МАГІСТЕРСЬКА РОБОТА

МР. ІПм - 11.00.00.000 ПЗ

Група ІПм-24-1

Грицак Тарас

2025

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Грицак Тарас Богданович

(прізвище, ім'я, по батькові)

УДК 004.9

(індекс)

МАГІСТЕРСЬКА РОБОТА

Моделі та методи захисту даних в мобільних додатках

(назва роботи)

Інженерія програмного забезпечення

(назва освітньої програми)

121 - Інженерія програмного забезпечення

(шифр і назва спеціальності)

Грицак Т.Б.

(підпис, ініціали та прізвище здобувача освітнього ступеня)

Науковий керівник Бандура Вікторія Валеріївна, к.т.н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Допущено до захисту

Завідувач кафедри

доц. Бандура В.В.

(посада) (підпис) (дата) (ініціали та прізвище)

Нормоконтроль

доц. Вовк Р.Б.

(посада) (підпис) (дата) (ініціали та прізвище)

Робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

Івано-Франківськ – 2025

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Освітній рівень магістр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:
Зав. кафедрою **ІПЗ**

доц. В.В. Бандура

“ 04 ” вересня 2025 р.

ЗАВДАННЯ

НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Грицаку Тарасу Богдановичу

(прізвище, ім’я, по-батькові)

1. Тема магістерської роботи “ Моделі та методи захисту даних в мобільних додатках”

керівник проекту (роботи) Бандура В.В., к.т.н., доцент

затверджені наказом закладу вищої освіти від “ 05 ” листопада 2025 р. № 695/7

2. Строк подання студентом проекту (роботи) 15 грудня 2025 р.

3. Вихідні дані до проекту (роботи) Теоретичні концепції та формальні моделі побудови та функціонування інформаційних технологій захисту даних

4. Зміст розрахунково - пояснювальної записки(перелік питань, які потрібно розробити)

1. Аналіз концепцій захисту мультимедійних даних
2. Аналіз досліджень щодо захисту даних в мобільних додатках
3. Методи захисту мультимедіа в мобільних додатках
4. Реалізація моделей та методології захисту мультимедійних даних в мобільних додатках

5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)

1. Приклад вимірювання розмиття на фото із визначенням суб’єкта і перехожого (рис. 1.1)
2. Кут тангажу (Pitch) як кут обертання голови навколо осі x (рис. 1.2)
3. Кут yaw - як кут обертання голови навколо осі y (рис. 1.3)
4. Графічне представлення роботи Darkly (рис. 1.5)
5. Стек програмного забезпечення Android (рис. 2.1)

6. Консультанти розділів проекту (роботи)

Розділ	Консультант	Підпис, дата
Перевірка на плагіат	доц., к.т.н. Вовк Р.Б.	

7. Дата видачі завдання 04 вересня 2025 р.

Керівник _____
(підпис)

Завдання прийняв до виконання_____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назви етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Підбір і вивчення літератури по темі магістерської роботи	15.09.2025	виконано
2	Аналіз концепцій захисту мультимедійних даних	29.09.2025	виконано
3	Аналіз досліджень щодо захисту даних в мобільних додатках	15.10.2025	виконано
4	Методи захисту мультимедіа в мобільних додатках	08.11.2025	виконано
5	Реалізація моделей та методології захисту мультимедійних даних в мобільних додатках	20.11.2025	виконано
6	Реалізація функціональності запропонованої інформаційної технології	01.12.2025	виконано
7	Затвердження пояснювальної записки роботи завідувачем кафедри	15.12.2025	виконано

Студент – магістр _____
(підпис)

Керівник роботи _____
(підпис)

АНОТАЦІЯ

Магістерська робота: 75 с., 20 рис., 10 табл., 37 джерел.

Тема: Моделі та методи захисту даних в мобільних додатках

Мета магістерської роботи - розробка та обґрунтування моделей і методів захисту мультимедійних даних у мобільних додатках, спрямованих на підвищення рівня конфіденційності користувачів та мінімізацію ризиків несанкціонованого доступу чи витоку інформації.

Об'єкт дослідження - процеси обробки та захисту мультимедійних даних у мобільних додатках.

Предмет дослідження - моделі, методи та інструменти захисту мультимедійних даних у мобільних додатках із використанням механізмів аналізу поведінки програм та класифікації зображень.

Результати дослідження

В роботі розроблено вдосконалену модель диференціації осіб «Ціль/Перехожий», яка враховує розширені біометричні ознаки для захисту конфіденційності випадкових осіб;

Висновок

Проведена розробка концепції системи захисту мультимедійних даних у мобільних додатках, архітектура якої включає модуль динамічного аналізу, класифікатор зображень на базі DCNN та механізми анонімізації даних

МОБІЛЬНІ ДОДАТКИ; ЗАХИСТ ДАНИХ; МУЛЬТИМЕДІЙНА ІНФОРМАЦІЯ; КОНФІДЕНЦІЙНІСТЬ; УПРАВЛІННЯ ДОЗВОЛАМИ; АНАЛІЗ ПОВЕДІНКИ ПРОГРАМ; КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ; ГЛИБОКЕ НАВЧАННЯ; КОМП'ЮТЕРНИЙ ЗІР.

ABSTRACT

Master Thesis: 75 pp., 20 fig., 10 tab., 37 sources.

Topic: Models and methods of data protection in mobile applications

Meta master's thesis - development and justification of models and methods of multimedia data protection in mobile applications aimed at increasing the level of user privacy and minimizing the risks of unauthorized access or information leakage.

The object of research is the processes of processing and protecting multimedia data in mobile applications.

The subject of research is models, methods and tools of multimedia data protection in mobile applications using mechanisms of program behavior analysis and image classification.

Research results

The work developed an improved model of differentiation of persons "Target/Passerby", which takes into account extended biometric features to protect the privacy of random persons;

Conclusion

The concept of a multimedia data protection system in mobile applications was developed, the structure of which includes a dynamic analysis module, an image classifier based on DCNN and data anonymization mechanisms.

MOBILE APPLICATIONS; DATA PROTECTION; MULTIMEDIA INFORMATION; PRIVACY; PERMISSION MANAGEMENT; PROGRAM BEHAVIOR ANALYSIS; IMAGE CLASSIFICATION; DEEP LEARNING; COMPUTER VISION.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

DCNN - Deep Convolutional Neural Network - Глибока згорткова нейронна мережа

HOG - Histogram of Oriented Gradients - Гістограма орієнтованих градієнтів

ReLU - Rectified Linear Unit - Випрямлений лінійний блок (функція активації)

GDBT - Gradient Boosting Decision Tree - Градієнтний бустинг дерев рішень

MLP - Multi-Layer Perceptron - Багатошаровий перцептрон

FFT - Fast Fourier Transform - Швидке перетворення Фур'є

AU - Action Unit - Одиниця дії

FACS- Facial Action Coding System - Фаціальна система кодування дій

OvA - One-vs-All - Один-проти-Всіх (стратегія класифікації)

OvO - One-vs-One - Один-проти-Одного (стратегія класифікації)

ВСТУП

Актуальність теми.

Сучасний розвиток інформаційних технологій зумовив стрімке зростання використання мобільних пристроїв, які перетворилися на універсальні інструменти для обробки, зберігання та передавання інформації. Значну частку даних, що циркулюють у мобільних додатках, становлять мультимедійні об’єкти – фотографії, відео, аудіо, які є надзвичайно чутливими з точки зору конфіденційності. Вони здатні містити не лише безпосередньо персональні відомості, а й додаткові ознаки, які можуть бути використані для ідентифікації, профілювання користувачів та несанкціонованого спостереження.

Зростання обсягів обробки мультимедійних даних, поєднане з поширенням хмарних сервісів і соціальних платформ, створює нові виклики у сфері інформаційної безпеки. Особливої актуальності набуває проблема захисту не лише основного користувача пристрою, але й випадкових осіб, які потрапляють у поле зйомки без їхньої згоди. Недосконалість існуючих моделей дозволів мобільних платформ та відсутність ефективних методів анонімізації зумовлюють потребу в нових підходах до забезпечення приватності.

Вирішення зазначених проблем потребує комплексного дослідження моделей і методів захисту мультимедійних даних у мобільних додатках із використанням інструментів аналізу поведінки програм, методів машинного навчання та комп’ютерного зору.

Актуальність дослідження обумовлюється тим, що мультимедійні дані стали основним джерелом витоків інформації та засобом прихованого збору персональних відомостей у мобільних середовищах. Існуючі механізми безпеки, зокрема моделі дозволів Android, мають низку недоліків: вони не забезпечують диференційованого контролю доступу, часто є складними для

звичайного користувача, а також не здатні попереджати приховані канали витоку.

Важливим аспектом є соціально-правова складова: сучасні мобільні додатки створюють ризики для конфіденційності третіх осіб, що формує новий рівень суспільних викликів. Водночас методи комп'ютерного зору та глибокого навчання відкривають перспективи автоматизованого розпізнавання, класифікації та анонімізації мультимедійних даних, що значно підвищує рівень захищеності.

Отже, розробка моделей і методів для захисту мультимедійної інформації в мобільних додатках є науково значущим і практично необхідним завданням у контексті розвитку безпечних інформаційних технологій.

Метою магістерської роботи є розробка та обґрунтування моделей і методів захисту мультимедійних даних у мобільних додатках, спрямованих на підвищення рівня конфіденційності користувачів та мінімізацію ризиків несанкціонованого доступу чи витоку інформації.

Об'єкт дослідження - процеси обробки та захисту мультимедійних даних у мобільних додатках.

Предмет дослідження - моделі, методи та інструменти захисту мультимедійних даних у мобільних додатках із використанням механізмів аналізу поведінки програм та класифікації зображень.

Завдання дослідження

- Провести аналіз предметної області захисту мультимедійних даних у мобільних додатках.

- Ідентифікувати основні вразливості та загрози, пов'язані з обробкою мультимедійної інформації.

- Дослідити існуючі моделі управління дозволами мобільних платформ та визначити їхні обмеження.

- Розробити методологію аналізу поведінки мобільних додатків для виявлення потенційних витоків даних.

- Створити класифікатор для ідентифікації конфіденційних мультимедійних даних із використанням методів машинного навчання.

- розробити архітектуру системи захисту мультимедійних даних та оцінити її ефективність у порівнянні з існуючими рішеннями.

Методи дослідження

- теоретичний аналіз наукових публікацій та практичних підходів до захисту даних у мобільних додатках;

- методи статичного та динамічного аналізу поведінки програмного забезпечення;

- методи машинного навчання, зокрема класифікація зображень на основі SVM та глибоких згорткових нейронних мереж (DCNN);

- інструменти комп'ютерного зору для виявлення та обробки біометричних ознак;

- експериментальне моделювання та порівняльний аналіз продуктивності розроблених методів.

Наукова новизна отриманих результатів

Удосконалено методологію виявлення потенційних витоків даних у мобільних додатках шляхом інтеграції статичного та динамічного аналізу і запропоновано автоматизований класифікатор конфіденційних мультимедійних даних на основі DCNN, що підвищує точність ідентифікації приватних зображень.

Практичне застосування результатів

Розроблені моделі та методи можуть бути інтегровані у мобільні додатки для підвищення рівня конфіденційності користувачів, а також використані у створенні спеціалізованих систем контролю доступу до мультимедійних даних. Запропоновані підходи можуть застосовуватись у сферах інформаційної безпеки, мобільної розробки, а також у правозахисних практиках, пов'язаних із захистом персональної інформації та приватності у публічному просторі.

Структура магістерської роботи. Робота складається зі вступу, трьох розділів та висновків. Загальний обсяг роботи становить 75 сторінок, і містить 20 рисунків, 10 таблиць, список використаних джерел із 37 найменувань.

РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ЗАХИСТУ МУЛЬТИМЕДІЙНИХ ДАНИХ

1.1. Забезпечення конфіденційності мультимедійних даних в умовах широкого розповсюдження мобільних пристроїв

Експоненційне зростання застосування камер у сучасних мобільних пристроях протягом останнього десятиліття спричинило безпрецедентне накопичення особистих мультимедійних даних (зображень) на популярних апаратних платформах. Однак ця тенденція одночасно загострила проблему конфіденційності цих даних. Вказані зображення можуть інкапсулювати конфіденційну інформацію, потенційно шкідливу в разі несанкціонованого доступу, зокрема персональні ідентифікатори, що містяться в посвідченнях особи чи юридичній документації.

1.1.1. Ідентифікація вразливостей та об'єкт дослідження

Поточні механізми безпеки в операційних системах iOS та Android демонструють вразливість, за якої сторонні застосунки, інстальовані з офіційних або альтернативних сховищ, можуть отримувати несанкціонований доступ до медіагалерей пристроїв, часто без явного сповіщення чи згоди користувача. Крім того, повсюдне використання камер смартфонів у публічному просторі створює ризик порушення приватності за рахунок ненавмисного або несанкціонованого захоплення зображень сторонніх осіб (перехожих).

1.1.2. Методологія та очікувані результати

З метою пом'якшення ризиків конфіденційності, спричинених ненадійним програмним забезпеченням та небажаною публічною фотографією, це дослідження охоплює три ключові науково-технічні напрямки:

1. Розробка двоетапного методу аналізу доступу.

Створення комбінованого методу, що інтегрує аналіз дозволів (permissions analysis) та аналіз системних викликів (system call analysis), для детермінування потенційної здатності сторонніх застосунків до неявного доступу до конфіденційних фотоматеріалів.

2. Створення автоматизованого класифікатора приватних даних.

Проєктування та імплементація автоматизованого класифікатора для ідентифікації та забезпечення захисту приватних зображень у сховищі медіа мобільного пристрою.

3. Розробка Моделі Комп'ютерного Зору для Виявлення Перехожих.

Створення високоточної моделі комп'ютерного зору для виявлення випадкових перехожих на зображеннях, що дозволить автоматично застосовувати техніки обфускації обличчя (наприклад, розмиття) або інші методи для захисту їхньої конфіденційності.

Очікується, що результати аналізу системних викликів сприятимуть підвищенню обізнаності громадськості щодо вразливостей, асоційованих із завантаженням ненадійних застосунків. Встановлено, що розроблений класифікатор приватних фотографій та модель виявлення випадкових перехожих демонструють прийнятний рівень точності на тестових наборах даних. Ці рішення можуть слугувати основою для майбутніх досліджень, спрямованих на впровадження працездатних систем захисту індивідуальної конфіденційності в умовах вищезазначених векторів загроз.

1.2. Вплив мультимедійних даних на конфіденційність персональної інформації

Широке впровадження камер у смартфони спровокувало масове поширення особистої фотографії серед мільйонів користувачів платформ iOS та Android. Хоча ця технологічна еволюція демократизувала доступ до

високоякісної фотозйомки, вона одночасно створила нові, суттєві виклики для захисту особистої інформації.

Однією з найбільш значущих вразливостей у сфері захисту конфіденційності є поточні протоколи надання дозволів (permissions protocols) у мобільних операційних системах.

1.2.1. Обмеження поточної моделі дозволів додатків

Згідно з чинною архітектурою, користувачі, які інсталюють сторонні застосунки, одноразово отримують запит на надання доступу до медіасховища та/або мережевого з'єднання пристрою. Після надання цей дозвіл є фактично безстроковим, надаючи застосунку постійну можливість доступу до файлів та їх передачі через мережу. Цей підхід, хоча й спрощує керування безпековими налаштуваннями для користувача, не враховує сценарії, де необхідно забезпечити гранульований контроль доступу до окремих фотографічних файлів, щоб запобігти їх несанкціонованому читанню та мережевій передачі застосунками.

Враховуючи недостатню прозорість поведінки більшості сторонніх програм, користувачі змушені покладатися на автоматизовані процеси скринінгу, впроваджені компаніями Apple (App Store) та Google (Google Play Store). Проте, неефективність цих процесів уже була продемонстрована серйозними інцидентами, такими як виявлення значної кількості інфікованих шкідливим програмним забезпеченням застосунків у Google Play [1] та випадки несанкціонованого викрадення даних у програмах App Store [2]. Крім того, користувачі, які отримують програми з альтернативних джерел, часто повністю позбавлені захисту від прихованої шкідливої активності. Неспроможність існуючих скринінгових механізмів у поєднанні з відсутністю контролю доступу на рівні окремих файлів створює чіткі можливості для зловмисників для отримання доступу та ексфільтрації приватних фотографій разом із вбудованою в них конфіденційною інформацією.

