КІБЕРБЕЗПЕКА: освіта, наука, техніка

CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE

№ 4 (12), 2021

ISSN 2663 - 4023

DOI 10.28925/2663-4023.2021.12.8595

УДК 004.056.5: 655.25

Возний Ярослав Васильович
аспірант кафедри інформаційних технологій видавничої справи
Національний університет "Львівська політехніка", Львів, Україна
ORCID ID: 0000-0002-5481-9973
voznyyy@outlook.com

Назаркевич Марія Андріївна
доктор технічних наук, професор, професор кафедри інформаційних технологій видавничої справи
Національний університет "Львівська політехніка", Львів, Україна
ORCID ID: 0000-0002-6528-9867
mariia.a.nazarkevych@lpnu.ua

Грицик Володимир Володимирович
доктор технічних наук, професор, професор кафедри автоматизованих систем управління
Національний університет "Львівська політехніка", Львів, Україна
ORCID ID: 0000-0002-9696-5805
volodymyr.v.hrytsyk@lpnu.ua

Лотошинська Наталія Дмитрівна
кандидат технічних наук, доцент, доцент кафедри інформаційних технологій видавничої справи
Національний університет "Львівська політехніка", Львів, Україна
ORCID ID: 0000-0002-6618-0070
nataliia.d.lotoshynska@lpnu.ua

Гавриш Богдана Михайлівна
кандидат технічних наук, доцент, доцент кафедри інформаційних технологій видавничої справи
Національний університет "Львівська політехніка", Львів, Україна
ORCID ID: 0000-0003-3213-9747
bohdana.m.havrysh@lpnu.ua

ПРОЕКТУВАННЯ СИСТЕМИ АВТЕНТИФІКАЦІЇ БІОМЕТРИЧНОГО ЗАХИСТУ НА ОСНОВІ МЕТОДУ К-СЕРЕДНІХ

Анотація. Розглянуто метод біометричної ідентифікації, призначений для забезпечення захисту конфіденційної інформації. Запропоновано метод класифікації біометричних відбитків за допомогою машинного навчання. Подано один із варіантів розв'язку задачі ідентифікації біометричних зображень на основі алгоритму к-середніх. Було створено позначені зразки даних для процесів навчання та тестування. Для встановлення особистості використовувались біометричні дані відбитків пальців. Нове сканування відбитків пальців, яке належить певній особі, порівнюється з даними, що зберігаються для цієї особи. Якщо вимірювання збігаються, твердження про те, що особа пройшла ідентифікацію, відповідає дійсності. Експериментальні результати вказують, що метод k-середніх є перспективним підходом до класифікації відбитків пальців. Розвиток біометрії призводить до створення систем безпеки з кращим ступенем розпізнавання і з меншою кількістю помилок, ніж система безпеки на традиційних носіях інформації. Машинне навчання проводили з використанням ряду зразків із відомої біометричної бази даних, а перевірку / тестування проводили із зразками з тієї самої бази даних, які не були включені до набору навчальних даних. Для встановлення особистості використовувались біометричні дані відбитків пальців на основі вільдоступної бази NIST Special Database 302, та показано результати навчання. Нове сканування відбитків пальців, яке належить певній особі, порівнюється з даними, що зберігаються для цієї особи. Якщо вимірювання збігаються, твердження про те, що особа пройшла ідентифікацію, відповідає дійсності. Система машинного навчання побудована на модульній основі, шляхом формування комбінацій окремих модулів бібліотека scikit-learn у середовищі python.

Ключові слова: відбитки пальців, біометричні зображення, машинне навчання

ВСТУП

Машинне навчання – великий підрозділ штучного інтелекту, що вивчає методи побудови алгоритмів, здатних навчатися. Машинне навчання має широкий спектр використання у різних технологіях. Нові завдання, що виникають практично щодня, призводять до появи нових напрямків машинного навчання. Багато успіхів у біометричному розпізнаванні відбитків пальців було зроблено завдяки машинному навчанні. У даній роботі представлені окремі підходи до машинного навчання, які поділяються на навчання без вчителя та навчання з учителем. Навчання з учителем - це навчання під керівництвом. Якщо машина тренується виконувати завдання з керівником, вона отримує розмічені дані. Таким чином, комп'ютеру вдається швидше видавати результати. Тут комп'ютер представлений бажаними входами, встановленими керівником, та бажаними виходами [1, 20-22].

