УДК 50.10.43 DOI: 10.31673/2412-9070.2022.023239

К. П. СТОРЧАК, доктор техн. наук, професор;
О. М. ТКАЛЕНКО, канд. техн. наук, доцент;
О. В. ПОЛОНЕВИЧ, канд. техн. наук, доцент;
А. М. ТУШИЧ, доктор філософії, доцент;
М. Л. УСИК, студент,
Державний університет телекомунікацій, Київ

ЗАСТОСУВАННЯ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ ОБЛИЧ ДЛЯ ЗАПОБІГАННЯ ІНСАЙДЕРСЬКИМ АТАКАМ

Розглянуто види біометричної автентифікації для запобігання інсайдерським атакам, здійснено їх порівняльний аналіз. Біометричні інновації є рішенням, що дає змогу надати достатній рівень безпеки, зручність роботи користувачів та оптимізацію бізнес-процесів. Біометричні технології полягають у вимірі та використанні унікальних фізичних рис чи поведінки людей для розрізнення їх один від одного. Біометричне розпізнавання формує міцний зв'язок між людиною та її особистістю як біометричною ознакою. Його не можна відокремити від людини, втратити чи дублювати. Отже, біометричне розпізнавання — це стійкіший до атак соціальної інженерії метод захисту порівняно з консервативними методами розпізнавання (паролями, токенами тощо). Запропоновано підхід до ідентифікації людських осіб та опис системи їх розпізнавання, що працює в режимі, близькому до реального часу, що дає змогу відстежувати голову суб'єкта, а потім розпізнає його, порівнюючи вихідні риси обличчя з раніше розпізнаними.

Ключові слова: біометрія; розпізнавання облич; автентифікація; інсайдерські атаки.

Вступ

Щорічно компанії витрачають значні обсяги коштів на захист своїх активів від зовнішніх загроз. Однак є серйозніша проблема, з якою стикаються компанії по всьому світу і від якої мало хто добре захищений, а саме інсайдерські атаки. Незважаючи на те, що лише небагато компаній розуміють і готуються до внутрішніх загроз, кількість інсайдерських атак збільшується з кожним роком. Понад 80% інцидентів у галузі інформаційної безпеки за останні чотири роки є результатом внутрішніх атак, і частота нападів постійно зростає. Середня вартість наслідків успішного інсайду майже в 50 разів вища, ніж втрати від зовнішньої атаки. Інсайдери можуть не тільки завдати шкоди інформаційним активам компанії, а й фізичних збитків, а також підірвати ділову репутацію компанії. Найбільшою проблемою щодо внутрішніх загроз є той факт, що більшість компаній до них не готові: їм не вистачає належного розуміння інсайдерів і, отже, вони не можуть їх вчасно ідентифікувати.

Існують різноманітні методи визначення діяльності інсайдера. Сьогодні доволі поширені системи Data Leak Prevention (DLP), які використовують для запобігання витоку конфіденційної інформації з інформаційної системи. Однак ці програмні комплекси не дають змоги виявляти потенційного зловмисника на ранній стадії, а також здійснювати моніторинг з'єднання користувачів, оскільки між користувачами та комп'ютерами з правами доступу їхнього рівня формується жорсткий зв'язок.

Основна частина

Керування правами доступу до інформаційних ресурсів підприємства зазвичай ґрунтується на величезній кількості паролів та інших облікових даних, що дають змогу ідентифікувати користувачів. Ситуація ускладнюється використанням на підприємствах комбінованих засобів ідентифікації: безконтактних (смарт-) карток, цифрових сертифікатів тощо. Тож, звичайному користувачеві, крім необхідності мати кілька різних облікових записів для доступу в різні операційні середовища, потрібно тримати при собі смарт-картку або інші ідентифікатори. Однак пароль та ім'я користувача можуть бути забуті або дискредитовані, цифровий сертифікат — украдений або зламаний і якимось чином модифікований, смарт-картка — передана іншим особам (добровільно або під примусом), уся інформація на ній може бути видалена або замінена фальшивими даними. Отже, з'являється можливість підробити особу користувача, отримати доступ до його облікового запису в інформаційній системі, вкрасти інформаційні чи комерційні ресурси та передати їх зацікавленим третім особам [1]. Останнім часом найчастіше використовуваними програмами для запобігання або виявлення атак є системи виявлення вторгнень.

З основних видів біометричних методів автентифікації найбільшого поширення набули статичні методи біометрії людини. До них належить ідентифікація за папілярним рисунком на пальцях, райдужною оболонкою та сітківкою ока, рисунком вен руки, геометрією рук, обличчям. До сімейства методів,

що використовують динамічні характеристики, можна віднести ідентифікацію за голосом, динамікою рукописного почерку, серцевим ритмом, ходою.

