

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.896

doi: 10.31498/2225-6733.41.2020.226118

© Проніна О.І.¹, Юхно Д.В.², Альошин С.В.³

РОЗПІЗНАВАННЯ ОБЛИЧ ЗА ДОПОМОГОЮ ЗГОРНУТОЇ НЕЙРОННОЇ МЕРЕЖІ

В статті розглядається тема розпізнавання обличчя за допомогою згорнутої нейронної мережі, яка може на основі проведеного тренування з обмеженою кількістю світлин розпізнати та ідентифікувати людину в реальному часі. В рамках роботи було проведено аналіз наукових публікацій існуючих методів розпізнавання та вилучення об'єктів на зображенні. Виявлено, що використання вже існуючого апарату класифікаторів Хаара має ряд недоліків, тому необхідно покращити та модифікувати цей класифікатор. Були визначені основні критерії, що необхідно модифікувати в стандартному класифікаторі для його покращення. Був виявлений загальний алгоритм для роботи зі зображенням при розпізнаванні. Для реалізації була побудована математична модель нейронної мережі та класифікатора. В експериментальних дослідженнях було проведено навчання нейронної мережі та її тестування на різних варіантах відображення обличчя на світлинах. Для визначення адекватності роботи розробленої згорнутої нейронної мережі було проведено тестування на визначення помилок.

Ключові слова: розпізнавання обличчя, нейронні мережі, алгоритм Хаара.

O.I. Pronina, D.V. Yuhno, S.V. Alosin. Face recognition using a neural network. Modern trends in security and development of information technologies push forward all spheres of human life. The task of isolating a human face in a natural or artificial setting and subsequent identification has always been among the highest priority tasks for researchers working in the field of machine vision systems and artificial intelligence. Besides, the task of recognition is very relevant in the field of security – just as for storing data and so for finding criminals on surveillance cameras, and so on. Besides, all recognition systems use neural networks to improve performance, increase efficiency and facilitate the process itself. However, at present, despite the similarity of the tasks and methods used in the development of alternative systems for biometric identification of a person, such as identification by fingerprints or by the image of the iris, the identification systems by the image of the face are significantly inferior to the above mentioned systems. Therefore, there many perfecting techniques for improving face recognition systems. In the work, an analysis of literary publications, existing algorithms used for face recognition and human identification has been carried out. The main method of face recognition is the use of a convolutional neural network, the selection of objects in the image is carried out using the Viola-Jones method, the AdaBoost machine learning algorithm is used, and the Haar classifier is most often used as a classifier. The article is devoted to the creation of software for face recognition using a convolutional neural network in real time. The software can recognize and identify a person with the head tilted, turned, and under different lighting conditions. In this case, sampling training for the model is car-

¹ канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, ORCID: 0000-0001-7085-8027, pronina.lelka@gmail.com

² магістр, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь

³ ст. викладач, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь

ried out on a limited number of photographs. Experimental studies were carried out to test the developed mathematical model and the real-time face recognition algorithm.

Keywords: *face recognition, neural networks, Haar algorithm.*

Постановка проблеми. Розпізнавання обличчя – це прогресивна технологія, що активно використовується в сучасному світі в різних сферах діяльності. Технологія розпізнавання осіб вимагає від штучного інтелекту чималих здібностей, а від розробника – постійного удосконалення алгоритмів. Розпізнавання обличчя вже застосовується в повсякденні, так, наприклад соціальна мережа Facebook вже може визначити ваших друзів на фотографії [1]. Надалі розпізнавання осіб дозволить вгадувати емоції людей, безпомилково вибирати їх на відео і знімках з натовпу, а також знаходити їх в будь-якій точці світу з вуличних відеокамер.

Розпізнавання осіб зазвичай використовується в системах безпеки і може бути порівняне з іншими біометричними системами (наприклад, розпізнавання відбитків пальців або очної діфрагми). Останнім часом воно також стало популярним в якості інструменту комерційної ідентифікації та реклами.