Розрізняють: а) навчання з учителем, де комп'ютер отримує неповний навчальний комплек, у якого відсутні численні цільові результати; б) активний навчальний комп'ютер може отримувати навчальні дані лише для обмеженого набору випадків, і він повинен оптимізувати свій вибір об'єктів; в) навчання в динаміці – навчальні дані подаються як зворотний зв'язок щодо програми в динамічному середовищі.

При навчанні без учителя алгоритм навчання не позначений тегами, тому він повинен знайти структуру у своєму введенні. Якщо алгоритми тренуються без учителя, їм доводиться самостійно аналізувати інформацію й шукати закономірності. Такий підхід може тривати довше, однак розробникам не потрібно готувати базу даних заздалегідь. Навчання без учителя використовується в представленні зображень, генерації відео, моделюванні мов, творчому письмі, синтезі мовлення.

Навчання з учителем використовується в розпізнаванні зображень, аналізі настроїв, розпізнаванні мови [2, 14, 23]. Підготовка моделей пальців або навіть сітківки (за допомогою спеціальних голограм) технічно можлива. Оскільки ідентифікація на основі руху очей використовує інформацію, яка виробляється здебільшого мозком (її неможливо наслідувати), підробити цей тип інформації набагато складніше [3, 15].

Біометричне узгодження - це "нечітке порівняння". Це пов'язано з тим, що біометричні характеристики, виявлені вдруге, ніколи не бувають точно такими ж, як у перший раз. Ця характеристика біометричної відповідності призвела до використання таких методів машинного навчання, як нейронні мережі, нечітка логіка, еволюційні обчислення тощо в біометричних алгоритмах та захисті даних [4, 16-17]. Машинне навчання має ключові властивості шумостійкості і може ефективно вирішувати складні проблеми розпізнавання образів. Крім того, машинне навчання є впорядкованою адаптацією і, як правило, має паралельну обчислювальну архітектуру, адаптивно моделюючи складні біометричні характеристики, не роблячи багатьох припущень, використовуючи точну математичну модель [5]. Структура та модель визначені в [6].

ТЕОРЕТИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ

СПОСІБ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧА

Типова біометрична система автентифікації складається з двох етапів (рис. 1). Під час автентифікації користувач (скажімо, Аліса) сканує свої біометричні дані, включаючи

відбитки пальців, з яких дані зчитуються та створюється шаблон, який зберігається або в центральній базі даних, або на мобільному пристрої.

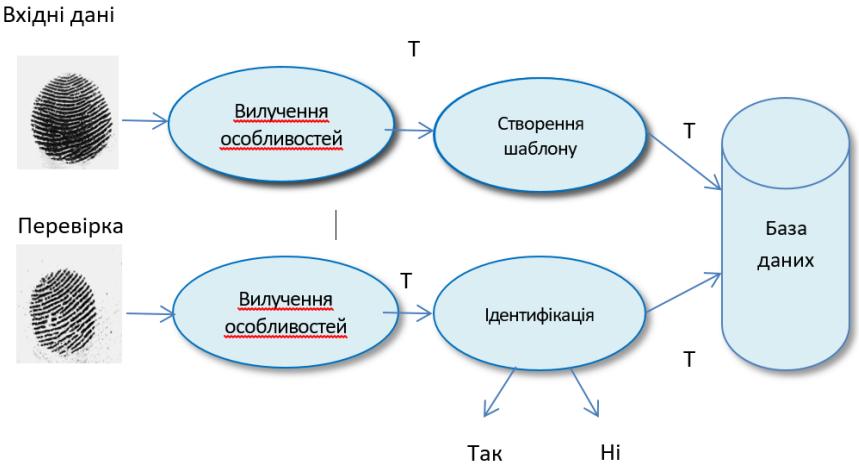

Рис. 1. Біометрична система аутентифікації

На етапі автентифікації користувач, який стверджує, що Аліса знову сканує свої біометричні дані, і ця сама функція використовується для визначення характеристик [7, 19].

Потім результат порівнюється із збереженим шаблоном. Якщо дані мають достатньо високу ступінь подібності, алгоритм відображає інформацію про те, що користувач справжній або користувач не справжній. Оскільки, важко замінити або скасувати біометричні дані, важливо забезпечити безпеку біометричних даних користувачів [8, 18], які вони використовують в цих системах автентифікації.

Найважливіший ключ до процесу автентифікації - це унікальність заходів безпеки.

