

Дадиверін В.В.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Фещенко І.О.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Потапова К.Р.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Тарасенко-Клятченко О.В.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Івасенко Д.В.

Національна академія внутрішніх справ

СИСТЕМА РОЗПІЗНАВАННЯ ОБЛИЧЧЯ З ВИКОРИСТАННЯМ ПОЄДНАННЯ МЕТОДІВ HAAR ТА SVM

У статті досліджено метод Haar та запропоновано варіант його поліпшення. На сьогоднішній день програми, що спираються на алгоритми розпізнавання обличчя, відіграють ключову роль у наших життєвих процесах, бо різного роду біометричні системи задовольняють потреби у аутентифікації та ідентифікації людей для подальшого використання у конструкціях, які гарантують різного роду безпекові питання: охорона місцевості та режимних об'єктів різного ґатунку секретності, гарантування безпеки фінансових операцій, забезпечення секретності інформації інтимного характеру тощо. Існує вже багато гарних прикладів комерційних продуктів, що стосуються приватних даних, які повинні бути максимально захищеними та в той же час достатньо комфортні у використанні, і на поміч у цій справі приходить технологія розпізнавання обличчя з інтеграцією у більшу систему алгоритмів обробки та аналізу. Місцевими прикладами таких успішних комерційних продуктів є фінансові операції за використанням додатку Приват 24 або його екосистеми, система ідентифікації у Дії, яка дає змогу верифікувати особистість максимально швидко та зручно. Також варто зазначити, що ця технологія стає невід'ємною частиною систем контролю правопорядку, яка може дуже допомогти слідчим органам притягнути обвинувачуваного до відповідальності, чи військових систем виявлення противника або локалізації району майбутнього ураження.

Чіткий алгоритм, висока швидкодія та точність є визначальними рисами методів, що обираються для використання у вищезазваних системах, бо в цьому питанні кількість допущених помилок у гео-метричні й прогресії погіршує можливі наслідки, саме тому в цій статті буде надано поліпшення існуючого алгоритму з вибудованою архітектурою.

Ключові слова: метод опорних векторів, Haar, система розпізнавання обличчя.

Постановка проблеми. Одним з головних видів верифікації та ідентифікації є системи розпізнавання обличчя. Зазвичай такого роду системи складаються з трьох основних етапів:

- отримання зображення і його початкова обробка;
- визначення характеристик та основних особливостей;
- класифікування.

Алгоритми машинного навчання відносять до останнього етапу і є одним з двох головних стовпів якості розпізнавання обличчя у різних формаціях та позиціях верифікації, а, отже, існує потреба у модифікації існуючих систем, використовуючи позитивні сторони такого роду навчання.

Аналіз останніх досліджень і публікацій. Системи розпізнавання обличчя зазвичай складаються з двох основних фаз: тренування або під-

готовка моделі та етап аналізу або тестування. В свою чергу ці дві фази поділяють на три етапи, що представлені трохи вище, перший етап фактично є першою фазою, а ось останні два етапи відносять до другої фази.

Далі будуть представлені роботи, які являються дуже корисними кроками у загальному розвитку такої технології як детектор обличчя або схожих за призначення систем.

У 2003 році вийшла робота під назвою «Face photo recognition using sketch» [3], розроблена вона була представниками Китайського Університету Гонконга Ксиаоу Тангом і Ксиагенг Вангом. У ній було представлено дороблений алгоритм «конвенціонального ейгенфесу», який в основі своїм мав трансформаційний алгоритм, що допомагав перетворити реальні фотографії у відповідний доробок для більш точного порівняння, використовуючи при обробці «ейганфейс особливості». Порівняння проводилося між трьома методами:

- Простий геометричний метод, що включав у себе 26 ідентифікаторів, які допомагають за ключовими особливостями такого об'єкту як обличчя визначити схожість фото із замальовкою;
- «Ейганфейс» метод
- Новий модифікований метод.

Їхня галерея для порівняння складалась з 100 фотографій та отриманий результат виправдав сподівання, було отримано приріст щонайменше на 20-30% від значень перших двох методів, таблицю результатів можна побачити нижче (рис. 1).

