

КРИВЕНЧУК Ю. П.

Національний університет "Львівська політехніка"

ORCID ID: 0000-0002-2504-5833

e-mail: yurkokryvenchuk@gmail.com

ЛАВРИК Ю. О.

Національний університет "Львівська політехніка"

e-mail: ylilav55547@gmail.com

СИСТЕМА РОЗПІЗНАВАННЯ ЕМОЦІЙНОЇ ХАРАКТЕРИСТИКИ ОБЛИЧЧЯ ЛЮДИНИ

У сучасному світі дуже багато речей і процесів, до яких ще не так давно ставились, як до чогось з області фантастики, все стрімкіше впроваджуються у реальне життя. Прикладом такого є і автоматизація процесу розпізнавання емоцій, яка вже на сьогодні не є недосяжною технологією та займає вагоме місце у галузі комп'ютерного зору. Технології розпізнавання емоцій виразу обличчя людини можуть слугувати для різних областей та цілей, що приваблює велику кількість науковців і розробників, деякі з яких вже реалізували такі системи, але їх кількість на сьогодні є невеликою і більшість таких продуктів вимагають спеціального обладнання та високих обчислювальних потужностей. Ще одними поширеними проблемами є вразливість до нечітких зображень та закритість систем, що не дає можливості проаналізувати та модифікувати розробку. Тому виникає необхідність оптимізації автоматизованого процесу розпізнавання емоцій виразу обличчя людини для створення більш стійкої, якісної та доступної системи. У роботі наведено опис та результати створеної системи розпізнавання емоцій, основою якої є згорток нейронна мережа. Зв'язок користувача з системою проводиться через простий інтерфейс, який надає можливість завантажити готове зображення для аналізу та увімкнути камеру для ідентифікації емоцій у режимі реального часу. Практична цінність цієї роботи полягає в тому, що її результатом є розроблена високоякісна системи розпізнавання емоційної характеристики обличчя людини, яка є вже готовою до використання та впровадження у відповідні сфери.

Ключові слова: розпізнавання емоцій, обличчя, згорток нейронна мережа, OpenCV.

YURI KRYVENCHUK, Y. LAVRYK

Lviv Polytechnic National University

CREATION OF FACIAL EMOTION RECOGNITION SYSTEM

In the modern world, a lot of things and processes, which not so long ago were treated as something from fiction, are increasingly being implemented in real life. An example of this is the automation of the process of recognizing emotions, which is no longer an unattainable technology and occupies an important place in the area of computer vision. Technologies for recognizing the emotions of facial expression can be used for various fields and purposes: from creating games and applications to more complex, such as analysis of people's mental state, discussions, marketing research, attention monitoring, driver status, robotics and so on. All this confirms the high academic and commercial potential of systems for recognizing the emotional characteristics of the person and, accordingly, brings a large number of scientists and developers, some of whom have already implemented such systems. But today their number is small and a lot of such products require special equipment and high computing performance. Another common problem is the vulnerability to blurred images and the closed nature of the system, which makes it impossible to analyze and modify the development. Therefore, the optimization of the automated process of recognizing the emotions of facial expression is needed in order to create a more stable, high quality and affordable system. The paper have description and results of the creation of system of emotion recognition, the basis of which is a convolutional neural network. The user can connect with the system through a simple interface that allows you to download the image for analysis or turn on the camera to identify emotions in real time. The practical value of this work is the developed high-quality system for recognizing the facial emotional characteristics which is ready for use and implementation in relevant fields.

Keywords: face, emotion recognition, CNN, OpenCV.

Постановка проблеми

Знання емоцій дає важливу інформацію, яка надає можливість глибше зрозуміти поведінку людини. Наприклад, емоції можуть впливати на прийняття рішення про те, чи купувати, чи ні, на вибір певних речей, на взаємодію з іншими. Таким чином технології розпізнавання емоційної характеристики обличчя людини можуть слугувати для різних областей та цілей: від створення ігор та додатків до більш складних і затребуваних, таких як аналіз психічного стану людей, дискусій, маркетингові дослідження, моніторинг уваги, визначення стану водія, робототехніка та інше. Наприклад, за допомогою системи розпізнавання емоцій можна проаналізувати реакцію споживачів на певну послугу чи продукт. Такий аналіз підкаже, на чому потрібно більше зосередитись постачальнику, і допоможе покращити продажі у майбутньому.