1.2.2. Проблема конфіденційності перехожих у публічному просторі

Окрім ризиків несанкціонованого доступу, поширення мобільної фотографії негативно вплинуло на конфіденційність осіб у громадських місцях. Значна частка фотографій, зроблених у публічних просторах, містить випадкових перехожих. Зважаючи на оцінену кількість цифрових фотографій, зроблених у всьому світі у 2024 році (близько 1.5×10^{14}) [3], обсяг випадково захоплених осіб є екстремально високим.

Наразі відсутня універсальна система для забезпечення конфідційності незнайомців на публічних фотографіях, хоча окремі корпорації (наприклад, Google у сервісі Street View) застосовують механізми автоматичного розмиття обличчя. Існує значний суспільний попит на механізми приховування або видалення їхніх облич із фотографій третіх осіб, особливо тих, що можуть бути опубліковані в соціальних мережах або інших загальнодоступних ресурсах. Це обумовлює необхідність розробки автоматизованої системи комп'ютерного зору для виявлення та обфускації облич, що не є цільовими об'єктами фотографії, з метою надання користувачам інструментів для захисту приватності випадкових перехожих.

1.3. Методології захисту мультимедійних даних на мобільних платформах

1.3.1. Методологія моніторингу поведінки застосунків Android

Розроблено методологію для моніторингу поведінки застосунків на платформі Android з метою детектування несанкціонованого доступу до фотографічних даних. Ефективність методу було продемонстровано шляхом аналізу найбільш популярних безкоштовних застосунків з каталогу Google Play Store.

Процес відстеження поведінки реалізовано через запис усіх системних викликів читання файлів у каталозі зберігання мультимедіа (фотографій) на емульованому пристрої Android. Будь-який застосунок, який ініціює

операцію читання в медіа-каталозі та має дозвіл на доступ до мережі, апріорі вважається таким, що має потенціал для несанкціонованого витоку фотографічних даних користувача. Важливо зазначити, що отримана інформація про поведінку не є остаточним доказом фактичної шкідливої активності, а слугує для демонстрації поширеності потенціалу витоків даних у сучасних екосистемах мобільних застосунків.

1.3.2. Розробка автоматизованого класифікатора приватних фотографій

Для пом'якшення ризиків, створюваних ненадійними програмами щодо приватних зображень, пропонується імплементація системи автоматизованого класифікатора фотографій. Для досягнення максимальної ефективності цю систему необхідно інтегрувати в майбутні версії операційних систем Android та iOS як невід'ємну складову архітектури безпеки, аналогічно поточним схемам дозволів застосунків.

Класифікатор фотографій відповідає за ідентифікацію зображень, які містять потенційно конфіденційну інформацію, як-от юридичні документи, посвідчення особи, або певні особисті знімки (селфі), які користувач може бажати захистити від несанкціонованого доступу. Детектування відбувається під час збереження будь-якого фотографічного файлу в медіа-каталозі пристрою.

Після ідентифікації приватної фотографії її ім'я файлу та шлях кешуються в захищеній області пам'яті. Операційна система пристрою повинна виконувати додаткову перевірку безпеки щоразу, коли встановлений застосунок ініціює запит на читання з медіа-каталогу. Ця перевірка полягає у зверненні до кешу для верифікації, чи є фотографія захищеною. У разі позитивного результату, запит на читання може бути припинено або скеровано на підтвердження користувачем.

Для оцінки продуктивності було розроблено та протестовано декілька прототипів класифікатора, вимірюючи точність на різних моделях

машинного навчання та алгоритмах. Впровадження цієї модифікації в iOS та Android призведе до суттєвого посилення захисту конфіденційності фотографій завдяки можливості надійного керування запитами на читання на рівні операційної системи.

1.3.3. Модель комп'ютерного зору для захисту конфіденційності перехожих

Вирішення проблеми конфіденційності випадкових перехожих на публічних фотографіях передбачає розробку автоматизованої системи виявлення на основі методів комп'ютерного зору. Ці методи можуть бути застосовані для розпізнавання та диференціації цільових об'єктів фотографії від незнайомців з прийнятною точністю.

Для ефективної роботи цих моделей необхідне використання набору відмінних ознак обличчя (facial features), таких як ширина, висота, ступінь розмитості та напрямок погляду (очей). У деяких випадках різні алгоритми машинного навчання демонструють покращену точність або продуктивність. Проведено детальне порівняння різних алгоритмів для виявлення їхніх переваг та недоліків.

Розпізнавання обличчя випадкових перехожих дозволяє подальше застосування методів обфускації (наприклад, розмивання обличчя), забезпечуючи збереження їхньої конфіденційності.

Основні науково-технічні внески цієї роботи можна резюмувати наступним чином:

1. Розроблено методологію моніторингу додатків Android з використанням трасування системних викликів для детектування неявних операцій читання зображень. Цей підхід є значним покращенням порівняно з простим аналізом статичних дозволів.

2. Реалізовано та оцінено кілька класифікаторів приватних фотографій з використанням сучасних архітектур нейронних мереж та гібридних

моделей, що поєднують ознаки гістограми орієнтованих градієнтів (HOG) з класифікатором опорних векторів (SVM).

3. Створено розширену та вдосконалену модель на основі виділення ознак для автоматичного розпізнавання облич цільових осіб та випадкових перехожих. На відміну від сучасної моделі, пропонована модель інтегрує нову ознаку відстеження погляду (gaze tracking) та переробляє ознаки позиції та розміру обличчя. Ці модифікації спрямовані на підвищення загальної точності моделі розпізнавання.

Існує евристичний підхід до розрізнення цільового суб'єкта (target) та випадкового перехожого (stranger) на фотографії на основі трьох ознак: посмішка, розмір обличчя та позиція обличчя. Також, ґрунтуючись на аналізі сотень фотографій, набір ознак розширюється для підвищення точності класифікації.

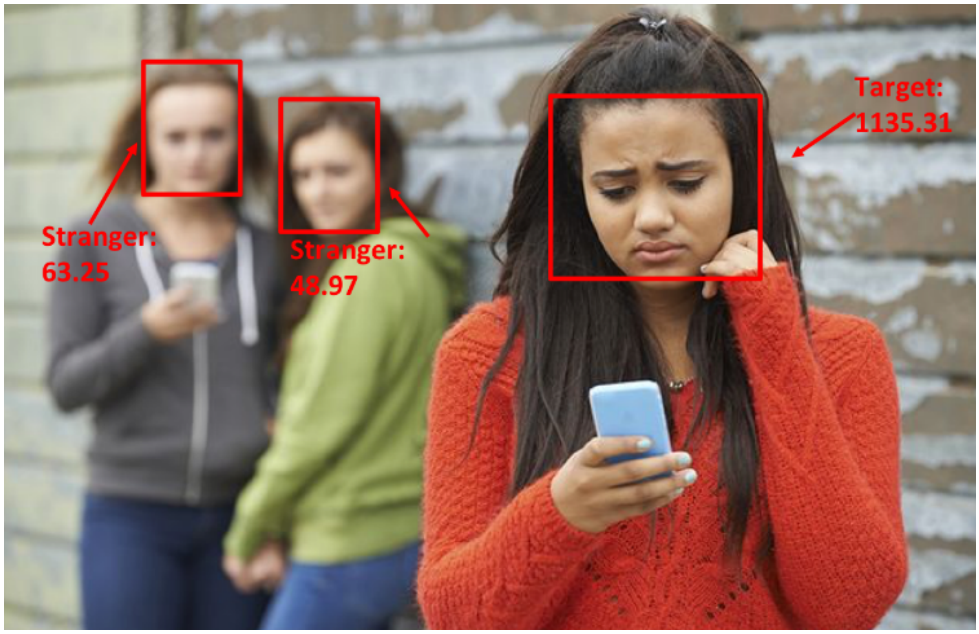

Рис. 1.1. Приклад вимірювання розмиття на фото із визначенням суб’єкта і перехожого

Часто додаються чотири нові метрики, пов'язані з якістю зображення та просторовою орієнтацією обличчя, а також модифікується існуюча ознака посмішки. Загалом класифікатор використовує сім ознак.

1. Розмитість (Blurriness)

Припускається, що цільовий суб'єкт сфокусований навмисно, і тому його обличчя, ймовірно, розташоване в зоні фокуса, тоді як обличчя перехожих частіше будуть розмитими.

Метод Вимірювання: Замість обчислення швидкого перетворення фур'є (FFT) та аналізу високочастотних компонентів (що є чутливим до визначення оптимального порогу), ми використовуємо варіацію оператора Лапласіана.

Варіація Лапласіана обчислюється як:

$$Var(Lap(x,y)) = Var(\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}),$$

де I(x,y) — інтенсивність пікселів.

Лапласіан вимірює другу похідну зображення. Висока дисперсія свідчить про швидкі зміни інтенсивності, що відповідає високій деталізації та фокусу. Як показано на рис. 1.1, обличчя цільового суб'єкта має значно більшу варіацію Лапласіана (меншу розмитість), ніж обличчя перехожих.

2. Кути орієнтації голови (Head Pose Angles)

Використовуються три кути обертання голови, які визначаються відносно осей камери.

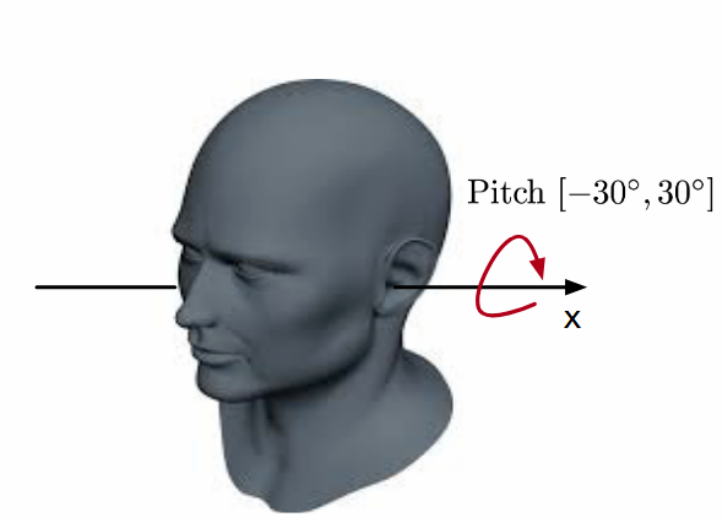

Рис. 1.2. Кут тангажу (Pitch) як кут обертання голови навколо осі x

Кут тангажу (Pitch) - визначається як кут обертання голови навколо осі x (рис. 1.2). Вимірюється в діапазоні [−30◦,+30◦], який охоплює більшість випадків надійного виявлення обличчя. Якщо виявлений кут перевищує цей діапазон, він обмежується відповідно ±30◦.

Кут Рискання (Yaw): Визначається як кут обертання голови навколо осі y (рис. 1.3). Діапазон вимірювання також встановлено [−30◦,+30◦] з аналогічними обмеженнями через чутливість виявлення обличчя.

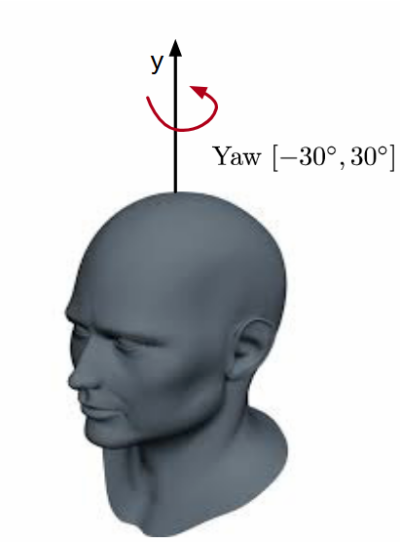

Рис. 1.3. Кут yaw - як кут обертання голови навколо осі y

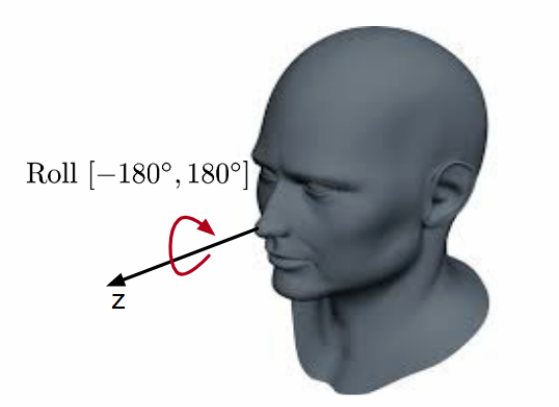

Рис. 1.4. Кут обертання (Roll) як кут обертання голови навколо осі z

Кут обертання (Roll): Визначається як кут обертання голови навколо осі z (рис. 1.4). Діапазон вимірювання становить [−180◦,+180◦].

3. Модифікація ознаки посмішки

Ознака посмішки, яка була бінарною змінною, тепер змінена на числову змінну з діапазоном значень [0,100], що дозволяє точніше відобразити інтенсивність виразу обличчя.

Фінальний бінарний класифікатор використовує ці сім ознак. Діапазони значень кожної ознаки підсумовано в таблиці 1.1.

Таблиця 1.1.

Діапазон значень кожної ознаки

Ознака	**Діапазон Значень**
Посмішка (Smiling)	[0,100]
Розмір Обличчя (Face Size)	(0,1]
Позиція Обличчя (Face Position)	-0,1
Розмитість (Blurriness)	[0,+∞)
Тангаж (Pitch)	[−30◦,+30◦]
Рискання (Yaw)	[−30◦,+30◦]
Обертання (Roll)	[−180◦,+180◦]

1.4 Аналіз досліджень щодо захисту даних в мобільних додатках

Даний розділ містить огляд існуючих робіт у сферах управління дозволами платформи Android та захисту конфіденційності фотографій, що слугує контекстом для позиціонування пропонованої системи.

1.4.1. Управління дозволами Android (Android Permissions)

Платформа Android пропонує користувачам два основні підходи до контролю дозволів.

1. До Android 6.0 (Marshmallow)

До виходу Android 6.0, застосунки були зобов'язані розкривати повний перелік необхідних ресурсів під час інсталяції. Користувач мусив надати згоду на всі запитувані дозволи, інакше встановлення переривалося. Дослідження [2, 3] показали, що користувачі не приділяють достатньої уваги

і, як правило, не розуміють значення дозволів, що надаються під час інсталяції (install-time permissions).

2. Після Android 6.0 (Marshmallow)

Починаючи з Android 6.0, користувачі надають дозволи лише в момент, коли застосунок вперше запитує доступ до чутливого ресурсу. Ця схема надає користувачам контекстуальні підказки щодо необхідності запитуваного ресурсу. Однак, вона не враховує той факт, що переваги користувача щодо подальших запитів дозволу можуть варіюватися залежно від контекстуальних обставин [2, 4, 5]. Дослідження продемонстрували, що витік конфіденційної інформації більш імовірний, коли користувачі не усвідомлюють мету запиту на доступ до певного чутливого ресурсу.

1.4.2. Системи управління дозволами третіх сторін

В [9] проаналізували AppOps — диспетчер дозволів, впроваджений у Android 4.3 та пізніше видалений у Android 4.4.2. AppOps дозволяв користувачам переглядати та змінювати дозволи після встановлення застосунків, що значно підвищувало обізнаність користувачів про ризики конфіденційності. Хоча Android 6.0 представив нову систему управління дозволами, вона залишається прихованою в налаштуваннях, що ускладнює її використання пересічними користувачами. Існують сторонні програми управління дозволами (наприклад, XPrivacy, DonkeyGuard), але їхня функціональність вимагає додаткових привілеїв (наприклад, розблокований завантажувач), що обмежує їхнє широке використання та необхідне для захисту системи дозволів від втручання зловмисних програм.

1.4.3. Захист конфіденційності фото (Photo Privacy)

Існує низка робіт, присвячених захисту конфіденційності фотографій, переважно у сфері обміну даними в соціальних мережах:

- P3 та Обмін Фотографіями. Є розроблена система P3 для захисту конфіденційності фотографій, якими обмінюються в онлайн-соціальних

мережах. Також запропонували підхід для захисту конфіденційності користувачів під час обміну фотографіями.

- Контекстна обфускація. В [6] представили систему Darkly, засновану на бібліотеці OpenCV, для захисту приватної інформації від застосунків, що безперервно здійснюють моніторинг. В дослідженні [7] імплементували PlaceAvoider для захисту візуальної конфіденційності шляхом ідентифікації чутливих місць у відеопотоках.

Система Darkly є механізмом захисту візуальної приватності, який використовує бібліотеку OpenCV для обробки зображень. Її основне призначення — захист користувачів від додатків, що здійснюють безперервне зчитування даних із сенсорів (наприклад, камер) або екрана, шляхом автоматичної обфускації (приховування) чутливої інформації в реальному часі.

