхніх смартфонів. Prove побудована на основі унікальної моделі Prove «Phone-Centric Identity», яка забезпечує безпечний, безперешкодний і багатоканальний доступ до веб-додатків і мобільних додатків на основі інформації, отриманої з мобільних телефонів користувачів. Prove Auth – це засіб для аутентифікації без пароля та OTP від Prove, який організації можуть використовувати для аутентифікації користувачів за допомогою біометричних даних, push-сповіщень або технології ідентифікації, орієнтованої на телефон, від Prove.

За допомогою Prove Auth організації можуть підтверджувати особистість кінцевих користувачів за допомогою біометрії на пристрої, як-от сканування обличчя або відбитків пальців, за допомогою push-сповіщень, надісланих через додаток аутентифікатора, або за допомогою моделі ідентифікації, орієнтованої на телефон Prove. Ця модель аналізує велику кількість сигналів про використання мобільних пристроїв, телекомунікацій і пристроїв зі смартфона користувача, щоб переконатися, що користувач володіє телефоном, що він володіє ним у режимі реального часу, і що спроба входу не демонструє жодних підозрілих або ризикованих дій із високим ризиком.

Використовуючи цю інформацію, Prove Auth надає або відмовляє в доступі, підвищуючи безпеку облікового запису та мінімізуючи ризик шахрайства, забезпечуючи при цьому оптимізований, інтуїтивно зрозумілий досвід входу та аутентифікації для кожного користувача. Організації, які хочуть ще більше «посилити» свою аутентифікацію, можуть впровадити багатофакторну аутентифікацію, використовуючи комбінацію трьох доступних методів, наприклад, використовуючи ідентифікацію, орієнтовану на телефон, як основний метод

аутентифікації, і вводячи біометричні дані, щоб посилити аутентифікацію для ризикованих спроб входу.

2. BehavioSec

Спеціалізуючись на поведінковій біометрії, її мультимодальне рішення перевіряє користувачів на основі їхніх звичок і моделей поведінки та може використовуватися як додатковий фактор у процесі MFA або для безперервної аутентифікації користувачів протягом усього життєвого циклу їхньої взаємодії з певною платформою. Це дозволяє організаціям впроваджувати як аутентифікацію на основі ризиків, так і підхід Zero Trust.

Програмний засіб BehavioSec працює безшумно у фоновому режимі, аутентифікуючи користувачів пасивно та без тертя, але з високим ступенем точності. Використовується машинне навчання для аналізу взаємодії з користувачами – наприклад, того, як вони друкують, взаємодіють зі смартфонами або рухають мишею – разом із різними контекстуальними факторами, щоб порівняти їх із минулою поведінкою та визначити «оцінку ризику» на основі того, наскільки добре вони збігаються. Адміністратори також можуть аналізувати активність аутентифікації та аналітику, а також оцінки ризику користувачів за допомогою простої у використанні інформаційної панелі.

Даний продукт рекомендований для великих підприємств, особливо в банківській сфері, електронній комерції та фінансах, які шукають пасивне, безперешкодне рішення для забезпечення високого рівня захисту від шахрайства.

3. BIO-key PortalGuard

Платформа дозволяє користувачам аутентифікувати свою особистість різними способами, у тому числі за допомогою надійних можливостей безпарольної біометричної аутентифікації BIO-key.

Використовуючи PortalGuard, користувачі можуть увійти в систему без пароля, замінюючи паролі біометричними сканами, або використовувати біометрію як частину процесу MFA. Щоб доповнити цю платформу, нещодавно запущений мобільний додаток BIO-key MobileAuth використовує технологію PalmPositive для перевірки особи користувача. Щоб аутентифікуватися за

допомогою цієї технології, користувачі мають сканувати долоні за допомогою камери пристрою. Потім програма аналізує сканування на сервері, а не на пристрої, і надає користувачеві доступ, якщо сканування відповідає його попередньо зареєстрованому біометричному шаблону. Для адміністраторів платформа також оснащена центральною інформаційною панеллю, з якої вони можуть налаштовувати політики доступу, переглядати звіти про активність у реальному часі тощо.

Даний продукт рекомендований організаціям у фінансовій, державній, медичній та комерційній галузях, які шукають перевіреного постачальника для надання безперебійних можливостей єдиного входу та багатофакторної аутентифікації без пароля за допомогою біометричної аутентифікації.

4. Верифікатор обличчя iProov і верифікатор долоні

iProov – це постачальник біометричної аутентифікації, який спеціалізується на верифікації обличчя та долоні для аутентифікації користувачів та запобігання спуфінгу. Його рішення Face Verifier і Palm Verifier створені з використанням запатентованої технології Genuine Presence Assurance, яка визначає, чи є користувач, який намагається увійти в систему, правильною особою, реальною особою, і проходить аутентифікацію в режимі реального часу. Обидва рішення також можна використовувати як первинну аутентифікацію, як частину процесу MFA або як поетапну аутентифікацію.

Біометричні засоби iProov використовують технології глибокого навчання, а також вбудовані функції запобігання повторним атакам, щоб забезпечити найвищий рівень безпеки та точності під час спроб аутентифікації. Щоб аутентифікуватися за допомогою Face Verifier або Palm Verifier, користувачі мають представити свої обличчя або долоні на передню камеру. Потім технологія порівнює цю спробу аутентифікації з попередньо зареєстрованим біометричним шаблоном цього користувача, надаючи доступ лише в тому випадку, якщо він збігається. iPortal від iProov також надає командам безпеки централізовану область звітності, де вони можуть контролювати та керувати адмініструванням користувачів, підготовкою, інтеграцією тощо.

Довіряючи організаціям державного сектору, урядам і банкам безпечну і точну перевірку особистості користувача, iProov рекомедовані для біометричної аутентифікації для великих підприємств, яким потрібен високий рівень впевненості в тому, що користувачі, які отримують доступ до їхніх систем, мають на це дозвіл.

5. TypingDNA Verify 2FA та ActiveLock

Дані засоби працюють завдяки використанню інженерії даних та алгоритмів глибокого навчання для аналізу поведінки та шаблонів набору тексту, щоб розпізнавати та перевіряти користувачів за лічені секунди. Використовуючи цю технологію, TypingDNA пропонує продукт для двофакторної аутентифікації Verify 2FA, а також рішення для неперервної аутентифікації ActiveLock.

Щоб спочатку зареєструватися для використання продуктів TypingDNA, користувачам потрібно надати лише один зразок типізації. На основі цього зразка система аналізує їхню поведінку під час набору, а потім може використовувати ці дані для розпізнавання їх під час майбутніх спроб аутентифікації. Verify 2FA працює, просячи користувачів ввести в систему лише чотири слова, щоб підтвердити свою особу. Якщо шаблон введення збігається з профілем, він отримує доступ за лічені секунди. Якщо це не так, але користувач виявився справжнім, він може отримати доступ до свого облікового запису за допомогою одноразового пароля на основі SMS. З іншого боку, ActiveLock працює у фоновому режимі, відстежуючи поведінку під час набору тексту під час роботи користувачів протягом дня та постійно аутентифікуючи їх. Якщо він виявляє шаблон набору тексту, який не збігається з зареєстрованим користувачем, він може автоматично заблокувати пристрій або активувати беззвучне сповіщення.

Даний продукт рекомендується компаніям у всіх галузях, особливо у фінансах, освіті та роздрібній торгівлі, які шукають універсальне рішення, яке може не лише перевіряти потрібних користувачів, але й виявляти, коли шахраї використовують пристрій, на використання якого вони не мають дозволу.

6. Голосова біометрія Veridas

Компанія Veridas, заснована спільно BBVA та Das-Nano у 2017 році, є постачальником біометричної автентифікації, який розробляє програмне забезпечення для розпізнавання обличь і голосу для компаній по всьому світу.

Veridas Voice Biometrics – це засіб для розпізнавання голосу, який базується на механізмі перевірки динаміків das-Peak і може аутентифікувати особу користувача за допомогою менш ніж трьох секунд аудіо.

Veridas Voice Biometrics – це просте у використанні хмарне рішення, яке заявляє точність 99,5% для своєї технології розпізнавання голосу.

1.4 Висновки та постановка задачі

Метою дослідження є удосконалення багатофакторної біометричної аутентифікації на основі CNN та LSTM.

Сформулюємо задачі для дослідження в рамках даної кваліфікаційної роботи:

– провести огляд і аналіз сучасного стану задачі «біометричної аутентифікації»;

– розглянути існуючі методи для вирішення задачі;

– розробити програмну реалізацію удосконаленої моделі;

– зробити оцінку якості роботи моделі на реальних даних;

– на основі отриманих даних зробити висновок про проведену роботу.

2 УДОСКОНАЛЕННЯ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІ З ВИКОРИСТАННЯМ CNN ТА LSTM

У даному розділі буде обґрунтовано необхідність удосконалення багатофакторної біометричної аутентифікації шляхом аналізу сучасних методів та визначення їхніх обмежень, таких як низька стійкість до атак та обмежена ефективність у складних умовах. Будуть розглянуті особливості вдосконалення цієї технології з використанням методів глибинного навчання, зокрема CNN та LSTM, які дозволяють поєднати статичні та динамічні ознаки для підвищення точності. Також буде розроблено алгоритм, що інтегрує ці підходи, описано його архітектуру та принципи роботи. У висновках розділу будуть підсумовані основні результати та визначено напрями подальшого впровадження запропонованих рішень.

2.1 Обґрунтування необхідності удосконалення багатофакторної біометричної аутентифікації

Багатофакторна автентифікація (MFA) – це перевірка особи користувача за допомогою кількох облікових даних для підвищення безпеки. MFA – це популярний спосіб нейтралізувати ризики, пов’язані з паролями, додавши другий, інший фактор до входу на основі пароля. Сьогодні користувачі знайомі з двофакторною або двоетапною автентифікацією, яка є підмножиною MFA.

Хоча рішення MFA не обов’язково має містити пароль, воно має містити принаймні два докази того, що ви є тим, за кого себе видаєте. Подібним чином ці фактори мають виходити принаймні з двох із трьох категорій прийнятих облікових даних [14].

У більшості випадків два фактори аутентифікації пропонують кращий захист від захоплення облікового запису, особливо якщо одним із цих факторів є пароль. MFA допомагає захистити облікові записи, коли один із факторів скомпрометовано. Крім того, MFA допомагає деяким компаніям відповідати нормам щодо надійної автентифікації клієнтів (SCA) . Недоліком є те, що це може ускладнити роботу користувача. Саме тут біометрія надає значну цінність, оскільки вона посилює

безпеку, не додаючи часу та ускладнюючи процес входу. Крім того, застосування біометричних даних із захистом від спуфінгу є надійнішим фактором, ніж такі варіанти, як одноразові паролі, надіслані через текстове повідомлення [17].

Для того, щоб мінімізувати ризики втрати особистих даних, потрібно удосконалювати процедуру багатофакторної аутентифікації. Це можна зробити за рахунок додавання більшої кількості біометричних факторів. Щоб це реалізувати можна використати нейронні мережі. В даному випадку, корисним буде використання CNN та LSTM.

Згорткова нейронна мережа (CNN) — це розширена версія штучних нейронних мереж (ШНМ), яка переважно використовується для вилучення функції з набору даних сіткоподібної матриці. Наприклад, візуальні набори даних, такі як зображення або відео, де шаблони даних відіграють велику роль [9].

Згорткова нейронна мережа складається з кількох рівнів, таких як вхідний шар, згортковий шар, шар пулінгу та повністю пов'язані шари (рис.2.1).

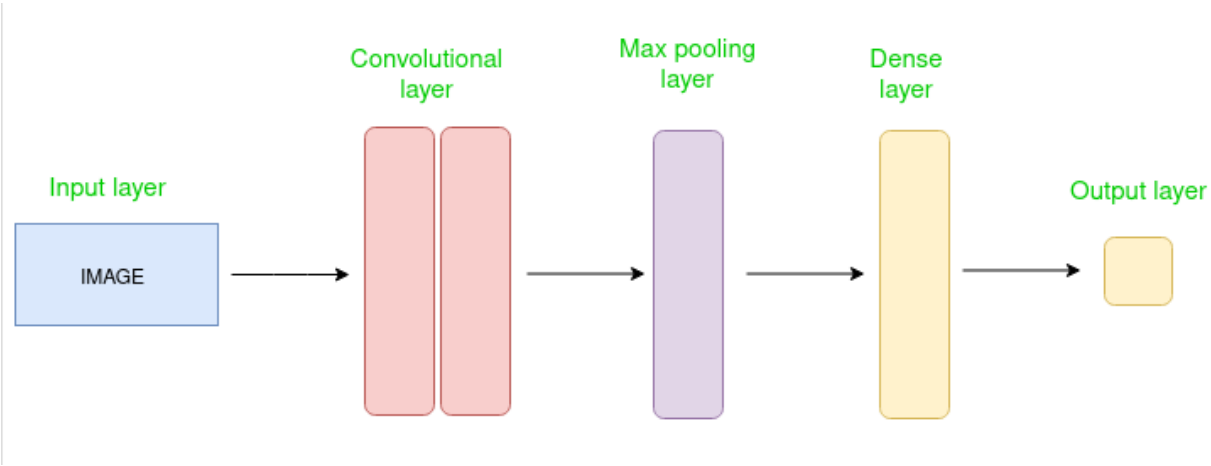


Рис. 2.1 - Проста архітектура CNN

Шар згортки застосовує фільтри до вхідного зображення для вилучення об'єктів, шар «Об'єднання» зменшує вибірку зображення, щоб зменшити кількість обчислень, а повністю підключений шар робить остаточне прогнозування. Мережа вивчає оптимальні фільтри за допомогою зворотного поширення і градієнтного спуску.

Згорткові нейронні мережі або ковнети – це нейронні мережі, які мають спільні параметри. Уявіть, що у вас є образ. Він може бути представлений у

вигляді кубоподібного м'яза, що має свою довжину, ширину (розмір зображення) та висоту (тобто канал, оскільки зображення зазвичай мають червоний, зелений та синій канали) (рис. 2.2).

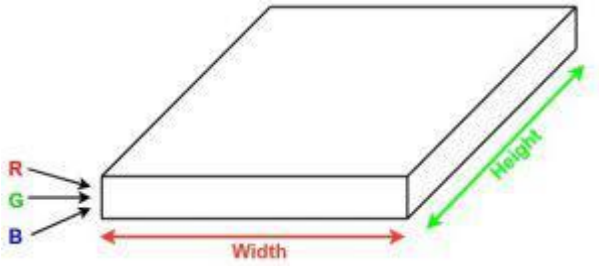


Рис. 2.2 - Згорткові шари

Якщо взяти невеликий фрагмент цього зображення і запустити на ньому невелику нейронну мережу, яка називається фільтром або ядром, з виходами, K і представленням їх вертикально. Поширивши дію нейронної мережі на все зображення, отримується ще одне зображення з різною шириною, висотою і глибиною. Ця операція називається згорткою. Якщо розмір патча збігається з розміром зображення, це буде звичайна нейронна мережа (рис. 2.3).

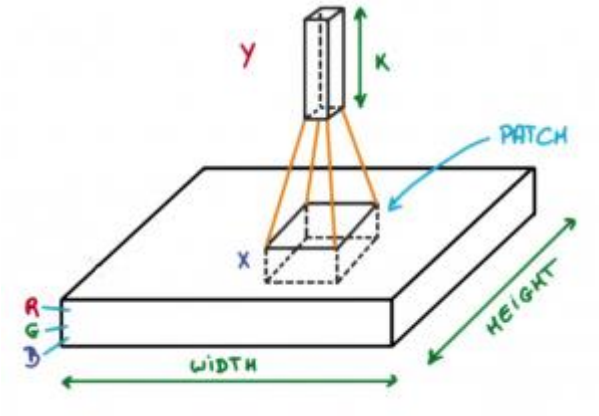


Рис. 2.3 - Згортка

Переваги CNN:

1. Добре виявляє закономірності та особливості на зображеннях, відео та аудіо сигналах [15].
2. Стійкість до інваріантності перекладу, обертання та масштабування.
3. Наскрізне навчання, немає необхідності в ручному вилученні функцій.
4. Може обробляти великі обсяги даних і досягати високої точності.

Недоліки CNN:

1. Обчислювально затратний для навчання і вимагає багато пам'яті.