Нині дактилоскопія (розпізнавання відбитків пальців) є найпопулярнішим методом біометричної ідентифікації особистості. Такі алгоритми використовують характерні точки на відбитках пальців: закінчення візерунка, розгалуження лінії, поодинокі точки. Крім того, використовується інформація про морфологічну структуру відбитка пальця: відносне розташування замкнених ліній папілярного візерунка, дугових та спіральних ліній. Незважаючи на низьку вартість сканерів і досить просту процедуру сканування, високою є ймовірність помилкових спрацьовувань. Деякі види сканерів не розпізнають відбиток пальця вологої руки. Багато сканерів некоректно працюють із сухою шкірою і, як наслідок, не розпізнають людей віком понад 45 років. Крім цього, мінімальна дія хімічних реактивів на пальці співробітників призводить до збоїв у роботі систем безпеки сканерів. Також є недостатня захищеність від фальсифікації зображення відбитка, частково зумовлена потужним поширенням методу.

Система ідентифікації особистості за райдужною оболонкою ока є одним із найточніших серед біометричних методів. Цей метод логічно поділяється на два основні етапи: захоплення зображення, його первинне оброблення та передавання обчислювачу; порівняння зображення із зображеннями в базі даних та передавання команди про допуск виконавчому пристрою.

До недоліків ідентифікації за райдужною оболонкою можна віднести: більш високу ціну порівняно із системами, що ґрунтуються на розпізнаванні пальця або на розпізнаванні особи, а також низьку доступність готових вирішень [2].

Розпізнавання особи — процес зіставлення облич, що потрапили в об'єктив камери, з базою даних раніше записаних та ідентифікованих зображень-еталонів, через аналіз співвідношення відстаней між точками обличчя, що легко визначаються. Особливо важливі характерні частини обличчя, які практично не змінюються з часом: верхні контури очних ямок, очі, ділянки навколо вилиці, кінчик носа, куточки рота. Перевага методу полягає в можливості розпізнавання людини на відстані, не попереджаючи її про сканування.

В основі двовимірного (2D) розпізнавання облич лежать плоскі двовимірні зображення. Для розпізнавання використовують антропометричні параметри обличчя. Оскільки основні бази даних ідентифікованих осіб двовимірні, і більшість уже встановленого у світі обладнання теж 2D, то найбільшим попитом користуються двовимірні системи розпізнавання осіб.

Головна перевага технології 2D-розпізнавання полягає в готових базах даних еталонних осіб та готовій інфраструктурі. На відміну від більшості біометричних методів, не потрібне дороге обладнання. Достатньо відповідного обладнання, яке уможливить розпізнавання на значних відстанях від камери. Система працює з відносно простим двовимірним зображенням, що помітно спрощує алгоритми та знижує інтенсивність обчислень. Однак така технологія характеризується більш високими коефіцієнтами помилкового пропуску та помилкової відмови порівняно з 3D-розпізнаванням осіб [3].

Технологія тривимірного розпізнавання здійснюється за реконструйованими тривимірними образами, здобутими за допомогою лазерних сканерів з оцінюванням віддалі елементів поверхні об'єкта, що розпізнається, або за допомогою сканерів із підсвічуванням поверхні об'єктів і математичним обробленням вигинів смуг. Також використовуються сканери, що обробляють фотограмметричним методом синхронні стереопари зображень облич.

Така технологія має більш якісне оцінювання, але потребує використання спеціальних камер для сканування, котрі в кілька разів дорожчі за звичайні 2D-рішення. Однак є змога збільшити точність і швидкість роботи класифікатора облич, який працює з двовимірними зображеннями, що допоможе значно знизити кількість успішно проведених інсайдерських атак, спростити розслідування інцидентів та мінімізувати витрати на розгортання системи. Зробити це можна завдяки сучасним методам машинного навчання та математичним підходам. Отже, у статті пропонується доповнити стандартні методи захисту від інсайдерських атак, такі як паролі, токени, смарт-картки, біометричними методами автентифікації, а саме технологією двовимірного розпізнавання осіб у реальному часі.

Більшість робіт з автоматизованого розпізнавання осіб ігнорує питання стосовно параметрів, важливих для ідентифікації людини, вважаючи, що вимір статистик є релевантним результатом. Можна припустити, що теоретичний підхід до кодування та декодування зображень облич, заснований на теорії інформації, може дати уявлення про дані, що містяться у фотографіях, підкреслюючи важливі особливості. Такі риси обличчя можуть мати або не мати прямого відношення до інтуїтивного уявлення про фізіологічні параметри людини, такі як очі, ніс, губи та волосся. Необхідно дістати відповідні дані із зображення, закодувати їх якомога ефективніше та порівняти кодування одного обличчя з базою даних моделей, закодованих у такий саме спосіб. Підхід до вилучення інформації, що міститься у фотографії, полягає у фіксації змін у наборі даних (незалежно від будь-якого припущення про особливості людського обличчя) використання цих даних для кодування та порівняння окремих прикладів облич.