У 2010 році вийшла робота під назвою «Selecting Optimal Orientations of Gabor Wavelet Filters for Facial Image Analysis» [3], розроблена вона була представниками Китайського Університету Гонконга Ксиаоу Тангом і Ксиагенг Вангом. У цьому дослідженні вони представили алгоритм дій для

зменшення кількості фільтрів Габора, при тому покращуючи точність класифікації. В роботі вони описують надмірність фільтрів Габора при використанні однаково розподілених орієнтацій, тому вони створили наступний алгоритм дій:

1. Аналіз гістограми орієнтованого градієнта середнього зображення обличчя в наборі даних з виключенням менш значущих комбінацій орієнтації;

2. Проводять повторні проходи з вибіркою інших комбінацій для точності перехресної перевірки. Процес вибору досить простий, але дуже ефективний і відповідає розподілу градієнта обличчя.

3. Виступ і узагальнення вибраних фільтрів Габора оцінюється за допомогою розпізнавання обличчя та класифікацією за статевою ознакою на різних наборах даних.

Результати показують, що вибрані фільтри Габора відповідають вимогам

оптимальний і може краще відображати зображення обличчя з меншими витратами на обчислення, також таблицю з результатами точності визначення (рис. 2).

У 2018 році вийшла робота під назвою «Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM algorithms» [2], розроблена вона була представниками Вищої національної інженерної школи Тунісу Зіедом Баннуром Лахавом, Джеккою Ессаїданом та Хессеном Седіком. У цій статті вони запропонували підходи до розпізнавання обличчя ДВТ (Дискретне вейвлет-перетворення) + АСК (Аналіз основних компонентів) + ПВМ (Підтримуюча векторна машина), ДВТ + НАК (Незалежний аналіз компонентів) + ПВМ і ДВТ + ЛДА (Лінійний дискримінантний аналіз) + ПВМ. Ефективність цих підходів оцінювалась шляхом моделювання в базі

Рівень	1	2	3	4	5	6	7	8	9	10
Геометричний метод	30	37	45	48	53	59	62	66	67	70
Ейганфейс метод	31	43	48	55	61	63	65	65	67	67
Новий метод	71	78	81	84	88	90	94	94	95	96

Рис. 1. Сукупна оцінка відповідності для трьох методів

Експеримент	Кількість фільтрів	Кількість фільтрів	Час	Точність
Розпізнавання обличчя	40	10000	0.36 с.	98.5%
	20	5000	0.18 с.	100%
Гендерна класифікація	40	4000	0.4 с.	96.17%
	20	2000	0.2 с.	97.02%
	10	1000	0.1 с.	96.17%

Рис. 2. Порівняння продуктивності оптимальних орієнтацій Габора та повного сімейства Габорів

даних АТ&Т і порівняльного дослідження. Така ефективність пояснюється попередньою обробкою ДВТ, яка дозволяє зберігати найважливішу інформацію у вхідних зображеннях. Крім того, ця попередня обробка підвищує швидкість розпізнавання підходами порівняно з іншими підходами на основі АСК + ПВМ, НАК + ПВМ або ЛДА + ПВМ.

Чітко було продемонстровано приріст точності визначення за запропонованим підходом, при найскромніших приростах вони отримали значення 1-2%, а при найкращому варіанті – коло 20%, що є досить непоганим результатом.

Формулювання цілей статті.

Алгоритми машинного є другим за важливістю форпостом у системах розпізнавання обличчя і на них буде побудовано все дослідження для того, щоб знайти найліпший та найоптимальніший за нашою версією метод.

Визначатися якість алгоритму буде за такими характеристиками :

F-міра, дає достатньо глибоке розуміння точності алгоритму, беручи до уваги повноту тесту;

Коефіцієнт правильного визначення.

Виклад основного матеріалу.

Наар Cascade Face Detector

Наар Cascade Face – це алгоритм для виявлення об’єктів, в нашому випадку обличчя людей, на основі машинного навчання.

Машинне навчання полягає в тому[8], що каскадна функція проходить тренування на великій кількості позитивних та негативних зображень, де позитивні – це ті зображення, на яких присутній об’єкт для пошуку, а негативні – це ті, де цього об’єкту немає.