2. Може бути схильний до перенавчання, якщо використовується недостатньо даних або правильна регуляризація.

3. Вимагає великих обсягів мічених даних.

4. Інтерпретація обмежена, важко зрозуміти, чому навчилася мережа.

Long Short-Term Memory

LSTM чудово справляється із завданнями прогнозування послідовностей, фіксуючи довготривалі залежності. Ідеально підходить для часових рядів, машинного перекладу та розпізнавання мовлення через залежність від порядку [19].

Long Short-Term Memory – це вдосконалена версія рекурентної нейронної мережі, розроблена компанією Hochreiter & Schmidhuber.

Традиційний RNN має єдиний прихований стан, який передається в часі, що може ускладнити вивчення мережею довгострокових залежностей. Моделі LSTM вирішують цю проблему шляхом введення комірки пам'яті, яка є контейнером, який може зберігати інформацію протягом тривалого періоду.

Архітектури LSTM здатні вивчати довготривалі залежності в послідовних даних, що робить їх добре придатними для таких завдань, як переклад мов, розпізнавання мови та прогнозування часових рядів.

LSTM також можна використовувати в поєднанні з іншими архітектурами нейронних мереж, такими як згорткові нейронні мережі (CNN) для аналізу зображень і відео [21].

Архітектура LSTM включає комірку пам'яті, яка керується трьома вентилями: вхідним вентилем, затвором забуття та вихідним вентилем. Ці вентилі вирішують, яку інформацію додавати, видаляти з неї та виводити з комірки пам'яті.

– Вхідний вентиль контролює, яка інформація додається в комірку пам'яті.

– Вентиль забуття контролює, яка інформація видаляється з комірки пам'яті.

– Вихідний вентиль контролює, яка інформація виводиться з комірки пам'яті.

Це дозволяє мережам LSTM вибірково зберігати або відкидати інформацію під час її проходження через мережу, що дозволяє їм вивчати довгострокові залежності.

LSTM підтримує прихований стан, який виступає в ролі короткочасної пам'яті мережі. Прихований стан оновлюється на основі вхідних даних, попереднього прихованого стану та поточного стану комірки пам'яті [22].

Двонаправлена модель LSTM

Двонаправлена LSTM (Bi LSTM/BLSTM) – це рекурентна нейронна мережа (RNN), яка здатна обробляти послідовні дані як у прямому, так і у зворотному напрямках. Це дозволяє Bi LSTM вивчати більш довготривалі залежності в послідовних даних, ніж традиційні LSTM, які можуть обробляти послідовні дані тільки в одному напрямку.

– Bi LSTM складаються з двох мереж LSTM, одна з яких обробляє вхідну послідовність у прямому напрямку, а інша — у зворотному напрямку.

– Потім виходи двох мереж LSTM об'єднуються для отримання кінцевого виходу.

Моделі LSTM, включаючи Bi LSTM, продемонстрували найсучаснішу продуктивність у різних завданнях, таких як машинний переклад, розпізнавання мови та підсумовування тексту.

Мережі в архітектурах LSTM можуть бути складені для створення глибоких архітектур, що дозволяє вивчати ще більш складні шаблони та ієрархії в послідовних даних. Кожен шар LSTM у складеній конфігурації фіксує різні рівні абстракції та часових залежностей у вхідних даних.

Архітектура LSTM має ланцюгову структуру, яка містить чотири нейронні мережі та різні блоки пам'яті, які називаються комірками.

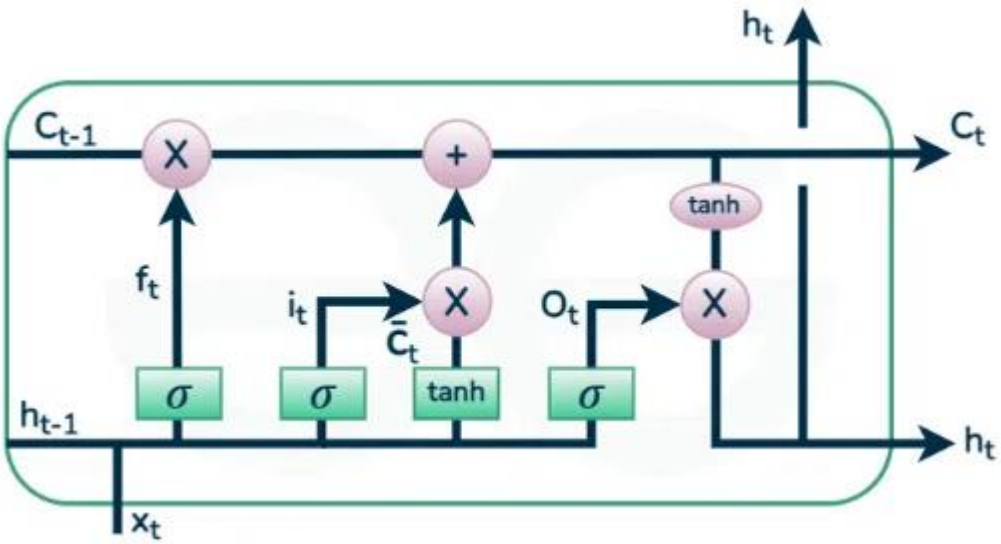


Рис. 2.4 - Архітектура LSTM

Деякі з відомих застосувань LSTM включають:

– Моделювання мови: LSTM використовувалися для завдань обробки природної мови, таких як моделювання мови, машинний переклад і підсумовування тексту. Їх можна навчити створювати зв'язні та граматично правильні речення, вивчаючи залежності між словами в реченні [25].

– Розпізнавання мовлення: LSTM використовуються для завдань із розпізнавання мовлення, таких як транскрибування мовлення в текст і розпізнавання голосових команд. Їх можна навчити розпізнавати закономірності в мові та зіставляти їх з відповідним текстом.

– Прогнозування часових рядів: LSTM використовувалися для завдань прогнозування часових рядів, таких як прогнозування цін на акції, погоди та споживання енергії. Вони можуть вивчати закономірності в даних часових рядів і використовувати їх для прогнозування майбутніх подій [26].

– Виявлення аномалій: LSTM використовувалися для завдань із виявлення аномалій, таких як виявлення шахрайства та вторгнення в мережу. Їх можна навчити виявляти закономірності в даних, які відхиляються від норми, і позначати їх як потенційні аномалії.

– Рекомендаційні системи: LSTM використовувалися для рекомендаційних завдань, таких як рекомендація фільмів, музики та книг. Вони

можуть вивчати закономірності в поведінці користувачів і використовувати їх для надання персоналізованих рекомендацій.

– Аналіз відео: LSTM використовувалися для завдань відеоаналізу, таких як виявлення об'єктів, розпізнавання активності та класифікація дій. Їх можна використовувати в поєднанні з іншими архітектурами нейронних мереж, такими як згорткові нейронні мережі (CNN), для аналізу відеоданих та вилучення корисної інформації [27].

Відповідно, використання CNN та LSTM дає можливість підвищити рівень безпеки особистих даних при здійснені багатофакторної аутентифікації.

2.2 Особливості удосконалення багатофакторної біометричної аутентифікації

Удосконалення багатофакторної біометричної аутентифікації є важливим етапом у підвищенні безпеки цифрових систем. Сучасні підходи до автентифікації користувачів орієнтовані на використання кількох факторів для підтвердження особи, що дозволяє значно знизити ризики компрометації облікових записів. У цьому контексті біометрія є потужним інструментом, оскільки вона поєднує високу надійність і зручність для користувача, забезпечуючи додатковий рівень захисту без значного ускладнення процесу автентифікації.

Біометричні фактори, на відміну від традиційних паролів, важко підробити або викрасти, оскільки вони базуються на фізичних або поведінкових характеристиках користувача, таких як відбитки пальців, сітківка ока, обличчя, голос чи навіть поведінкові патерни під час введення даних. Це робить біометрію надзвичайно привабливим варіантом для підвищення рівня безпеки в системах багатофакторної автентифікації (MFA). Крім того, інтеграція біометрії в MFA дозволяє зменшити ризики, пов'язані з втратами або викраденням паролів, адже навіть у разі компрометації одного з факторів, інші залишаються захищеними.

Одним із важливих напрямків удосконалення біометричних систем є розвиток технологій, які забезпечують високу точність і стійкість до спуфінгу, тобто до підробки біометричних даних. Для цього розробляються нові методи аналізу та

верифікації біометричних ознак, що дозволяють підвищити надійність автентифікації. Крім того, використання біометрії в поєднанні з іншими технологіями, такими як штучні нейронні мережі, дозволяє ще більше покращити ефективність систем. Зокрема, використання згорткових нейронних мереж (CNN) для аналізу зображень, таких як обличчя чи відбитки пальців, разом з рекурентними нейронними мережами довгої короткочасної пам'яті (LSTM) для обробки послідовних даних, може значно поліпшити точність та швидкість розпізнавання. Такі гібридні підходи дозволяють обробляти великі обсяги даних, забезпечуючи високий рівень захисту і адаптуючи систему до нових викликів, таких як спроби обходу за допомогою 3D-моделей або зловживання поведінковими патернами.

Важливою перевагою біометричних систем є їх здатність працювати в реальному часі, забезпечуючи користувачу зручність без потреби в додаткових зусиллях для запам'ятовування паролів чи введення кодів. Однак одним із викликів є забезпечення конфіденційності та захисту особистих даних користувачів. Тому удосконалення біометричної автентифікації також включає розвиток методів криптографічного захисту, що дозволяють надійно зберігати та обробляти біометричні дані без ризику їх компрометації.

У загальному підсумку, удосконалення багатофакторної біометричної аутентифікації дозволяє не тільки підвищити рівень безпеки, але й зробити процес входу в систему більш зручним та ефективним, що відповідає сучасним вимогам до користувацького досвіду та захисту особистих даних.

2.3 Розробка алгоритму удосконалення багатофакторної біометричної аутентифікації

Для реалізації удосконалення багатофакторної біометричної аутентифікації можна скористатись наступним алгоритмом

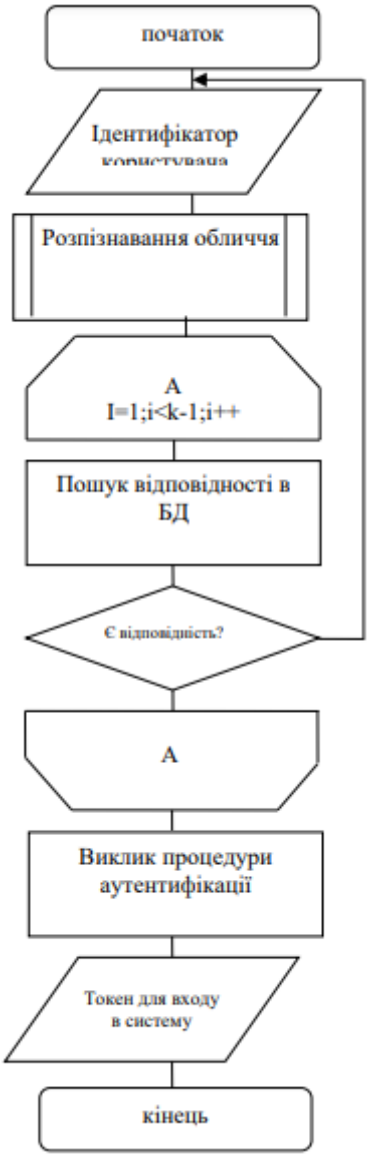


Рис. 2.5 - Алгоритм здійснення біометричної аутентифікації

Процес аутентифікації користувача в системі здійснюється згідно наступної послідовності дій:

1) На вхід підсистеми збору даних надходить біометричний образ користувача (опрацьовується зображення), де він піддається оцифровці та голосовий фрагмент.

2) Оцифрований біометричний образ передається на вхід підсистеми обробки сигналів, де він піддається операції нормалізації та виділення біометричні параметру вектор(контрольні точки рис обличчя).

3) Аналізатор голосу виявляє збіжності між отриманими даними та даними в базі.

4) Вектор біометричних параметрів надходить до модуля аутентифікації, який виконує перевірку введеного ідентифіатора та ідентифікатора що в БД відповідає біометричним даним.

5) Введення пароля.

На рисунку 2.6 приведена спрощена схема біометричної аутентифікації, що дозволяє відобразити зв’язки компонентів системи.

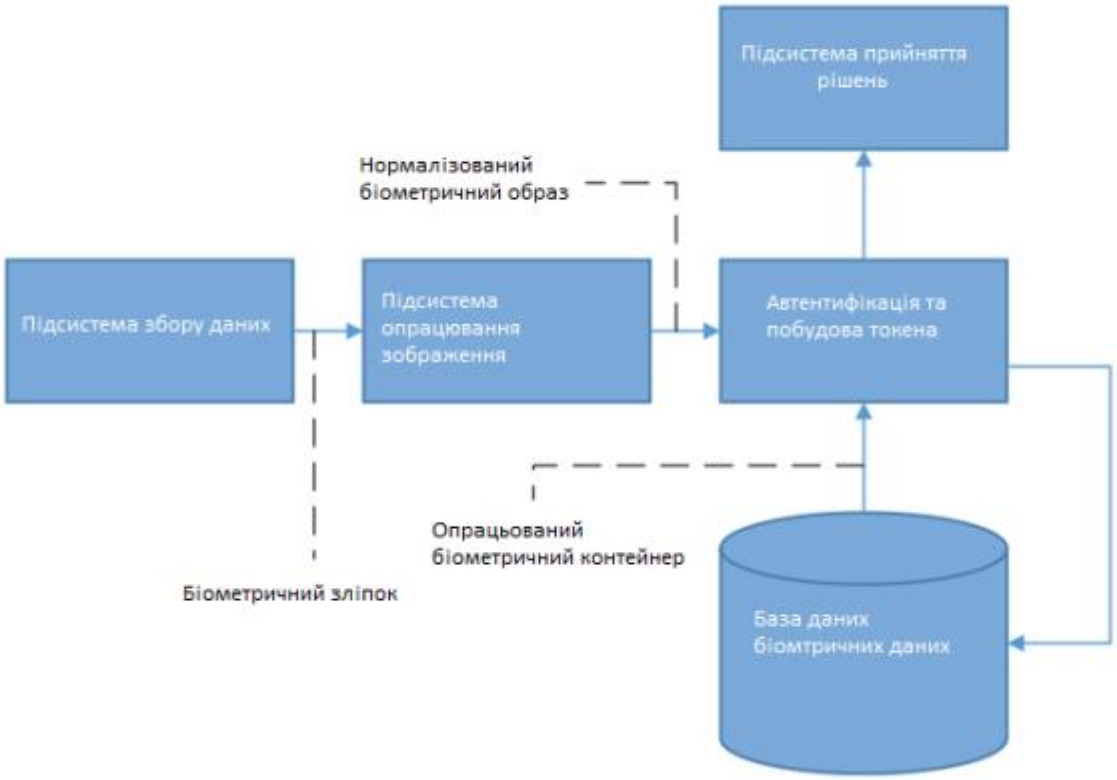


Рис. 2.6 - Спрощена схема біометричної аутентифікації

Основна ідея запропонованого рішення базується на отриманні ідентифікатора, який згодом перевіряється з ідентифікатором введеним користувачем. Для отримання ідентифікатора у процесі реєстрації, образ, що надходить на вхід підсистеми аутентифікації, також подається модуль розпізнавання обличчя та голосу.

Діаграма послідовності для процесу біометричної аутентифікації користувача приведена на рисунку 2.7.