За останні кілька років дослідниками було створено безліч алгоритмів розпізнавання обличчя. Запропоновано інший підхід, як-от нейронні мережі, мережі радіально-базисної функції одиниці грані. Визначимо три алгоритми, які використовують вилучення ознак. Перші два алгоритми — Eigenface і Fisherface, і третій алгоритм — Elastic Bunch Graph Matching [4].

З математичного погляду потрібно знайти основні компоненти розподілу облич або власні вектори матриці набору фотографій. Ці власні вектори можна визначати як набір ознак, котрі в сукупності характеризують різницю між обличчями. Кожне зображення робить більш-менш вагомий внесок у кожний власний вектор у такий спосіб, що його можна відображати у вигляді так званого власного обличчя. Кожне зображення обличчя в тренувальній вибірці може бути подано як лінійна комбінація власних поверхонь. Кількість можливих власних поверхонь дорівнює кількості зображень облич у тренувальному наборі. Однак обличчя можуть бути апроксимовані тільки за допомогою «найкращих» власних поверхонь — тих, які мають найбільші власні значення і тому є причиною найбільшої дисперсії набору зображень. Основною причиною використання меншої кількості власних поверхонь є обчислювальна ефективність. Кращі власні M' поверхні охоплюють M'-вимірне з усіх можливих зображень. Оскільки синусоїди різної частоти і фази є базисною функцією декомпозиції Фур'є, власні поверхні є базисними векторами розкладання власної поверхні.

Було запропоновано Eigenface. Основна ідея власного обличчя полягає в тому, щоб дістати ознаки в математичному сенсі замість фізичних ознак обличчя за допомогою власних поверхонь [10].

Перший етап — етап навчання. На цьому етапі велика група окремих облич виступає як набір для навчання. Ці зображення мають добре відображати всі обличчя, з якими можна зіткнутися. Розмір, орієнтація та інтенсивність світла мають бути стандартизовані. Наприклад, усі зображення мають розмір 125×125 пікселів і всі є фронтальними зображеннями обличчя.

Кожне із зображень у навчальному наборі подано вектором розміром $N\times N$, де N — розмір зображення. За допомогою навчальних зображень набір власних векторів визначається за допомогою аналізу головних компонентів (PCA).

Основна ідея PCA полягає в тому, щоб скористатися перевагами надлишковості, наявної в навчальному наборі, для представлення набору в більш компактний спосіб. Використовуючи PCA, можна подати зображення за допомогою M власних векторів, де M — кількість використаних власних векторів. Нехай навчальний набір зображень обличчя буде $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M$. Тоді середнє обличчя набору визначатиметься за формулою

$$\psi = \frac{1}{M}\sum_{n-1}^{M}\Gamma_M.$$

Кожна грань відрізняється від середньої на вектор $\varphi_n = \Gamma_n - \psi$. Потім цей набір дуже великих векторів підлягає аналізу головних компонент, який шукає набір M ортонормованих векторів, μ_n, що найкраще описує розподіл даних. Вектор k, μ_k, вибирається так, що

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}\left(\mu_k^T\varphi_n\right)^2$$

є максимумом за умови

$$\mu_l^T\mu_k = \begin{cases} 1, & l = k, \\ 0, & otherwise. \end{cases}$$

Вектори μ_k і значення λ_k є відповідно власними векторами та власними значеннями коваріаційної матриці:

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T = AA^T,$$

де матриця $A = [\Phi_1\Phi_2\ldots\Phi_M]$. Однак матриця C має розмір $N^2\times N^2$, а значення N^2 власних векторів і власних значень є складним завданням для типових розмірів зображення. Щоб знайти ці власні вектори, потрібен придатний для обчислень метод [7].

Якщо кількість точок даних у просторі зображення менша за розмірність простору ($M < N^2$), буде лише $M - 1$, а не N^2 значущих власних векторів (решта власних векторів матимуть асоційовані власні значення нуль). Тож, у цьому разі ми можемо розв'язати N^2-вимірні власні вектори, спочатку визначивши власні вектори матриці $M\times M$, наприклад, розв'язавши матрицю 15×15, а потім взявши відповідні лінійні комбінації зображень обличчя Φ_n. Розглянемо власні вектори v_n, якщо A^TA такі, що:

$$A^TAv_n = \lambda_n v_n.$$

Попередньо помноживши обидві сторони на A, дістанемо:

$$AA^TAv_n = \lambda_n Av_n,$$

звідки випливає, що Av_n є власними векторами $C = AA^T$.