Одержання відбитків пальців складаються з етапів навчання та локалізації. Проводиться навчальний етап для створення бази даних [9], що складається з набору опорних точок (RP) з відомими координатами та RSS з доступних точок доступу. На етапі локалізації знаходиться найближча відповідність між вимірюваним RSS та тим, що зберігається в базі даних. Тому ми використовуємо неконтрольоване навчання для розпізнавання відбитків пальців.

Наприклад, відбиток пальця, показаний на Рис. 2.

Рис. 2. Відбиток пальця з бази даних NIST Special Database 302

На відбитку можна виділити спеціальні орієнтири для отримання рис. 3.

Рис. 3. Відбиток пальця з виділенням спеціальних опорних точок. Зображення становить 800x750 пікселів і вагою 585,9 кБ.

НАВЧАННЯ БЕЗ УЧИТЕЛЯ

Навчання без учителя належить до процесу побудови моделі машинного навчання, яка не вимагає позначених навчальних даних. Експеримент навчання без учителя у вирішенні проблеми розпізнавання образів можна сформулювати як проблему кластерного аналізу. Вибірка об'єктів розділена на підмножини, які називаються кластерами, так що кожен кластер складається з подібних об'єктів, а об'єкти різних кластерів суттєво відрізняються. Інформація про джерело представлена у вигляді матриці відстані.

Алгоритми навчання без учителя намагаються побудувати моделі, які можуть знаходити підгрупи в даному наборі даних за допомогою різних метрик подібності.

Розглянемо, як формулюється навчальне завдання, якщо воно виконується без учителя. Коли ми маємо набір даних, який не пов’язаний з якимись мітками, ми вважаємо, що ці дані генеруються під впливом прихованих змінних, які контролюють їх розподіл. У цьому випадку процес навчання може слідувати певній ієрархічній схемі, використовуючи окремі точки даних на початковому етапі [10].

КЛАСИФІКАЦІЯ ДАНИХ ІЗ ВИКОРИСТАННЯМ МЕТОДУ K-СЕРЕДНІХ

Кластеризація використовується для аналізу даних та виділення кластерів серед них. Використовуються різні методи подібності для пошуку скупчень, таких як евклідова відстань, для виділення підгруп даних. Використовуючи міру подібності, можна оцінити зв’язок кластера. Отже, кластеризація - це процес організації даних у підгрупи, елементи яких подібні між собою згідно з деякими критеріями. Наше завдання - виявити приховані властивості точок даних, які визначають їх приналежність до однієї і тієї ж підгрупи. Не існує універсальних метричних характеристик подібності, які працювали б у всіх випадках. Все визначається специфікою завдання. Наприклад, нам може бути цікаво знайти репрезентативну точку даних для кожної підгрупи або невідомого. Залежно від ситуації ми вибираємо ту чи іншу метрику, яка, на наш погляд, найбільш повно показує специфіку проблеми.

Метод k-середніх значень також є добре відомим алгоритмом кластеризації. Його використання передбачає, що кількість кластерів відома заздалегідь. Потім дані сегментуються на k-підгрупи, застосовуючи різні атрибути даних. Аналіз починається із запису кількості кластерів та класифікації даних на основі цього. У цьому алгоритмі вибір початкового місця розташування центроїдів відіграє дуже важливу роль, оскільки

він безпосередньо впливає на кінцеві результати. Одна з стратегій - розмістити центроїди якомога далі один від одного. Основний метод k-середніх відповідає випадковому розташуванню центроїдів, тоді як у вдосконаленій нами версії машинного навчання (метод k-середніх++) ці точки вибираються алгоритмічно із вхідного списку точок даних. На початку процесу робиться спроба розмістити центри скупчень на великій відстані один від одного, щоб забезпечити швидку конвергенцію. Потім відбувається ітерація набору навчальних даних та вдосконалення початкової кластеризації шляхом присвоєння кожної точки найближчому центру кластера.

Як правило, системи машинного навчання будуються на модульній основі.

Конкретна кінцева мета досягається шляхом формування відповідних комбінацій окремих модулів. Бібліотека scikit-learn містить функції, що дозволяють поєднувати різні модулі в єдині конвеєри. Конвеєр може бути сформований з модулів, які виконують різноманітні функції, такі як вибір функцій, попередня обробка даних, побудова випадкових лісів, кластеризація тощо.