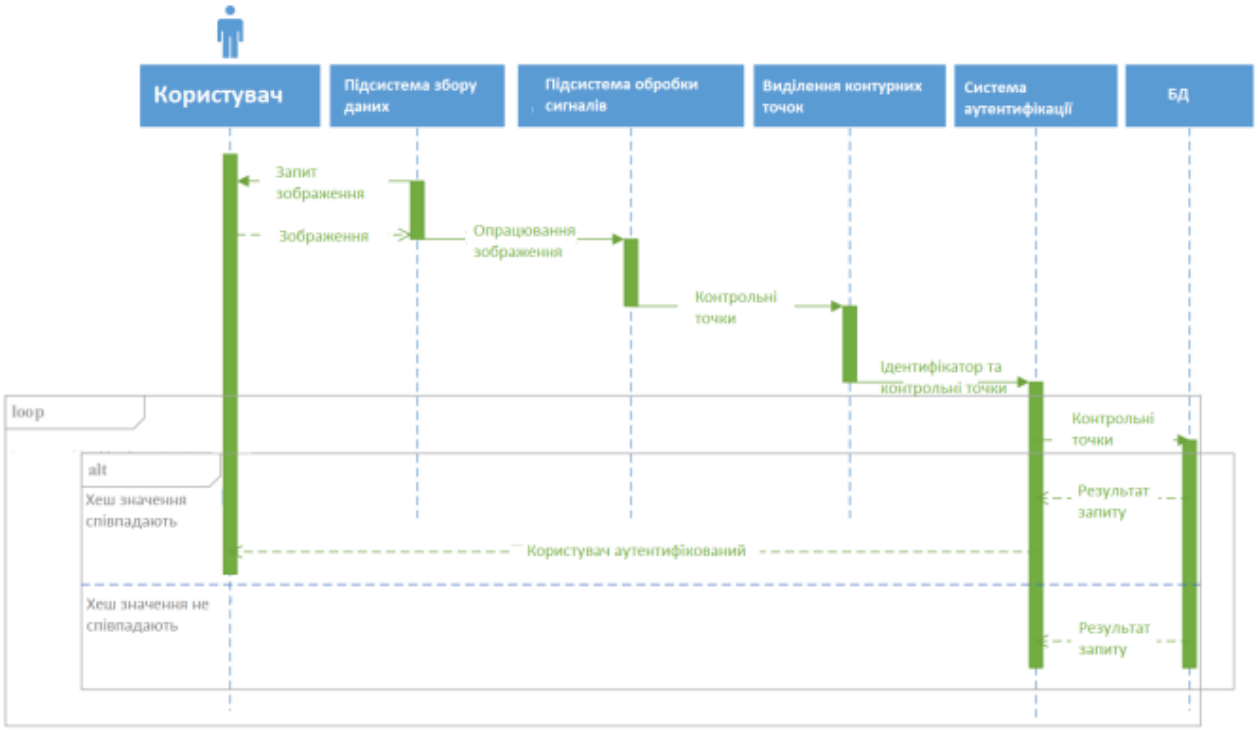


Рис. 2.7 - Процес аутентифікації користувача

Процес аутентифікації користувача включає послідовні етапи: збирання біометричних даних, їх порівняння з еталонними зразками в базі даних та прийняття рішення про відповідність. Такий підхід забезпечує надійність і точність ідентифікації, знижуючи ризик несанкціонованого доступу.

2.4 Висновки до розділу

В результаті роботи було розглянуто теоретичні аспекти представленої теми. Проведений аналіз літературних джерел, дав можливість виявити основні закономірності у функціонуванні багатофакторної біометричної аутентифікації

Також було виявлено, що представлені методи та алгоритми є ефективними для вирішення поставленої задачі, хоча можливі подальші покращення та оптимізації.

3. ПРОГРАМНА РЕАЛІЗАЦІЯ УДОСКОНАЛЕННЯ БАГАТОФАКТОРНОЇ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ

У розділі 3 буде виконано програмну реалізацію удосконаленої системи багатофакторної біометричної аутентифікації. Спочатку буде обґрунтовано вибір мови програмування, зокрема Python, з поясненням її переваг та сфер застосування, які забезпечують ефективність у роботі з методами глибинного навчання. Далі будуть описані ключові фрагменти програмного коду, які реалізують алгоритм удосконалення системи, з детальним поясненням їх функціональності. На етапі тестування буде перевірено працездатність створеного програмного забезпечення, оцінено його ефективність та відповідність поставленим вимогам. У висновках розділу буде підсумовано результати реалізації та тестування, а також визначено перспективи вдосконалення програмного продукту.

3.1 Обґрунтування вибору мови програмування

Python – це інтерпретована мова програмування загального призначення високого рівня, яка має простий синтаксис, який легко освоїти, і підкреслює легкість читання. Його в основному використовують професійні програмісти та розробники в різних галузях, включаючи розробку Інтернету та програмного забезпечення, машинне навчання, штучний інтелект, великі дані та складну математику. Як і всі інші мови програмування, Python також має свої плюси та мінуси.

Інтерпретований – інтерпретатор обробляє вихідний файл під час виконання, він читає рядки коду один за одним і виконує те, що сказано. Подібно до Perl і PHP, Python не вимагає компіляції програми перед її виконанням. Отже, не потрібно викликати компілятор. Замість запуску компілятора, який допомагає перетворити вихідні файли на скомпільовані файли класів, просто запускається файл .py. Компіляція байт-коду Python є автоматичною та повністю неявною.

Високорівневий – Python спирається на легкочитані структури, які згодом перекладаються на мову низького рівня, оригінальний код, який виконується на центральному процесорі (CPU) комп’ютера. Мова високого рівня призначена для

використання програмістом, а написаний код далі інтерпретується мовою низького рівня. Як і C++ або Java, перед запуском Python потрібно обробити. Це забезпечує портативність Python — він може працювати на різних типах комп’ютерів майже без модифікацій.

Загального призначення – Python можна використовувати майже для всього. Він застосовний майже в усіх галузях для різноманітних завдань. Будь то виконання таких короткотермінових завдань, як тестування програмного забезпечення чи довгострокова розробка продукту, що передбачає планування дорожньої карти, Python добре працює для всіх них, він застосовний усюди. Його ролі необмежені. Він популярний не тільки серед інженерів-програмістів, а й серед фахівців інших галузей: математики, аналізу даних, науки, бухгалтерського обліку та мережевої інженерії.

Об’єктно-орієнтоване — ця парадигма програмування дає загальну орієнтацію на сценарії та потужне структурування коду. Цей об'єктно-орієнтований підхід дозволяє мислити проблеми в термінах класів і об'єктів. Потім об’єкти компонуються таким чином, щоб складати складні комп’ютерні програми. Окрім об’єктно-орієнтованого програмування, Python також підтримує процедурну парадигму. Оскільки ООП є лише одним із варіантів, ви можете зробити програмування на Python більш просунутим, вибравши підхід до об’єктно-орієнтованого програмування. Розробники можуть створювати шаблони коду для повторного використання, таким чином зменшуючи надмірність у проектах розробки.

У різних галузях існує велика різноманітність варіантів використання Python. Звичайно, перше, що спадає на думку, коли ми думаємо про найпоширеніші способи використання Python, це для створення веб-додатків, мобільних і настільних програм, а також для їх тестування. Але Python — це мова, яка служить багатьом цілям. Загалом, Python ідеально підходить для таких сфер використання:

1. Розробка веб-додатків
2. Наука про дані
3. Сценарії

4. Програмування бази даних

5. Швидке створення прототипів

Python підходить для всіх форм програмування, що сприяє швидкому зростанню бази користувачів. Скрипти міжплатформної оболонки, швидка автоматизація, проста веб-розробка, аналіз і візуалізація даних, штучний інтелект і машинне навчання – це деякі приклади.

Часто фахівці використовують Python для кращого виконання різноманітних завдань у різних дисциплінах. Кращої продуктивності, серед іншого, можна досягти за допомогою автоматизації. Фінанси, страхування та маркетинг є основними сферами, у яких люди стикаються з необхідністю виконувати повторювані завдання: переглядати, копіювати, перейменовувати та завантажувати файли на сервер, завантажувати веб-сайти чи аналізувати дані. Натомість програміст може написати сценарій на Python і автоматизувати все це.

Крім того, не обов’язково бути розробником програмного забезпечення, щоб використовувати Python. Мова дозволяє полегшити аналіз і візуалізацію даних. Він має багату екосистему, що включає ефективні бібліотеки для обробки даних і, отже, допомагає спеціалістам із обробки даних у виконанні складних числових обчислювальних операцій.

Не дарма найбільші компанії – світу використовують Python. Pixar використовує його для створення фільмів, Google для сканування сторінок, Netflix для доставки вмісту та Spotify для рекомендації пісень. Ця мова має багато переваг, і є кілька вагомих причин:

Простота**.** Простий і зрозумілий синтаксис Python спонукає початківців вивчати цю мову сценаріїв. Його код легко зрозуміти, поширювати та підтримувати. Немає багатослівності, мова легко вивчається.

Потужний інструментарій**.** За своєю суттю програми на Python є текстовими файлами, що містять інструкції для інтерпретатора та написані в текстовому редакторі або IDE. IDE є повнофункціональними та пропонують такі вбудовані інструменти, як перевірка синтаксису, налагоджувачі та браузери коду, текстові редактори зазвичай не включають функції IDE, але їх можна налаштувати. Python

також має величезний набір сторонніх пакетів, бібліотек і фреймворків, які полегшують процес розробки. Таким чином, ці можливості оптимізації роблять Python чудовим для великомасштабних проектів.

Швидкість розвитку. Тут мається на увазі швидкість бізнесу та показник часу виходу на ринок. Python це динамічна мова сценаріїв, тому вона не призначена для написання програм з нуля, а в першу чергу призначена для підключення компонентів. Компоненти призначені для повторного використання, а інтерфейси між компонентами та сценаріями чітко визначені. Усе це прискорює розробку програмного забезпечення завдяки Python, що робить мову надзвичайно лаконічною та продуктивною.

Гнучкість**.** Хоча Python робить наголос на простоті та читабельності коду, а не на гнучкості, у цій мові це все одно є. Python можна використовувати в різних проектах. Це дозволяє розробникам вибирати між об’єктно-орієнтованим і процедурним режимами програмування. Python також гнучкий у типі даних. Їх 5: число, рядок, список, кортеж і словник, і кожен тип підданих відповідає одному з цих кореневих типів. У результаті дослідницький аналіз даних стає легше проводити завдяки гнучкості Python.

Портативність**.** Python створено для переносимості. Його програми підтримуються на будь-якій сучасній комп'ютерній ОС. Завдяки високорівневому характеру мови сценарій Python інтерпретується, тому його можна написати для подальшої інтерпретації однаково добре в Linux, Windows, Mac OS і UNIX, не вимагаючи коригувань. Програми на Python також дозволяють реалізувати портативні GUI.

Сильна громада**.** Python має швидко зростаючу базу користувачів і фактично є репрезентативним прикладом сильної спільноти. Є тисячі учасників потужного інструментарію Python — Pythonists. В онлайн-сховище вже завантажено майже 200 000 програмних пакетів, створених на замовлення . Усе це означає, що велика підтримуюча спільнота є як причиною, так і наслідком попиту на мову.

Той факт, що Python має репутацію зручної для програмування мови, якій віддають перевагу розробники, безсумнівний, але час від часу Python порівнюють

з іншими мовами програмування, включаючи Java, C#, PHP і Ruby on Rails. Однак порівняння дійсне, якщо взяти до уваги продуктивність, функціональність та всі інші адекватні показники обговорюваної пари.

Усі мови програмування мають свої недоліки. Незважаючи на всі переваги, які пропонує Python як мова програмування, є недоліки, якими слід скористатися:

– Швидкість як інтерпретована мова. Хороша новина полягає в тому, що цей недолік можна виправити з появою PyPy, яка обіцяє приріст продуктивності.

– Динамізм Python запобігає виявленню семантичних помилок заздалегідь. Але такі інструменти, як PyChecker, можуть перевіряти наявність помилок, що робив би компілятор таких мов, як C або Java.

– Потоковість є менш продуктивною в Python, ніж в інших мовах. Багатопотоковість може стати можливою з Jython, але незмінність не надто важлива в Python, тому однопотоковий паралелізм працює добре.

– Залежність від сторонніх бібліотек і фреймворків. Існує чимало широко використовуваних ресурсів сторонніх розробників, які по суті не є Pythonic. Це фактично суперечить девізу Python.

3.2 Опис фрагментів лістингу програми

Для того, щоб охарактеризувати роботу по удосконаленню багатофакторної біометричної аутентифікації потрібно розглянути лістинг програми.

```
# import libraries
import tensorflow
from tensorflow import keras
train_images = r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Training Images"
## Image Preprocessing using keras
# As we know deep-learning is hungry for data, the data we have is only limited.
# so lets perform **Image Agumentation** to create different versions
# of the original image, which leads to a better model, since it learns
# on the good and bad mix of images.
from keras.preprocessing.image import ImageDataGenerator
train_gen = ImageDataGenerator(
  shear_range=0.1,
  zoom_range=0.1,
```

```
  horizontal_flip=True
)
# No transformations are made on the test data
test_gen = ImageDataGenerator()
```

Далі потрібно створити архітектуру CNN, провести компіляцію моделі та навчання. Створення архітектури CNN

```
# Generating training data
training_data = train_gen.flow_from_directory(
  train_images,
  target_size = (100,100),
  batch_size = 30,
  class_mode = 'categorical'
)

# generating test data
testing_data = test_gen.flow_from_directory(
  train_images,
  target_size = (100,100),
  batch_size = 30,
  class_mode = 'categorical'
)
```

Компіляція моделі

```
        # Printing class labels for each face
testing_data.class_indices
from keras.models import Sequential
from keras.layers import Conv2D, MaxPool2D, Flatten, Dense
'''Initializing the Convolutional Neural Network'''
Model = Sequential()

''' STEP--1 Convolution
# Adding the first layer of CNN
# we are using the format (100,100,3) because we are using TensorFlow backend
# It means 3 matrix of size (100x100) pixels representing Red, Green and Blue components of pixels
'''
```

Використання класифікатора для прогнозування на невидимих тестових зображеннях

```
        import time
# Measuring the time taken by the model to train
```

```
StartTime=time.time()
'''# Model Training'''
Model.fit_generator(training_data,
        epochs = 30,
        validation_data=testing_data,
             callbacks=call)


Endtime = time.time()
print('Total Training Time taken: ',round((Endtime-StartTime)/60),'Minutes')
```

```
        import numpy as np
from keras.preprocessing import image
import os
import glob
```

Реалізація алгоритму LSTM.

Для реалізації потрібно виконати наступне:

Крок 1. Імпортування необхідних бібліотек

```
import os
import glob
```

numpy: ця бібліотека використовується для чисельних обчислень. Зокрема, це допомагає в обробці матричних і векторних операцій, які є важливими для нейронних мереж.

Крок 2: Визначення класу LSTM

```
        class LSTM:
  def __init__(self, input_size, hidden_size):
    self.input_size = input_size
    self.hidden_size = hidden_size
```

- Клас LSTM : цей клас представляє модель LSTM.
- input_size: кількість вхідних ознак (розмірність кожного вхідного вектора).
- hidden_size**:** кількість прихованих одиниць у LSTM (розмір прихованого стану).

```
    # Initialize weights and biases for forget, input, cell, and output gates
    self.Wf = np.random.randn(hidden_size, hidden_size + input_size)
    self.Wi = np.random.randn(hidden_size, hidden_size + input_size)
    self.WC = np.random.randn(hidden_size, hidden_size + input_size)
    self.Wo = np.random.randn(hidden_size, hidden_size + input_size)
```

Вагові коефіцієнти (Wf, Wi, WC, Wo) : це вагові матриці для чотирьох воріт LSTM (забути шлюз, вхідний шлюз, шлюз комірки та вихідний шлюз). Розміри цих матриць обумовлені (hidden_size, hidden_size + input_size)тим, що вхідні дані для кожного вентиля є конкатенацією попереднього прихованого стану та поточного вхідного вектора.

- Wf: Ваги для Forget Gate.
- Wi: ваги для Input Gate.
- WC: Ваги для Cell Gate .
- Wo: ваги для Output Gate.
- Зміщення (bf, bi, bC, bo) : це вектори зміщення, пов’язані з кожним вентилем, ініціалізовані нулями. Розміри пояснюються (hidden_size, 1)тим, що кожна брама має власний ухил для кожного прихованого блоку.

Крок 4: Визначення функцій активації

```
Model.fit_generator(training_data,
epochs = 30,
validation_data=testing_data,
callbacks=call)

Endtime = time.time()
print('Total Training Time taken: ',round((Endtime-StartTime)/60),'Minutes')
import numpy as np
from keras.preprocessing import image
import os
import glob
```

- sigmoid(z): Ця функція обчислює сигмоїдну активацію, яка використовується в forget, input, output,gates. Сигмоїдна функція стискає значення в діапазоні від 0 до 1.
- tanh(z): ця функція обчислює активацію гіперболічного тангенса, яка використовується для оновлення стану комірки. Функція tanh виводить значення в діапазоні від -1 до 1, що ідеально підходить для керування станом комірки.