Після цього аналізу будуємо матрицю $M \times M$, $L = A^TA$, де $L_{mn} = \Phi_m^T\Phi_n$, і знаходимо M власних векторів v_n для L. Ці вектори визначають лінійні комбінації M зображень граней навчального набору для формування власних граней μ_n:

$$\mu_n = \sum_{k=1}^{M} v_{nk}\Phi_k = Av_n, \quad n = 1, ..., M.$$

За допомогою цього аналізу обчислення значно скорочуються, від порядку кількості пікселів у зображеннях N^2 до порядку кількості зображень у навчальному наборі M. На практиці навчальний набір зображень обличчя буде відносно невеликим ($M < N^2$), і обчислення стають цілком керованими. Пов'язані власні значення дають змогу ранжувати власні вектори відповідно до їх корисності для характеристики варіації між зображеннями.

Найвідомішим *DA* є лінійний дискримінантний аналіз (*LDA*). Коли *LDA* використовується для пошуку підпросторового подання набору зображень обличчя, результуючі базисні вектори, що визначають цей простір, відомі як Fisherfaces [9].

Запропоновано кілька різних підходів до аналізу інформації, отриманої з кількох джерел. Найпростішим методом є формування розширеного вектора даних (ознак), що містить інформацію з усіх джерел, і розглядання цього вектора як векторного виходу з одного джерела. Зазвичай у таких системах усі подібності та відстані вимірюються в евклідовому розумінні. Цей підхід є успішним лише тоді, коли всі джерела мають подібні статистичні характеристики та порівнянну надійність. Альтернативний підхід передбачає, що кожна проєкція вхідного шаблону на дискримінантний вектор *ui* створює вісь рішень із певним рівнем надійності та потужності розрізнення. Рівень значущості або надійності рішень, заснованих на *ui*, безпосередньо пов'язаний із поділом класів уздовж цієї осі, який дорівнює відповідному (нормалізованому) власному значенню в *LDA*:

$$\forall(\lambda_i, U_i) \in \left(\Lambda^{(m)} \times U^{(m)}\right): a_i = \frac{\lambda_i}{\sum_{i=1}^{m}\lambda_i}. \qquad (1)$$

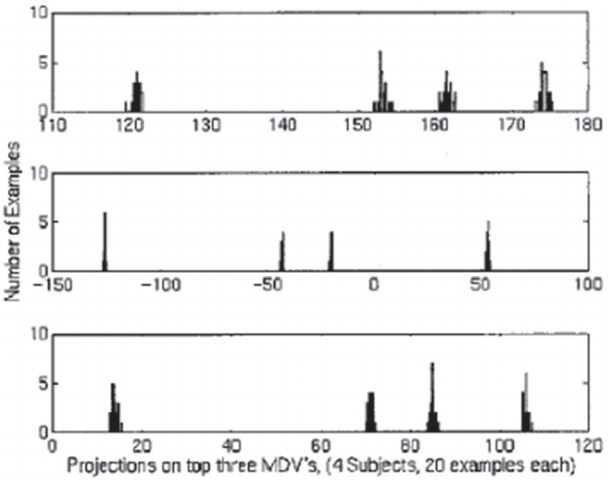


Рис. 1. Розподіл проєкційних коефіцієнтів уздовж трьох дискримінантних векторів із різними рівнями дискримінаційної сили для кількох поз чотирьох різних суб'єктів

Розподіл коефіцієнтів проєкції на три дискримінантні вектори показано на рис. 1. Для будь-якого тестового векторизованого зображення обличчя f проєктуємо зображення на кожен із верхніх дискримінантних векторів u. На основі відстаней між отриманими коефіцієнтами $f(u)$ і коефіцієнтами наявних шаблонів, що зберігаються в базі даних, оцінюється рівень подібності вхідного зображення до кожного відомого суб'єкта на рис. 2:

$$\forall_u \in U^{(m)}: \Phi(u) = \langle \Phi, u \rangle, \qquad (2)$$

$$\forall_u \in \bar{S}: d_u(\varphi, s) = |\varphi(u) - \psi_u^s|, \qquad (3)$$

$$\pi_u(\varphi, s) = 1 - \frac{d_u(\varphi, s)}{\sum_{s \in \bar{S}} d_u(\varphi, s)}, \qquad (4)$$

де $\pi_u(\varphi, s)$ характеризує відносний рівень подібності між вхідними даними φ та суб'єктом s відповідно до джерела, яке має надійність a_u. Визначивши нашу вісь рішень і надійності, ми можемо застосувати ймовірнісну або доказову схему аналізу даних із багатьох джерел до об'єднання в м'яке рішення, прийняті на основі окремих неточних джерел, щоб здобути більш точний і надійний кінцевий результат. Нормалізовані показники подібності ($\pi's$) вказують пропорції доказів, запропонованих різними джерелами. Їх можна інтерпретувати як так звану базову масу доказів або їх можна використовувати як приблизні оцінки апостеріорних імовірних здібностей за кожним вимірюванням. Починаючи з цього етапу, для комбінування базових м'яких рішень можна застосувати апробабілістичний або доказовий підхід [8].