Проблема розпізнавання осіб розглядалася ще на ранніх стадіях комп'ютерного зору. Ряд компаній протягом більше 40 років активно розробляють автоматизовані, а зараз і автоматичні системи розпізнавання людських облич: Smith & Wesson (система ASID – Automated Suspect Identification System); ImageWare (система FaceID); Imagis, Epic Solutions, Spillman, Miros (система Trueface); Vissage Technology (система Vissage Gallery); Visionics (система FaceIt) та тощо.

На даний момент існує велика кількість алгоритмів і методів, які використовуються при розпізнаванні обличчя. У найзагальнішому випадку алгоритм вирішення задачі виявлення та ідентифікації людини по зображенню його особи складається з наступних очевидних кроків: виявлення факту присутності людини на аналізованій сцені; виділення фігури людини; виділення голови; визначення ракурсу спостереження голови (анфас, профіль); виділення особи; порівняння з еталонами і ідентифікація.

Метою даної роботи є використання вже існуючих алгоритмів розпізнавання особи з метою їх удосконалення шляхом додавання нейронної мережі. Розпізнавання обличчя буде відбуватися на основі декількох широко відомих алгоритмів, таких як виділення об'єктів на зображенні Віоли-Джонса, активні моделі зовнішнього вигляду і каскаду Хаара.

Аналіз останніх досліджень і публікацій. Технології розпізнавання осіб дозволяють виробляти автоматичний пошук і розпізнавання осіб в графічних файлах і відео потоці. Як зазначено у статті [2], усі системи складаються з трьох програм: основне розпізнавання обличчя; faceId – збереження даних лиця; визначення незнайомця – порівняння лиця з базою даних.

В роботі [3] представлено підхід до розпізнавання облич на основі комбінованого каскаду нейромережних класифікаторів, методу головних компонент та згорткової нейронної мережі. Наведено результати експериментальних досліджень етапів виявлення, визначення ракурсу та розпізнавання облич порівняно з відомими методами.

В роботі [4] розглянуто використання штучного інтелекту для розпізнавання обличчя у рамках роботи органів правопорядку; вони добиваються 100% точності системи за допомогою спеціальних дерев, які зберігають у собі перелік спеціальних ознак обличчя, потрібних для підвищення точності розпізнавання. Також, автор демонструє позбавлення зображення від шумів, тому що вони заважають у певній мірі досягати точності.

У статті [5] представлений метод розпізнавання осіб з низькою роздільною здатністю, заснований на дискримінантному кореляційному аналізі (DCA). Пропонований метод обчислювально ефективний і може застосовуватися для складних додатків реального часу, таких як розпізнавання декількох осіб, що з'являються в перевантаженому кадрі відеоспостереження.

Аналізуючи існуючі методи, можна визначити загальний алгоритм для роботи зі зображенням при розпізнаванні: необхідно мати на вході максимально якісне зображення обличчя (чи мати можливість покращити його); необхідно знайти обличчя на цьому зображенні; знайти ключові точки на зображенні (точність буде залежати від їх кількості та вірності їх пошуку відносно реального їх місцезнаходження); необхідно знайти усі взаємні відстані між контрольними точками; порівняти отримані дані з раніше отриманими. У даній роботі буде розглянуто всі пункти.

Першочерговим завданням в роботі є використання вже існуючого апарату класифікаторів Хаара [6], при цьому були виявлені основні недоліки, а саме – обмежена кількість точок, що використовуються, для знаходження обличчя та його подальшого розпізнання; обмежена кількість облич, що можуть бути розпізнані; обмеження щодо самих зображень – пошук та розпізнання обличчя задовільно працює лише у тих випадках, коли обличчя знаходиться анфас.

Маючи цю інформацію, було розроблено схему щодо вдосконалення класифікатора Хаара наступним чином: збільшити кількість точок, що використовуються для пошуку та розпізнання обличчя (не тільки положення очей та відстань між ними, положення роту та інші, а й використання брів, контуру обличчя, форми очей, положення носу, ніздрі та інше); покращити розпізнавання обличчя у різних положеннях (не тільки анфас, а й у профіль).