Крок 5: Прохід вперед через LSTM (рис. 3.9)

```
x_t = np.random.randn(input_size, 1)
h_prev = np.zeros((hidden_size, 1))
C_prev = np.zeros((hidden_size, 1))
```

forward()метод : цей метод обчислює наступний прихований стан (h_t) і стан комірки (C_t) на основі поточного введення, попереднього прихованого стану та попереднього стану комірки.

– x: поточний вхідний вектор на кроці часу t.

– h_prev: Прихований стан із попереднього кроку часу (t-1).

– C_prev: стан клітинки з попереднього кроку часу (t-1).

– np.vstack((h_prev, x)): Прихований стан і вхідні дані об’єднані вертикально. Це пояснюється тим, що кожен шлюз приймає як прихований стан, так і вхід як об’єднаний вхід.

3.3 Тестування програмного засобу

Для реалізації розпізнавання голосу було підібрано аудіо дані на 40 різних мовах. Відповідно, діаграма об’єму аудіо даних має вигляд:

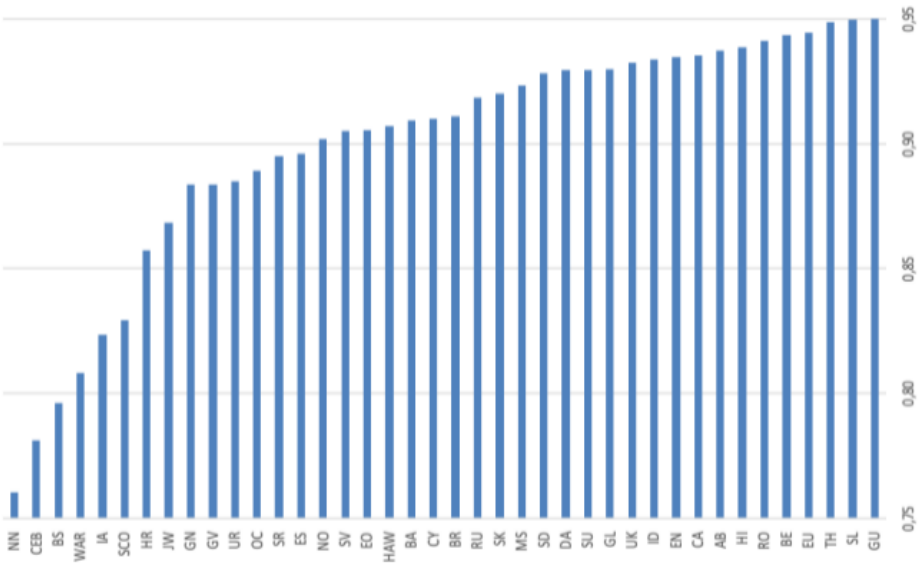


Рис. 3.1 – Діаграма аудіоданих

Рис. 3.1 ілюструє діаграму аудіоданих, яка відображає структуру та характерні особливості звукового сигналу. Вона демонструє ключові параметри, такі як амплітуда, частота та час, які використовуються для аналізу та обробки аудіо в процесі аутентифікації.

Розмір навчальних даних для окремих мов (рис. 3.2):

Код	Зареєстровано, год.	Перевірено, год.	Кількість голосів
SL	14	11	146
TH	420	171	8
EU	159	105	1
BE	1 632	1 586	8 205
RO	44	19	408
HI	20	14	396
AB	85	60	400
CA	3 328	2 554	35 062
EN	3 347	2 532	88 904
ID	64	29	516
UK	105	94	1
GL	61	38	1
DA	13	12	243
MS	10	3	128
SK	27	22	216
RU	260	228	3
BR	26	12	196
CY	155	122	1 800
BA	268	258	912
EO	1 897	1 432	1 682
SV	54	45	808
ES	2 188	526	25 338
SR	6	4	144
OC	13	2	141
UR	221	63	316
GN	25	4	140
IA	17	14	66
NN	2	2	32

Рис. 3.2 – Розмір навчальних даних для окремих мов

Матриця заплутаності мов низької точності має вигляд (рис. 3.3):

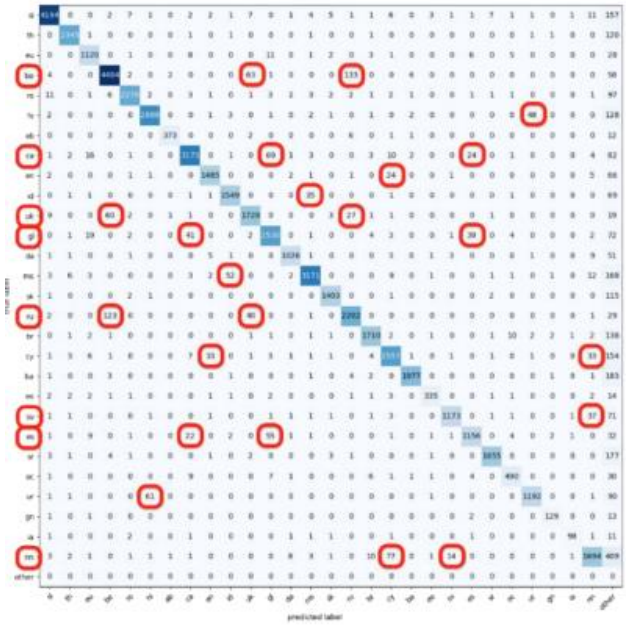


Рис. 3.3 – Матриця заплутаності мов низької точності

Рис. 3.3 представляє матрицю заплутаності для мов із низькою точністю, яка відображає результати класифікації моделей. Вона показує співвідношення між

фактичними та передбаченими класами, вказуючи на кількість правильних і помилкових визначень, що допомагає оцінити ефективність алгоритму та виявити джерела помилок.

Розмір навчальних наборів даних для окремих мов (рис. 3.4):

Мова	Розмір		Набір даних
	год.	екз.	
CS	60,0	23 215	Vox107
	25,5	19 358	CV
	12,3	9 656	CV-Cl
	3,5	2 729	CV-Bal
	1,4	1 128	CV-Cl-Bal
SK	35,8	13 727	Vox107
	3,5	3 276	CV
	1,4	1 389	CV-Cl
	3,5	3 276	CV-Bal
	1,4	1 389	CV-Cl-Bal
SV	30,6	11 475	Vox107
	8,4	7 584	CV
	4,9	4 638	CV-Cl
	0,5	496	CV-Bal
	—	1	CV-Cl-Bal
NO	96,0	37 408	Vox107
	0,5	407	CV
	—	1	CV-Cl
	0,5	407	CV-Bal
	—	1	CV-Cl-Bal
UK	47,2	16 976	Vox107
	22,5	19 024	CV
	10,6	9 313	CV-Cl

Мова	Розмір		Набір даних
	год.	екз.	
UK	22,5	19 024	CV-Bal
	10,6	9 313	CV-Cl-Bal
RU	65,6	21 412	Vox107
	37,7	26 328	CV
	24,8	17 789	CV-Cl
	22,5	15 854	CV-Bal
	10,6	7 487	CV-Cl-Bal
ES	34,7	11 475	Vox107
	450,8	311 392	CV
	297,3	206 270	CV-Cl
	16,5	10 057	CV-Bal
	4,9	2 987	CV-Cl-Bal
CA	79,2	30 460	Vox107
	1743,8	1 142 607	CV
	821,0	546 025	CV-Cl
	16,5	9 625	CV-Bal
	4,9	2 942	CV-Cl-Bal
GL	65,0	24 958	Vox107
	16,5	12 688	CV
	4,9	3 945	CV-Cl
	16,5	12 688	CV-Bal
	4,9	3 945	CV-Cl-Bal

Рис. 3.4 – Розмір навчальних наборів даних

Результати роботи алгоритму CNN (рис. 3.5):

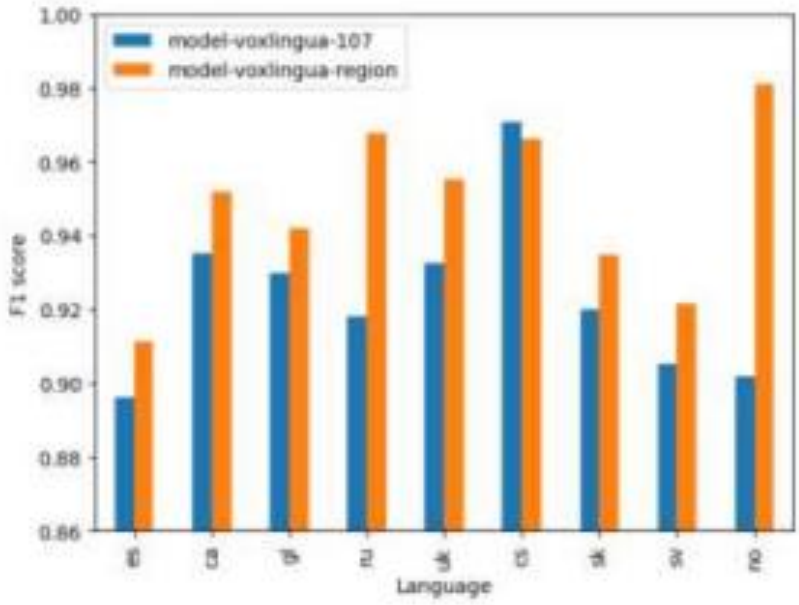


Рис. 3.5 – Діаграма точності для моделі, порівняно з CNN

Рис. 3.5 ілюструє діаграму точності, яка порівнює результати роботи моделі з використанням CNN та іншими методами. Вона демонструє рівень точності

класифікації для кожного підходу, що дозволяє оцінити переваги використання CNN у процесі аутентифікації.

Розглянемо роботу алгоритму LSTM.

Для ідентифікації було обрано наступне зображення (рис. 3.6):

Рисунок 3.6 – Зображення для тестування

Результат роботи алгоритму розпізнавання обличчя (рис. 3.7)

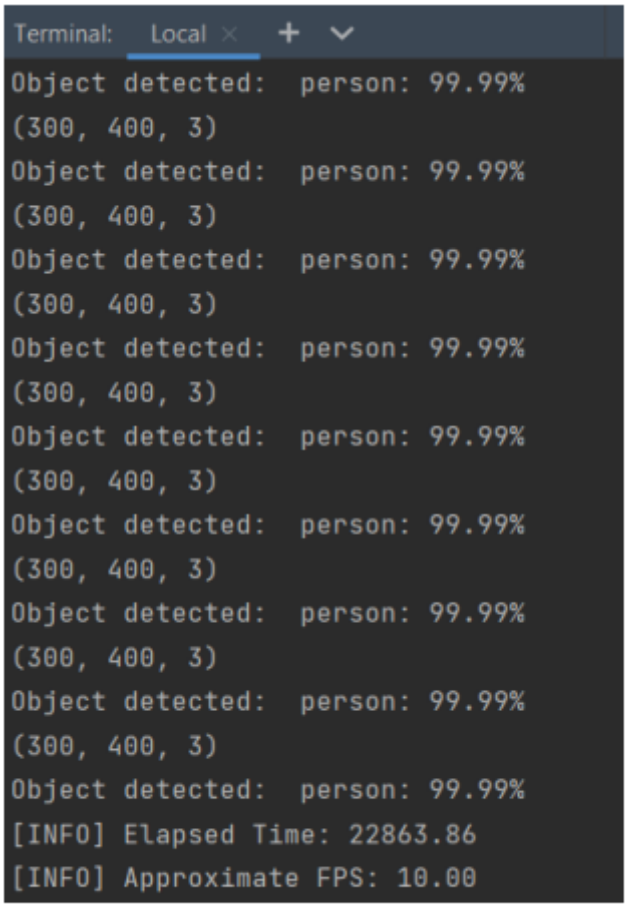


Рис. 3.7 - Результат роботи програми

Відображення результату сканування (рис. 3.8)

Рис. 3.8 - Візуальне відображення роботи алгоритму LSTM

Порівняння результатів найсучасніших методів на контрольних наборах даних (рис. 3.9)

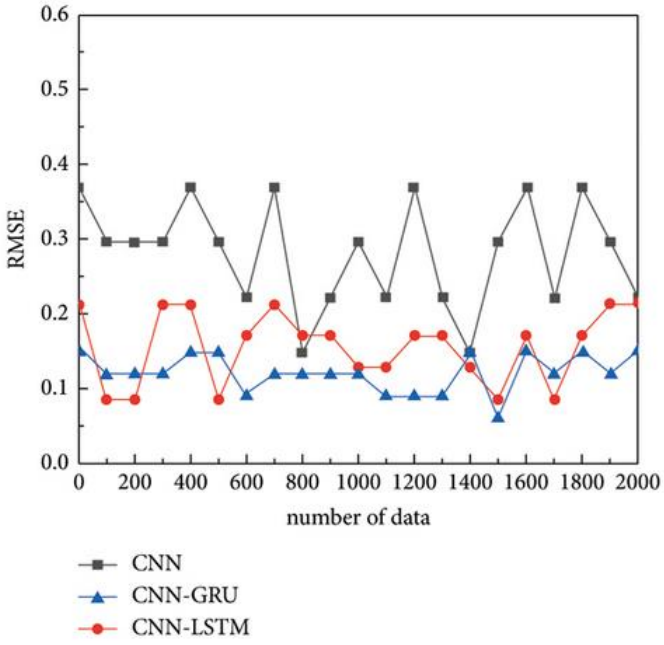


Рис. 3.9 - Порівняння даних роботи різних моделей біометричної аутентифікації

Оцінка точності моделі в порівняння з іншими (рис. 3.10)

	Fingerprint identification	Image recognition	Voice recognition
CNN	0.35	0.58	0.38
CNN-GRU	0.17	0.31	0.22
CNN-LSTM	0.07	0.19	0.15

Рис. 3.10 - Порівняння точності результатів

Отже, судячи з результатів можна зробити висновок що тестування пройшло успішно.

3.4 Висновки до розділу

Було обґрунтовано вибір мови Python для реалізації програмного коду покращення моделі біометричної аутентифікації. Проведено огляд основних бібліотек, які можуть знадобитись при створені програми. Приведені фрагменти коду, які безпосередньо стосуються розробки програмного засобу. За допомогою даних фрагментів було проведено демонстрацію по аналізу та підготовці даних біометричної аутентифікації, виявлення аномалій. Також була проведена оцінка застосування реалізованої моделі в порівняння з іншими.

4. ЕКОНОМІЧНА ЧАСТИНА

У розділі 4 буде оцінено економічний аспект науково-технічної розробки. Спочатку буде проведено оцінювання комерційного потенціалу розробки, визначено її можливості для впровадження на ринку. Далі буде здійснено прогнозування витрат на виконання науково-дослідної роботи та розраховано економічну ефективність розробки, враховуючи потенційні доходи та витрати. Оцінено ефективність вкладених інвестицій і період їх окупності, що дозволить визначити фінансову привабливість проекту. Висновки до розділу підсумують економічні результати та перспективи подальшої реалізації розробки.

4.1 Оцінювання комерційного потенціалу розробки

Метою проведення комерційного та технологічного аудиту є оцінювання комерційного потенціалу інформаційної веб-системи аналізу динаміки географічної структури зовнішньої торгівлі товарами України.

Для проведення технологічного аудиту було залучено 3-х незалежних експертів Вінницького національного технічного університету кафедри системного аналізу та інформаційних технологій: к.т.н., доц. Козачко О.М., к.т.н., доц. Крижановський Є. М., к.т.н., доц. Варчук І. В. Для проведення технологічного аудиту було використано таблицю 4.1 в якій за п’ятибальною шкалою використовуючи 12 критеріїв здійснено оцінку комерційного потенціалу.