У такий саме спосіб працюючи з відстанями як мірами відмінності, можна поєднати основне м'яке рішення та додати надійність кожного джерела, щоб визначити розумну міру відстані в просторі ознак. Хоча найпоширенішою мірою, яка використовується в літературі, є Евклідова відстань, як більш

розумну міру запропоновано середньозважену абсолютну квадратичну відстань, вагові коефіцієнти якої базуються на ступенях розрізнення. Інакше кажучи,

$$\delta_u(\varphi,s)=\frac{d_u(\varphi,s)}{\sum_{s\in\overline{S}}d_u(\varphi,s)}, \tag{5}$$

$$D(\varphi,s)=\sum_{u\in U}(m)[\delta_u(\varphi,s)\times a_u]. \tag{6}$$

Отже, для даного вхідного параметра найкращим збігом є його міра вірогідності:

$$s^0=\arg\min_{s\in\overline{S}}\{D(\varphi,s)\}, \tag{7}$$

$$\mathrm{conf}\left(\varphi,s^0\right)=1-\frac{D\left(\varphi,s^0\right)}{D\left(\varphi,s'\right)}, \tag{8}$$

де s' — другий найкращий кандидат; conf — міра впевненості. У цій структурі введення додаткової інформації або попередніх знань і очікувань із контексту стає дуже простим і логічним.

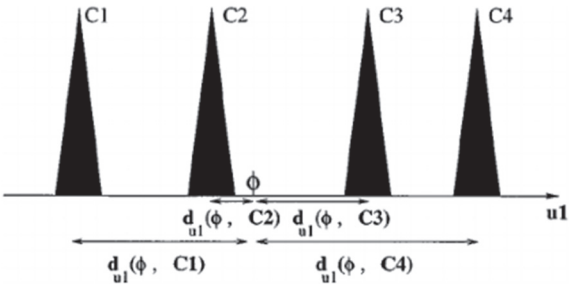


Рис. 2. Необроблені відстані між кожним тестовим прикладом і відомими кластерами вздовж кожної дискримінантної осі зумовлюють м'яке рішення вздовж цієї осі

Task	No. of Examples	No. of Features	Recognition Rate (%) (Training Set)	Recognition Rate (%) (Test Set)
Face recognition	2000	4	100	99.2
Gender classification	400	1	100	95

Рис. 3. Підсумкові показники розпізнавання

Все, що потрібно зробити, це розглядати кожен із них як додаткове джерело інформації, що відповідає осі рішення з певною надійністю, і додати його в процес ухвалення рішення [5]. Підсумкові показники розпізнавання надано на рис. 3.

Завдяки принципу алгоритму *LDA* відомо, що алгоритм мінімізує міжкласову відстань і максимізує міжкласову відстань після зіставлення даних зразка з іншим простором ознак. Алгоритм *LDA* може використовуватися для зменшення розмірності. Принцип алгоритму дуже схожий на алгоритм *PCA*, тому алгоритм *LDA* може бути застосовний у сфері розпізнавання осіб. Алгоритм розпізнавання облич із використанням алгоритму *PCA* називається методом ознаки обличчя, а алгоритм розпізнавання облич із застосуванням алгоритму *LDA* — методом обличчя Фішера. Оскільки алгоритм *LDA* дуже нагадує алгоритм *PCA*, ми просто порівнюємо їх. Подібності між алгоритмами *LDA* та *PCA*: у зменшенні розмірності обидва використовують розкладання матричного елемента; обидва припускають, що дані відповідають розподілу Гаусса.

Різниця між *LDA* та *PCA*: *LDA* — це контрольований метод зменшення розмірності, тоді як *PCA* не контролюється; якщо дані є k-вимірними, то *LDA* може бути зменшено лише до $(k-1)$ вимірювань і *PCA* не підпадає під це обмеження; з математичного погляду *LDA* вибирає напрямок проєкції з найкращими характеристиками класифікації, тоді як *PCA* вибирає напрямок, в якому точка проєкції зразка має найбільшу дисперсію.