Darkly не обмежує доступ до сенсорів, але гарантує, що дані, які потрапляють до застосунків, є неінформативними щодо конфіденційних об'єктів. Це досягається за допомогою інструментів комп'ютерного зору, наданих OpenCV:

- Darkly використовує алгоритми виявлення об'єктів (object detection) та розпізнавання обличчя (face recognition) (що легко реалізується через OpenCV's модулі Haar Cascades або DNN), щоб визначити точне розташування конфіденційних даних у відеопотоці або на знімку екрана.

- Це можуть бути обличчя людей, документи, екрани, що містять паролі, або інші особисті дані.

- Після ідентифікації чутливі області піддаються обфускації (obfuscation) — процесу навмисного приховування інформації.

Найчастіше застосовується пікселізація (pixelation), розмиття за Гаусом (Gaussian blur) або заповнення однотонним кольором (маскування). Усі ці фільтри та операції трансформації зображень ефективно виконуються за допомогою функцій OpenCV.

Система працює на рівні ОС або як спеціалізований проксі-сервіс, який перехоплює запит застосунку на отримання зображення чи відеопотоку. Замість оригінальних даних застосунку повертається обфускована копія.

Система Darkly вирішує проблему, коли користувач свідомо надає застосунку доступ до камери або екрана (наприклад, для відеодзвінка чи демонстрації екрана), але не бажає, щоб застосунок бачив всю інформацію в кадрі. Забезпечує захист конфіденційності, не вимикаючи сенсори повністю, дозволяючи застосункам продовжувати функціонувати.

Завдяки OpenCV, Darkly має доступ до високооптимізованих та швидких алгоритмів для роботи з зображеннями, що дозволяє проводити обфускацію в реальному часі і мінімізувати затримки в роботі системи.

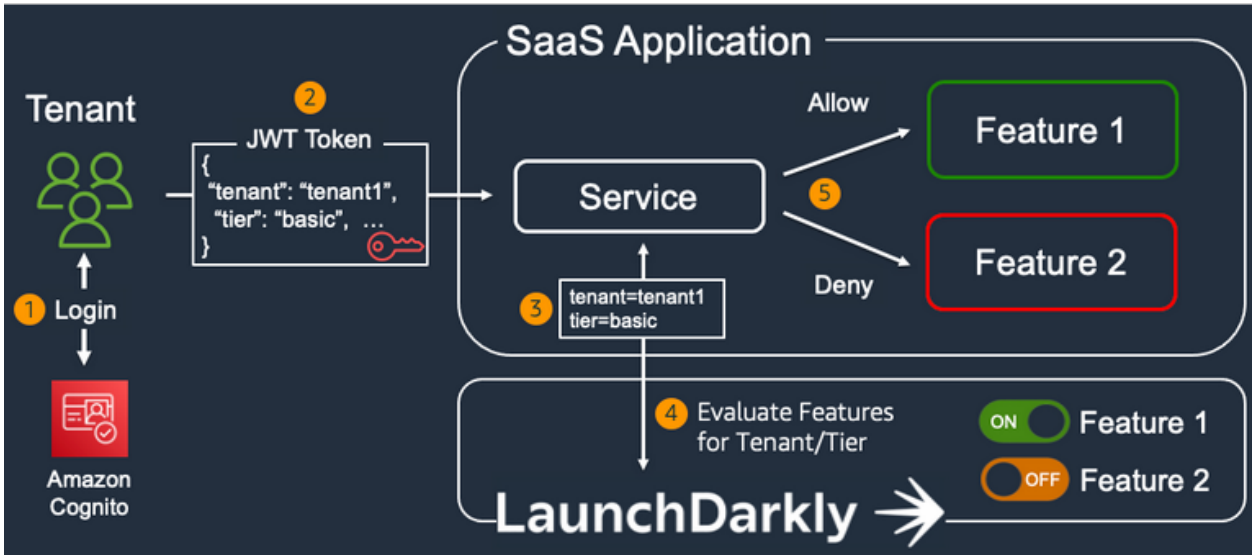

Рис. 1.5. Графічне представлення роботи Darkly

Архітектура Darkly зазвичай позиціонується як посередницький модуль або проксі-сервіс між джерелом візуальних даних та застосунком-споживачем.

1. Джерело Даних (Data Source)

Вхід - відеопотік (від камери) або знімок екрана. Дані є "сирими" і містять конфіденційну інформацію.

2. Модуль Перехоплення (Interception Module)

Функція - перехоплює запит від стороннього застосунку на доступ до візуальних даних.

3. Модуль Обробки OpenCV (OpenCV Processing Module)

Цей блок є ядром Darkly і використовує функції OpenCV:

Виявлення чутливих об'єктів - застосовує алгоритми виявлення облич, текстів документів, чи інших визначених конфіденційних об'єктів.

Генерація маски - створює бінарну маску, що покриває лише ідентифіковані чутливі області.

Обфускація - застосовує до областей, покритих маскою, обрану техніку приховування (наприклад, розмиття за Гаусом або пікселізацію).

4. Вихід обфускованих даних

Створюється модифікований (обфускований) потік/зображення.

5. Повернення до застосунку

Модуль перехоплення повертає обфусковані дані запитуючому застосунку замість оригінальних.

Ця архітектура забезпечує прозору обфускацію для застосунку: він отримує візуальні дані, як і очікував, але ці дані вже не містять конфіденційної інформації.

- Контроль доступу на мобільних пристроях - розробили схему контролю доступу для захисту приватних фотографій на мобільних телефонах, але вона обмежена лише попередньо збереженими обличчями, що надає лише частковий захист. На противагу, пропонована система надає захист для широкого спектра типів приватних фотографій.

- Автоматична класифікація конфіденційності – в [20] зібрали набір даних фотографій з Flickr з мітками (public, private або undecided), вилучили низькорівневі ознаки та навчили модель SVM для ідентифікації приватних фотографій. Розробили моделі навчання для оцінки адекватних налаштувань конфіденційності для спільних фотографій на основі того ж набору даних. Виявили, що існує значна розбіжність між фактичними налаштуваннями конфіденційності у Facebook та бажаним рівнем захисту, оскільки

користувачі часто використовують налаштування конфіденційності за замовчуванням, які є нижчими, ніж бажаний рівень захисту.

Пропонована система розвиває ці дослідження, інтегруючи контекстну обізнаність (стан блокування, стан застосунку) з автоматизованою класифікацією контенту для забезпечення динамічного та своєчасного контролю доступу до приватних фотографій, що є значним кроком вперед порівняно з існуючими статичними або обмеженими методами.

Висновки до розділу

У результаті аналізу предметної області встановлено, що мультимедійні дані у мобільних додатках є найбільш чутливим типом інформації через високі ризики несанкціонованого доступу. Досліджено існуючі моделі управління дозволами, які виявилися недостатньо ефективними для забезпечення конфіденційності користувачів та випадкових осіб. Визначено перспективні напрями розвитку — моніторинг поведінки додатків, автоматизована класифікація приватних зображень та застосування комп’ютерного зору для анонімізації.

Досліджено сучасні підходи до ідентифікації вразливостей мобільних застосунків, які пов’язані з доступом до камер, мікрофонів, геолокаційних сервісів та сховищ мультимедійного контенту. Визначено, що існуючі моделі дозволів у мобільних операційних системах, зокрема Android, не завжди забезпечують належний рівень контролю над доступом до конфіденційних ресурсів, що створює передумови для несанкціонованих витоків даних.

Виявлено проблему конфіденційності випадкових осіб (перехожих), які потрапляють у зону зйомки мобільних пристроїв без їхньої згоди, що формує новий рівень соціально-етичних і правових викликів у сфері захисту даних. Даний аспект обумовлює необхідність розробки автоматизованих методів розпізнавання та обробки облич на зображеннях з метою мінімізації ризиків вторгнення в особисте життя.

Окрему увагу приділено аналізу існуючих методологій захисту мультимедійних даних на мобільних платформах, серед яких виділяються: моніторинг поведінки застосунків для виявлення аномальної активності, класифікація приватних фотографій із застосуванням методів машинного навчання, а також моделі комп’ютерного зору для анонімізації даних третіх осіб. Ці підходи мають потенціал підвищення рівня захисту даних, проте характеризуються обмеженнями у продуктивності, точності та масштабованості.

РОЗДІЛ 2. МЕТОДИ ЗАХИСТУ МУЛЬТИМЕДІА В МОБІЛЬНИХ ДОДАТКАХ

2.1. Архітектура операційної системи Android та механізми безпеки

Операційна система Android є відкритою програмною платформою (open-source software stack), переважно орієнтованою на мобільні пристрої. В її основі лежить модифіковане Ядро Linux, яке забезпечує фундаментальні системні функції, такі як управління пам'яттю, процесами та базові механізми безпеки [4].

2.1.1. Стек програмного забезпечення Android та трасування системних викликів

Компоненти програмного стеку Android ієрархічно структуровані, як проілюстровано на рис. 2.1.

Системні виклики (system calls) є ключовим інтерфейсом взаємодії між процесами, що виконуються в просторі користувача, та Ядром Linux. Ці виклики є запитами до ядра на виконання операцій низького рівня, включаючи апаратні взаємодії. Наприклад, операція читання файлу ініціюється викликом Linux з сигнатурою: ssize_t read(int fd, void *buf, size_t count).

Розробники зазвичай використовують бібліотеки вищого рівня або API (Application Programming Interfaces) замість прямого маніпулювання системними викликами. Однак, усі високорівневі функції зрештою транслюються в базові системні виклики. Таким чином, моніторинг активності процесу на рівні системних викликів забезпечує повну прозорість щодо фактичної поведінки застосунку.

Оскільки Android базується на Ядрі Linux, усі виконувані системні виклики є стандартними викликами Linux. Ця архітектурна подібність дозволяє застосовувати інструменти командного рядка Linux.

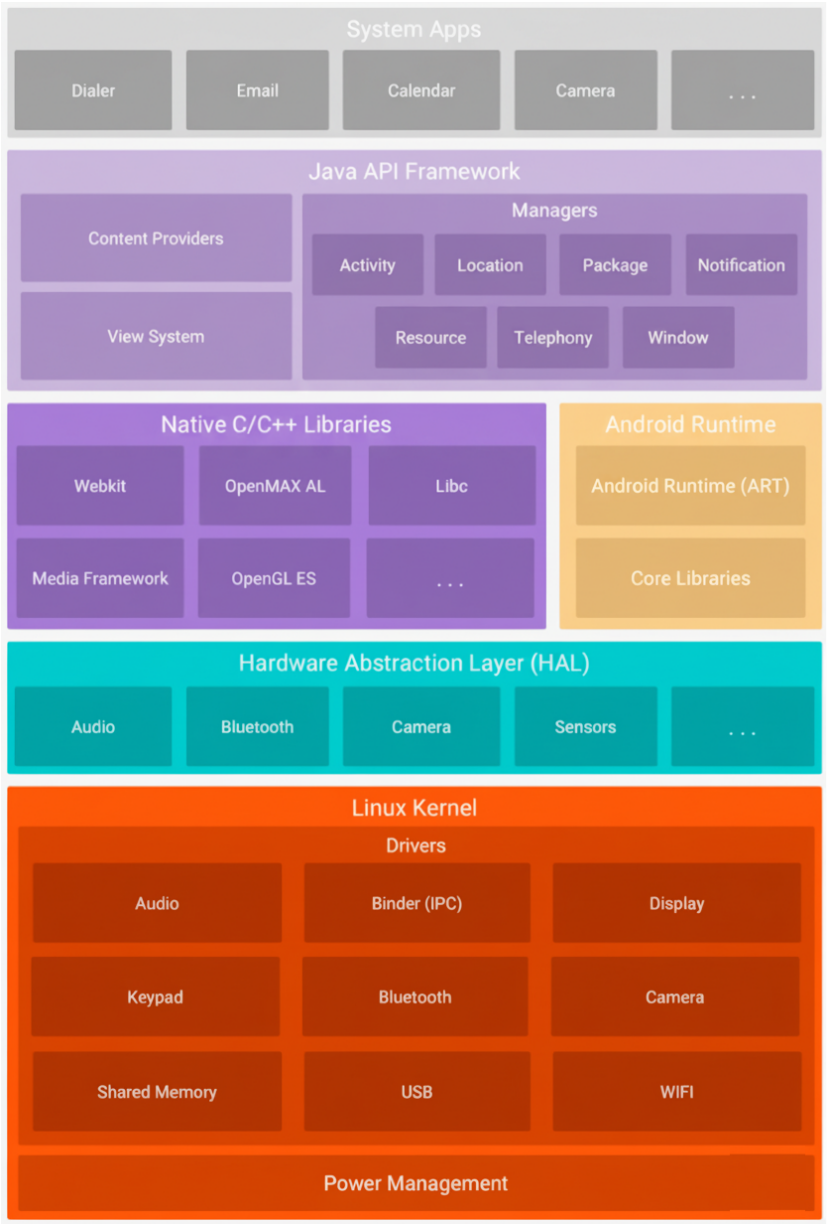

Рис. 2.1. Стек програмного забезпечення Android

Зокрема, утиліта Strace використовує функціонал ptrace ядра [5] для трасування та запису системних викликів практично будь-якого процесу. Використання Strace дозволяє фіксувати системні виклики основного процесу встановленого застосунку з моменту його запуску до завершення, охоплюючи також період його роботи у фоновому режимі.

2.1.2. Механізм дозволів безпеки застосунків Android

Кожен застосунок Android повинен декларувати необхідні дозволи (permissions) у своєму файлі маніфесту (manifest file). Дозволи

класифікуються відповідно до рівня потенційної загрози конфіденційності кінцевого користувача. Основними класами дозволів Android є:

- Нормальні дозволи (Normal permissions).
- Дозволи підпису (Signature permissions).
- Небезпечні дозволи (Dangerous permissions).

За чинною схемою, лише небезпечні дозволи вимагають явної згоди користувача під час виконання програми (runtime). До цієї категорії належать дозволи, які можуть спричинити значну шкоду у випадку компрометації (наприклад, надсилання SMS-повідомлень, доступ до камери або мікрофона пристрою) [6].

Нормальні дозволи, навпаки, можуть бути згруповані в маніфесті та автоматично надаються операційною системою під час встановлення. Доступ до мережі (Інтернет-дозвіл) включено до класу нормальних дозволів.

Для доступу до медіа-галереї пристрою застосунок повинен запитувати небезпечний дозвіл READ_EXTERNAL_STORAGE. Комбінація цього дозволу та дозволу на доступ до Інтернету теоретично надає застосунку можливість не лише читати файли з медіа-галереї користувача, але й передавати будь-які отримані дані через мережеве з'єднання пристрою. Це створює суттєвий ризик для приватності.

Детальний аналіз поширеності цього вектора загроз був проведений в [7], де досліджено сотні популярних застосунків Android для визначення кількості тих, що одночасно володіють дозволами READ_EXTERNAL_ STORAGE та доступом до Інтернету. Висновки підкреслили значну поширеність потенціалу витоку даних у комерційних застосунках.

2.2. Методи класифікації зображень для захисту мультимедійних даних

Класифікація фотографій за допомогою алгоритмів машинного навчання є актуальною та добре дослідженою сферою у комп'ютерному зорі,

яка інтенсивно підтримується великими технологічними корпораціями з метою розробки високоточних моделей для загального розпізнавання.

2.2.1. Класифікатори зображень та архітектура моделей глибокого навчання

Сучасні класифікатори зображень зазвичай реалізовані на базі глибоких згорткових нейронних мереж (CNN). Ці архітектури здатні автоматично витягувати мільйони ієрархічних ознак (features) із зображень, що забезпечує їхню високу ефективність у задачах класифікації. Згорткова нейронна мережа (CNN) являє собою тип багатошарової штучної нейронної мережі. Як проілюстровано на рис. 2.1, її архітектура складається з чергування згорткових шарів та шарів зменшення розмірності (наприклад, пулінгу). Після цих послідовних етапів до мережі додаються шари повного з'єднання (fully connected layers), що функціонально нагадують структуру багатошарового перцептрона. Далі буде розглянуто детальний опис кожного з цих шарів.