Таблиця 4.1 – Рекомендовані критерії оцінювання комерційного потенціалу розробки та їх можлива бальна оцінка

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
Кри-терій	0	1	2	3	4
Технічна здійсненність концепції:					
1	Достовірність концепції не підтверджена	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено робото здатність продукту в реальних умовах
Ринкові переваги (недоліки):					
2	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3	Ціна продукту значно вища за ціни аналогів	Ціна продукту дещо вища за ціни аналогів	Ціна продукту приблизно дорівнює цінам аналогів	Ціна продукту дещо нижче за ціни аналогів	Ціна продукту значно нижче за ціни аналогів
4	Технічні та споживчі властивості продукту значно гірші, ніж в аналогів	Технічні та споживчі властивості продукту трохи гірші, ніж в аналогів	Технічні та споживчі властивості продукту на рівні аналогів	Технічні та споживчі властивості продукту трохи кращі, ніж в аналогів	Технічні та споживчі властивості продукту значно кращі, ніж в аналогів
5	Експлуатаційні витрати значно вищі, ніж в аналогів	Експлуатаційні витрати дещо вищі, ніж в аналогів	Експлуатаційні витрати на рівні експлуатаційних витрат аналогів	Експлуатаційні витрати трохи нижчі, ніж в аналогів	Експлуатаційні витрати значно нижчі, ніж в аналогів
Ринкові перспективи					
6	Ринок малий і не має позитивної динаміки	Ринок малий, але має позитивну динаміку	Середній ринок з позитивною динамікою	Великий стабільний ринок	Великий ринок з позитивною динамікою
7	Активна конкуренція великих компаній на ринку	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкурентів немає
Практична здійсненність					
8	Відсутні фахівці як з технічної, так і з комерційної реалізації ідеї	Необхідно наймати фахівців або витрачати значні кошти та час на навчання наявних фахівців	Необхідне незначне навчання фахівців та збільшення їх штату	Необхідне незначне навчання фахівців	Є фахівці з питань як з технічної, так і з комерційної реалізації ідеї

Продовження таблиці 4.1

9	Потрібні значні фінансові ресурси, які відсутні. Джерела фінансування ідеї відсутні	Потрібні незначні фінансові ресурси. Джерела фінансування відсутні	Потрібні значні фінансові ресурси. Джерела фінансування є	Потрібні незначні фінансові ресурси. Джерела фінансування є	Не потребує додаткового фінансування
10	Необхідна розробка нових матеріалів	Потрібні матеріали, що використовуються у військово-промисловому комплексі	Потрібні дорогі матеріали	Потрібні досяжні та дешеві матеріали	Всі матеріали для реалізації ідеї відомі та давно використовуються у виробництві
11	Термін реалізації ідеї більший за 10 років	Термін реалізації ідеї більший за 5 років. Термін окупності інвестицій більше 10-ти років	Термін реалізації ідеї від 3-х до 5-ти років. Термін окупності інвестицій більше 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій від	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій менше 3-х років

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
Кри-терій	0	1	2	3	4
				3-х до 5-ти років	
12	Необхідна розробка регламентних документів та отримання великої кількості дозвільних документів на виробництво та реалізацію продукту	Необхідно отримання великої кількості дозвільних документів на виробництво та реалізацію продукту, що вимагає значних коштів та часу	Процедура отримання дозвільних документів для виробництва та реалізації продукту вимагає незначних коштів та часу	Необхідно тільки пові-домлення відповідним органам про виробництво та реалізацію продукту	Відсутні будь-які регламентні обмеження на виробництво та реалізацію продукту

Таблиця 4.2 – Рівні комерційного потенціалу розробки

Середньоарифметична сума балів СБ, розрахована на основі висновків експертів	Рівень комерційного потенціалу розробки
0-10	Низький
11-20	Нижче середнього
21-30	Середній
31-40	Вище середнього
41-48	Високий

В таблиці 4.3 наведено результати оцінювання експертами комерційного потенціалу розробки.

Таблиця 4.3 – Результати оцінювання комерційного потенціалу розробки

Критерії	Прізвище, ініціали, посада експерта		
	Козачко О.М.	Крижановський Є.М.	Варчук І.В.
	Бали, виставлені експертами:		
1	1	2	4
2	4	2	3
3	3	3	2
4	2	5	2
5	4	3	2
6	4	3	2
7	1	5	2
8	1	3	1
9	1	3	2
10	2	1	4
11	5	3	5
12	3	3	3
Сума балів	СБ1=31	СБ2=36	СБ3=32
Середньоарифметична сума балів $\overline{СБ}$	$\overline{СБ} = \frac{\sum_1^3 СБ_i}{3} = \frac{31+36+32}{3} = 33$		

Середньоарифметична сума балів, розрахована на основі висновків експертів склала 33 бали, що згідно таблиці 4.2 вважається, що рівень комерційного потенціалу проведених досліджень є вище середнього.

Підвищення захищеності IoT мереж на основі аналізу мережевого трафіку і машинного навчання з використанням вдосконаленого алгоритму глибокого виявлення аномалій із застосуванням адаптивного порогу виявлення та алгоритму навчання з підкріпленням Deep Q-Net, а саме програма, яка допомагає проводити моніторинг, аналізувати та порівнювати дані мережевого трафіку, виявляти аномальний трафік та реагувати на потенційні загрози. Дана програма буде цікава різним організаціям та компаніям які прагнуть підвищити рівень захисту внутрішньої інформації від можливих витоків та кібератак.

4.2 Прогнозування витрат на виконання науково-дослідної роботи

Витрати, пов’язані з проведенням науково-дослідної роботи групуються за такими статтями: витрати на оплату праці, витрати на соціальні заходи, матеріали, паливо та енергія для науково-виробничих цілей, витрати на службові відрядження, програмне забезпечення для наукових робіт, інші витрати, накладні витрати.

1. Основна заробітна плата кожного із дослідників $З_0$, якщо вони працюють в наукових установах бюджетної сфери визначається за формулою:

$$З_0 = \frac{М}{Т_р} * t \text{ (грн)}, \quad (4.1)$$

де M – місячний посадовий оклад конкретного розробника (інженера, дослідника, науковця тощо), грн.;

T_P – число робочих днів в місяці; приблизно $T_P \approx 21...23$ дні;

t – число робочих днів роботи дослідника.

Для розробки програмні засоби необхідно залучити програміста з посадовим окладом 10000 грн. Кількість робочих днів у місяці складає 40, а кількість робочих днів програміста складає 22. Зведемо сумарні розрахунки до таблиця 4.4.

Таблиця 4.4 – Заробітна плата дослідника в науковій установі бюджетної сфери

Найменування посади	Місячний посадовий оклад, грн.	Оплата за робочий день, грн.	Число днів роботи	Витрати на заробітну плату грн.
Керівник	20000	909,1	5	4545
Програмний інженер	15000	681,8	40	27272
Всього				31817

2. Розрахунок додаткової заробітної плати робітників

Додаткова заробітна плата $З_д$ всіх розробників та робітників, які приймали устають в розробці нового технічного рішення розраховується як 10 - 12 % від основної заробітної плати робітників.

На даному підприємстві додаткова заробітна плата начисляється в розмірі 12% від основної заробітної плати.

$$З_д = (З_о + З_р) * \frac{Н_{дод}}{100\%} \quad (4.2)$$

3. Нарахування на заробітну плату $Н_{ЗП}$ дослідників та робітників, які брали участь у виконанні даного етапу роботи, розраховуються за формулою:

$$Н_{ЗП} = (З_о + З_д) * \frac{\beta}{100}, \quad (4.3)$$

$$З_д = 0{,}1 * 31817 = 3181 \text{ (грн)}.$$

3. Нарахування на заробітну плату $Н_{ЗП}$ дослідників та робітників, які брали участь у виконанні даного етапу роботи, розраховуються за формулою:

$$Н_{ЗП} = (З_о + З_д) * \frac{\beta}{100}, \quad (4.3)$$

де $З_о$ – основна заробітна плата розробників, грн.;

$З_д$ – додаткова заробітна плата всіх розробників та робітників, грн.;

β – ставка єдиного внеску на загальнообов'язкове державне соціальне страхування, % .

Дана діяльність відноситься до бюджетної сфери, тому ставка єдиного внеску на загальнообов'язкове державне соціальне страхування буде складати 22%, тоді**:**

$$Н_{ЗП} = (31817 + 3181\ ) * \frac{22}{100} = 7700 \text{ (грн)}.$$

4. Витрати на комплектуючі вироби, які використовують при виготовленні одиниці продукції, розраховуються, згідно їх номенклатури, за формулою:

$$К = \sum_{i=1}^{n} Н_i * Ц_i * К_i , \quad (4.5)$$

де $Н_i$ – кількість комплектуючих *i*-го виду, шт.;

$Ц_i$ – покупна ціна комплектуючих *i*-го найменування, грн.;

$К_i$ – коефіцієнт транспортних витрат (1,1…1,15).

Таблиця 4.5 – Комплектуючі, що використані на розробку

Найменування матеріалу	Ціна за одиницю, грн.	Витрачено	Вартість витраченого матеріалу, грн.
Папір	150	1	150
Ручка	20	1	20
CD-диск	30	1	30
Флешка	150	1	150
Всього			350
З врахуванням коефіцієнта транспортування			400

5. Програмне забезпечення для наукової роботи включає витрати на розробку та придбання спеціальних програмних засобів і програмного забезпечення необхідного для проведення дослідження.

Для написання магістерської роботи використовувалися редактор VsCode та онлайн база даних MongoDB.

6. Амортизація обладнання, комп’ютерів та приміщень, які використовувались під час виконання даного етапу роботи

Дані відрахування розраховують по кожному виду обладнання, приміщенням тощо.

$$A = \frac{Ц*Т}{Т_{кор}*12}, \quad (4.6)$$

де Ц – балансова вартість даного виду обладнання (приміщень), грн.;

$Т_{кор}$ – час користування;

Т – термін використання обладнання (приміщень), цілі місяці.

Згідно пункта 137.3.3 Податкового кодекса амортизація нараховується на основні засоби вартістю понад 2500 грн. В нашому випадку для написання магістерської роботи використовувався персональний комп'ютер вартістю 30000 грн.

$$А = \frac{30000 \cdot 1}{2 \cdot 12} = 1150 \text{ (грн)}.$$

7. До статті «Паливо та енергія для науково-виробничих цілей» відносяться витрати на всі види палива й енергії, що безпосередньо використовуються з технологічною метою на проведення досліджень.

$$В_е = \sum_{i=1}^{n} \frac{W_{yt} \cdot t_i \cdot Ц_е \cdot К_{впі}}{\eta_і}, \qquad (4.7)$$

де W_{yt} – встановлена потужність обладнання на певному етапі розробки, кВт;

t_i – тривалість роботи обладнання на етапі дослідження, год;

$Ц_е$ – вартість 1 кВт-години електроенергії, грн;

$К_{впі}$ – коефіцієнт, що враховує використання потужності, $К_{впі} < 1$;

$\eta_і$ – коефіцієнт корисної дії обладнання, $\eta_і < 1$.

Для написання магістерської роботи використовується персональний комп'ютер для якого розрахуємо витрати на електроенергію.

$$В_е = \frac{0{,}3 \cdot 200 \cdot 1{,}68 \cdot 0{,}5}{0{,}8} = 63 \text{ (грн)}.$$

Витрати на службові відрядження, витрати на роботи, які виконують сторонні підприємства, установи, організації та інші витрати в нашому досліджені не враховуються оскільки їх не було.

Накладні (загальновиробничі) витрати Внзв охоплюють: витрати на управління організацією, оплата службових відряджень, витрати на утримання, ремонт та експлуатацію основних засобів, витрати на опалення, освітлення, водопостачання, охорону праці тощо. Накладні (загальновиробничі) витрати Внзв можна прийняти як (100…150)% від суми основної заробітної плати розробників та робітників, які виконували дану МКНР, тобто:

$$В_{нзв} = (З_о + З_р) \cdot \frac{Н_{нзв}}{100\%}, \quad (4.8)$$

де $Н_{нзв}$ – норма нарахування за статтею «Інші витрати».

$$В_{нзв} = 31817 \cdot \frac{100}{100\%} = 31817(грн).$$

Сума всіх попередніх статей витрат дає витрати, які безпосередньо стосуються даного розділу МКНР

$$В = 31817+3181+7700+400+1150+63+31817=76128 \text{ (грн)}.$$

Прогнозування загальних втрат ЗВ на виконання та впровадження результатів виконаної МКНР здійснюється за формулою:

$$ЗВ = \frac{В}{\eta}, \quad (4.9)$$

де η – коефіцієнт, який характеризує стадію виконання даної НДР.

Оскільки, робота знаходиться на стадії науково-дослідних робіт, то коефіцієнт β = 0,9.

Звідси:

$$ЗВ = \frac{76128}{0,9} = 84586,7(грн).$$

4.3 Розрахунок економічної ефективності науково-технічної розробки

У даному підрозділі кількісно спрогнозуємо, яку вигоду, зиск можна отримати у майбутньому від впровадження результатів виконаної наукової роботи. Розрахуємо збільшення чистого прибутку підприємства $\Delta П_i$, для кожного із років, протягом яких очікується отримання позитивних результатів від впровадження розробки, за формулою

$$\Delta П_i = \sum_1^n (\Delta Ц_o * N * Ц_o * \Delta N)_i * \lambda * \rho * \left(1 - \frac{\upsilon}{100}\right), \quad (4.10)$$

де $\Delta Ц_0$ – покращення основного оціночного показника від впровадження результатів розробки у даному році.

N – основний кількісний показник, який визначає діяльність підприємства у даному році до впровадження результатів наукової розробки;

ΔN – покращення основного кількісного показника діяльності підприємства від впровадження результатів розробки:

$Ц_o$ – основний оціночний показник, який визначає діяльність підприємства у даному році після впровадження результатів наукової розробки;

n – кількість років, протягом яких очікується отримання позитивних результатів від впровадження розробки:

λ – коефіцієнт, який враховує сплату податку на додану вартість. Ставка податку на додану вартість дорівнює 20%, а коефіцієнт λ = 0,8333.

ρ – коефіцієнт, який враховує рентабельність продукту. ρ = 0,25;

υ – ставка податку на прибуток. У 2022 році – 18%.

Припустимо, що ціна за програмний продукт зросте на 600 грн. Кількість одиниць реалізованої продукції також збільшиться: протягом першого року на 45 шт., протягом другого року – на 35 шт., протягом третього року на 25 шт. Реалізація продукції до впровадження розробки складала 1 шт., а її ціна до складає 10000 грн. Розрахуємо прибуток, яке отримає підприємство протягом трьох років.

$$\Delta П_1 = [600 \cdot 1 + (10000 + 600) \cdot 45] \cdot 0{,}833 \cdot 0{,}25 \cdot \left(1 + \frac{18}{100}\right) = 117815.59 \text{ (грн)}.$$

$$\Delta П_2 = [600 \cdot 1 + (10000 + 600) \cdot (45 + 35)] \cdot 0{,}833 \cdot 0{,}25 \cdot \left(1 + \frac{18}{100}\right) = 208983.28 \text{ (грн)}.$$

$$\Delta П_3 = [600 \cdot 1 + (10000 + 600) \cdot (45 + 35 + 25)] \cdot 0{,}833 \cdot 0{,}25 \cdot \left(1 + \frac{18}{100}\right) = 274103.05 \text{(грн)}.$$

4.4 Розрахунок ефективності вкладених інвестицій та періоду їх окупності

Розрахуємо основні показники, які визначають доцільність фінансування наукової розробки певним інвестором, є абсолютна і відносна ефективність вкладених інвестицій та термін їх окупності.

Розрахуємо величину початкових інвестицій PV, які потенційний інвестор має вкласти для впровадження і комерціалізації науково-технічної розробки.

$$PV = k_{\text{інв}} \cdot \text{ЗВ}, \qquad (4.11)$$

де $k_{\text{інв}}$ – коефіцієнт, що враховує витрати інвестора на впровадження науково-технічної розробки та її комерціалізацію. Це можуть бути витрати на підготовку

приміщень, розробку технологій, навчання персоналу, маркетингові заходи тощо ($k_{\text{інв}} = 2…5$).

$$PV = 2 \cdot 84586{,}7 = 169173.4(\text{грн}).$$

Розрахуємо абсолютну ефективність вкладених інвестицій $E_{абс}$ згідно наступної формули:

$$\text{Е}_{\text{абс}} = (\text{ПП} - PV), \tag{4.12}$$

де ПП – приведена вартість всіх чистих прибутків, що їх отримає підприємство від реалізації результатів наукової розробки, грн.;

$$\text{ПП} = \sum_1^T \frac{\Delta\text{П}_\text{і}}{(1+\tau)^t}\ , \tag{4.13}$$

де $\Delta\text{П}_\text{і}$ – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої НДЦКР, грн.;

Т – період часу, протягом якою виявляються результати впровадженої НДДКР, роки;

τ – ставка дисконтування, за яку можна взяти щорічний прогнозований рівень інфляції в країні; для України цей показник знаходиться на рівні 0,2;

t – період часу (в роках).

$$\text{ПП} = \frac{117815.59}{(1+0{,}2)^1} + \frac{208983.28}{(1+0{,}2)^2} + \frac{274103.05}{(1+0{,}2)^3} = 401931.37\ (\text{грн}).$$

$$\text{Е}_{\text{абс}} = (401931.37 - 169173.4) = 232757.97\ (\text{грн}).$$

Оскільки $Е_{абс} > 0$, то вкладання коштів на виконання та впровадження результатів НДДКР може бути доцільним.