Згенеруємо позначені зразки даних для навчальних та тестових процесів. Scikit-learn включає вбудовану функцію для цього. Наступний рядок коду створює 22 точки даних, кожна з яких є 4D-вектором (рис. 4). Кожна точка даних містить шість інформативних функцій і не містить жодних надлишкових. Використовуємо концепцію найближчих сусідів, суть якої полягає у пошуку тих точок даного набору, які розташовані на найближчій відстані від даної. Цей підхід часто використовується для створення систем, які класифікують точку даних на основі її близькості до різних класів. Розглянемо приклад пошуку найближчих сусідів даної точки даних. Обираємо зразок двовимірних точок даних:

[480, 333], [511, 354], [504, 367], [360, 422], [344, 432], [600, 432], [583, 391], [698, 462], [556 , 308],[551, 284], [603, 273], [267, 472], [710, 223], [668, 180], [512, 183], [517, 219], [482, 213], [473, 236], [377, 362], [371, 281], [310, 450], [321 , 483]

Надалі визначимо кількість найближчих сусідів, які знаходяться поруч наших особливих точок.

У процесі повторення описаних кроків центри скупчень поступово переходять у свої стабільні положення. Після виконання певної кількості ітерацій центри скупчень перестануть зміщуватися. Це вкаже на те, що ми досягли стабільного розташування центрів кластерів. Отримані k-центроїди представляють остаточну модель k-середніх значень, яка буде використана для висновку. Щоб побачити, як працює метод кластеризації k-середніх, застосуємо його до двовимірних даних. Ми використовуємо дані, виділені як спеціальні опорні точки. У цьому файлі кожен рядок містить два числа, розділені комою.

Другий знімок екрану показує межі k-середніх (рис. 4).

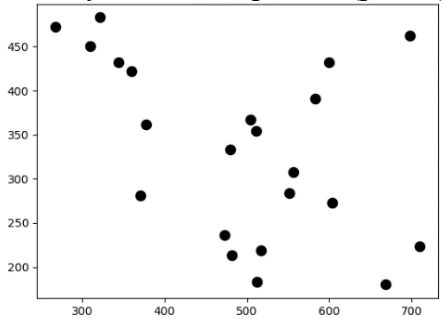

Рис. 4 Вхідні дані спеціальних опорних точок біометричного зображення

ISSN 2663 - 4023

Завершення описаного переліку всіх точок набору даних означає закінчення першої ітерації. На цьому етапі точки групуються на основі початкових положень центрів скупчень [11]. Далі нам потрібно повторно розрахувати положення центроїдів, починаючи з нових кластерів, отриманих в кінці першої ітерації. Отримавши новий набір k-центрів, ми повторюємо весь процес (рис.5). Суцільні чорні кола позначають центри скупчень.

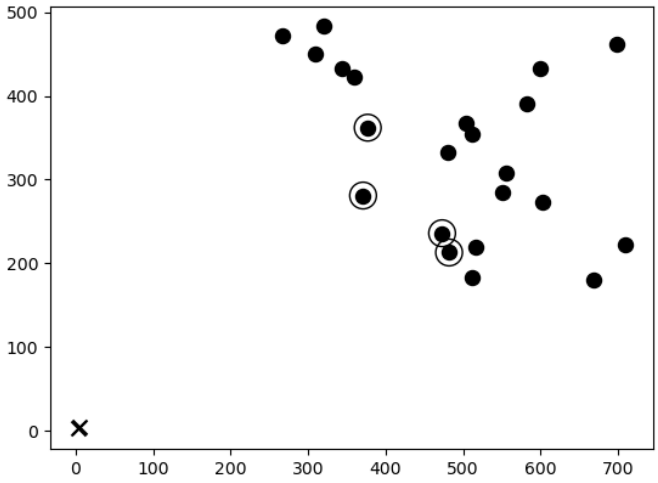

Рис. 5. Візуалізація центрів скупчень на біометричному зображенні

Кластеризація, як правило, спрямована на групування даних контрольних точок у біометричні зображення, якими обмінюються.

Алгоритм спрямований на мінімізацію дисперсії всередині кластера та максимізацію внутрішньої дисперсії кластера [12, 24].

Методика передбачає визначення кількості кластерів і випадкове присвоєння кластеру центроїдів кожному кластеру з цілих наборів даних. Цей крок - це ініціалізація кластерних центроїдів. Відстань між кожною точкою у цілочисельних наборах даних та кожним кластером визначається k-середніми. Потім центроїд обчислюється за допомогою метрики відстані (наприклад, евклідова відстань) [13, 25-26]. Надалі для кожної точки даних визначається мінімальна відстань, і ця точка призначається найближчому кластеру. Цей крок називається цільовим кластером і повторюється доти, поки всі точки даних не будуть призначені одному з кластерів. Середнє значення для кожного кластера обчислюється на основі накопичених балів у кожному кластері та кількості балів у цьому кластері. Потім ці інструменти призначаються як новий кластер центроїдів, і процес пошуку відстані між кожною точкою та новим центроїдом повторюється, де оцінки перепризначаються новим найближчим кластерам. Процес повторюється протягом фіксованої кількості разів або до тих пір, поки точки в кожному кластері не перестануть рухатися до різних кластерів.