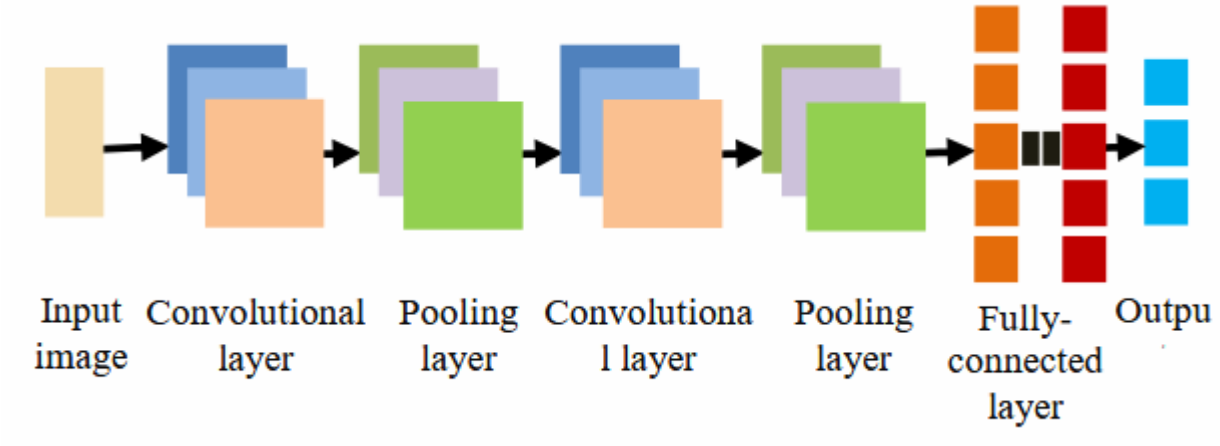

Рис. 2.1. Архітектура CNN

Convolutional Layers (Згорткові шари) відповідають за автоматичне вилучення ієрархічних ознак (hierarchical features) із вхідних даних шляхом застосування набору навчальних фільтрів (ядер).

Reduction Layers (Шари зменшення розмірності) найчастіше представлені шарами пулінгу (максимального або середнього), які зменшують просторові розміри карти ознак, знижуючи обчислювальне навантаження та забезпечуючи інваріантність моделі до невеликих зсувів та деформацій вхідних даних.

Fully Connected Layers (Шари повного з'єднання) розташовані наприкінці мережі, вони отримують високорівневі ознаки з останнього згорткового/пулінгового шару, перетворюють їх і виконують остаточну класифікацію або регресію.

Розробка таких моделей з нуля вимагає значних обчислювальних ресурсів (багатьох годин роботи GPU) та величезних обсягів розмічених даних для адекватного навчання узагальненому представленню ознак. Це позиціонує розробку передових класифікаторів як сферу, переважно доступну великим дослідницьким організаціям (наприклад, підрозділу TensorFlow компанії Google, який випустив такі моделі, як Inception-v3).

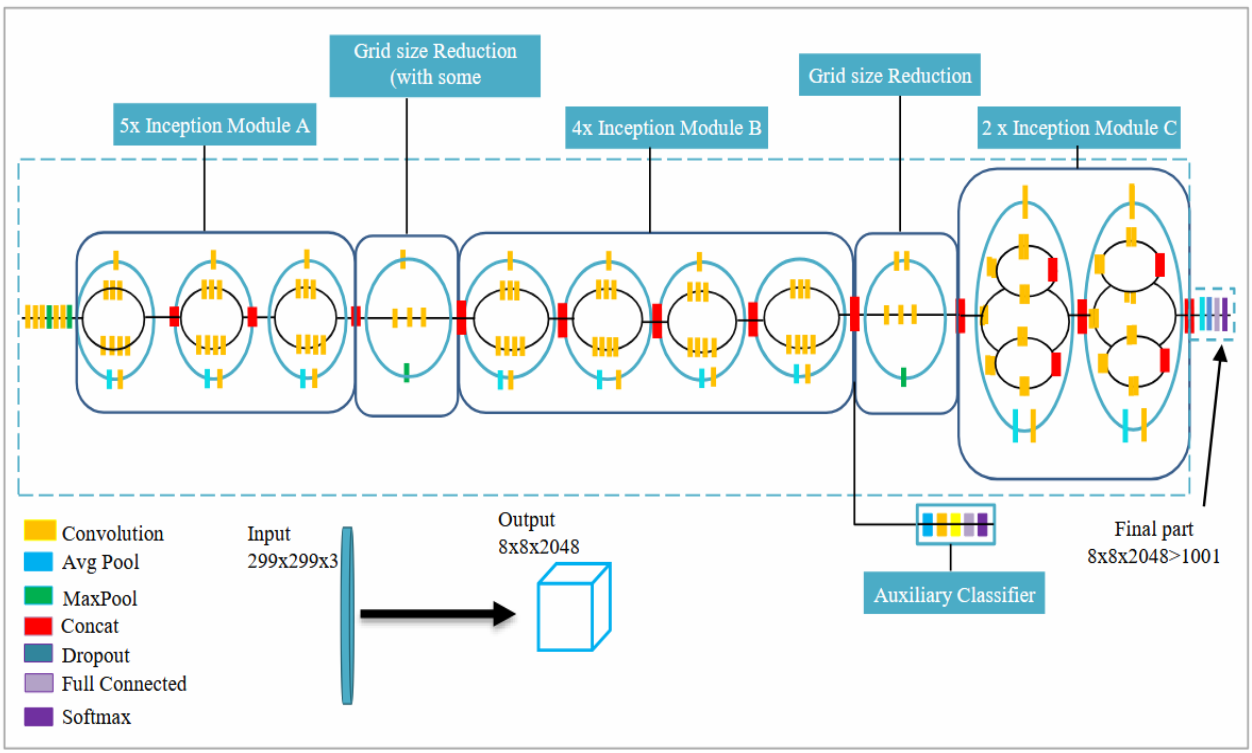

Рис. 2.2. Архітектура Inception-v3

Рисунок 2.2 ілюструє загальну структуру мережі, яка включає початкові шари згортки та пулінгу (пре-інцепшн блоки), за якими слідують серії Модулів Inception різних типів (наприклад, Type 1, Type 2, Type 3).

Розглянемо ключові елементи, відображені на рисунку 2.2.

Мережа починається з кількох стандартних згорткових шарів (Conv) і шарів максимального пулінгу (MaxPool), часто з використанням Batch Normalization.

Модулі Inception - це повторювані блоки, які є "серцем" архітектури. На схемі вони представлені як горизонтальні групи паралельних операцій, що включають:

- Згортки 1×1 (для зменшення розмірності).

- Факторизовані згортки (наприклад, 5×5 замінено на два 3×3, або 3×3 замінено на 1×n і n×1).

- Операції середнього (AvgPool) або максимального пулінгу (MaxPool).

- Виходи всіх паралельних гілок конкатенуються перед передачею на наступний модуль.

Зменшення розмірності (Grid Reduction) це спеціальні блоки, які ефективно зменшують просторову розмірність карт ознак, використовуючи поєднання пулінгу та згорток зі кроком 2.

Допоміжний класифікатор (Auxiliary Classifier) як правило, на схемі відображається один або два допоміжні класифікатори, підключені до проміжних шарів. Вони мають власну гілку, що закінчується шарами AvgPool, Conv, Fully Connected та Softmax.

Мережа завершується шаром глобального середнього пулінгу (Global AvgPool), за яким іде повністю з'єднаний шар (Fully Connected) та кінцевий шар Softmax для отримання ймовірностей класів (наприклад, 1000 класів ImageNet).

Продуктивність загальних класифікаторів часто оцінюється в рамках змагань, таких як ImageNet Large Visual Recognition Challenge (ILSVRC). Помилки вимірюються за метрикою "Top-5", де припущення вважається

неправильним, якщо жоден з п'яти найбільш імовірних прогнозів моделі не відповідає істинному класу. Наприклад, модель Inception-v3 досягла точності Top-5 на рівні 96.54% [8].

Для адаптації цих потужних загальних класифікаторів до специфічних завдань (наприклад, розпізнавання конфіденційних документів, посвідчень особи або селфі) традиційно потрібні значні часові та обчислювальні витрати. Однак, технологія перенесення навчання (Transfer Learning) суттєво зменшує ці вимоги.

Перенесення навчання ґрунтується на припущенні, що нейронна мережа, навчена на великому загальному наборі даних (наприклад, ImageNet), вже засвоїла універсальні ознаки зображень (лінії, кути, текстури). Цю здатність витягувати корисні ознаки можна перенести на нові, менші набори даних шляхом перенавчання лише верхніх шарів CNN. Як показано на рис. 2.2, лише класифікаційні шари (відсіювання, повністю з'єднаний шар та softmax) піддаються навчанню. Це призводить до суттєвого скорочення обчислювального часу за рахунок зменшення кількості параметрів для оновлення. Цей метод дозволяє досягти відмінної точності з мінімальними витратами ресурсів.

2.2.2. Альтернативні методи витягування ознак зображень

Крім CNN, існують також ефективніші, але менш обчислювально інтенсивні методи для витягування дискретних ознак зображень (image feature extraction). Ці методи розроблені для отримання інформації про візуальні патерни, такі як плями, краї, кути та гребені.

Трансформація ознак, інваріантна до масштабу (SIFT - Scale Invariant Feature Transform) обчислює ключові точки на зображенні, які є інваріантними до масштабу та обертання, що дозволяє використовувати їх для робустного розпізнавання об'єктів.

Ключова перевага SIFT полягає в його інваріантності до масштабу, обертання та частково до змін освітлення та шуму, що робить його

надзвичайно надійним для широкого спектру завдань, включаючи розпізнавання об'єктів та зшивання зображень.

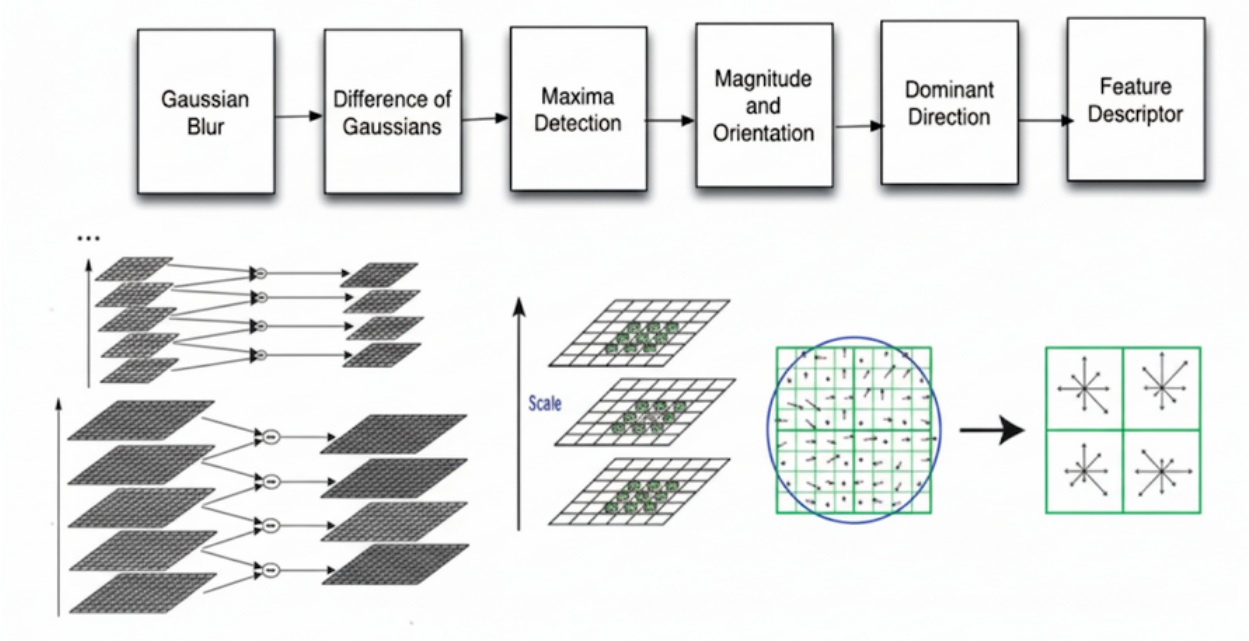

Рис. 2.3. Кроки SIFT алгоритму

На рисунку 2.3 показана послідовність, через яку алгоритм SIFT обробляє зображення:

1. Виявлення екстремумів масштабного простору (Scale-Space Extrema Detection). На цьому етапі демонструється створення Піраміди зображень та обчислення Різниці Гауссів (DoG) для виявлення потенційних ключових точок.

2. Локалізація ключових точок (Keypoint Localization). Показано, як уточнюється місце розташування екстремумів і відфільтровуються нестабільні точки на ребрах.

3. Призначення орієнтації (Orientation Assignment): Ілюструється, як навколо ключової точки обчислюються градієнти, створюється гістограма орієнтацій, і призначається домінантний напрямок, що забезпечує інваріантність до обертання.

4. Створення дескриптора (Keypoint Descriptor). Візуалізовано формування вектора ознак — типово сітки 4×4 з 8-напрямковими гістограмами градієнтів у кожній комірці, що дає 128-вимірний дескриптор.

Прискорений робочий алгоритм (SURF - Speeded Up Robust Features) - алгоритм, натхненний SIFT, який значно прискорює процес витягування ознак та пропонує високу стійкість до трансформацій зображення.

Графічне представлення методу Speeded Up Robust Features, SURF, який є оптимізованою альтернативою SIFT показано на рис. 2.4.

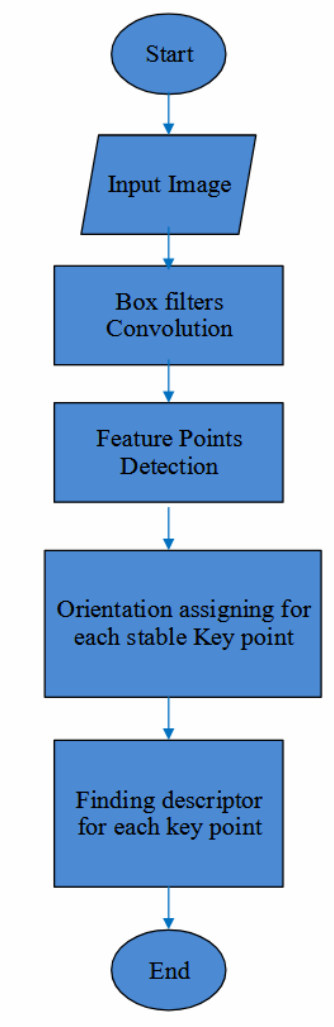

Рис. 2.4. Алгоритм SURF

Хоча детальна структура SURF схожа на SIFT, його ключові відмінності, спрямовані на прискорення обчислень, візуалізуються наступним чином:

1. Використання Інтегральних Зображень (Integral Images):

- SURF значно прискорює операції згортки, використовуючи інтегральні зображення. Це дозволяє швидко обчислювати суму пікселів у будь-якій прямокутній області за чотири операції, що є критичним для швидкої оцінки фільтрів.

2. Виявлення ключових точок через апроксимацію Гессіана (Approximation of Hessian Matrix):

- Замість Різниці Гауссів (DoG) у SIFT, SURF використовує детектор Гессіана для виявлення ключових точок.

- Фільтри Гессіана (ядра 3×3) апроксимуються за допомогою прямокутних фільтрів. Це, знову ж таки, дозволяє застосовувати інтегральні зображення для дуже швидкого обчислення визначника матриці Гессіана в різних масштабах.

3. Побудова масштабного простору (Scale Space):

На відміну від SIFT, який створює піраміду, зменшуючи розмір зображення, SURF підтримує постійний розмір зображення, але збільшує розмір прямокутних фільтрів Гессіана. Це забезпечує ефективне дослідження простору масштабу без необхідності субдискретизації.

4. Призначення орієнтації на основі хвильових перетворень Хаара (Haar Wavelet Response):

- Для забезпечення інваріантності до обертання SIFT використовує гістограми градієнтів. SURF використовує відгуки ґрадієнтів Хаара (Haar wavelet responses) у круговій області навколо ключової точки.

- Використання хвильових перетворень Хаара, обчислених за допомогою інтегральних зображень, є швидшим, ніж обчислення гістограм градієнтів.

5. Генерація Дескриптора (Descriptor Generation).

Дескриптор SURF (зазвичай 64-вимірний) є коротшим, ніж 128-вимірний дескриптор SIFT. Він формується шляхом обчислення суми

відгуків градієнтів Хаара (dx,dy, | dx | , | dy |) у підрегіонах (4×4 комірок) навколо ключової точки.

Таким чином, графічна схема SURF відображає його основну філософію: досягнення надійності, порівнянної з SIFT, але зі значно вищою швидкістю обробки завдяки застосуванню інтегральних зображень та апроксимованих фільтрів Гессіана.

Дескриптор ознак гістограми орієнтованих градієнтів (HOG) - цей дескриптор функціонує шляхом поділу зображення на невеликі просторові комірки та обчислення гістограми напрямків градієнтів у кожній комірці [9]. Це ефективно фіксує локальну інформацію про краї та форму (рис. 2.5).