Розрахуємо відносну (щорічну) ефективність вкладених в наукову розробку інвестицій $Е_в$. Для цього користуються формулою:

$$Е_в = \sqrt[Т_ж]{1 + \frac{Е_{абс}}{PV}} - 1 , \quad (4.14)$$

де $Т_ж$ – життєвий цикл наукової розробки, роки.

$$Е_в = \sqrt[3]{1 + \frac{232757.97}{169173.4}} - 1 = 0,26 = 26\%.$$

Визначимо мінімальну ставку дисконтування, яка у загальному вигляді визначається за формулою:

$$\tau = d + f , \quad (4.15)$$

де d – середньозважена ставка за депозитними операціями в комерційних банках; в 2022 році в Україні d = (0,14…0,2);

f – показник, що характеризує ризикованість вкладень; зазвичай, величина $f =$ (0,05…0,1).

$$\tau_{min} = 0,16 + 0,05 = 0,21.$$

Так як $E_в > \tau_{min}$ то інвестор може бути зацікавлений у фінансуванні даної наукової розробки.

Розрахуємо термін окупності вкладених у реалізацію наукового проекту інвестицій за формулою:

$$Т_{ок} = \frac{1}{Е_в} . \quad (4.16)$$

$$T_{ок} = \frac{1}{0,26} = 3,8 \text{ (роки)}.$$

Так як $T_{ок} \leq 3...5$-ти років, то фінансування даної наукової розробки в принципі є доцільним.

4.5 Висновки до розділу

Проведено оцінку комерційного потенціалу удосконалення багатофакторної біометричної аутентифікації з використання CNN та LSTM.

Прогнозування витрат на виконання науково-дослідної роботи по кожній з статей витрат складе 76128 грн. Загальна ж величина витрат на виконання та впровадження результатів даної НДР буде складати 84586,7 грн.

Вкладені інвестиції в даний проект окупляться через 3,8 роки, приведена вартість всіх чистих прибутків, що їх отримає підприємство від реалізації результатів наукової розробки склала 401931,37 грн.

ВИСНОВКИ

В результаті виконання кваліфікаційної роботи було зроблено наступне.

В першому розділі роботі було розглянуто теоретичні аспекти поняття біометрична аутентифікація, описані основні складові роботи таких систем, описані приклади застосування в різних сферах життєдіяльності. Також в рамках даного розділу були розглянуті фундаментальні методи реалізаці біометричної аутентифікації мереж.

В другому розділі роботи було розглянуто методологічні аспекти проблеми біометричної аутентифікації на прикладі застосування CNN та LSTM. Виділено основні моделі, які можуть застосовуватись для ідентифікації. Але, варто зазначити, що не існує однозначного та чіткого підходу до формалізації безпеки біометричної аутентифікації. Всі представлені способи відображають тільки окремі аспекти, а не цілісний погляд на проблему.

В третьому розділі роботи було обґрунтовано вибір мови програмування Python, описані основні переваги та недоліки, ключові бібліотеки, які допоможуть реалізувати представлені кроки алгоритму. Також, були представлені фрагменти коду майбутнього алгоритму з необхідними поясненнями. В кінці розділу було наведено деякі результати реалізації роботи алгоритму.

В четвертому розділу було приведене економічне обґрунтування розробки алгоритму підвищення захищеності IoT мереж. Розраховані всі економічні показники. За даними показниками, період окупності даного продукту складає 3,8 року, а отриманий прибуток складе 401931,37 грн.

Таким чином, питання знаходження удосконалення багатофакторної біометричної аутентифікації є постійно актуальним та вимагає впровадження нових засобів, методів, поєднання вже існуючих методів з їх подальшою модернізацією, для оперативного реагування на виникаючі загрози.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Y. Taigman, M. Yang, Marc'Aurelio Ranzato, Lior Wolf. DeepFace: "Closing the Gap to HumanLevel Perfomance in Face Verification" IEEE Conference on 2014, Computer Vision and Pattern Recognition – Columbus, OH, USA.

2. Аналіз існуючих підходів до розпізнавання URL: https://habr.com/ru/company/synesis/blog/238129 (дата звернення 25.10.2021)

3. Lubchenko, V., & Podvalnyi, Y. (2021). Finding and classification of objects.

4. Ciaparrone, G., Sánchez, F. L., Tabik, S., Troiano, L., Tagliaferri, R., & Herrera, F. (2020). Deep learning in video multi-object tracking: A survey. Neurocomputing, 381, pp. 61-88.

5. Зайцева, Е. В., & Степанова, А. Л. (2013). Классификация современных методов трекинга объектов в интеллектуальных системах видеонаблюдения. Горный информационно-аналитический бюллетень (научно-технический журнал), (5).

6. Lubchenko, V., & Podvalnyi, Y. (2021). Tools for following (tracking) objects.

7. Schulter, S., Vernaza, P., Choi, W., & Chandraker, M. (2017). Deep network flow for multi-object tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition - construct an end-to-end deep learning min-cost network flow, pp. 2730–2739.

8. Chu, Q., Ouyang, W., Li, H., Wang, X., Liu, B., & Yu, N. (2017). Online multi-object tracking using CNN-based single object tracker with spatial-temporal attention mechanism. In Proceedings of the IEEE international conference on computer vision, pp. 4836-4845.

9. Jiang, M. X., Deng, C., Pan, Z. G., Wang, L. F., & Sun, X. (2018). Multiobject tracking in videos based on lstm and deep reinforcement learning. Complexity, 2018.

10. Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., & Qu, R. (2019). A survey of deep learning-based object detection. IEEE access, 7, 128837-128868.

11. Bullinger, S., Bodensteiner, C., & Arens, M. (2017, September). Instance flow based online multiple object tracking. In 2017 IEEE International Conference on Image Processing (ICIP) IEEE, pp. 785-789.

12. Sheng, H., Zhang, Y., Chen, J., Xiong, Z., & Zhang, J. (2018). Heterogeneous association graph fusion for target association in multiple object tracking. IEEE Transactions on Circuits and Systems for Video Technology, 29(11), pp. 3269-3280.

13. Chen, J., Sheng, H., Zhang, Y., & Xiong, Z. (2017). Enhancing detection model for multiple hypothesis tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 18-27.

14. Wojke, N., Bewley, A., & Paulus, D. (2017, September). Simple online and realtime tracking with a deep association metric. In 2017 IEEE international conference on image processing (ICIP). IEEE, pp. 3645-3649.

15. Kim, S. J., Nam, J. Y., & Ko, B. C. (2018). Online tracker optimization for multi-pedestrian tracking using a moving vehicle camera. IEEE Access, 6, pp. 48675-48687.

16. Fang, K., Xiang, Y., Li, X., & Savarese, S. (2018, March). Recurrent autoregressive networks for online multi-object tracking. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, pp. 466-475.

17. Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN: towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence, 39(6), pp. 1137-1149.

18. Ren, S., He, K., GirshicN, R., & Sun, J. (2017). 9Faster RYCNN: towards RealY tiMe Object Detection with Region proPosal NetworNs, 9 in IEEE transactions on pattern analysis and machine Intelligence, vol. 39, no. 6. PP. 1137Y1149, 1.

19. Lu, Y., Lu, C., & Tang, C. K. (2017). Online video object detection using association LSTM. In Proceedings of the IEEE International Conference on Computer Vision, pp. 2344-2352.

20. Sharma, S., Ansari, J. A., Murthy, J. K., & Krishna, K. M. (2018, May). Beyond pixels: Leveraging geometry and shape cues for online multi-object tracking. In 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 3508-3515.

21. Ma, C., Yang, C., Yang, F., Zhuang, Y., Zhang, Z., Jia, H., & Xie, X. (2018, July). Trajectory factory: Tracklet cleaving and re-connection by deep siamese bi-gru for

multiple object tracking. In 2018 IEEE International Conference on Multimedia and Expo (ICME). IEEE, pp. 1-6

22. Alvar, S. R., & Bajić, I. V. (2018, August). MV-YOLO: Motion vectoraided tracking by semantic object detection. In 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP). IEEE, pp. 1-5.

23. Hossain, S., & Lee, D. J. (2019). Deep learning-based real-time multipleobject detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices. Sensors, 19(15), 3371.

24. Алгоритми детекції в відеопотоці, 2018. URL:https://masters.donntu.org/2018/fknt/konoshenko/library/modifitsirova nnyy-algoritm-detektsii-lits-v-videopotoke-i-ego-programmnaya realizatsiya.pdf

25. Bernardin, K., Elbs, A., & Stiefelhagen, R. (2006, May). Multiple object tracking performance metrics and evaluation in a smart room environment. In Sixth IEEE International Workshop on Visual Surveillance, in conjunction with ECCV (Vol. 90, No. 91). Citeseer

26. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision. Springer, Cham, pp. 740-755.

27. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788.

28. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

29. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

30. Zhou, X., Wang, D., & Krähenbühl, P. (2019). Objects as points. arXiv preprint arXiv:1904.07850

31. Продан Т.І., Івасьєв С.В., Сучасні методи біометричної ідентифікації. Збірник матеріалів проблемно-наукової міжгалузевої конференції «Автоматизація та комп'ютерно – інтегровані технології» (АКІТ -2022), Тернопіль, 2022. 62-65 с.

32. Кузик В.М., Продан Т.І., Івасьєв С.В., Слєпцова О.Я. Біометрична система аутентифікації з використанням голосових даних. Збірник матеріалів науково-практичної конференції молодих вчених, аспірантів та студентів «Кібербезпека та компютерно-інтегровані технології»(КБКІТ-2020), Тернопіль, 2020. – С.56-59.

33. Шкіра Ю.Р. , Гевко Н.І., Гавриків Н.Г., Осадчук О.Я. Алгоритми та засоби автоматичної детекції обличчя в відео потоці. Збірник матеріалів науковопрактичної конференції молодих вчених, аспірантів та студентів «Кібербезпека та компютерно-інтегровані технології»(КБКІТ-2020), Тернопіль, 2020. – С.19.

ДОДАТКИ

Додаток А. Технічне завдання

Вінницький національний технічний університет

Факультет менеджменту та інформаційної безпеки

Кафедра менеджменту та безпеки інформаційних систем

ЗАТВЕРДЖУЮ
Голова секції "Управління інформаційною безпекою" кафедри МБІС
д.т.н., професор
_____ Юрій ЯРЕМЧУК
"20" вересня 2024 р.

ТЕХНІЧНЕ ЗАВДАННЯ

до магістерської кваліфікаційної роботи на тему:

Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-team memory (LSTM) для поєднання статичних і динамічних ознак

08-72.МКР.000.00.095.ТЗ

Керівник магістерської кваліфікаційної роботи
к.ф.м.н., доцент Шиян А.А.
_____

Вінниця – 2024 р.

1. Найменування та область застосування

Програмний засіб для Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-team memory (LSTM) для поєднання статичних і динамічних ознак. Область застосування: захист інформаційних ресурсів від несанкціонованого доступу у системах безпеки.

2.Підстава для розробки

Розробка виконується на основі наказу ректора ВНТУ №310 від 17. 09. 2024 р.

3. Мета та призначення розробки

3.1 Мета розробки: розробка удосконаленого алгоритму біометричної аутентифікації

3.2 Призначення: розроблений програмний засіб покращує захист особистих даних.

4. Джерела розробки

4.1. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

4.2. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934

4.3. Zhou, X., Wang, D., & Krähenbühl, P. (2019). Objects as points. arXiv preprint arXiv:1904.07850

5. Вимоги до програми

5.1 Вимоги до функціональних характеристик:

5.1.1 Програмний засіб повинен мати зручний, легкий у використанні інтерфейс користувача;

5.1.2 Реалізація методу не повинна вимагати спеціальних ліцензійних програмних додатків;

5.1.3 Програмний засіб повинен виконувати процес автентифікації користувачів у системі.

5.2 Вимоги до надійності:

5.2.1 Програмний засіб повинен працювати без помилок, у випадку виникнення критичних ситуацій необхідно передбачити виведення відповідних повідомлень;

5.2.2 Бази даних повинні бути налаштовані на автоматичне створення резервних копій;

5.2.3 Програмний засіб повинен виконувати свої функції.

5.3 Вимоги до складу і параметрів технічних засобів:

– процесор – Pentium 1500 МГц і подібні до них;

– оперативна пам’ять – не менше 512 Mb;

– середовище функціонування – операційна система сімейство Windows;

– вимоги до техніки безпеки при роботі з програмою повинні відповідати існуючим вимогам та стандартам з техніки безпеки при користуванні комп’ютерною технікою.

6. Вимоги до програмної документації

6.1 Обов’язкова поетапна інструкція для майбутніх користувачів, наведена у пункті 3.4

7. Вимоги до технічного захисту інформації

7.1 Необхідно забезпечити захист розроблюваного програмного засобу від несанкціонованого використання.

7.2 Неможливість отримання доступу незареєстрованих користувачів до інформаційних ресурсів.

8. Техніко-економічні показники

8.1 Цінність результатів використання даного проекту повинна перевищувати витрати на його реалізацію.

8.2 Має бути реалізований таким чином, щоб підходити для використання широкого загалу.

9. Стадії та етани розробки

№ з/п	Назва етапів магістерської кваліфікаційної роботи	Початок	Закінчення
1	Визначення напрямку магістерської роботи, формулювання теми		
2	Аналіз предметної області обраної теми		
3	Апробація отриманих результатів		
4	Розробка алгоритму роботи		
5	Написання магістерської роботи на основі розробленої теми		
6	Розробка економічної частини		
7	Передзахист магістерської кваліфікаційної роботи		
8	Виправлення, уточнення, корегування магістерської кваліфікаційної роботи		
9	Захист магістерської кваліфікаційної Роботи		

10. Порядок контролю та прийому

10.1 До приймання магістерської кваліфікаційної роботи надається:

– ПЗ до магістерської кваліфікаційної роботи;

– програмний додаток;

– презентація;

– відзив керівника роботи;

– відзив опонента

Технічне завдання до виконання прийняв _____ Кухарук П.О.