Виклад основного матеріалу. Виділення об'єктів на зображенні здійснюється за методом Віюлі-Джонса [7]. В оригінальній версії алгоритму Віюлі-Джонса використовувалися тільки примітиви без поворотів, а для обчислення значення ознаки сума яскравостей пікселів однієї підобласті віднімалася з суми яркостей іншої підобласті. У розвитку методу були запропоновані примітиви з нахилом на 45 градусів і несиметричних конфігурацій. Також замість обчислення звичайної різниці, було запропоновано приписувати кожній підобласті певне політичне значення і значення ознаки обчислювати як зважену суму пікселів різнотипних областей [8]:

$$\text{feature} = \sum_{i \in I=1, \dots, N} w_i \times \text{RectSum}(r_i). \quad (1)$$

За основу було обрано саме цю функцію (1) обчислення ознак. Спираючись на те, що в роботі використовується алгоритм машинного навчання AdaBoost [9], формула класифікатора набуває вигляду (2):

$$f(x) = \sum_{k=1}^m \alpha_k g_k(x), \quad (2)$$

де α_k та g_k – «слабкі» класифікатори за формулою (1).

Для кожної ознаки слабкий класифікатор визначає оптимальну порогову функцію класифікації, що забезпечує мінімальне число неправильно класифікованих прикладів. Слабкий класифікатор h , таким чином, складається з значень ознаки об'єкта f , порога p і паритету θ , з вказівкою напрямку знаку нерівності. В результаті на кожній ітерації формується простий класифікатор, загальний вид якого наведено у формулі (3):

$$h_j(z) = \begin{cases} 1, & \text{якщо } p_j f_j(z) < p_j \theta_j \\ 0, & \text{інакше} \end{cases}, \quad (3)$$

де p_j – значення порога, показує напрямок знаку нерівності; $f_j(z)$ – обчислене значення ознаки; z – вікно зображення розміром 20×20 пікселів.

Кількість ознак є дуже великою, тому обчислити повну безліч ознак є надмірно трудомістким завданням. Було виявлено, що невелика кількість цих ознак може бути об'єднана для формування ефективного класифікатора. Знаходження цих ознак є головним завданням в розробленому алгоритмі. Знаходження ознак відбувається наступним чином. Нехай є набір слабких класифікаторів (4) та кількість ітерацій T :

$$S = \{(x_1, y_1), \dots, (x_N, y_N)\}. \quad (4)$$

Далі слід проініціалізувати набір, формула (5):

$$d_n^{(1)} = 1/N, \text{ для всіх } n = 1, \dots, N. \quad (5)$$

Наступним кроком для кожної ітерації ($t = 1, \dots, T$) слід тренувати класифікатор у відношенні до зваженої вибірки $\{S, d^{(t)}\}$, після чого слід обчислити гіпотенузу за формулою (6):

$$h_t: x \rightarrow \{-1, +1\}, h_t = \lambda(S, d^{(t)}). \quad (6)$$

Далі розраховується зважена помилка навчання (7):

$$\varepsilon_t = \sum_{n=1}^N d^{(t)} I(y_n \neq h_t(x_n)). \quad (7)$$

Після чого слід наповнити вибірку новими значеннями (8):

$$\alpha_t = \frac{1}{2} \log \frac{1 - \varepsilon_t}{\varepsilon_t}. \quad (8)$$

Далі оновлюються ваги:

$$d^{(t+1)} = \frac{d_n^{(t)} \exp\{-\alpha_t y_n h_t(x_n)\}}{Z_t} d_n^{(t)}, \quad (9)$$

де Z_t – нормалізована константа, така як $\sum_{n=1}^N d_n^{(t+1)} = 1$.

Після чого необхідно переривати ітерації, якщо виконується умова (10):

$$\epsilon_t=0, \text{ або } \epsilon_t \geq \frac{1}{2} \text{ та } T=t-1. \quad (10)$$

У результаті буде отримано наступний класифікатор:

$$f(x) = \sum_{t=1}^T \frac{\alpha_t}{\sum_{r=1}^T \alpha_r} h_t(x). \quad (11)$$

Для класифікації зображень при розпізнаванні використовується згорнута нейронна мережа. Накладаються обмеження на мережу: швидкість відповіді – не більше 1 секунди; точність розпізнавання – не менше 70%. Загальна топологія мережі наведена на рис. 1.