Ці експериментальні результати показують в середньому 40 кластерів з 800 порожніх кластерів після першої ітерації традиційного алгоритму k-середнього значення із випадковими вихідними точками. Подальший аналіз показує, що більшість початкових моментів, створених їх 40 кластерами, походять від кластерних викидів. Випадки скупчень належать до відрізків послідовності, що знаходяться далеко від центрів.

ВИСНОВКИ

Було створено позначені зразки даних для процесів навчання та тестування. Для встановлення особистості використовувались біометричні дані відбитків пальців. Нове сканування відбитків пальців, яке належить певній особі, порівнюється з даними, що зберігаються для цієї особи. Якщо вимірювання збігаються, твердження про те, що особа пройшла ідентифікацію, відповідає дійсності.

Оскільки аутентифікація відбувається досить швидко, можливе шахрайство з персональними даними. Зловмисник може обійти систему біометричної автентифікації і продовжувати безперешкодно працювати. Викрадені біометричні дані в системі є складною проблемою. На відміну від паролів або смарт-карт, які можна змінити або перевидати, за відсутності серйозного медичного втручання відбиток пальця змінити не можна.

Після того, як зловмисник успішно створив ці характеристики, кінцевий користувач повинен бути повністю відключений від системи, збільшуючи ймовірність величезних ризиків безпеки та / або витрат на повторне впровадження. Статичні фізичні характеристики можна дублювати цифровим способом. Наприклад, обличчя можна копіювати за допомогою фотографії, звуку голосу за допомогою звукозапису та відбитків пальців за допомогою різних методів підробки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Hrytsyk, V., Grondzal, A., & Bilenkyj, A. (2015, September). Augmented reality for people with disabilities. In 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT) (pp. 188-191). IEEE.
2. Zhang, Q., Zhou, D., & Zeng, X. (2017). Machine learning-empowered biometric methods for biomedicine applications. AIMS Medical Science, 4(3), 274-290.
3. Kasprowski, P., & Ober, J. (2004, May). Eye movements in biometrics. In International Workshop on Biometric Authentication (pp. 248-258). Springer, Berlin, Heidelberg.
4. Medykovskyy, M., Lipinski, P., Troyan, O., & Nazarkevych, M. (2015, September). Methods of protection document formed from latent element located by fractals. In 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT) (pp. 70-72). IEEE.
5. Syazana-Itqan, K., Syafeeza, A. R., Saad, N. M., Hamid, N. A., & Saad, W. H. B. M. (2016). A review of finger-vein biometrics identification approaches. Indian J. Sci. Technol, *9*(32), 1-9.
6. Tsmots, I., Skorokhoda, O., & Rabyk, V. (2016, September). Structure and software model of a parallel-vertical multi-input adder for FPGA implementation. In 2016 XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT) (pp. 158-160). IEEE.
7. Kovalskyi, B., Dubnevych, M., Holubnyk, T., Pysanchyn, N., & Havrysh, B. (2019). Development of a technology for eliminating color rendering imperfections in digital photographic images. Восточно-Европейский журнал передовых технологий, (1 (2)), 40-47.
8. Nazarkevych, M., Oliiarnyk, R., Nazarkevych, H., Kramarenko, O., & Onyshschenko, I. (2016, August). The method of encryption based on Ateb-functions. In 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP) (pp. 129-133). IEEE.
9. Nazarkevych, M., Logoyda, M., Troyan, O., Vozniy, Y., & Shpak, Z. (2019, September). The Ateb-Gabor Filter for Fingerprinting. In International Conference on Computer Science and Information Technology (pp. 247-255). Springer, Cham.
10. Nazarkevych, M., Lotoshynska, N., Brytkovskyi, V., Dmytruk, S., Dordiak, V., & Pikh, I. (2019, September). Biometric Identification System with Ateb-Gabor Filtering. In 2019 XIth International Scientific and Practical Conference on Electronics and Information Technologies (ELIT) (pp. 15-18). IEEE.
11. Sheketa, V., Zorin, V., Chupakhina, S, Kyrsta, N., Pasyeka, M., Pasieka, N. (2020). "Empirical Method of Evaluating the Numerical Values of Metrics in the Process of Medical Software Quality Determination,&quot. У *International Conference on Decision Aid Sciences and Application (DASA)*, (с. 22- 26). https://doi.org/10.1109/DASA51403.2020.9317218.