Додаток Б. Лістинг програми

```
        # import libraries
import tensorflow
from tensorflow import keras

train_images = r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Training Images"

## Image Preprocessing using keras

# As we know deep-learning is hungry for data, the data we have is only limited.
# so lets perform **Image Agumentation** to create different versions
# of the original image, which leads to a better model, since it learns
# on the good and bad mix of images.

from keras.preprocessing.image import ImageDataGenerator
train_gen = ImageDataGenerator(
   shear_range=0.1,
   zoom_range=0.1,
   horizontal_flip=True
)
# No transformations are made on the test data
test_gen = ImageDataGenerator()

# Generating training data
training_data = train_gen.flow_from_directory(
   train_images,
   target_size = (100,100),
   batch_size = 30,
   class_mode = 'categorical'
)

# generating test data
testing_data = test_gen.flow_from_directory(
   train_images,
   target_size = (100,100),
   batch_size = 30,
   class_mode = 'categorical'
)
```

```
# Printing class labels for each face
testing_data.class_indices

        from keras.models import Sequential
from keras.layers import Conv2D, MaxPool2D, Flatten, Dense

'''Initializing the Convolutional Neural Network'''
Model = Sequential()

''' STEP--1 Convolution
# Adding the first layer of CNN
# we are using the format (100,100,3) because we are using TensorFlow backend
# It means 3 matrix of size (100x100) pixels representing Red, Green and Blue components of pixels
'''

Model.add(Conv2D(16, kernel_size=(5,5), strides=(1,1), input_shape = (100,100,3),activation='relu'))

'''# Maxplooing layer'''

Model.add(MaxPool2D(pool_size=(2,2)))

'''Adding additional layers  of convolution and
    maxpooling for better model accuracy and performance'''

Model.add(Conv2D(32,kernel_size=(3,3),strides=(1,1),activation='relu'))
Model.add(MaxPool2D(pool_size=(2,2)))

'''# Add a flatten layer to convert the vector to one dimensional'''
Model.add(Flatten())

'''Add dense layers and Initialize weights using
        kernal initializer for better learing of image features and classification'''

Model.add(Dense(64,activation='relu'))
Model.add(Dense(Output_Neurons,activation='softmax'))

'''Perform Model Compilation'''

Model.compile(loss='categorical_crossentropy',optimizer = 'adam',metrics = ['Accuracy'])
```

```
'''# Using Early stopping to reduce the training time'''

from keras.callbacks import EarlyStopping
call = EarlyStopping(
                min_delta=0.005,
                patience=5,
                 verbose=1
                )

import time
# Measuring the time taken by the model to train
StartTime=time.time()

'''# Model Training'''
Model.fit_generator(training_data,
        epochs = 30,
        validation_data=testing_data,
              callbacks=call)

Endtime = time.time()
print('Total Training Time taken: ',round((Endtime-StartTime)/60),'Minutes')
             import numpy as np
from keras.preprocessing import image
import os
import glob

'''########### Making single predictions ###########'''

ImagePath=r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Training
Images\face4\image_0054_Face_1.jpg"
test_image=image.load_img(ImagePath,target_size=(100, 100))
test_image=image.img_to_array(test_image)

test_image=np.expand_dims(test_image,axis=0)

result=Model.predict(test_image,verbose=0)
#print(training_set.class_indices)
```

```
print('####'*10)
print('Prediction is: ',Result_class[np.argmax(result)])

'''############ Making multiple predictions ###########'''

## Loading all the image paths from final testing folder for prediction
main_ = r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Testing Images"
img_paths = glob.glob(os.path.join(main_,'**','*.jpg'))

print(img_paths[0:5]) # every image will be a PIL object
print('*'*50)

for path in img_paths:
    test_image = image.load_img(path,target_size=(100,100))
    test_image = image.img_to_array(test_image)
    test_image = np.expand_dims(test_image,axis =0)
    result = Model.predict(test_image,verbose=0)
    print('Prediction: ',Result_class[np.argmax(result)])

            import numpy as np

class LSTM:
    def __init__(self, input_size, hidden_size):
        self.input_size = input_size
        self.hidden_size = hidden_size

        # Initialize weights and biases for forget, input, cell, and output gates
        self.Wf = np.random.randn(hidden_size, hidden_size + input_size)
        self.Wi = np.random.randn(hidden_size, hidden_size + input_size)
        self.WC = np.random.randn(hidden_size, hidden_size + input_size)
        self.Wo = np.random.randn(hidden_size, hidden_size + input_size)

        self.bf = np.zeros((hidden_size, 1))
        self.bi = np.zeros((hidden_size, 1))
        self.bC = np.zeros((hidden_size, 1))
        self.bo = np.zeros((hidden_size, 1))

    def sigmoid(self, z):
```

```
        return 1 / (1 + np.exp(-z))

    def tanh(self, z):
        return np.tanh(z)

    def forward(self, x, h_prev, C_prev):
        # Concatenate hidden state and input
        concat = np.vstack((h_prev, x))

        # Forget gate
        f_t = self.sigmoid(np.dot(self.Wf, concat) + self.bf)

        # Input gate
        i_t = self.sigmoid(np.dot(self.Wi, concat) + self.bi)
        C_tilda = self.tanh(np.dot(self.WC, concat) + self.bC)

        # Cell state update
        C_t = f_t * C_prev + i_t * C_tilda

        # Output gate
        o_t = self.sigmoid(np.dot(self.Wo, concat) + self.bo)

        # Hidden state update
        h_t = o_t * self.tanh(C_t)

        return h_t, C_t
# Example usage
input_size = 3  # input feature size
hidden_size = 5  # size of the LSTM's hidden state
lstm = LSTM(input_size, hidden_size)

# Dummy input data
x_t = np.random.randn(input_size, 1)
h_prev = np.zeros((hidden_size, 1))
C_prev = np.zeros((hidden_size, 1))
# Forward pass
h_t, C_t = lstm.forward(x_t, h_prev, C_prev)
print("Hidden state:", h_t)
print("Cell state:", C_t)
```

Додаток В. Ілюстративний матеріал

Вінницький національний технічний університет

Кваліфікаційна робота освітнього ступеня магістр
125 «Кібербезпека»
125 «Кібербезпека інформаційних технологій та систем»
на тему:

«Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-term memory (LSTM) для поєднання статичних і динамічних ознак»

Виконав: ст. гр. КІТС-23м
Кухарук П.О.
Керівник: к.т.н., доц. Шиян А.А.

м. Вінниця, 2024

Актуальність теми

Біометрична аутентифікація – це процес безпеки, який спирається на унікальні біологічні характеристики людей, щоб підтвердити, що вони є тими, за кого себе видають. Біометричні системи аутентифікації порівнюють фізичні чи поведінкові риси зі збереженими підтвердженими автентичними даними в базі даних. Якщо обидва зразки біометричних даних збігаються, аутентифікація підтверджується.

Як правило, біометрична аутентифікація використовується для керування доступом до фізичних і цифрових ресурсів, таких як будівлі, кімнати та комп'ютерні пристрої.

Мета роботи, об’єкт та предмет дослідження

Метою кваліфікаційної роботи є розробка заходів по вдосконаленню багатофакторної біометричної аутентифікації на основі використання convolution neural networks (CNN) та long short-term memory (LSTM) для поєднання статичних і динамічних ознак.

Об’єктом кваліфікаційної роботи є методи біометричної багатофакторної аутентифікації.

Предметом кваліфікаційної роботи є впровадження багатофакторної біометричної аутентифікації з використанням CNN та LSTM

Основні задачі для досягнення мети досліджень :

–Розглянути поняття біометричної аутентифікації та підходи до її реалізації;

–Проаналізувати основні методи реалізації багатофакторної біометричної аутентифікації;

–Провести аналіз існуючих реалізацій методів біометричної ідентифікації;

–Представити обгрунтування необхідності удосконалення багатофакторної біометричної аутентифікації;

–Розглянути особливості використання CNN та LSTM

–Провести розробку алгоритму удосконалення багатофакторної біометричної аутентифікації на основі використання CNN та LSTM;

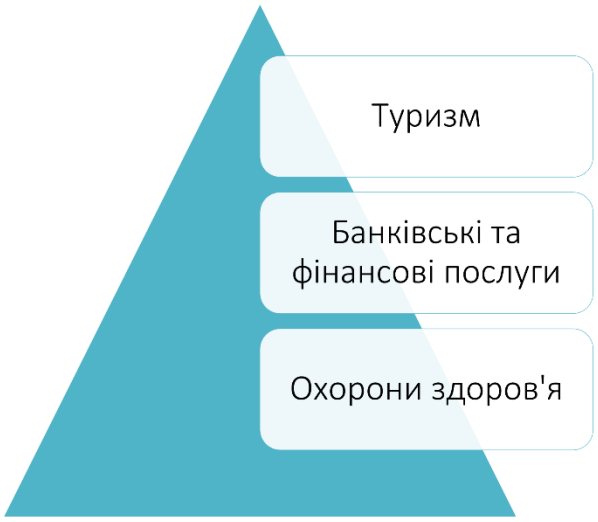
–Провести тестування розробленого алгоритму та навести результати його тестування;

–Оцінити економічну доцільність реалізації даного проекту.

Методи біометричної аутентифікації

- Розпізнавання відбитків пальців
- Розпізнавання облич
- Розпізнавання очей
- Розпізнавання голосу

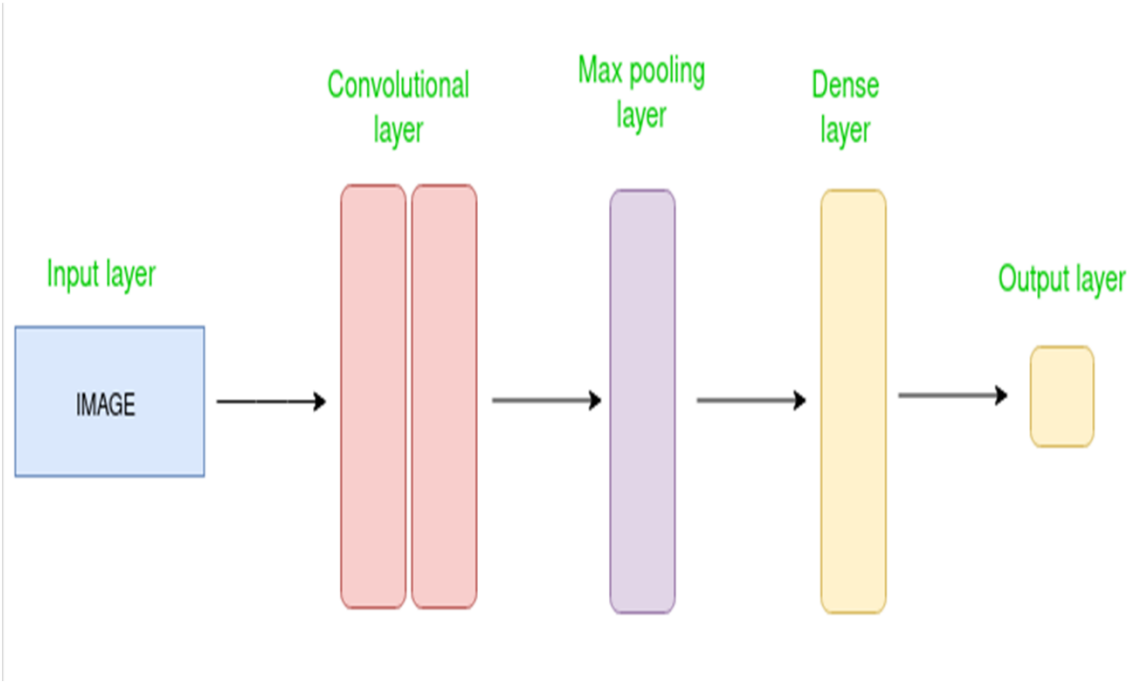
Сфери застосування біометричної аутентифікації



Методи розпізнавання облич

- Методи узгодження геометричних об’єктів
- Зіставлення графіків
- Підхід (Line Edge Map)
- Support Vector Machine (SVM)

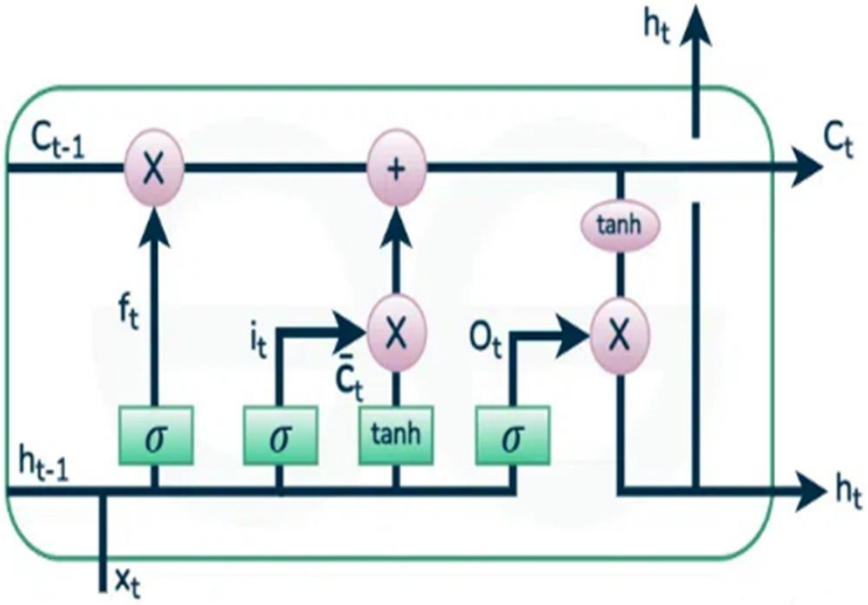
Архітектура CNN



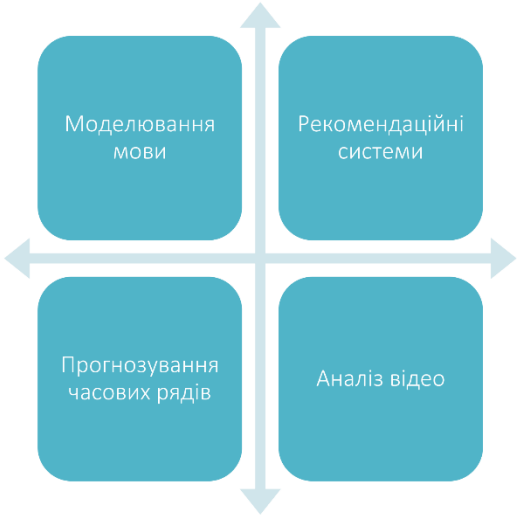
Переваги CNN

Недоліки CNN

Архітектура LSTM

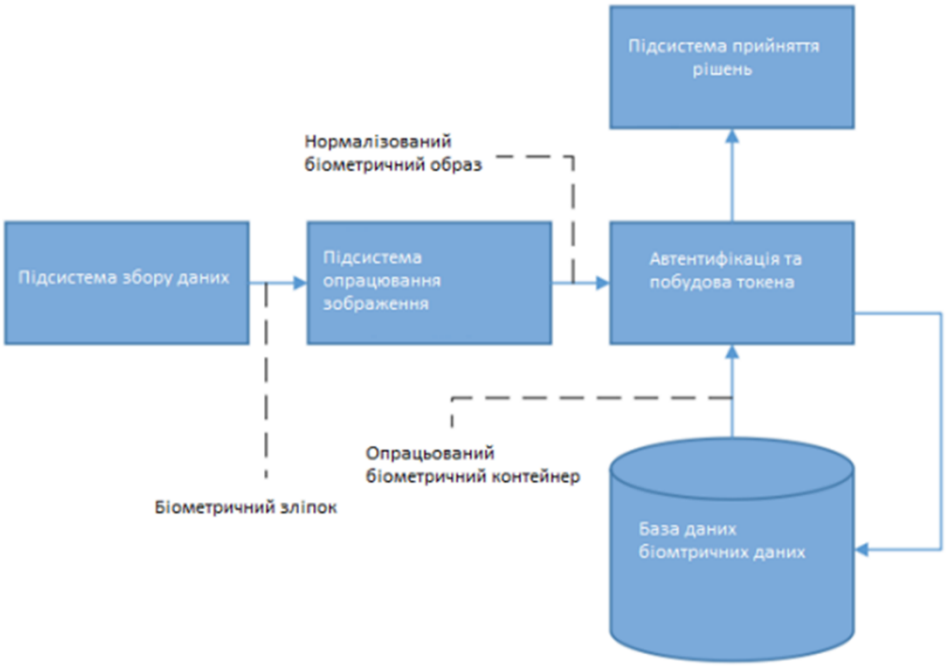


Області застосування LSTM



Алгоритм удосконалення багатофакторної біометричної аутентифікації

Спрощена схема біометричної аутентифікації



UML-діаграма

Розробка алгоритму удосконалення біометричної аутентифікації

Підготовка даних

```
# import libraries
import tensorflow
from tensorflow import keras

train_images = r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Training Images"

## Image Preprocessing using keras


# As we know deep-learning is hungry for data, the data we have is only limited.
# so lets perform **Image Agumentation** to create different versions
# of the original image, which leads to a better model, since it learns
# on the good and bad mix of images.

from keras.preprocessing.image import ImageDataGenerator
train_gen = ImageDataGenerator(
    shear_range=0.1,
    zoom_range=0.1,
    horizontal_flip=True
)
# No transformations are made on the test data
test_gen = ImageDataGenerator()

# Generating training data
training_data = train_gen.flow_from_directory(
    train_images,
    target_size = (100,100),
    batch_size = 30,
    class_mode = 'categorical'
)

# generating test data
testing_data = test_gen.flow_from_directory(
    train_images,
    target_size = (100,100),
    batch_size = 30,
    class_mode = 'categorical'
)

# Printing class labels for each face
testing_data.class_indices
```