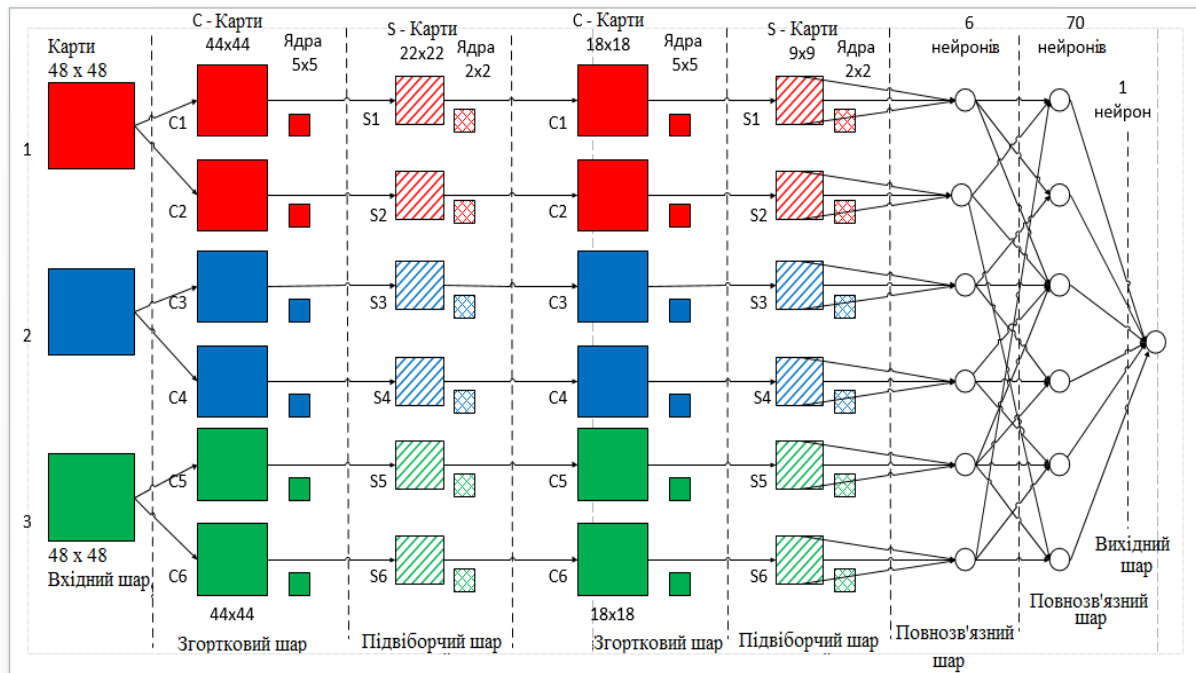


Рис. 1 – Топологія згорнутої нейронної мережі

Вхідні дані являють собою кольорові зображення типу JPEG, розміром 48×48 пікселів. Якщо розмір буде занадто великий, то обчислювальна складність підвищиться, відповідно, обмеження на швидкість відповіді будуть порушені, визначення розміру в даній задачі вирішується методом підбору. Якщо вибрати розмір занадто маленький, то мережа не зможе виявити ключові ознаки осіб. Кожне зображення розбивається на 3 канали: червоний, синій, зелений. Таким чином виходить 3 зображення розміром 48×48 пікселів.

Одним з етапів розробки нейронної мережі є вибір функції активації нейронів. Вид функції активації багато в чому визначає функціональні можливості нейронної мережі і метод навчання цієї мережі. У даній роботі в якості функції активації в прихованих і вихідному шарах застосовується гіперболічний тангенс, в згортальних шарах застосовується ReLU.

Був проведений експеримент для перевірки розробленої нейронної мережі. Для проведення експерименту на першому етапі було проведено налаштування розробленої згорнутої нейронної мережі; для цього до нейронної мережі було завантажено світліну зі зображенням двадцяти трьох людей. Далі відбувалось тренування нейронної мережі, щоб покращити показник визначення обличчя на світліні. Після того, як нейронна мережа розпізнала всі обличчя, вона була готова для подальшого використання, приклад рис. 2. Також ця світліна була завантажена для перевірки на стандартному класифікаторі Хаара.