Ці вектори об'єктів, отримані за допомогою алгоритму *LDA*, є FisherFace. Подальший процес розпізнавання осіб такий самий, як і в попередньому. Потрібно лише змінити модель об'єктів на модель FisherFace. Код, який необхідно змінити, складається з одного рядка (рис. 4).

```
model = cv2.face.FisherFaceRecognizer_create()
#model = cv2.face.EigenFaceRecognizer_create()
```

Рис. 4. Зміна моделі об'єктів на модель FisherFace

Експериментально використана суміш двох баз даних. Розпочато з бази даних, наданої Olivetti Research Ltd, яка містить 10 різних зображень кожного із 40-ка різних суб'єктів. Усі зображення зроблено на однорідному фоні, а деякі ще й у різний час. База даних містить фронтальні зображення вертикальних облич із незначними змінами освітленості, виразу обличчя (відкриті чи закриті очі, усмішка або без усмішки), деталі обличчя (окуляри чи без окулярів) і деякі рухи збоку. Потім для збільшення розміру бази даних додано кілька сегментованих вручну зображень обличчя з бази даних FERRET. Додано дзеркальне зображення та шумові версії кожного прикладу обличчя, щоб розширити набір даних і підвищити стійкість продуктивності розпізнавання до викривлень зображення. Загальна кількість зображень, використаних під час навчання та тестування, приблизно відповідно 1500 і 500. Кожне обличчя подано 50×60 пікселями 8-бітовим зображенням із сірим рівнем, яке для експериментів було зменшено до 25×30. Базу даних розділено на два непересічні навчальні та тестові набори.

Скориставшись цією зведеною базою даних, ми здійснили кілька тестів на гендерну класифікацію та розпізнавання облич.

Перший тест стосувався класифікації за статтею з використанням підмножини бази даних, яка містила кілька фронтальних проєкцій 20-ти чоловіків і 20-ти жінок різних рас. *LDA* застосовано до даних і найбільш дискримінантний шаблон було вилучено. Власний шаблон і розподіл коефіцієнтів проєкції для всіх зображень у наборі наведено на рис. 5. Як показано на рис. 5, за допомогою лише однієї функції можна досягти дуже хорошого розділення.

Класифікаційні тести на несумісному наборі тестів також дали 95% точності. Крім того, застосування цього дискримінантного шаблону до набору нових облич індивідів за межами навчального набору зменшило точність до 92%. Можна також застосувати *LDA* для вейвлет-перетворень зображень облич і витягти найбільш дискримінантні вектори кожного, трансформувати компонент і об'єднати результати багатомасштабної класифікації за допомогою запропонованого методу інтеграції м'якого рішення.

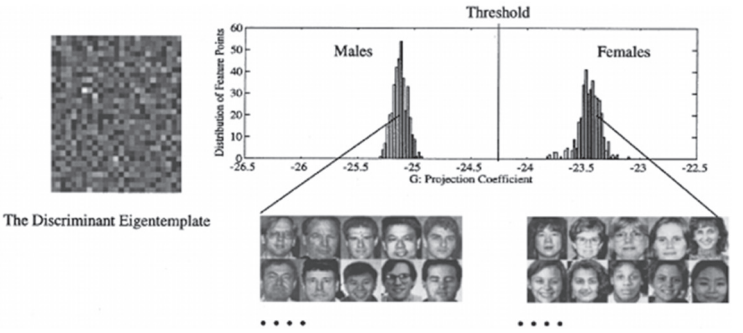


Рис. 5. Розподіл балів ознак для чоловічих і жіночих прикладів у базі даних

Потім *LDA* було застосовано до бази даних із 1500 облич, де 60 класів відповідають 60-м особам. Дискримінаційну силу 40 найкращих власних векторів, вибраних відповідно до *PCA* та *LDA*, зображено на рис. 6. Як показано на рис. 6, інформація про класифікацію основних компонентів не зменшується монотонно з їх енергією. Іншими словами, є багато випадків, коли низькоенергетичний компонент має вищу дискримінаційну здатність, ніж високоенергетичний компонент. На рис. 6 також показано, що кілька верхніх дискримінантних векторів із *LDA* містять майже всю інформацію про класифікацію, вбудовану у вихідний простір зображення. Поділ кластерів для десяти поз чотирьох різних індивідів, отриманих за допомогою двох найбільш дискримінаційних власних векторів або власних зображень, унаочнює рис. 7. Як показано на рис. 7, відмінності між класами (індивідами) підкреслюються, тоді як варіації одного обличчя в різних позах припиняються [6].

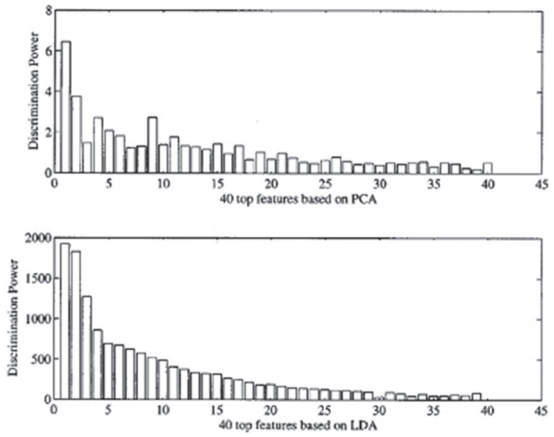


Рис. 6. Порівняння DP 40 найкращих вибраних власних векторів на основі *PCA* та *LDA*