12. Varetskyy, Y., Rusyn, B., Molga, A., & Ignatovych, A. (2010). A new method of fingerprint key protection of grid credential. In Image Processing and Communications Challenges 2 (pp. 99-103). Springer, Berlin, Heidelberg.
13. Ahmed, S., Lee, Y., Hyun, S. H., & Koo, I. (2018). Covert cyber assault detection in smart grid networks utilizing feature selection and euclidean distance-based machine learning. *Applied Sciences*, *8*(5), 772.
14. Yaacob, R., Pritam, H. M. H., Hassan, N. F. N.,. Ooi, C. D., Ibrahim, H., Othman, P. J. (2017). Computer assisted segmentation of palmprint images for biometric research. У *7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)* (с. 273–277). https://doi.org/10.1109/ICCSCE.2017.8284418.
15. Mishra, M., Bhattacharya, A., Singh, A., Dutta, M. K. (2018). A Lossless Model for Generation of Unique Digital Code for Identification of Biometric Images. У *4th International Conference on Computational Intelligence & Communication Technology (CICT)*, (с. 1-5). https://doi.org/10.1109/CIACT.2018.8480297.
16. Jaswal, G., Nath, R., Nigam, A. (2017). Deformable multi-scale scheme for biometric personal identification. У *IEEE International Conference on Image Processing (ICIP)*, (с. 3555-3559). https://doi.org/10.1109/ICIP.2017.8296944.
17. Do, H., Truong, V., George, K., Shirke, B. (2019). EEG-Based Biometrics Utilizing Image Recognition for Patient Identification. У *IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, (с. 0591-0595). https://doi.org/10.1109/UEMCON47517.2019.8992962.
18. Bychkov, O., Merkulova, K., Zhabska, Y. (2019). Software Application for Biometrical Person's Identification by Portrait Photograph Based on Wavelet Transform. У *IEEE International Conference on Advanced Trends in Information Theory (ATIT)*, (с. 253-256). https://doi.org/10.1109/ATIT49449.2019.9030462.
19. Setiawan, H., Yuniarno, E. M. (2017). Biometric Recognition Based on Palm Vein Image Using Learning Vector Quantization. У *5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)*, (с. 95-99). https://doi.org/10.1109/ICICI-BME.2017.8537770.
20. Abdufattokhov, S., Muhiddinov, B. (2019). Stochastic Approach for System Identification using Machine Learning. *Dynamics of Systems, Mechanisms and Machines (Dynamics)*, (с. 1-4). https://doi.org/10.1109/Dynamics47113.2019.8944452.
21. Yang, B., Liu, D. (2019). Research on Network Traffic Identification based on Machine Learning and Deep Packet Inspection. *У IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, (с. 1887-1891). https://doi.org/10.1109/ITNEC.2019.8729153.
22. Hammad I., El-Sankary, K. (2020).Using Machine Learning for Person Identification through Physical Activities. У *IEEE International Symposium on Circuits and Systems (ISCAS)*, (с. 1-5). https://doi.org/10.1109/ISCAS45731.2020.9181231.
23. Jadav, S. (2018). Voice-Based Gender Identification Using Machine Learning. У *4th International Conference on Computing Communication and Automation (ICCCA)*, (с. 1-4). https://doi.org/10.1109/CCAA.2018.8777582.
24. Deng, Y., Wu, F., Du, L., Zhou, R., Cao, L. (2019). EEG-Based Identification of Latent Emotional Disorder Using the Machine Learning Approach. У *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, (с. 2642-2648). https://doi.org/10.1109/ITNEC.2019.8729424.
25. Traboulsi, A., Barbeau, M. (2019). Recognition of Drone Formation Intentions Using Supervised Machine Learning,". У *International Conference on Computational Science and Computational Intelligence (CSCI)*, (с. 408-411). https://doi.org/10.1109/CSCI49370.2019.00079.
26. Mokni, R., Elleuch, M., Kherallah, M. (2016). Biometric Palmprint identification via efficient texture features fusion. У *International Joint Conference on Neural Networks (IJCNN)*, (с. 4857-4864). https://doi.org/10.1109/IJCNN.2016.7727838.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 4 (12), 2021
ISSN 2663 - 4023