Це навчання відбувається за наступним алгоритмом :

– Дасмо тестові зображення $(x_1, y_1), \dots, (x_n, y_n)$ where $y_i = 0, 1$ для негативного або позитивного результату відповідно;

– Ініціалізуємо значення ваги $w_{1,i} = \frac{1}{2^m}, \frac{1}{2^l}$ для $y_i = 0, 1$ відповідно, де m та l – це кількість негативних та позитивних відповідно;

– Далі ми заходимо у цикл $= 1, \dots, T$:

1. Нормалізуємо вагу, $w_{1,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$, так що w_t – це розподіл ймовірностей;

2. Для кожної ознаки j , треба «потренувати» класифікатор h_j , який обмежується використання однієї ознаки. Похибка оцінюється з урахуванням $w_t, \epsilon_{1,i} = \sum_i w_i |h_j(x_i) - y_i|$;

3. Обираємо класифікатор h_j з відповідним найменшим значенням помилки ϵ_t ;

4. Перераховуємо значення ваги: $w_{t+1,i} = w_{t+1,i} \beta_t^{1-e_i}$,

де $e_i = 0$, якщо x_i класифіковано правильно та $e_i = 1$, якщо класифікована неправильно та в цьому випадку $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

– Остаточний вирішальний класифікатор буде таким:

$$h(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq 0.5 \sum_{t=1}^T \alpha_t \\ 0, & \text{якщо протилежне першому} \end{cases}, \text{ де } \alpha_t = \log \frac{1}{\beta_t}$$

Переносячись від алгоритму у текст, як було сказано раніше навчання проводиться на двох типах зображень і в нашому випадку позитивними є зображення обличчя, а негативними – зображення без граней, це нам дає змогу підготувати класифікатор. Вся основна робота проходить за допомоги так званого «фільтра», використовуючи який вилучаються особливості із поточного зображення, також варто сказати, що ця система є подібною до згорткового ядра.

Ці фільтри представлені на рисунку 3.

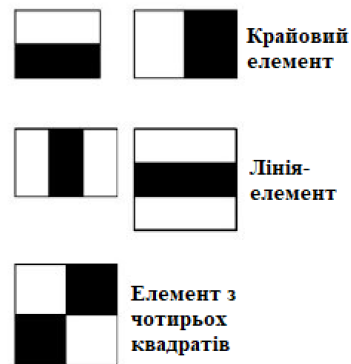


Рис. 3. Вигляд фільтрів Наар

Представлені вище фільтри накладаються на зображення перевіряючи один сектор (вікно) і далі переходять на інший. Далі для кожного сектору іде обрахунок інтенсивності кожного пікселя за його відповідністю чорному та білому кольорам. Наступним кроком отримується значення витягнутої ознаки, що обраховується як різниця двох сум отриманих на минулому етапі. В найкращому випадку значення ознаки повинно показувати її релевантність.

Далі ми стикаємося з великою кількістю повернутих обрахунків з секторів, які треба оптимізувати. І тут на допомогу приходить концепція Integral Image – це особлива структура даних та алгоритм підрахунку значень у прямокутній підмножині сітки, який спровокує зменшення кількості обчислень для отримання потрібного результату.

На наступному кроці оптимізації ми повертаємося до класифікаторів і завдяки технології Adaboost

створюється так званий «сильний» класифікатор з декількох «слабких», а об'єднання проводиться за спільними найкращими характеристиками.

Останнім кроком оптимізації виступає наша каскадність, яка існує для правильного групування ознак за різними стадіями класифікаторів і подальшого їх застосування по черзі у відповідних секторах. Це дає змогу відкинути відповідні сектори і застосування означень для них, а отже це створює прецедент більш економічного використання пам'яті.