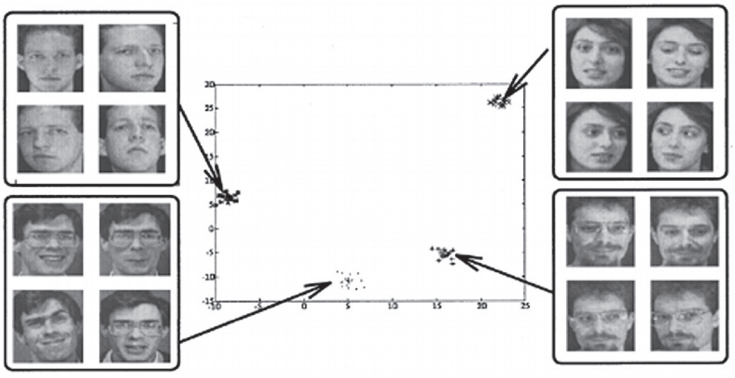
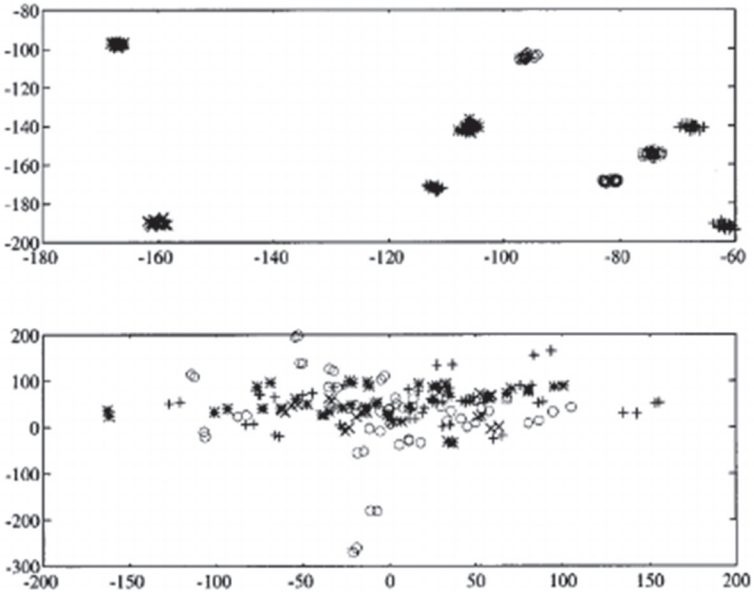


Рис. 7. Розділення кластерів у вибраному двовимірному просторі ознак (чотири кластери відповідають варіаціям облич чотирьох різних суб'єктів у базі даних)

Розділення досягається, незважаючи на всі варіації зображення, що є результатом різних поз кожного суб'єкта. Розподіл кластерів для 200 зображень 10-ти суб'єктів у найкращому просторі двовимірних дискримінантних ознак і в найкращому двовимірному просторі на основі *PCA* зображено на рис. 8.

Для кожного прикладу тестової грані спочатку його спроєктували на вибрані власні вектори та знайшли відстань від відповідної точки в чотиривимірному просторі ознак до всіх попередньо збережених екземплярів. Усі відстані було виміряно відповідно до формули (6), було вибрано найкращий збіг. Для наведеної бази даних було досягнуто відмінної (тобто 99,2%) точності на тестовому наборі. Щоб оцінити узагальнення вилучення ознак за межами оригінальних навчальних і тестових наборів, перевірили результати класифікації на фотографіях нових осіб, жодного з яких не було в навчальному наборі. Через уведені обмеження щодо доступності була змога скористатися лише десятьма новими предметами з двома зображеннями на суб'єкт: одне, збережене в базі даних як шаблон, і два для тестування. Як і очікувалося, застосування шаблонів проєкцій до цих абсолютно нових облич призвело до зниження

точності класифікації до 90%. Це зниження було очікуваним з огляду на той факт, що не було дуже великого навчального набору. Виокремлення дискримінантних рис обличчя з великого навчального набору з різноманітними прикладами має покращити узагальнення та продуктивність системи для розпізнавання суб'єктів поза навчальним набором. Результати невеликих варіацій пози та шуму показують, що запропонована схема є хорошим альтернативним підходом до розпізнавання обличчя. Забезпечує висококонкурентоспроможні результати із значно меншою складністю з використанням малорозмірних розмірів функцій.



а

б

Рис. 8. Розділення кластерів у найкращому двовимірному просторі ознак: *а* — на основі LDA; *б* — на основі PCA

Висновки

Розглянутий підхід до розпізнавання облич дає можливість збільшити швидкість роботи класифікатора, допомагає системі адаптуватися з часом, що зменшує кількість помилкових спрацьовувань. Аналіз відео в реальному часі дозволяє зменшити ймовірність обману системи за допомогою фотографії іншої людини завдяки перетворенню аналізованої ділянки з фотографії особи на ділянку, що змінюється в часі.