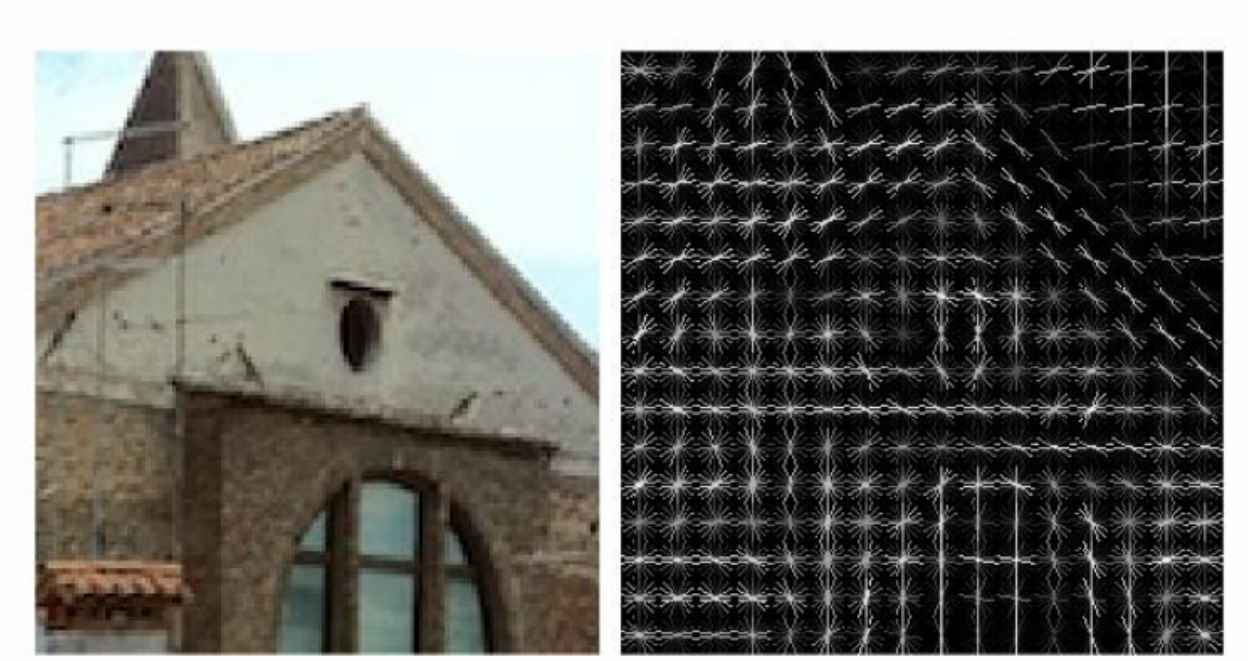

Рис. 2.5. Візуалізація HOG дескрипторів

Дескриптор гістограми орієнтованих градієнтів (Histogram of Oriented Gradients, HOG) є потужним алгоритмом вилучення ознак (feature extraction), який широко використовується в комп'ютерному зорі, зокрема для розпізнавання об'єктів і, що найбільш відомо, для виявлення людей. Основна ідея методу полягає в тому, що локальний вигляд і форма об'єкта можуть бути точно описані розподілом інтенсивності локальних градієнтів та

напрямків країв. Алгоритм HOG оперує на локальних ділянках зображення і формує вектори ознак, інваріантні до геометричних та фотометричних перетворень (окрім обертання та зміни перспективи). HOG, у поєднанні з SVM, став золотим стандартом для виявлення людей на статичних зображеннях та у відео, а його відносна простота та швидкість роблять його придатним для використання на пристроях з обмеженими ресурсами.

Ці традиційні алгоритми витягування ознак можуть слугувати альтернативою складним нейронним мережам для задач розпізнавання об'єктів. Витягнуті ознаки можуть бути використані для навчання будь-якого алгоритму навчання з учителем (supervised learning), включаючи:

- Машини опорних векторів (SVM).
- Лінійна регресія.
- Дерева рішень.

Кожен з цих алгоритмів пропонує унікальні переваги щодо швидкості навчання, інтерпретованості та стійкості до різних типів даних. Для завдань, критичних до швидкості та обмежених обчислювальними ресурсами мобільних пристроїв, гібридні підходи (наприклад, HOG + SVM) можуть бути більш практичним рішенням, ніж використання повномасштабних CNN.

2.3. Науковий аналіз диференціації осіб: виявлення цільових суб'єктів та випадкових перехожих

Розпізнавання обличчя (Face Recognition) є динамічною галуззю комп'ютерного зору, що охоплює автоматизоване виявлення та ідентифікацію людей на основі вилучених біометричних ознак обличчя. Хоча існують численні дескриптори, що використовуються для унікальної ідентифікації, дане дослідження зосереджено на більш загальному завданні: простому виявленні облич (face detection) та, головне, їхній подальшій диференціації за контекстом.

2.3.1. Обмеження базового виявлення облич

Базове виявлення обличчя, яке полягає лише в ідентифікації присутності обличчя на зображенні, може бути реалізоване за допомогою загальних алгоритмів вилучення ознак, таких як SIFT, SURF або HOG, які були описані раніше. Ці алгоритми функціонують ідентично, як і при розпізнаванні будь-якого іншого об'єкта. Однак використання такої базової моделі не дозволяє вилучати специфічну для обличчя інформацію (face-specific features). Такі параметри, як висота, ширина, орієнтація (кут рискання, кут нахилу, кут обертання) та інші дескриптори, є критично важливими для встановлення відмінностей між виявленими об'єктами.

2.3.2. Модель диференціації «Ціль/Перехожий»

Основна мета цього етапу дослідження полягає в розробці моделі машинного навчання, здатної диференціювати людей, які є фокусом фотографії (цільові суб'єкти), від випадкових незнайомців (перехожих), які ненавмисно потрапили в кадр. Це специфічне контекстуальне застосування виявлення облич є значно менш дослідженим, ніж задачі унікальної ідентифікації.

Ранній прототип такої моделі був розроблений в [7] де розроблено метод вилучення набору релевантних ознак обличчя, які є важливими для розрізнення двох категорій:

1. Метричні ознаки: Розмитість (blurriness), розмір обличчя та позиція обличчя на зображенні.

2. Орієнтаційні/Мімічні ознаки: Кут нахилу (pitch), кут рискання (yaw), кут обертання (roll), наявність посмішки.

Використовуючи ці комбіновані ознаки, було досягнуто 93.27% точності класифікації за допомогою алгоритму градієнтного бустингу дерев рішень (Gradient Boosting Decision Tree). Цей результат підтвердив ефективність даної концепції для диференціації цільових осіб від випадкових перехожих.

2.3.3. Вдосконалення існуючої моделі

У даній магістерській роботі пропонується покращення існуючої моделі, шляхом:

- Модифікації обчислення параметрів: внесення змін у методи обчислення кількох існуючих параметрів, зокрема розміру обличчя та його просторового розташування на зображенні.

- Інтеграції нової ознаки: додавання нової, високоінформативної ознаки, яка вимірює відносний напрямок погляду очей суб'єкта відносно об'єктива камери.

Рис. 2.6. Приклад фото з цілями та перехожими

На рис. 2.6 демонструється типовий сценарій: цільові суб'єкти (Принц Гаррі та Меган Маркл) знаходяться на передньому плані та є повністю сфокусованими. Натомість, випадкові перехожі розташовані на задньому плані, їхні обличчя мають менший відносний розмір і є значно більш розмитими.

Ефективна диференційна модель повинна використовувати ці метричні та просторові відмінності для надійного розрізнення двох груп. Це дозволить в подальшому застосовувати механізми обфускації обличчя (наприклад, розмиття або маскування) для випадкових перехожих з метою підвищення їхньої індивідуальної конфіденційності.

Висновки до розділу

У цьому розділі проаналізовано архітектуру Android та механізми безпеки, що дозволило виявити їхні сильні сторони й обмеження у захисті мультимедійних даних. Досліджено методи класифікації зображень, включно з традиційними підходами та глибокими згортковими нейронними мережами. Показано, що саме DCNN забезпечують найвищу точність при виявленні та диференціації чутливих мультимедійних об’єктів.

РОЗДІЛ 3. МОДЕЛІ ТА МЕТОДОЛОГІЯ ЗАХИСТУ МУЛЬТИМЕДІЙНИХ ДАНИХ В МОБІЛЬНИХ ДОДАТКАХ

3.1. Методологія динамічного аналізу поведінки мобільних додатків детектування потенційних витоків даних

Розроблено двоетапну методологію для аналізу поведінки застосунків на платформі Android, спрямовану на виявлення несанкціонованого доступу до медіа-каталогу та ідентифікації потенційних каналів витоку даних користувача. Цей процес забезпечує механізм для розкриття точних операцій читання зображень, ініційованих сторонніми застосунками, та виокремлення тих, які вимагають додаткового дослідження на предмет шкідливої активності.

Для демонстрації ефективності методології було відібрано 15 найбільш популярних безкоштовних застосунків з Google Play Store для комплексного тестування. Цей підхід є вдосконаленням існуючого аналізу дозволів, шляхом інтеграції динамічного аналізу поведінки.

3.1.1. Етап статичного аналізу дозволів

Аналіз дозволів є першим кроком і призначений для встановлення теоретичної здатності застосунку до витоку даних, виходячи з його конфігурації безпеки.

1. Отримання APK-файлів.

Для доступу до файлу маніфесту (AndroidManifest.xml) кожного застосунку необхідно отримати його інсталяційний файл APK. Оскільки офіційні репозиторії, як-от Google Play Store, не надають прямого доступу до цих файлів, було використано перевірений сторонній репозиторій (наприклад, APKMirror.com). Для забезпечення автентичності версії APK проводиться порівняння контрольних сум з офіційними релізами.

2. Витягнення Маніфесту.

За допомогою інструментів аналізу (наприклад, Android Studio APK Analyzer) з APK-файлу витягується AndroidManifest.xml у читабельному форматі.

3. Оцінка Вразливості.

Застосунок вважається потенційно вразливим до витоку зображень, якщо він одночасно вимагає два ключові дозволи:

- INTERNET (Нормальний дозвіл, надається автоматично).

- READ_EXTERNAL_STORAGE (Небезпечний дозвіл, вимагає згоди користувача).

Згідно з проведеним аналізом, лише один із протестованих застосунків (The Weather Channel) не мав необхідної комбінації дозволів INTERNET і READ_EXTERNAL_STORAGE для потенційного витоку медіа-даних.

3.1.2. Етап динамічного аналізу поведінки через системні виклики

Друга частина дослідження фокусується на динамічному моніторингу активності основного процесу застосунку в реальному часі.

1. Середовище тестування.

Експеримент проводився на емульованому пристрої (віртуальний Google Pixel з Android). Через обмеження системних образів емулятора, що не дозволяють запускати налагоджувальні інструменти з функціональністю Google Play Store, застосунки встановлювалися вручну за допомогою утиліти Android Debug Bridge (ADB).

2. Моніторинг процесів.

Після встановлення та запуску застосунку, ідентифікатор основного процесу (PID) визначався за допомогою стандартної утиліти Linux top через оболонку ADB.

3. Трасування системних викликів.

Для запису системних викликів використовувалася утиліта Strace, запущена з оболонки ADB. Була налаштована фільтрація, щоб реєструвати

лише ті системні виклики, які запитують операції читання у медіа-каталозі пристрою.

4. Експериментальні стани.

Кожен тестований застосунок працював у трьох послідовних станах протягом 30 хвилин у кожному:

- Робота на передньому плані (foreground).
- Робота у фоновому режимі (background).
- При заблокованому пристрої.

Таблиця 3.1.

Записана активність доступу до галереї фотографій

Назва додатку	**Активність на передньому плані**	**Активність у фоновому режимі**	**Активність при заблокованому телефоні**
Facebook Messenger	ТАК	НІ	НІ
Instagram	ТАК	НІ	НІ
Snapchat	ТАК	НІ	НІ
WhatsApp	ТАК	НІ	НІ
Netflix	НІ	НІ	НІ
YouTube	ТАК	НІ	НІ
Wish	ТАК	НІ	НІ
Spotify	НІ	НІ	НІ
Cash App	ТАК	НІ	НІ
Walmart	ТАК	НІ	НІ
Amazon	ТАК	НІ	НІ
PayPal	ТАК	НІ	НІ
Venmo	ТАК	НІ	НІ
SoundCloud	НІ	НІ	НІ
The Weather Channel	НІ	НІ	НІ

Результати динамічного аналізу показали, що жоден із протестованих застосунків не ініціював запити на читання до медіа-каталогу, коли вони

працювали у фоновому режимі або коли пристрій був заблокований. Більшість застосунків, які мали необхідні дозволи, здійснювали доступ до медіа-папки лише під час роботи на передньому плані (стан ТАК).

Така поведінка, ймовірно, пов'язана з користувацькими функціями (наприклад, завантаження фотографій для профілю чи обміну), і в контексті відомих корпорацій, залучених у тестування, ця активність узгоджується з їхніми політиками конфіденційності. Це не обов'язково вказує на шкідливу поведінку.

3.1.3. Значимість методології

Незважаючи на відсутність підозрілої активності серед популярних застосунків, описаний метод аналізу є легко відтворюваним і має високу цінність для скринінгу менш відомих або ненадійних застосунків. Виявлення будь-яких запитів на читання медіа у фоновому режимі або при заблокованому телефоні, хоча й не є остаточним доказом шкідливості, беззаперечно вимагає подальшого детального дослідження для встановлення причини такої неявної активності.

3.2. Розробка автоматизованого класифікатора для ідентифікації конфіденційних даних мультимедіа

Для забезпечення ефективного захисту приватної мультимедійної інформації та запобігання несанкціонованому доступу критично необхідна автоматизована ідентифікація конфіденційних фотографій. Ручне сортування великих масивів даних є неефективним та нереалістичним.

У цьому розділі представлено реалізацію та оцінку класифікатора зображень на основі методів машинного навчання, здатного прогнозувати наявність певних конфіденційних мультимедійних даних на фотографії з високою точністю.

3.2.1. Формування навчального набору даних

Першим етапом у створенні спеціалізованого класифікатора є формування репрезентативного навчального набору даних. Через конфіденційний характер вмісту, який класифікується, публічно доступні набори даних приватних фотографій практично відсутні. З огляду на це, було створено спеціалізований набір даних, що включає п'ять класів:

- Публічні (Public)
- Селфі (Selfies)
- Посвідчення особи (ID Documents)
- Документи (General Documents)
- Сімейні портрети (Family Portraits)

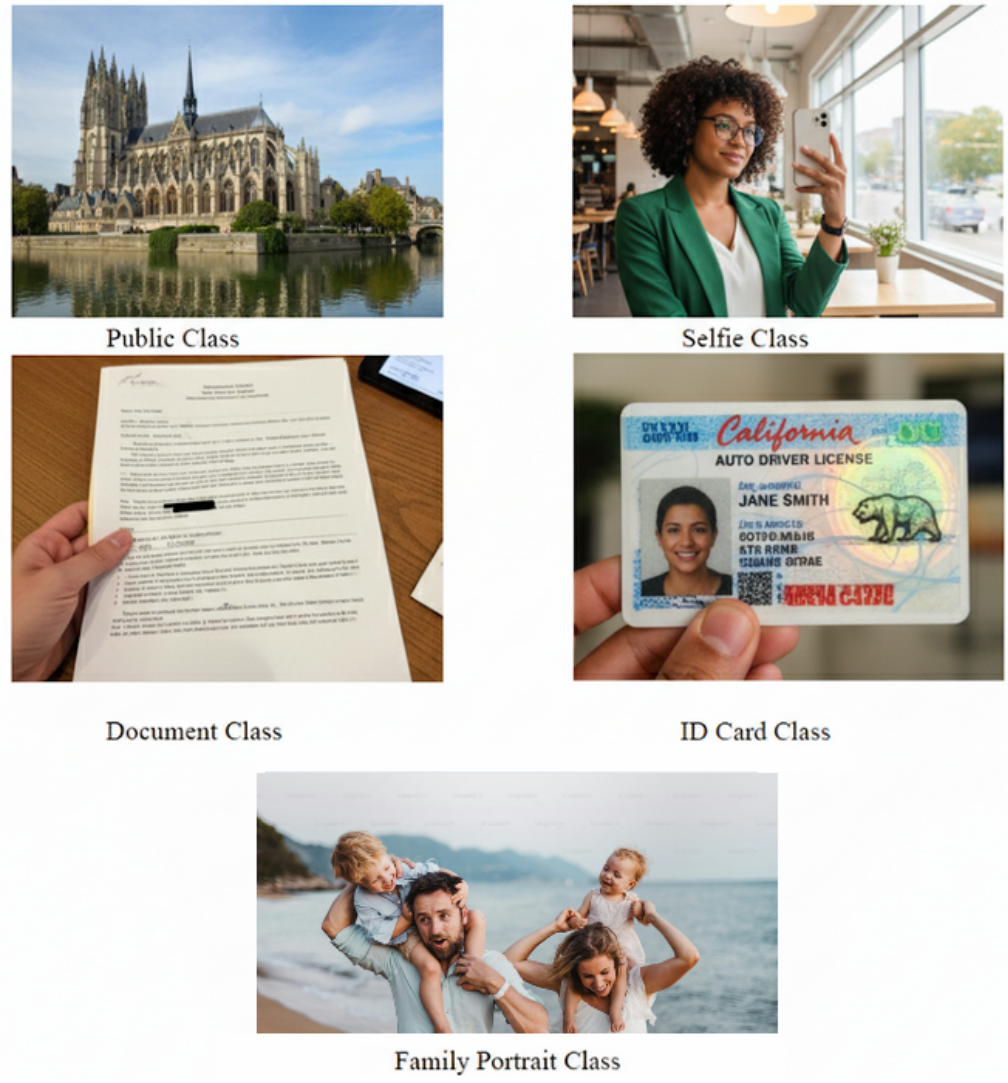

Рис. 3.1. Класи фото вибрані для процесу навчання

Зразки для кожного класу були зібрані з відкритих джерел, таких як Google Images, з метою створення набору даних обсягом приблизно 1000 зображень (по ≈200 зображень на кожен клас). Хоча такий обсяг є відносно невеликим порівняно з деякими сучасними великомасштабними навчальними корпусами, він виявився достатнім для тренування декількох точних моделей.