Метод опорних векторів

SVM – алгоритм навчання з учителем. Головна мета SVM як класифікатора – знайти рівняння роздільної гіперплощини $w_1x_1 + w_2x_2 + \dots + w_nx_n + w_0 = 0$ у просторі R^n , яка розділила б два класи якимось оптимальним чином [5]. Загальний вид перетворення F об'єкта x у мітку класу Y: $F(x) = \text{sign}(w^T x - b)$. Пам'ятатимемо, що ми позначили $w = (w_1, w_2, \dots, w_n)$, $b = -w_0$. Після налаштування ваг алгоритму w і b (навчання), всі об'єкти, які потрапляють по одну сторону від побудованої гіперплощини, передбачатимуть як перший клас, а об'єкти, що потрапляють по інший бік – другий клас.

Усередині функції sign() стоїть лінійна комбінація ознак об'єкта з вагами алгоритму, саме тому SVM відноситься до лінійних алгоритмів. Розділяючу гіперплощину можна побудувати різними способами, але в SVM ваги w і b налаштовуються таким чином, щоб об'єкти класів лежали якнайдалі від роздільної гіперплощини. Іншими словами, алгоритм максимізує відступ (англ. margin) між гіперплощиною та об'єктами класів, які розташовані найближче до неї. Такі об'єкти називають опорними векторами (рис. 4). Звідси й назва алгоритму.

Архітектура системи та запропоноване вдосконалення

Для початку обираються негативні фотографії, що не містять жодних обличчя в цілому та позитивні з нормальними зображеннями, далі все проводиться за алгоритмом описаним вище і ми отримуємо так званий потужний класифікатор. Далі до справи доєднається метод опорних векторів, який робить свій вклад у алгоритмізацію машинного

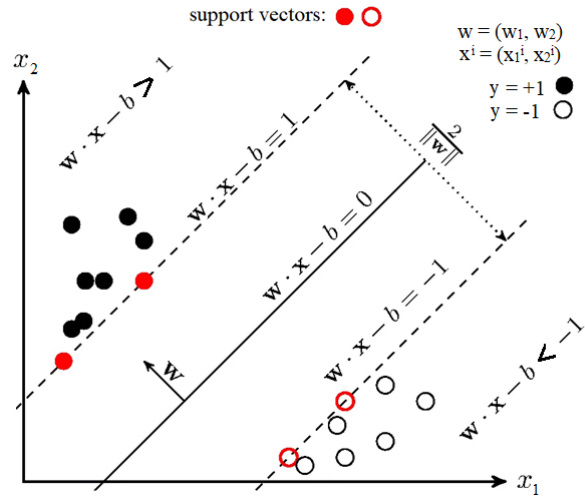


Рис. 4. Візуалізація методу опорних векторів

навчання для вирішення проблем класифікації, що часто виникають у таких випадках. В нашому випадку це буде використовуватись у вигляді додавання нових прикладів до даних навчання, класифікуючи одержані екземпляри зображень у різні категорії. Тому тут використовується так зване навчання без контролю, де буде відбуватися пошук кластеризації у групах, а потім відображатися у вигляді даних для визначених підгруп. Структура роботи системи можна побачити на рисунку 5 нижче.

Системи оцінювання та бази даних

Щоб оцінити систему розпізнавання обличчя було обрано дві бази даних:

- FEI Face Database;

База даних обличчя FEI – це бразильська база даних обличчя, яка містить набір зображень обличчя, зроблених у період з червня 2005 року по березень 2006 року в Лабораторії штучного інтелекту FEI в Сан-Бернарду-ду-Кампо, Сан-Паулу, Бразилія. Є 14 зображень для кожного з 200 осіб, загалом 2800 зображень. Усі зображення кольорові та зроблені на білому однорідному фоні у вертикальному фронтальному положенні з поворотом профілю приблизно до 180 градусів. Масштаб може відрізнятись приблизно на 10%, а вихідний розмір кожного зображення становить 640x480 пікселів.