Пропонується основувати розпізнавання облич на невеликому наборі образів, які найкраще описують набір відомих облич, не вимагаючи, щоб вони відповідали інтуїтивним уявленням про частини обличчя та особливості людини. Даний підхід швидкий, відносно простий і застосовний до роботи в дещо обмеженому середовищі.

Список використаної літератури

1. ***Основи*** *теорії процесів в інформаційних системах: підручник (у 2-х кн.). Кн. 1. Аналіз детермінованих процесів / М. Б. Гумен, В. М. Співак, С. К. Мещанінов [та ін.]. 2-е вид., зі змінами і доповн. Київ: Кафедра, 2017. 281 с.*

2. ***Іващенко П. В.*** *Основи теорії інформації: навч. посіб. Одеса: ОНАЗ ім. О. С. Попова, 2015. 53 с.*

3. ***Комп'ютерні*** *мережі: навч. посіб. для техн. спеціальностей вищих навч. закладів / А. Г. Микитишин, М. М. Митник, П. Д. Стухляк, В. В. Пасічник. Львів: Магнолія 2006, 2017. 256 с.*

4. ***Технологія*** *розпізнавання осіб // greenvision [Електронний ресурс]. URL: https://greenvision.ua/ua/blog/articles/Tekhnologiya_raspoznavaniya_lits (дата звернення: 26.10.2022).*

5. ***Чеславський Л.*** *Як працює технологія розпізнавання облич? // spilno [Електронний ресурс]. URL: https://spilno.org/article/tekhnolohiya-rozpiznavannya-oblych-yak-tse-pratsyuye (дата звернення: 26.10.2022).*

6. ***Face*** *recognition using local binary patterns (LBP) / Sh. Azam, N. Hossain, A. Rahim, T. Wahid // globaljournals [Електронний ресурс]. URL: https://globaljournals.org/GJCST_Volume13/1-Face-Recognition-using-Local.pdf (дата звернення: 27.10.2022).*

7. ***Bhadauria S. S., Jadon R. S., Jaiswal S.*** *Comparison between face recognition algorithm Eigenfaces, Fisherfaces and elastic bunch graph matching // rroij [Електронний ресурс]. URL:*

https://www.rroij.com/open-access/comparison-between-face-recognition-algorithmeigenfaces-fisherfaces-and-elastic-bunch-graph-matching-187-193.pdf (дата звернення: 27.10.2022).

8. ***Chellappa R., Etemad K.*** *Discriminant analysis for recognition of human faces // researchgate [Електронний ресурс]. URL:*

https://www.researchgate.net/publication/225123273_Discriminant_analysis_for_recognition_of_human_faces (дата звернення: 27.10.2022).

9. ***Martinez A.*** *Fisherfaces // scholarpedia [Електронний ресурс]. URL:*

http://www.scholarpedia.org/article/Fisherfaces (дата звернення: 27.10.2022).

10. ***Turk M., Zhang Sh.*** *Eigenfaces // scholarpedia [Електронний ресурс]. URL:*

http://www.scholarpedia.org/article/Eigenfaces (дата звернення: 27.10.2022).

K. Storchak, O. Tkalenko, O. Polonevych, A. Tuschych, M. Usyk

APPLICATION OF FACIAL RECOGNITION TECHNOLOGY TO PREVENT INSIDER ATTACKS

The article discusses the types of biometric authentication to prevent insider attacks, their comparative analysis. There are various methods of determining insider activity. Today, Data Leak Prevention (DLP) systems, which are used to prevent the leakage of confidential information from the information system, are widespread. However, these software complexes do not allow detecting a potential attacker at an early stage, as well as monitoring the connection of users, since a tight connection is formed between users and computers with access rights of their level. Biometric innovations are a solution that allows providing a sufficient level of security, user convenience and optimization of business processes. Among the main types of biometric authentication methods, static methods of human biometrics have become the most widespread. These include identification based on the papillary pattern on the fingers, the iris and retina of the eye, the pattern of hand veins, the geometry of the hands, and the face. The family of methods using dynamic characteristics include: voice identification, handwriting dynamics, heart rate, and gait. Biometric technologies consist in measuring and using unique physical or behavioral traits of people to distinguish them from each other. Biometric recognition forms a strong connection between a person and his personality as a biometric feature. It cannot be separated from a person, lost or duplicated. So, biometric recognition is a protection method more resistant to social engineering attacks compared to conservative recognition methods (passwords, tokens, etc.).Also the approach to identification of human faces and description of the system of their recognition working in a mode close to real time which allows to trace a subject's head, and then distinguishes it, comparing initial features of a face with earlier recognized is offered.

Keywords: biometrics; face recognition; authentication; insider attacks.