3.2.2. Порівняння підходів класифікації

З метою визначення оптимального алгоритму та архітектури, що забезпечують найвищу точність і продуктивність, було порівняно кілька різних підходів у галузі машинного навчання.

Глибокі згорткові нейронні мережі (DCNN)

Було навчено та протестовано кілька сучасних архітектур DCNN. Очікується, що ці просунуті моделі забезпечать високу точність завдяки їхній здатності автоматично вивчати комплексні ієрархічні ознаки.

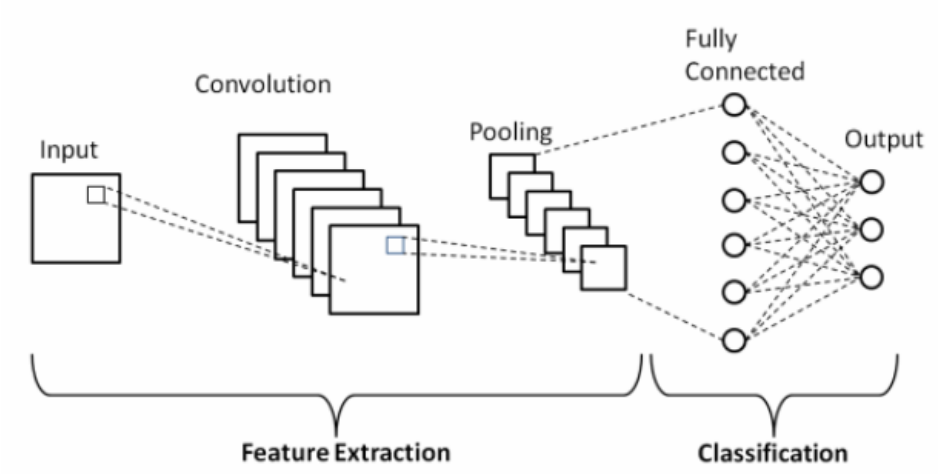

Рис. 3.2. Компоненти архітектури глибокої згорткової нейронної мережі для класифікації зображень

На цій схемі (рис. 3.2) візуалізовані ключові компоненти, які утворюють DCNN для класифікації зображень:

1. Згортковий Базис (Feature Extraction)

Цей розділ, як правило, складається з кількох повторюваних блоків, що виконують вилучення ознак:

- Вхідне зображення (input image): зображення подається у вигляді тривимірного тензора (ширина × висота × канали).

- Згортковий шар (conv layer): представлений як куб, що створює карти ознак. На схемі показано, як застосування фільтрів зменшує просторовий розмір, але збільшує глибину (кількість каналів/фільтрів).

- Шар ReLU (Rectified Linear Unit): застосовує нелінійність до виходу згорткового шару.

- Шар пулінгу (Pooling Layer): відображається як зменшення просторового розміру куба, що допомагає узагальнювати ознаки та зменшувати обчислювальне навантаження.

2. Класифікаційна "голова" (Classification Head)

- Згладжування (Flatten / Global Pooling): Перетворює останню тривимірну карту ознак на одновимірний вектор.

- Шари повного з'єднання (Fully Connected Layers / FC): Відображаються як послідовність щільно з'єднаних нейронів (як у традиційному перцептроні), що виконують високоабстрактне міркування.

- Вихідний шар (Output Layer): Кінцевий шар із функцією Softmax, який видає ймовірності для кожного класу.

Традиційний підхід на основі дескрипторів

Для встановлення базового рівня точності (baseline) використано більш традиційну методологію. Вона передбачає застосування багатокласової машини опорних векторів (SVM), навченої на витягнутих вручну ознаках.

Графічно SVM найкраще ілюструється її основною концепцією — максимізацією поля розділення (margin).

На двовимірній схемі, де точки даних (вектори ознак) представляють зображення (рис. 3.3):

1. Точки даних: зображення представляються як точки у багатовимірному просторі ознак.

2. Розділяюча гіперплощина: SVM знаходить оптимальну гіперплощину, яка найкраще розділяє класи.

3. Опорні вектори (Support Vectors): це ті найближчі до гіперплощини точки (зображення), які визначають її положення. Вони мають вирішальне значення для моделі.

4. Поле (Margin): SVM прагне максимізувати відстань між гіперплощиною та найближчими опорними векторами.

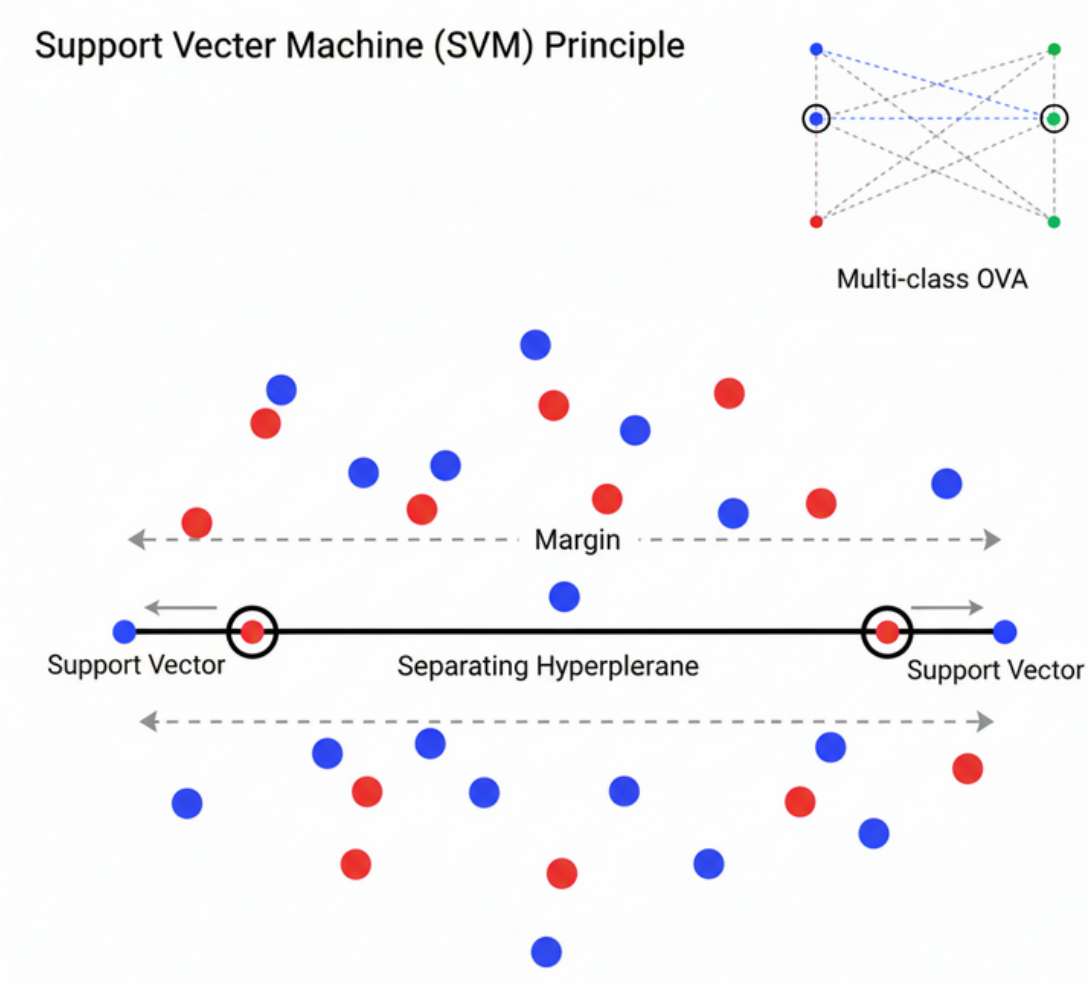

Рис. 3.3. Візуалізація принципу роботи SVM

У багатокласовому випадку OvA це уявлення просто повторюється для кожної пари "клас проти решти". Створюється 2N(N−1) бінарних SVM-

класифікаторів. Кожен класифікатор тренується, щоб відрізняти лише дві конкретні пари класів. Прогноз визначається голосуванням - кожен класифікатор видає один голос на користь одного з двох класів, які він розрізняє, і обирається клас із найбільшою кількістю голосів.

Як дескриптор обрано гістограму орієнтованих градієнтів (HOG) через її доведену ефективність у застосуваннях, пов'язаних із розпізнаванням об'єктів та форм.

3.2.3. Критерії оцінки

Усі реалізовані методи будуть порівнюватися за двома ключовими метриками: точність класифікації та продуктивність на мобільних пристроях, що є критично важливим для вибору найбільш ефективної моделі для розгортання в середовищі кінцевого користувача.

3.3. Порівняльний аналіз класифікаторів зображень для ідентифікації конфіденційних даних

3.3.1. Еталонна продуктивність: SVM з дескрипторами HOG

Першим етапом дослідження була оцінка базової продуктивності традиційного алгоритму — машини опорних векторів (SVM), навченої на ознаках гістограми орієнтованих градієнтів (HOG).

Навчальний набір даних було розділено на навчальну (приблизно 90%) та тестову (решта 10%) вибірки.

Для забезпечення сумісності з реалізацією HOG у бібліотеці OpenCV, всі зображення були стандартизовані до єдиного розміру 500×500 пікселів. Був застосований метод інтерполяції за співвідношенням площі пікселів; очікується, що інші методи інтерполяції дадуть порівнянні результати.

Після обчислення ознак HOG, було проведено навчання багатокласового SVM з подальшим налаштуванням гіперпараметрів (C та гамма).

Фінальна точність класифікації на тестовій вибірці для моделі SVM + HOG склала 81.3%. Цей показник встановлює еталон для порівняння з більш сучасними методами.

3.3.2. Оцінка глибоких згорткових нейронних мереж (DCNN)

Для порівняння були протестовані декілька сучасних архітектур DCNN за допомогою бібліотеки TensorFlow. Моделі були розділені на дві категорії на основі їхніх обчислювальних вимог.

Таблиця 3.2.

Тестовані моделі DCNN та їх класифікація продуктивності

Категорія	**Моделі DCNN, що тестувалися**
Орієнтовані на мобільні пристрої	MobileNet, NASNet-A (мобільний)
Обчислювально вимогливі	Inception-v3, Inception-v4, NASNet-A (великий)

Через обмежений обсяг навчального набору даних використовувався підхід перенесення навчання (Transfer Learning). Це дозволило мережам скористатися раніше вивченими ваговими коефіцієнтами, донавчивши їх розпізнавати п'ять цільових класів (публічні, селфі, посвідчення особи, документи, сімейні портрети).

Ефективність кожної моделі була оцінена на випадково вибраній тестовій вибірці. У таблиці 3.3 і рисунку 3.4 наведено середні результати за 10 оцінок.

Таблиця 3.3.

Порівняння точності класифікації тестованих моделей DCNN та моделі SVM

Модель	Середня Точність Класифікації
NASNet-A (великий)	94.1%
Inception-v4	91.4%
Inception-v3	91.3%

NASNet-A (мобільний)	89.7%
MobileNet-v2	88.5%
SVM + особливості HOG	81.3%

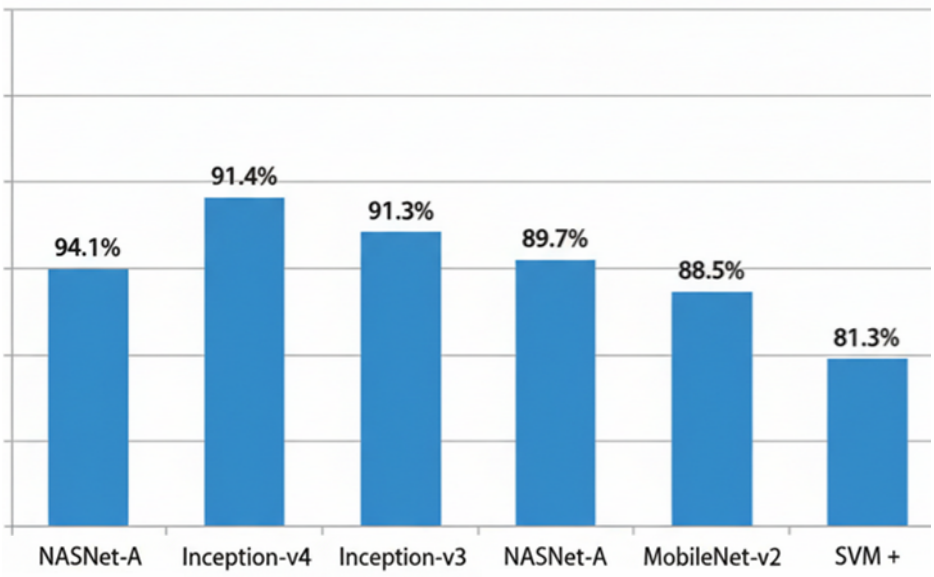

Рис. 3.4. Графічне представлення точності класифікації

3.3.3. Аналіз точності та обчислювальної ефективності

Обчислювально вимогливі моделі продемонстрували вищу точність порівняно з мобільними та базовою SVM. Модель NASNet-A (великий) показала найкращу продуктивність, досягнувши точності 94.1%.

Незважаючи на високу точність, модель NASNet-A (великий) є найменш придатною для розгортання на мобільних пристроях з архітектурою ARM через значні обчислювальні вимоги (23.4 мільярда операцій множення-додавання на одне зображення).

Модель NASNet-A (мобільний), хоча й має зниження точності на 4.4% порівняно з лідером (89.7% vs 94.1%), вимагає в середньому лише 564 мільйони операцій множення-додавання на зображення [11]. Це являє собою скорочення обчислювальної складності на 97.6% і робить NASNet-A (мобільний) найбільш бажаним вибором для застосувань, що функціонують на апаратному забезпеченні з обмеженими ресурсами.

3.4. Вдосконалення моделі диференціації осіб через інтеграцію розширених біометричних ознак

Покращення існуючої моделі [7] для розпізнавання цільових суб'єктів та випадкових перехожих на фотографіях вимагає перегляду методів обчислення ознак та інтеграції нових високоінформативних дескрипторів. Оригінальна модель базувалася на таких ознаках: посмішка, розмір обличчя, позиція обличчя, кути орієнтації обличчя (рискання, нахил, обертання) та розмитість [7]. У пропонованому підході зберігаються найбільш ефективні ознаки (наприклад, розмитість за Гаусом), але вносяться суттєві зміни у розрахунок розміру та розташування обличчя, а також додається нова, критично важлива метрика.

3.4.1. Модифікація існуючих ознак

1. Удосконалення обчислення розміру обличчя

Оригінальний метод [7] використовував площу квадратної обмежувальної рамки, наданої Android SDK FaceDetector, що часто призводило до спотворення справжнього розміру обличчя через надмірний або нерівномірний обхват області.