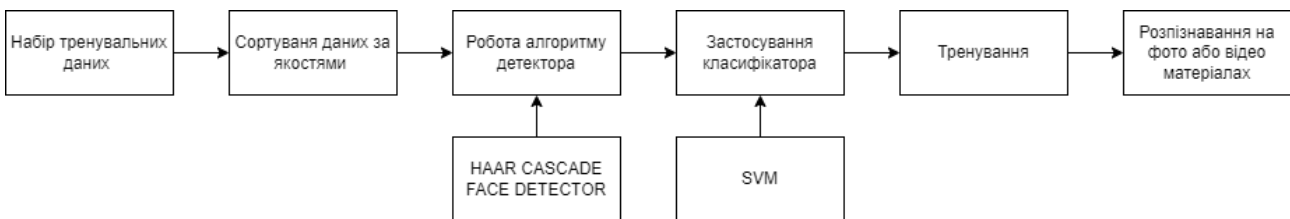


Рис. 5. Архітектура запропонованої системи

Усі обличчя в основному представлені студентами та співробітниками FEI віком від 19 до 40 років із виразною зовнішністю, зачіскою та прикрасами. Кількість суб'єктів чоловічої і жіночої статі абсолютно однакова і дорівнює 100.

- Senthill IRTT Face Database Version 1.1.

Ця база даних містить зображення обличчя студентів IRTT (Інститут дорожніх і транспортних технологій) як кольорові, так і зображення обличчя в градаціях сірого. Є 317 зображень обличчя для 13 студентів IRTT. Вони мають однаковий віковий фактор приблизно від 23 до 24 років. Зображення разом із фоном знімаються цифровою камерою Canon із роздільною здатністю 14,1 мегапікселя. Фактичний розмір обрізаних обличчя становить 550x780, і вони додатково змінюються до коефіцієнта зменшення 5. З 13, 12 чоловіків і одна жінка. Кожен суб'єкт має різні вирази обличчя, невеликий макіяж, шарф, пози та капелюх.

Системи оцінювання за двома параметрами продуктивності:

- Коефіцієнт вдалого розпізнання або ККД розпізнавача.

Визначається як відношення вдалих розпізнавань до загальної кількості спроб.

- F-міра.

Визначається вона через два основні параметри – повнота та влучність [1], де першим є наш коефіцієнт вдалого розпізнавання, а другим – відношення вірно визначених позитивних результатів до всіх позитивних з визначеними неправильно, формула виглядає наступним чином (рис. 6).

Загальним результатом успішності є сума цих двох параметрів.

Беремо наші методи машинного навчання та дві бази даних і отримуємо графіки з наступними результатами (рис. 7 та 8).

$$F_1 = \frac{2}{\text{повнота}^{-1} + \text{влучність}^{-1}} = 2 \cdot \frac{\text{влучність} \cdot \text{повнота}}{\text{влучність} + \text{повнота}} = \frac{2 \cdot \text{ІП}}{\text{ІП} + \frac{1}{2}(\text{ХП} + \text{ХН})}$$