Стандартний класифікатор Хаара на світліні з рис. 2 зміг знайти лише 10 облич, оскільки деякі обличчя є в нахилі чи перекриті іншими обличчями. Тобто розроблений класифікатор прокрашує результати визначення облич.

Для визначення адекватності роботи розробленої згорнутої нейронної мережі було проведено тестування на визначення помилок. Існує два види помилок: помилки першого роду та

помилки другого роду. Помилки першого роду – це коли об’єкту не має на цій світліні, а мережа його там розпізнає, та помилки другого роду – коли об’єкт є на світліні, а нейронна мережа його не розпізнає.

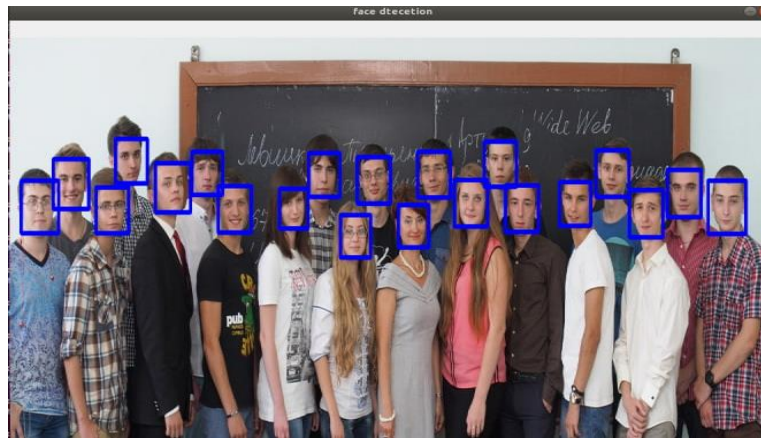


Рис. 2 – Налаштована нейронна мережа, що знайшла обличчя

Для розрахунку помилок першого роду використовується формула (12), для розрахунку помилок другого роду, використовується формула (13) [10].

$$d_{\alpha} = 1 - \sum_{i=k}^n S_n^i (1-\alpha)^i \cdot \alpha^{n-i}; \quad (12)$$

$$d_{\beta} = \sum_{i=k}^n S_n^i (1-\beta)^{n-i} \cdot \beta^i; \quad (13)$$

де α – помилки першого роду; β – помилки другого роду; d_{α} – ймовірності помилок першого роду, відмова більш, ніж $(n-k)$ -показників; S_n^i – ситуація, яка виникла; k – кількість показників, яка необхідна для коректної роботи; i – номер поточного показника; n – можлива загальна кількість показників в системі.

Для цього було сформовано три різних вибірки. Вибірki були сформовані випадковим чином та мають де окремі світліни людей, де групові світліни. В кожній вибірці своя кількість об’єктів, що треба класифікувати, та кількість класів, до яких їх слід віднести. Вибірka 3 має більшу кількість об’єктів та класів та має не зовсім точні обличчя та перехрестя облич, тим самим вона завідомо більш складніша. Навчання відбувалось для всіх трьох вибірок. Графічні результати дослідження наведені на рис. 3.

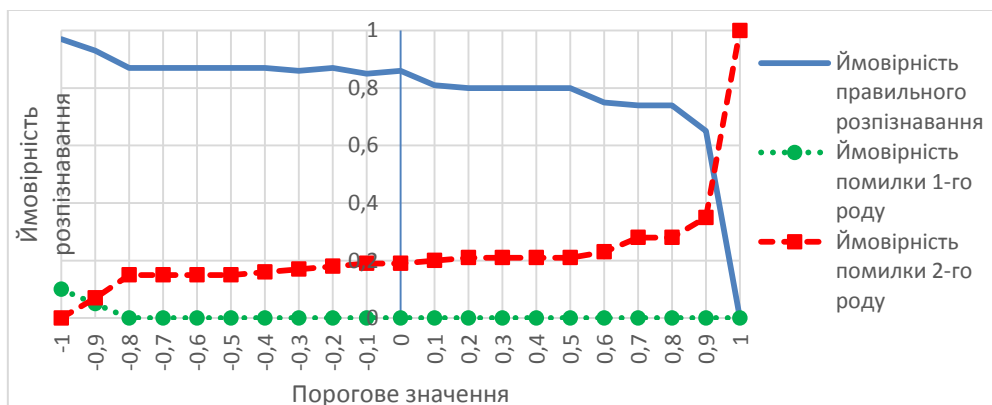


Рис. 3 – Залежність достовірності розпізнавання обличчя від порогу згорнутої нейронної мережі для вибірки

Результати дослідження, наведені на рисунку 3, демонструють, що при більш складних зображеннях результати розпізнавання є менш точнішими та ймовірність виявлення помилок як першого так і другого роду зростає.