У даній роботі використовується фреймворк OpenFace, відомий своїми передовими алгоритмами виявлення та вилучення ознак. OpenFace — це потужний і відкритий (open-source) фреймворк комп'ютерного зору, розроблений для аналізу та розуміння людських облич у реальному часі. Він широко використовується в академічних дослідженнях та прикладних системах завдяки своїм найсучаснішим алгоритмам для виявлення обличчя, розпізнавання, відстеження ключових точок, оцінки орієнтації (пози) та аналізу виразу обличчя .

OpenFace дозволяє витягувати набір ключових точок навколо обличчя (рис. 3.5) , що уможливлює точне обчислення висоти та ширини обличчя на

основі координат цих точок. Цей метод замінює менш точний розрахунок площі квадрата.

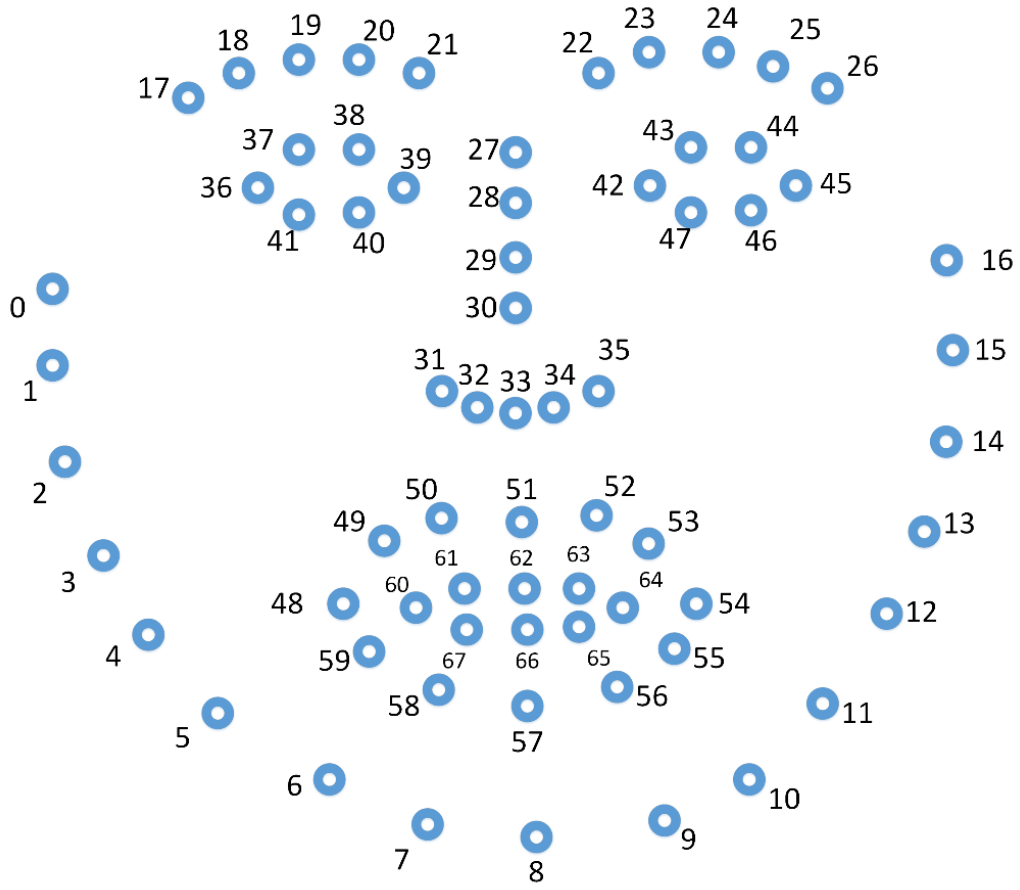

Рис. 3.5. Діаграма всіх вихідних орієнтирів обличчя, отриманих за допомогою OpenFace

2. Перегляд метрики розташування обличчя

Оригінальна модель використовувала бінарну класифікацію позиції, розділяючи зображення на ліву, середню та праву секції, де обличчя за межами середньої секції, ймовірно, не було цільовим суб'єктом.

Новий метод. Бінарна евристика замінена на параметр відстані, що вимірює абсолютне піксельне відхилення центру виявленої голови від геометричного центру зображення (по координатах x та y).

Перехід до спектрального (дистанційного) значення дозволяє створити неперервну ймовірнісну функцію, де ймовірність того, що обличчя є

цільовим, безперервно зростає зі зменшенням виміряного відхилення. Це підвищує гнучкість моделі, уникаючи жорсткості бінарних правил.

3. Додавання нової ознаки - напрямок погляду очей

Ключовим нововведенням є інтеграція вимірювання напрямку погляду очей (eye gaze direction).

Для цього використовується модуль відстеження погляду OpenFace.

Модуль вимірює вектор погляду кожного ока відносно сенсора камери. Велика простежена різниця від походження камери вказує на те, що суб'єкт, ймовірно, дивиться вбік, тоді як мала різниця свідчить про прямий погляд на об'єктив.

Приклади оброблених зображень із відстеженням погляду та вилученням ознак представлені на рис. 3.6 та 3.7.

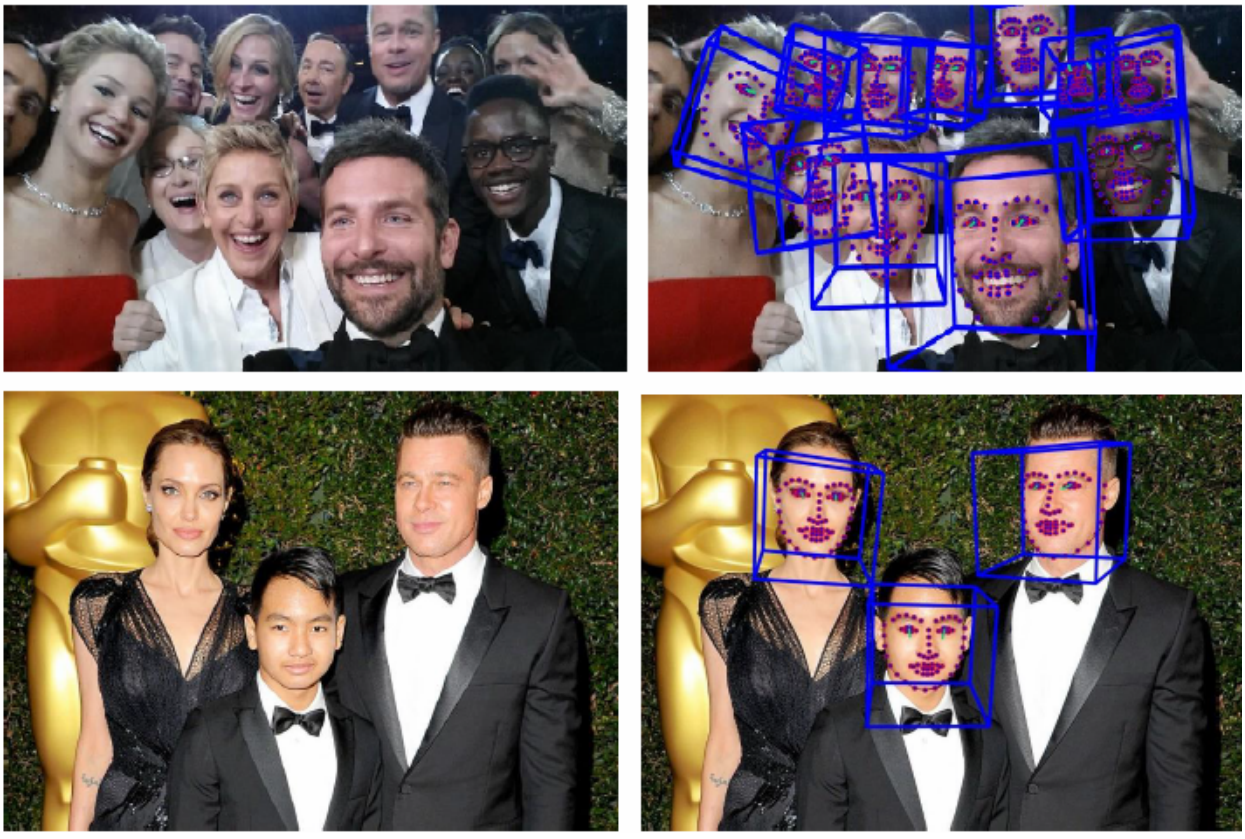

Рис. 3.7. Приклад оброблених зображень за допомогою OpenFace з розпізнаванням погляду, показаних яскраво-зеленими векторами

Більшість об'єктів зйомки дивляться на сенсор камери або поблизу.

Рис. 3.8. Приклад зображення, обробленого за допомогою OpenFace, де об'єкти зйомки не дивляться на сенсор камери

Хоча окремі перехожі можуть випадково дивитися на камеру, а цілі — вбік, ця метрика в комбінації з іншими (розмитість, центральна позиція) посилює диференціацію. Очікується, що цільові суб'єкти частіше матимуть сфокусовані обличчя, розташовані в центрі, що пом'якшує випадки, коли незнайомці дивляться прямо на камеру.

3.4.2. Методологія навчання та тестування

Для навчання та валідації класифікаторів використовувався розширений набір даних. До оригінального набору даних (≈200 фотографій) було додано ≈100 вручну відібраних фотографій, збільшивши загальний розмір до ≈300 зображень. Невеликий розмір набору даних обумовлений високою трудомісткістю його створення: корисні фотографії повинні містити одночасно цільових суб'єктів та випадкових перехожих, а кожне обличчя має бути ручно анотовано.

Для порівняння та оцінки покращень використано ті ж алгоритми машинного навчання, що і в роботі [7]:

- Градієнтний бустинг дерев рішень (Gradient Boosting Decision Tree)
- Машина опорних векторів (Support Vector Machine, SVM)
- Випадковий ліс (Random Forest)

- Проста нейронна мережа (Simple Neural Network)

Всі алгоритми були реалізовані та навчені за допомогою бібліотеки Scikit-learn на базі Python, що забезпечило надійну та уніфіковану платформу для експериментальної оцінки.

3.5. Імплементація моделей і методів для розробки системи захисту даних мультимедіа

3.5.1. Концепція та цілі проєктування

Система розроблена з метою захисту приватних фотографій користувача від несанкціонованого доступу з боку мобільних застосунків, при цьому не змінюючи стандартні механізми доступу до медіа-даних та користувацькі звички зберігання фотографій. Ключовою вимогою є забезпечення мінімального впливу на юзабіліті застосунків та користувацький досвід, що вимагає прийняття рішень щодо контролю доступу в реальному часі (real-time response).

Основний принцип роботи системи полягає у впровадженні контекстно-залежного контролю доступу. При запиті застосунку на доступ до конкретного фото, користувач повинен бути свідомий цього запиту. Наївний підхід, який передбачає постійне виведення запитів на підтвердження, є неприйнятним через значне погіршення юзабіліті.

Система вирішує цю проблему шляхом двокомпонентного автоматизованого аналізу:

1. Автоматична класифікація контенту - визначення, чи містить фотографія конфіденційну інформацію.

2. Контекстна обсвідомленість користувача (Awareness Check) - оцінка, чи є користувач свідомим про запит на доступ на основі поточного стану системи та застосунку.

Контекстні правила свідомості користувача системи формуються таким чином:

- Якщо пристрій заблокований (locked) або застосунок працює у фоновому режимі (background), користувач, імовірно, не усвідомлює запиту доступу. У цих сценаріях запит слід автоматично відхилити.

- Якщо пристрій розблокований і застосунок працює на передньому плані (foreground), запит, як правило, ініційовано користувачем. У цьому випадку система переходить до перевірки контенту.

3.5.2. Архітектура системи

Як проілюстровано на рис. 3.9, система складається з чотирьох основних модулів та функціонує як системний сервіс, інтегрований у ядро Android для забезпечення необхідних привілеїв перехоплення доступу.

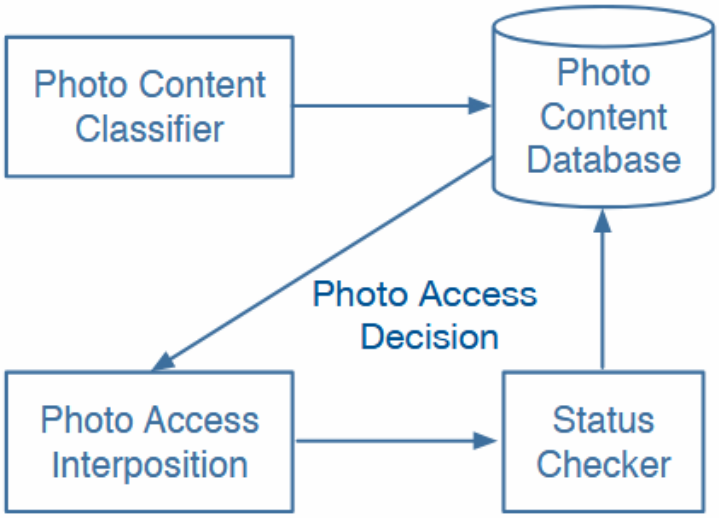

Рис. 3.9. Архітектура системи

Основні модулі:

1. Модуль перехоплення доступу до фотографій (Photo Access Interposition) - відповідає за переривання операції застосунку, коли він запитує доступ до фото.

2. Перевірка статусу (Status Checker) - використовує Android APIs (KeyguardManager та ActivityManager) для визначення стану пристрою (заблоковано/розблоковано) та стану застосунку (передній/фоновий план).

3. Класифікатор контенту фотографій (Photo Content Classifier) - попередньо навчена модель глибокого навчання, яка ідентифікує тип контенту (наприклад, "публічний", "ID-документ", "сімейний портрет").

4. База даних контенту фотографій (Photo Content Database) - імплементована з використанням SQLite, вона кешує результати класифікації для забезпечення роботи в реальному часі. Кожен запис має вигляд кортежу (photo id, content type).

3.5.3. Робочий процес (Workflow)

Робочий процес (рис. 3.10) поділяється на три ключові етапи:

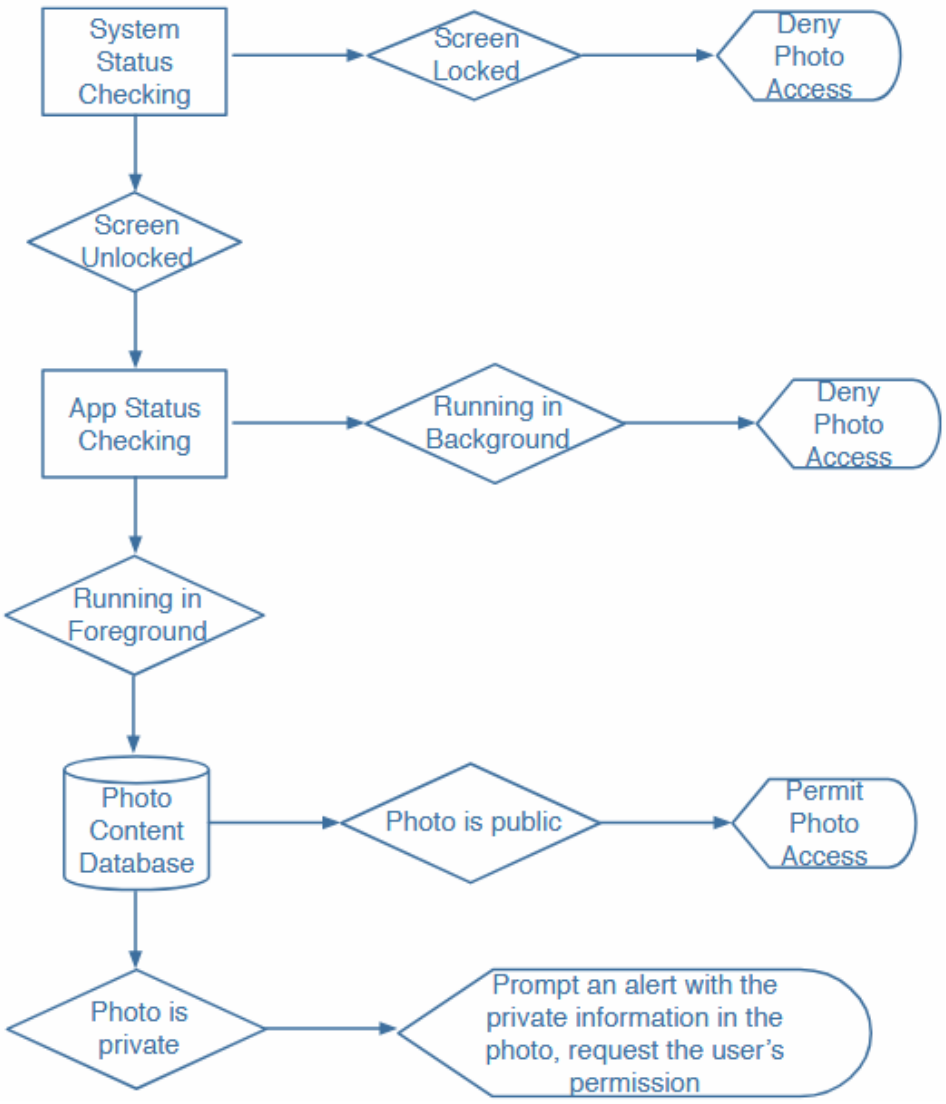

Рис. 3.10. Основні потоки в системі

1. Ініціалізація системи

При ініціалізації системи (запуску пристрою) навчений класифікатор обробляє всі збережені фотографії.