Voznyi Yaroslav
graduate student of the Department of Information Technologies of Publishing
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-5481-9973
voznyyy@outlook.com

Nazarkevych Mariia
Doctor of Technical Sciences, Professor, Professor of the Department of Information Technologies of Publishing
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-6528-9867
mariia.a.nazarkevych@lpnu.ua

Hrytsyk Volodymyr
Doctor of Technical Sciences, Professor, Professor of the Department of Automated Control Systems
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-9696-5805
volodymyr.v.hrytsyk@lpnu.ua

Lotoshynska Nataliia
Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Information Technologies of Publishing
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-6618-0070
nataliia.d.lotoshynska@lpnu.ua

Havrysh Bohdana
Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Information Technologies of Publishing
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0003-3213-9747
bohdana.m.havrysh@lpnu.ua

DESIGN OF BIOMETRIC PROTECTION AUTHENTIFICATION SYSTEM BASED ON K-AVERAGE METHOD

Abstract. The method of biometric identification, designed to ensure the protection of confidential information, is considered. The method of classification of biometric prints by means of machine learning is offered. One of the variants of the solution of the problem of identification of biometric images on the basis of the k-means algorithm is given. Marked data samples were created for learning and testing processes. Biometric fingerprint data were used to establish identity. A new fingerprint scan that belongs to a particular person is compared to the data stored for that person. If the measurements match, the statement that the person has been identified is true. Experimental results indicate that the k-means method is a promising approach to the classification of fingerprints. The development of biometrics leads to the creation of security systems with a better degree of recognition and with fewer errors than the security system on traditional media. Machine learning was performed using a number of samples from a known biometric database, and verification / testing was performed with samples from the same database that were not included in the training data set. Biometric fingerprint data based on the freely available NIST Special Database 302 were used to establish identity, and the learning outcomes were shown. A new fingerprint scan that belongs to a particular person is compared to the data stored for that person. If the measurements match, the statement that the person has been identified is true. The machine learning system is built on a modular basis, by forming combinations of individual modules scikit-learn library in a python environment.

Key words: fingerprints, biometric images, machine learning

REFERENCES

1. Hrytsyk, V., Grondzal, A., & Bilenkyj, A. (2015, September). Augmented reality for people with disabilities. In 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT) (pp. 188-191). IEEE.
2. Zhang, Q., Zhou, D., & Zeng, X. (2017). Machine learning-empowered biometric methods for biomedicine applications. AIMS Medical Science, 4(3), 274-290.
3. Kasprowski, P., & Ober, J. (2004, May). Eye movements in biometrics. In International Workshop on Biometric Authentication (pp. 248-258). Springer, Berlin, Heidelberg.
4. Medykovskyy, M., Lipinski, P., Troyan, O., & Nazarkevych, M. (2015, September). Methods of protection document formed from latent element located by fractals. In 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT) (pp. 70-72). IEEE.
5. Syazana-Itqan, K., Syafeeza, A. R., Saad, N. M., Hamid, N. A., & Saad, W. H. B. M. (2016). A review of finger-vein biometrics identification approaches. Indian J. Sci. Technol, 9(32), 1-9.
6. Tsmots, I., Skorokhoda, O., & Rabyk, V. (2016, September). Structure and software model of a parallel-vertical multi-input adder for FPGA implementation. In 2016 XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT) (pp. 158-160). IEEE.
7. Kovalskyi, B., Dubnevych, M., Holubnyk, T., Pysanchyn, N., & Havrysh, B. (2019). Development of a technology for eliminating color rendering imperfections in digital photographic images. Vostochno-Evropeiskyi zhurnal peredovыkh tekhnolohyi, (1 (2)), 40-47.
8. Nazarkevych, M., Oliiarnyk, R., Nazarkevych, H., Kramarenko, O., & Onyshschenko, I. (2016, August). The method of encryption based on Ateb-functions. In 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP) (pp. 129-133). IEEE.
9. Nazarkevych, M., Logoyda, M., Troyan, O., Vozniy, Y., & Shpak, Z. (2019, September). The Ateb-Gabor Filter for Fingerprinting. In International Conference on Computer Science and Information Technology (pp. 247-255). Springer, Cham.
10. Nazarkevych, M., Lotoshynska, N., Brytkovskyi, V., Dmytruk, S., Dordiak, V., & Pikh, I. (2019, September). Biometric Identification System with Ateb-Gabor Filtering. In 2019 XIth International Scientific and Practical Conference on Electronics and Information Technologies (ELIT) (pp. 15-18). IEEE.
11. Sheketa, V., Zorin, V., Chupakhina, S, Kyrsta, N., Pasyeka, M., Pasieka, N. (2020). "Empirical Method of Evaluating the Numerical Values of Metrics in the Process of Medical Software Quality Determination,&quot. U International Conference on Decision Aid Sciences and Application (DASA), (s. 22- 26). https://doi.org/10.1109/DASA51403.2020.9317218.
12. Varetskyy, Y., Rusyn, B., Molga, A., & Ignatovych, A. (2010). A new method of fingerprint key protection of grid credential. In Image Processing and Communications Challenges 2 (pp. 99-103). Springer, Berlin, Heidelberg.
13. Ahmed, S., Lee, Y., Hyun, S. H., & Koo, I. (2018). Covert cyber assault detection in smart grid networks utilizing feature selection and euclidean distance-based machine learning. Applied Sciences, 8(5), 772.
14. Yaacob, R., Pritam, H. M. H., Hassan, N. F. N.,. Ooi, C. D., Ibrahim, H., Othman, P. J. (2017). Computer assisted segmentation of palmprint images for biometric research. U 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE) (s. 273–277). https://doi.org/10.1109/ICCSCE.2017.8284418.
15. Mishra, M., Bhattacharya, A., Singh, A., Dutta, M. K. (2018). A Lossless Model for Generation of Unique Digital Code for Identification of Biometric Images. U 4th International Conference on Computational Intelligence & Communication Technology (CICT), (s. 1-5). https://doi.org/10.1109/CIACT.2018.8480297.
16. Jaswal, G., Nath, R., Nigam, A. (2017). Deformable multi-scale scheme for biometric personal identification. U IEEE International Conference on Image Processing (ICIP), (s. 3555-3559). https://doi.org/10.1109/ICIP.2017.8296944.
17. Do, H., Truong, V., George, K., Shirke, B. (2019). EEG-Based Biometrics Utilizing Image Recognition for Patient Identification. U IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), (s. 0591-0595). https://doi.org/10.1109/UEMCON47517.2019.8992962.