Висновки

Робота є актуальною, оскільки на сьогоднішній день питання ідентифікації використовується в багатьох сферах як засіб безпеки, ідентифікації при використанні пристроїв, тощо. В статті були розглянуті основні методи розпізнавання обличчя.

На основі проведеного літературного огляду було вирішено розробляти нейронну мережу для розпізнавання та ідентифікації обличчя в режимі реального часу на основі класичних алгоритмів. Для цього було покращено існуючий класифікатор Хаара та удосконалено методи верифікації обличчя-кандидатів. Метод базується на властивості згорнутої нейронної мережі, що дозволяє обробляти вхідне зображення за один етап і формувати навчальну вибірку згорнутої нейронної мережі, що використовує параметричну адаптацію структури активної навчальної вибірки. Запропоновано враховувати ракурс обличчя при розпізнаванні. Було розроблено програмний продукт та експериментально досліджено, що алгоритм ідентифікації та класифікації обличчя з використанням згорнутої нейронної мережі та покращеного класифікатора працює достовірно.

Розроблена система вирішує найрозповсюдженіші проблеми розпізнавання обличчя, такі як чутливість до зашумленого фону, чутливість до освітлення, відмінного від того, що було присутнє на навчальній вибірці, нестійкість до присутності декількох людей на зображенні.

Розроблене програмне забезпечення можна використовувати в реальному часі, здійснювати ідентифікацію обличчя, тобто використовувати в системах безпеки як дешифратор ключа, оснований на розпізнаванні володаря мобільного пристрою, тощо. Такий підхід збільшить надійність систем, що розробляються зі графічним ключом, та може бути використаний в будь-якій сфері.

Перелік використаних джерел:

1. Facebook's facial recognition now looks for you in photos you're not tagged in [Electronic resource]: [Website]. – Electronic data. – Mode of access: <https://www.theverge.com/facebook-facial-recognition-tagging-photos>.
2. Andrew Heinzman. How Does Facial Recognition Work? [Electronic resource]: [Website]. – Electronic data. – Режим доступу: <https://www.howtogeek.com/427897/how-does-facial-recognition-work/>.
3. Палій І.О. Нейромережний підхід до комп'ютерного розпізнавання обличчя / І.О. Палій, А.О. Саченко, С.Г. Антошук, Т.О. Бурак // Штучний інтелект. – 2010. – № 3. – С. 378-387.
4. Error Rates in Users of Automatic Face Recognition Software / D. White, J.D. Dunn, A.C. Schmid, R.I. Kemp // PLOS ONE. – 2015. – Vol. 10(10). – Pp. 1-14. – Mode of access: [DOI: 10.1371/journal.pone.0139827](https://doi.org/10.1371/journal.pone.0139827).
5. Haghghat M. Low Resolution Face Recognition in Surveillance Systems Using Discriminant Correlation Analysis / M. Haghghat, M. Abdel-Mottaleb // 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). – 2017. – Pp. 912-917. – Mode of access: [DOI: 10.1109/FG.2017.130](https://doi.org/10.1109/FG.2017.130).
6. Cootes T. Active appearance models / T. Cootes, G. Edwards, C. Taylor // Proceedings of the European Conference on Computer Vision. – 1998. – Vol. 2. – Pp. 484-498.
7. Convolutional Neural Networks for Visual Recognition [Electronic resource]: [Website]. – Electronic data. – Mode of access: <https://www.cs231n.github.io>.
8. Pat. 2002/0102024A1 US, Int. Cl. G 06 K 9/62. Method and system for object detection in digital images / M.J. Jones, P. Viola. – № 09/992,795; filed 12.11.2001; pub. date 01.08.2002. – 19 p.
9. Boosting and AdaBoost for Machine Learning [Electronic resource]: [Website]. – Electronic data. – Mode of access: <https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/>.
10. Каргин А.А. Введение в интеллектуальные машины. Кн. 1. Интеллектуальные регуляторы / А.А. Каргин. – Донецк : Норд-Пресс, ДонНУ, 2010. – 526 с.