Рис. 6. Формула F-міри

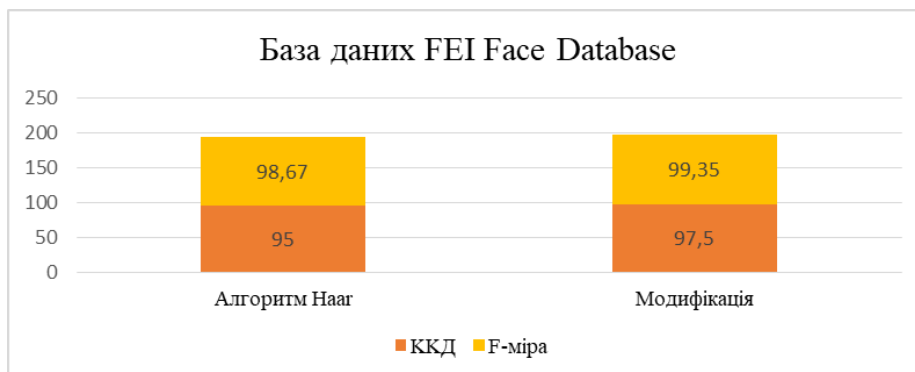


Рис. 7. Графік для першої бази даних

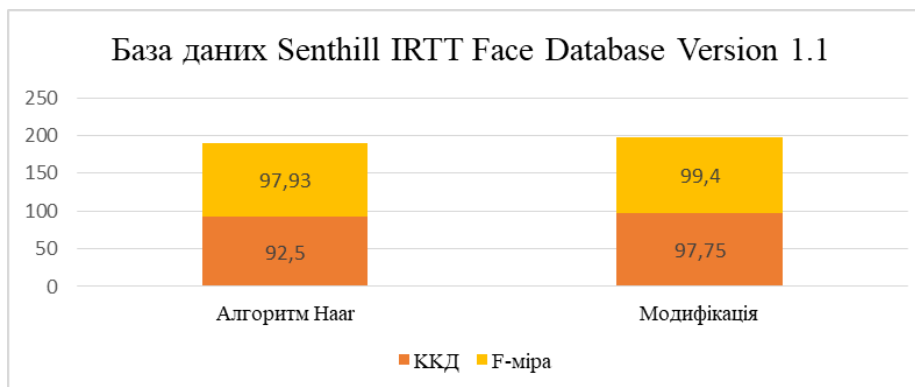


Рис. 8. Графік для другої бази даних

Висновки. У цій статті розкрито принципи роботи алгоритму Хаар та модифікацію у вигляді поліпшення класифікації та проблем, що можуть виникати з нею. Для модифікації було використано метод опорних координат, який показав дуже непоганий результат. Чітко видно прирости у якості ідентифікації на кожній базі даних з великою кількістю різноманітних фотозображень,

в першому випадку споглядається невелике збільшення на 3,17 відносних одиниць при досить гідному значенню розпізнавання до модифікації, а от у другому випадку ситуація є ще кращою і виражається вона у 6,72 відносних одиниць. Також варто зазначити створену архітектуру системи розпізнавання за її структуру та чітку послідовність дій.

Список літератури:

1. Гущин І. В., Сич Д. О., Аналіз впливу попередньої обробки тексту на результати текстової класифікації, 264–266, 2018.
2. Z. B. Lahaw, D. Essaidani, and H. Seddik. “Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms”, 2018, pages 1–5.
3. T. Zhang, Lu Bao-Liang “Selecting optimal orientations of Gabor wavelet filters for facial image analysis”, 2010, pages 218–227.
4. X. Tang, X. Wang, “Face photo recognition using sketch”, 2003, pages 712–715.
5. К. В. Воронцов, “Лекции по методу опорных векторов”, 16 с., 2007.
6. Freund Y., Schapire R. E., «Large margin classification using the perceptron algorithm». Machine Learning., 1999, pages 277–296.
7. Zhu R, Zeng D, Kosorok MR, «Reinforcement Learning Trees». Journal of the American Statistical Association, 2015, pages 1770–1784.
8. Дадиверін В. В., Фещенко І. О., Потапова К. Р., Про організацію та використання детектора обличчя. Modern research in world science, Львів, 4-6 вересня 2022, 226–232с.

Dadyverin V.V., Potapova K.R., Feshchenko I.O., Tarasenko-Kliatchenko O.V., Ivasenko D.V. FACE RECOGNITION SYSTEM USING COMBINATION OF HAAR TA SVM METHODS

The article examines the Haar method and offers an option to improve it. Today, programs based on face recognition algorithms play a key role in our life processes, because various types of biometric systems satisfy the needs of authentication and identification of people for further use in structures that guarantee various types of security issues: protection of the area and regime objects of various levels of secrecy, guaranteeing the security of financial transactions, ensuring the secrecy of information of an intimate nature, etc. There are already many good examples of commercial products related to private data, which should be maximally protected and at the same time comfortable enough to use, and facial recognition technology with integration into a larger system of processing and analysis algorithms comes to the rescue. Local examples of such successful commercial products are financial transactions using the Privat 24 application or its ecosystem, the identification system in Diya, which allows you to verify your identity as quickly and conveniently as possible. It is also worth noting that this technology is becoming an integral part of law enforcement systems, which can greatly help investigative agencies to bring the accused to justice, or military systems for detecting the enemy or locating the area of future damage.

A clear algorithm, high speed and accuracy are the defining features of the methods chosen for use in the above-mentioned systems, because in this matter the number of mistakes made in geometric and progression worsens the possible consequences, that is why this article will provide an improvement of the existing algorithm with a built-in architecture.

Key words: support vector method, Haar, face recognition system.