Створення архітектури CNN

```
from keras.models import Sequential
from keras.layers import Conv2D, MaxPool2D, Flatten, Dense

'''Ініціалізація згорткової нейронної мережі''''
 Model = Sequential()


''' КРОК--1 Згортка
# Додавання першого рівня CNN
# ми використовуємо формат (100,100,3), оскільки ми використовуємо бекенд TensorFlow
# Це означає 3 матриці розміру (100x100) пікселів, що представляють червоний, зелений і синій компоненти пікселів
'''

 Model.add(Conv2D( 16 , kernel_size=( 5 , 5 ), strides=( 1 , 1 ), input_shape = ( 100 , 100 , 3 ), activation= 'relu' ))

'''# Maxplooing шар'''

 Model.add(MaxPool2D(pool_size=( 2 , 2 )))

'''Додавання додаткових шарів згортки та
       maxpooling для кращої точності та продуктивності моделі''''

 Model.add(Conv2D( 32 ,kernel_size=( 3 , 3 ), strides=( 1 , 1 ),activation= 'relu' ))
Model.add(MaxPool2D(pool_size=( 2 , 2 )))

'''# Додати згладжений шар для перетворення вектора в одновимірний''
 Model.add(Flatten())

'''Додати щільні шари та ініціалізувати ваги використання
              ініціалізатора ядра для кращого вивчення функцій зображення та класифікації'''

 Model.add(Dense( 64 ,activation= 'relu' ))
Model.add(Dense(Output_Neurons,activation= 'softmax' ))
```

Компіляція моделі

```
'''Виконати компіляцію моделі'''

 Модель. компілювати (loss= 'categorical_crossentropy' ,optimizer = 'adam' ,metrics = [ 'Accuracy' ])


''''# Використання ранньої зупинки для скорочення часу навчання''''

з keras.callbacks import EarlyStopping
call = EarlyStopping(
                    min_delta= 0,005 ,
                    patience= 5 ,
                     verbose= 1
                     ) час

імпорту
# Вимірювання часу, необхідного моделі для навчання
 StartTime=час.time()

''''# Модель навчання''''
 Model.fit_generator(training_data,
          епохи = 30 ,
          validation_data=testing_data,
                   callbacks=call)

Endtime = time.time()
print ( 'Загальний час навчання: ' , раунд ((Endtime-StartTime)/ 60 ), 'Хвилини' )
```

Класифікатор зображень

```
import numpy as np
from keras.preprocessing import image
import os
import glob

'''########### Making single predictions ###########'''

ImagePath=r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Training Images\face4\image_0054_Face_1.jpg"
test_image=image.load_img(ImagePath,target_size=(100, 100))
test_image=image.img_to_array(test_image)

test_image=np.expand_dims(test_image,axis=0)

result=Model.predict(test_image,verbose=0)
#print(training_set.class_indices)

print('####'*10)
print('Prediction is: ',Result_class[np.argmax(result)])

'''############ Making multiple predictions ###########'''

## Loading all the image paths from final testing folder for prediction
main_ = r"C:\Users\RAGUWING\Desktop\Resume\Additional Documents\Face Images\Final Testing Images"
img_paths = glob.glob(os.path.join(main_,'**','*.jpg'))

print(img_paths[0:5]) # every image will be a PIL object
print('*'*50)

for path in img_paths:
    test_image = image.load_img(path,target_size=(100,100))
    test_image = image.img_to_array(test_image)
    test_image = np.expand_dims(test_image,axis =0)
    result = Model.predict(test_image,verbose=0)
    print('Prediction: ',Result_class[np.argmax(result)])
```

Реалізація LSTM

```
class LSTM:
    def __init__(self, input_size, hidden_size):
        self.input_size = input_size
        self.hidden_size = hidden_size


# Initialize weights and biases for forget, input, cell, and output gates
self.Wf = np.random.randn(hidden_size, hidden_size + input_size)
self.Wi = np.random.randn(hidden_size, hidden_size + input_size)
self.WC = np.random.randn(hidden_size, hidden_size + input_size)
self.Wo = np.random.randn(hidden_size, hidden_size + input_size)

self.bf = np.zeros((hidden_size, 1))
self.bi = np.zeros((hidden_size, 1))
self.bC = np.zeros((hidden_size, 1))
self.bo = np.zeros((hidden_size, 1))
```

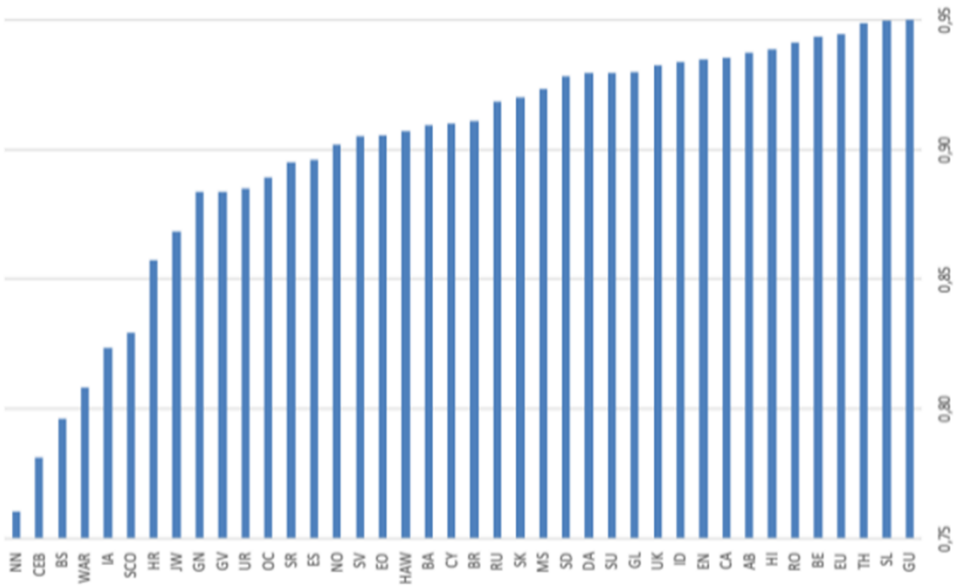
```
def sigmoid(self, z):
    return 1 / (1 + np.exp(-z))

def tanh(self, z):
    return np.tanh(z)


def forward(self, x, h_prev, C_prev):
    # Concatenate hidden state and input
    concat = np.vstack((h_prev, x))
```

Результати тестування

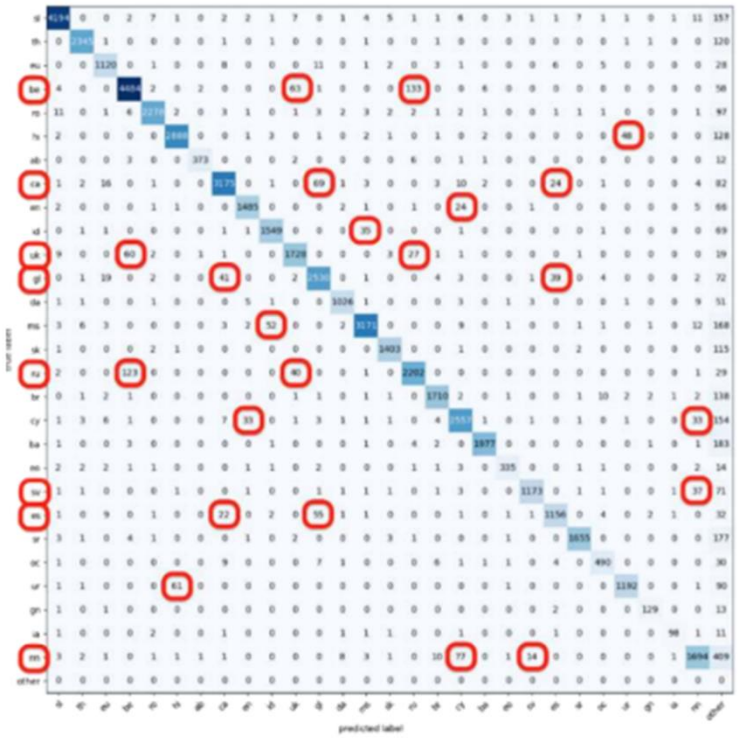
Зібрані аудіодані



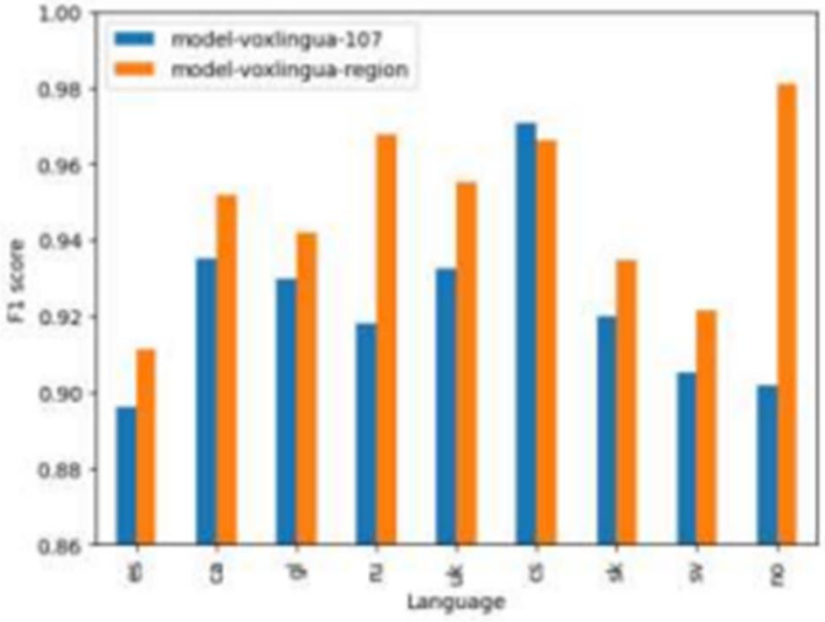
Розмір навчальних даних для окремих мов

Код	Зареєстровано, год.	Перевірено, год.	Кількість голосів
SL	14	11	146
TH	420	171	8
EU	159	105	1
BE	1 632	1 586	8 205
RO	44	19	408
HI	20	14	396
AB	85	60	400
CA	3 328	2 554	35 062
EN	3 347	2 532	88 904
ID	64	29	516
UK	105	94	1
GL	61	38	1
DA	13	12	243
MS	10	3	128
SK	27	22	216
RU	260	228	3
BR	26	12	196
CY	155	122	1 800
BA	268	258	912
EO	1 897	1 432	1 682
SV	54	45	808
ES	2 188	526	25 338
SR	6	4	144
OC	13	2	141
UR	221	63	316
GN	25	4	140
IA	17	14	66
NN	2	2	32

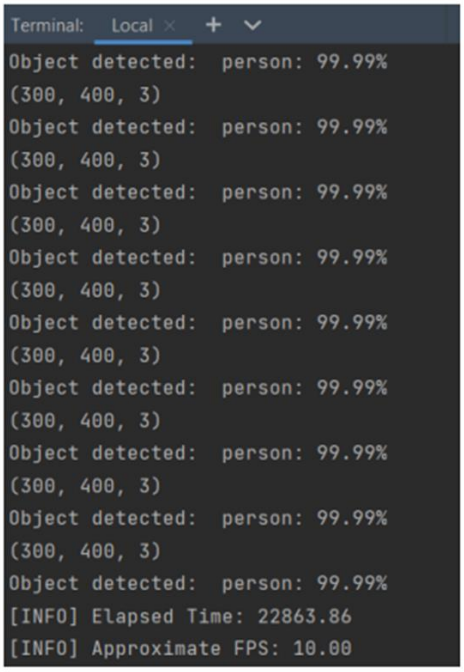
Матриця заплутаності



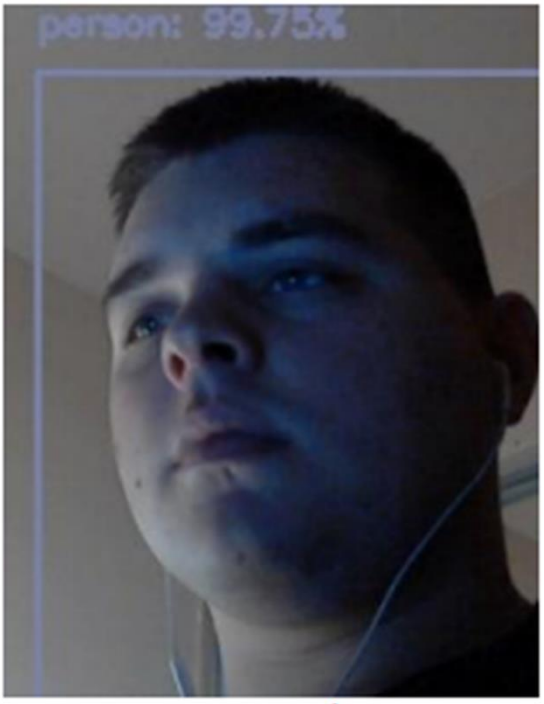
Результати роботи CNN



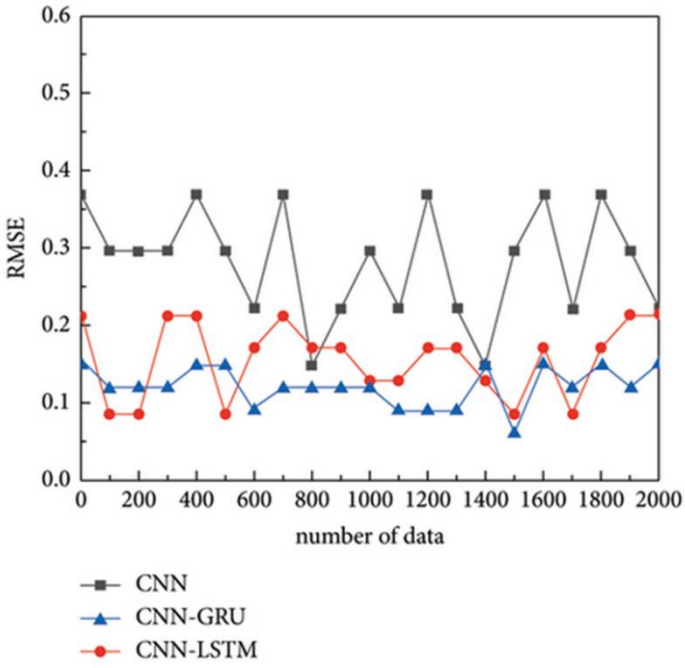
Результати точності ідентифікації зображення з використанням LSTM



Візуалізація біометричної аутентифікації з використанням LSTM



Порівняння роботи різних моделей біометричної аутентифікації



Висновки

Потреба в удосконаленні багатофакторної біометричної аутентифікації

Проблеми та вимоги реалізації біометричної аутентифікації

Створення алгоритму на основі використання CNN та LSTM

Аналіз отриманих результатів

Додаток Г. Протокол перевірки на антиплагіат

ПРОТОКОЛ ПЕРЕВІРКИ НАВЧАЛЬНОЇ (КВАЛІФІКАЦІЙНОЇ) РОБОТИ

Назва роботи:

Удосконалення багатофакторної біометричної аутентифікації на основі convolution neural networks (CNN) та long short-term memory (LSTM) для поєднання статичних і динамічних ознак

Тип роботи: магістерська кваліфікаційна робота

Підрозділ Кафедра менеджменту на безпеки інформаційних систем
Факультет менеджменту та інформаційної безпеки
Гр. КІТС-23м
Керівник доцент Шиян А.А.

Показники звіту подібності

Turnitin	
Оригінальність	78%
Загальна схожість	22%

Аналіз звіту подібності (відмітити потрібне)

- ☐ Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.
- ☐ Виявлені у роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і відсутності самостійності її автора. Роботу направити на доопрацювання.
- ☐ Виявлені у роботі запозичення є недобросовісними і мають ознаки плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень.

Заявляю, що ознайомлений (-на) з повним звітом подібності, який був згенерований Системою щодо роботи (додається)

Автор _____ Кухарук П.О.
(підпис) (прізвище, ініціали)

Опис прийнятого рішення

- ☐ **Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.**

Особа, відповідальна за перевірку _____ Коваль Н.П.
(підпис) (прізвище, ініціали)

Експерт _____ _____
(за потреби) (підпис) (прізвище, ініціали, посада)