References:

1. Facebook's facial recognition now looks for you in photos you're not tagged in Available at: <https://www.theverge.com/facebook-facial-recognition-tagging-photos> (accessed 15 June 2020).
2. Andrew Heinzman. How Does Facial Recognition Work? Available at:

- <https://www.howtogeek.com/427897/how-does-facial-recognition-work/> (accessed 10 June 2020)
3. Paliy I.O., Sachenko A.O., Antoshchuk S.H., Burak T.O. Neyromerezhniy pidkhdid do komp'yuternoho rozpiznavannya oblych [Neural network approach to computer facial recognition]. *Shtuchnyy intelekt – Artificial Intelligence*, 2010, no. 3, pp. 378-387.
 4. White D., Dunn J.D., Schmid A.C., Kemp R.I. Error Rates in Users of Automatic Face Recognition Software. *PLOS ONE*, 2015, vol. 10 (10), pp. 1-14. doi: 10.1371/journal.pone.0139827.
 5. Haghghat M., Abdel-Mottaleb M. Low Resolution Face Recognition in Surveillance Systems Using Discriminant Correlation Analysis. *12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 912-917. doi: 10.1109/FG.2017.130.
 6. Cootes T., Edwards G., Taylor C. Active appearance models. *Proceedings of the European Conference on Computer Vision*, 1998, vol. 2, pp. 484-498.
 7. Convolutional Neural Networks for Visual Recognition Available at: <https://www.cs231n.github.io> (accessed 10 August 2020).
 8. Viola P., Jones M.J. Method and system for object detection in digital images. Patent 2002/0102024A1, 2002.
 9. Boosting and AdaBoost for Machine Learning Available at: <https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/> (accessed 20 July 2020).
 10. Kargin A.A. *Vvedeniye v intellektual'nyye mashiny. Kniga 1. Intellektual'nyye regulatory* [Introduction to intelligent machines. Book 1. Intellectual regulators]. Donetsk, Nord-Press, DonNU Publ., 2010. 526 p. (Rus.)

Рецензент: О.Є. П'ятикоп
канд. техн. наук, доц., ДВНЗ «ПДТУ»

Стаття надійшла 15.11.2020

УДК 004.78

doi: 10.31498/2225-6733.41.2020.226119

© П'ятикоп О.Є.¹, Мініна К.Г.², Безуглов Р.О.³

ФОРМУВАННЯ РЕКОМЕНДАЦІЙ ФІЛЬМІВ НА ОСНОВІ ГІБРИДНОГО ПІДХОДУ

Стаття присвячена темі створення системи рекомендацій, яка зможе проаналізувати дані, отримані від користувача про вподобання фільмів та використати їх для прогнозування нових рекомендацій. В рамках роботи було проведено аналіз наукових видань, існуючих методів та алгоритмів, які використовуються для побудови системи рекомендацій та їх гібридизацію. Визначено, що гібридизація може бути досягнута різними підходами завдяки поєднанню декількох алгоритмів. Об'єднання методів для вирішення проблеми в більшості випадків дає більш ефективні і точні результати. Для реалізації гібридного підходу обрано два метода: фільтрацію на основі вмісту та спільну (колаборативну) фільтрацію, кожен з яких відповідає певному сценарію рекомендацій. У експериментальних дослідженнях було випробувано здатність системи запропонувати фільми, спираючись на дані про подібність користувачів. Подібність користувачів розраховувалась на основі вимірювання ступеня лінійної залежності – коефіцієнт кореляції Пірсона.

Ключові слова: рекомендаційна система, колаборативна фільтрація, контентна фільтрація, схожість елементів.

¹ канд. техн. наук, доцент, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, ORCID: 0000-0002-7731-3051, pee_pstu@ukr.net

² магістр, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь

³ магістр, ДВНЗ «Приазовський державний технічний університет», м. Маріуполь