Результати класифікації ((photo id, content type)) кешуються у базі даних контенту. Це забезпечує можливість миттєвої відповіді на запити доступу в подальшому.

2. Перехоплення та контекстна перевірка

Коли застосунок запитує доступ до фото:

- Перехоплення: Модуль Photo Access Interposition перериває операцію.

- Перевірка Статусу Екрана:

Якщо пристрій заблоковано (Screen Locked), доступ автоматично відхиляється.

- Перевірка статусу застосунку (при розблокованому екрані):

Якщо застосунок працює у фоновому режимі (Running in Background), доступ автоматично відхиляється.

3. Прийняття рішення та підтвердження

Якщо застосунок працює на передньому плані (Running in Foreground), система переходить до перевірки контенту:

- Запит до БД: швидко отримується тип контенту фотографії з Photo Content Database.

- Класифікація:

  - Якщо фото класифіковано як "публічне", дозвіл на доступ автоматично надається.

  - Якщо фото містить "приватну інформацію", генерується сповіщення (Alert), яке інформує користувача про тип конфіденційного контенту та запитує явне підтвердження дозволу.

- Фіналізація: Доступ надається лише у випадку довіри та підтвердження користувачем; інакше — доступ до мультимедійних даних відхиляється.

Цей механізм мінімізує вплив на юзабіліті, оскільки користувацьке втручання необхідне лише у високоризикових, але контрольованих сценаріях (застосунок на передньому плані + конфіденційний контент).

3.5.4. Модуль класифікації конфіденційних мультимкдійних даних на основі DCNN

Цей модуль має на меті автоматичне виявлення специфічної конфіденційної інформації у вхідному зображенні, що має заздалегідь визначені розміри. Для вирішення цієї задачі використовується глибока згорткова нейронна мережа (DCNN), спеціально навчена для класифікації приватного контенту фотографій.

Виконаємо формалізацію завдання класифікації. Нехай P є множиною вхідних зображень. Для довільного зображення x є P, його категоріальна мітка y є {1,2,3,4,5} кодує наступні класи контенту: ’public’, ’photo id’, ’legal document’, ’family’.

Метою є знаходження оптимальної функції прийняття рішень $f(\theta^\top x)$, де H — простір гіпотез можливих функцій рішень, а $\theta=\{\theta 1,\theta 2,\ldots,\theta N\}$ — вагові коефіцієнти нейронної мережі.

Функція втрат $L(f(\theta^\top x),y)$ кількісно оцінює помилку між передбаченим та істинним класами. Нехай E(L) є очікуваною втратою по всій множині вхідних даних P. У даній роботі для оцінки втрат використовується **крос-ентропія**. Таким чином, завдання оптимізації зводиться до мінімізації очікуваної втрати крос-ентропії:

$$f = \arg\min_{f \in \mathbb{H}} \mathbb{E}(L)$$

Для кожного вхідного зображення x, відповідний результат класифікації f(x) дозволяє визначити точність (acc):

$$acc(x, y) = \begin{cases} +1 & y = f(x) \\ 0 & \text{otherwise} \end{cases}$$

Основна складність навчання DCNN для ідентифікації приватних фотографій полягає у недостатній кількості доступних приватних зображень, оскільки глибокі мережі вимагають великих обсягів навчальних даних для ефективної роботи.

Для подолання цієї проблеми застосовується методологія перенесення навчання (Transfer Learning):

1. Попереднє навчання (Pre-training): спочатку модель DCNN попередньо навчається на великому загальнодоступному наборі даних, зокрема ImageNet, який містить 1.2 мільйона зображень, розподілених за 1000 категоріями.
2. Тонке Налаштування (Fine-tuning): після попереднього навчання, параметри вихідного шару моделі (відповідального за класифікацію) доналаштовуються (fine-tuned) на невеликій кількості приватних фотографій, специфічних для нашого завдання. Параметри всіх інших шарів (які вилучають загальні ознаки, такі як краї та текстури) залишаються незмінними.

Цей підхід дозволяє ефективно використовувати потужність DCNN навіть за умов обмеженого доступу до конфіденційних навчальних даних.

3.6. Експериментальна оцінка та порівняння ефективності класифікаторів

Була проведена емпірична оцінка впливу модифікованого набору ознак на продуктивність чотирьох різних алгоритмів машинного навчання. Кожен алгоритм навчався та оцінювався на розширеному наборі даних 10 разів, при цьому фінальна точність оцінювання визначалася як середнє арифметичне

цих ітерацій. Усі процеси навчання та оцінювання виконувались за допомогою функціоналу бібліотеки Scikit-learn.

У таблиці 3.4 представлено порівняння точності алгоритмів, навчених на новому (модифікованому) наборі ознак (що включає удосконалені метрики розміру, позиції та погляду очей), з точністю тих же алгоритмів, навчених на оригінальному наборі ознак.

Таблиця 3.4.

Порівняння точності класифікації алгоритмів

Алгоритм Класифікації	**Точність (Оригінальні Ознаки)**	**Точність (Модифіковані Ознаки)**	**Зміна Точності (Δ)**
Багатошаровий Перцептрон (MLP)	91.4%	95.3%	+3.9%
Градієнтний Бустинг Дерев Рішень (GDBT)	93.4%	93.8%	+0.4%
Випадковий Ліс (Random Forest)	92.8%	93.2%	+0.4%
Машина Опорних Векторів (SVM)	91.7%	92.1%	+0.4%

Модель багатошарового перцептрона (MLP) продемонструвала найвищу ефективність, досягнувши середньої точності 95.3%. Це являє собою значне покращення на 3.9% порівняно з моделлю MLP, навченою на оригінальному наборі ознак.

Критично важливим висновком є те, що кожен із протестованих алгоритмів (GDBT, Random Forest, SVM та MLP) продемонстрував підвищення точності при використанні модифікованого набору ознак. Хоча зростання продуктивності для GDBT, Random Forest та SVM було помірним (≈0.4%), універсальне покращення свідчить про те, що новий набір ознак посилив інформативну цінність моделі для навчання з учителем.

Алгоритм GDBT, який був найефективнішим у попередньому дослідженні [7] (точність 93.3% на оригінальному меншому наборі даних), був перевершений як моделями GDBT (досягнувши 93.8%), так і, особливо, MLP (досягнувши 95.3%), навченими на новому наборі даних та модифікованому наборі ознак.

Таким чином, інтеграція вдосконалених метрик обличчя та нового показника погляду очей суттєво підвищила здатність класифікаційних моделей до диференціації цільових суб'єктів від випадкових перехожих, причому багатошаровий перцептрон виявився найбільш чутливим до цих покращень.

Висновки до розділу

Розроблено методологію інтегрованого статичного й динамічного аналізу мобільних додатків для виявлення потенційних витоків даних. Запропоновано класифікатор приватних зображень на основі глибоких нейронних мереж, а також удосконалену модель диференціації осіб «Ціль/Перехожий». Експериментальні результати підтвердили ефективність запропонованих методів та їх придатність до практичної реалізації у системах захисту мультимедійних даних.

ВИСНОВКИ

У магістерській роботі здійснено комплексне дослідження моделей та методів захисту мультимедійних даних у мобільних додатках, що дало змогу сформувати теоретичне підґрунтя та запропонувати практичні підходи до підвищення рівня конфіденційності персональної інформації користувачів.

На основі аналізу предметної області встановлено, що мультимедійні дані (фото, відео, аудіо) є найбільш чутливим типом інформації в сучасних мобільних додатках, оскільки містять як прямі, так і непрямі ідентифікаційні характеристики. Визначено, що традиційні моделі управління дозволами мобільних платформ (зокрема Android) демонструють низку обмежень: відсутність гнучкої диференціації рівнів доступу, складність для користувачів у прийнятті рішень, а також недостатній захист від прихованих витоків даних.

У процесі дослідження виявлено низку актуальних загроз для конфіденційності, серед яких:

- несанкціонований збір мультимедійних даних додатками;
- можливість витоку приватних фотографій у хмарні сервіси;
- відсутність контролю над обробкою даних випадкових осіб, що потрапляють у поле зйомки.

Для розв’язання зазначених проблем було розглянуто та апробовано методи захисту мультимедійних даних, серед яких: моніторинг поведінки мобільних додатків, класифікація приватних зображень на основі методів машинного навчання, а також застосування моделей комп’ютерного зору для анонімізації перехожих на фотографіях. Проведений аналіз показав, що найбільш перспективними є методи, засновані на глибоких згорткових нейронних мережах, які забезпечують високу точність виявлення та класифікації об’єктів.

У роботі розроблено методологію динамічного аналізу поведінки мобільних додатків, що поєднує статичний аналіз дозволів із трасуванням

системних викликів. Запропонований підхід дозволяє виявляти потенційні канали витоку даних на ранніх етапах використання програмного забезпечення. Також створено автоматизований класифікатор конфіденційних мультимедійних даних, який продемонстрував ефективність при ідентифікації фотографій приватного характеру.

Додатково було розроблено та вдосконалено модель диференціації осіб типу «Ціль/Перехожий» із використанням розширених біометричних ознак. Така модель підвищує рівень захисту конфіденційності випадкових осіб, які можуть бути зафіксовані на знімках без їхньої згоди.

Практичним результатом дослідження стала розробка концепції системи захисту мультимедійних даних у мобільних додатках, архітектура якої включає модуль динамічного аналізу, класифікатор зображень на базі DCNN та механізми анонімізації даних. Експериментальна оцінка підтвердила доцільність інтеграції запропонованих методів, оскільки вони забезпечують баланс між точністю класифікації, швидкодією та ресурсною ефективністю.

Таким чином, у роботі:

- здійснено системний аналіз сучасних підходів до захисту мультимедійних даних у мобільних додатках;

- ідентифіковано основні вразливості та загрози конфіденційності;

- розроблено методологію аналізу поведінки мобільних додатків для виявлення витоків;

- створено та апробовано класифікатор конфіденційних мультимедійних даних на основі глибокого навчання;

- удосконалено модель диференціації осіб для підвищення рівня приватності у публічному просторі;

- сформовано архітектурну концепцію інтегрованої системи захисту мультимедійних даних.

Отримані результати мають як наукове, так і практичне значення, оскільки вони не лише розширюють теоретичну базу в галузі захисту

інформації, але й можуть бути використані у створенні реальних мобільних додатків та сервісів, орієнтованих на безпечну обробку та зберігання мультимедійних даних користувачів.

ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Felt, A. P., Chin, E., Hannah, S., Balebako, R., Wethington, H., Rutenbar, R. A., & Bosh, L. F. "Android Permissions: Are They Effective?" USENIX Security Symposium, 2012.
2. Stevens, K., & Wulf, W. "The Problem with Mobile Permissions: Users' Awareness and the Path to Improvement." IEEE Security & Privacy, 2015.
3. Kelley, P. G., Consolvo, S., & Cranor, L. F. "A Contextual Permission Model for Mobile Devices." ACM Conference on Computer and Communications Security (CCS), 2011.
4. Ren, J., Liu, P., & Zhu, S. "PR-Droid: A Framework for Policy-Respecting Android Apps." International Symposium on Research in Attacks, Intrusions and Defenses (RAID), 2017.
5. Backes, M., Klee, B., & Stoepel, M. "Measuring the Effectiveness of Contextual Permission Systems in Android." ACM Conference on Data and Application Security and Privacy (CODASPY), 2018.
6. Shih, P. C., Han, K., & Lee, D. "Privacy Leaks and User Awareness: When and How Privacy Information is Compromised." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), 2019.
7. Almuhimedi, H., Alabdan, A., & Alabdan, A. "The AppOps Manager: A User Study of Permission Management in Android." Symposium on Usable Privacy and Security (SOUPS), 2015.
8. R. V. E. "XPrivacy: A Privacy Manager for Android." Github Repository/Software Documentation, 2014.
9. M. S. "DonkeyGuard: Android Runtime Permission Management." F-Droid Repository/Technical Report, 2016.
10. P. M. "Permission Manager: Control Your App Permissions." Third-Party Android Application Technical Documentation, 2015.

11. H. W. L. "Privacy Guard: Enhanced Permission Control on Custom ROMs." XDA Developers/Technical Blog Post,
12. Subrahmanyam, P., Murthy, C., & Srinivas, C. "Image Blur Assessment using Frequency Domain Analysis." International Journal of Computer Vision and Image Processing, 2017.
13. Pertuz, S., Puig, D., & Garcia, M. A. "Analysis of focus measure operators for shape-from-focus." Pattern Recognition, 2013.
14. Baltrusaitis, T., Robinson, P., & Morency, L. P. "OpenFace 2.0: Facial Behavior Analysis Toolkit." IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2018.
15. Z. G. R. S. L. T. S. A. G. B. "OpenFace Gaze Tracking Module: Technical Specifications and Implementation Details." Technical Documentation, University of Cambridge, 2017.
16. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research, 2011.
17. Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. "How transferable are features in deep neural networks?" Neural Information Processing Systems (NIPS), 2014.
18. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. "ImageNet: A Large-Scale Hierarchical Image Database." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.
19. Zerr, S., Schaer, P., & Siersdorfer, S. "Finding Private Content in Shared Collections." ACM International Conference on Information and Knowledge Management (CIKM), 2012.
20. Tan, H., & Liu, J. "Protecting Private Photos on Mobile Phones through Access Control and Face Recognition." International Conference on Mobile Computing and Networking, 2015.

21. Ra, M., Almuhimedi, H., Alabdan, A., & Alabdan, A. "P3: Privacy-Preserving Photo Sharing on Online Social Networks." IEEE Transactions on Mobile Computing, 2016.
22. He, T., Chen, C., & Wang, Y. "A Privacy-Preserving Approach for Photo Sharing in Mobile Cloud Computing." IEEE International Conference on Cloud Computing, 2014.
23. Jana, S., Hage, R., & Sarma, A. D. "Darkly: Protecting User Privacy from Continuously-Sensing Applications." USENIX Security Symposium, 2017.
24. Templeman, R., Hu, H., & Padhye, J. "PlaceAvoider: Preventing Visual Privacy Leakage in Video Streams." ACM Conference on Computer and Communications Security (CCS), 2018.
25. Squicciarini, A. C., Carminati, B., & Ben-David, A. "A Learning-Based Approach for Estimating Privacy Settings of Shared Photos." IEEE Transactions on Knowledge and Data Engineering, 2014.
26. Liu, Y., Lu, Y., & Liu, A. X. "Investigating User Privacy Preferences on Facebook Photo Sharing." Symposium on Usable Privacy and Security (SOUPS), 2016.
27. Dalal, N., & Triggs, B. "Histograms of Oriented Gradients for Human Detection." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.
28. D. L. "OpenCV: Open Source Computer Vision Library." Software Documentation, 2010.
29. S. G. P. "TensorFlow: A System for Large-Scale Machine Learning." OSDI, 2016.
30. V. K., Z. A., E. G., & S. B. "Inception-v3 Architecture." arXiv preprint arXiv:1512.00567, 2015.
31. S. S. S. E. "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning." AAAI Conference on Artificial Intelligence, 2017.

32. Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. "Learning Transferable Architectures for Scalable Image Recognition." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. (NASNet-A).
33. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision." arXiv preprint arXiv:1704.04861, 2017.
34. Cortes, C., & Vapnik, V. "Support-vector networks." Machine learning, 1995.
35. Freund, Y., & Schapire, R. E. "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting." European Conference on Computational Learning Theory, 1995.
36. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. "Learning representations by back-propagating errors." Nature, 1986.
37. Google. "Android Developers Documentation: KeyguardManager and ActivityManager." Technical API Documentation, 2020.