18. Bychkov, O., Merkulova, K., Zhabska, Y. (2019). Software Application for Biometrical Persons Identification by Portrait Photograph Based on Wavelet Transform. U IEEE International Conference on Advanced Trends in Information Theory (ATIT), (s. 253-256). https://doi.org/10.1109/ATIT49449.2019.9030462.
19. Setiawan, H., Yuniarno, E. M. (2017). Biometric Recognition Based on Palm Vein Image Using Learning Vector Quantization. U 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), (s. 95-99). https://doi.org/10.1109/ICICI-BME.2017.8537770.
20. Abdufattokhov, S., Muhiddinov, B. (2019). Stochastic Approach for System Identification using Machine Learning. Dynamics of Systems, Mechanisms and Machines (Dynamics), (s. 1-4). https://doi.org/10.1109/Dynamics47113.2019.8944452.
21. Yang, B., Liu, D. (2019). Research on Network Traffic Identification based on Machine Learning and Deep Packet Inspection. U IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), (s. 1887-1891). https://doi.org/10.1109/ITNEC.2019.8729153.
22. Hammad I., El-Sankary, K. (2020).Using Machine Learning for Person Identification through Physical Activities. U IEEE International Symposium on Circuits and Systems (ISCAS), (s. 1-5). https://doi.org/10.1109/ISCAS45731.2020.9181231.
23. Jadav, S. (2018). Voice-Based Gender Identification Using Machine Learning. U 4th International Conference on Computing Communication and Automation (ICCCA), (s. 1-4). https://doi.org/10.1109/CCAA.2018.8777582.
24. Deng, Y., Wu, F., Du, L., Zhou, R., Cao, L. (2019). EEG-Based Identification of Latent Emotional Disorder Using the Machine Learning Approach. U IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), (s. 2642-2648). https://doi.org/10.1109/ITNEC.2019.8729424.
25. Traboulsi, A., Barbeau, M. (2019). Recognition of Drone Formation Intentions Using Supervised Machine Learning,". U International Conference on Computational Science and Computational Intelligence (CSCI), (s. 408-411). https://doi.org/10.1109/CSCI49370.2019.00079.
26. Mokni, R., Elleuch, M., Kherallah, M. (2016). Biometric Palmprint identification via efficient texture features fusion. U International Joint Conference on Neural Networks (IJCNN), (s. 4857-4864). https://doi.org/10.1109/IJCNN.2016.7727838.