Більшість системи, які виконують завдання розпізнавання емоцій на обличчі людини, вимагають високої обчислювальної потужності чи спеціального обладнання. Окрім цього, такі розробки застосовуються і є достатньо поширеними за кордоном, що не можна сказати про нашу країну, одним з показників чого є невелика кількість публікацій на цю тему.

Вище наведена інформація доводить те, що є потреба у створенні системи автоматизованого процесу розпізнавання емоцій, яка не буде потребувати значних ресурсів для коректної роботи, та її впровадження в Україні. Такий продукт зможе широко застосовуватись у теперішній час і принести велику користь багатьом сферам. Це все підтверджує актуальність та значний академічний і комерційний потенціал цієї роботи.

Метою цього дослідження є створення високоякісної, швидкодіючої системи розпізнавання емоцій на обличчі людини з можливістю аналізу готового зображення та роботи у режимі реального часу.

Аналіз останніх досліджень

В [1] надано загальний опис та аналіз сучасних підходів до розпізнавання емоцій виразів обличчя. Проведено огляд деяких баз даних та основних етапів аналізу зображень. Визначено, що найкращими методами класифікації для визначеної задачі є штучні нейронні мережі.

У публікації [2] велика увага приділяється численним можливим сферам застосування системи розпізнавання емоцій на обличчі людини та її потенційній користі, що підтверджує актуальність роботи. Описано використання згорткової нейронної мережі, швидкості навчання якої в 1,5 рази швидша, ніж при використанні традиційного класифікатора, що ще раз доводить правильність та доцільність використання нейронної мережі у роботі. Недоліками є те, що система у [2] є вразливою до позиції обличчя, поганої якості зображення та досліджена лише при роботі зі статичними зображеннями.

У роботі [3] теж пропонується згорткова нейронна мережа, але у поєднанні з новим гібридним методом фільтрації, який використовується для видалення будь-яких шумів, присутніх на зображеннях. Таке поєднання дозволило підняти точність з 65 % до 85 % для датасету FER2013. У статті досліджено кілька архітектур нейронної мережі, описано їх переваги та недоліки. Інформація у [3] є допоміжною при вирішенні проблеми вразливості системи до шумів зображень та при виборі архітектури мережі при розробленні власної системи.

Одним з обмежень [4] є те, що модель навчена окремо для різних наборів даних. Таким чином, модель, навчена на певному наборі даних, може надто підлаштуватись до цих даних і показувати погані результати на нових даних. Схожа проблема є в [5], де також використані дані з мережі, які є однотипними і недостатньо великими для роботи з нейронною мережею. Погано класифікується емоція суму, причиною чого може бути нерівномірний розподіл навчальних даних і відповідно невелика кількість зображень з цією емоцією.

У [6] описано сформовану архітектуру згорткової нейронної мережі, на основі паралельного вилучення ознак. Така архітектура перевершує багато методів за точністю і містить значно меншу кількість параметрів, ніж інші архітектури. Це полегшує проблему високої обчислювальної потужності та великого об'єму пам'яті, що вимагають більшість методів для коректної роботи.

[7] описує реалізацію розпізнавання емоцій на обличчі за допомогою використання каскад Хаара для процесу виявлення обличчя та згорткової нейронної мережі для процесу розпізнавання емоцій. Стаття демонструє важливість етапу розпізнавання обличчя перед класифікацією емоції. Недоліком є недостатня кількість навчальних даних, серед яких лише зображення обличчя з фронтальною перспективою.

У [8] наведено процес розроблення системи для розпізнавання емоцій на основі і мовленнєвої, і мімічної інформації. Обмеження в неякісній класифікації емоцій суму та злості. Детально описано та показано на прикладах важливість попереднього налаштування гіперпараметрів моделі.

Виклад основного матеріалу

Огляд наборів даних. Проаналізувавши роботи у численних джерелах, було прийнято рішення для навчання моделей нейронної мережі обрати кілька наборів даних та об'єднати їх частини. Оскільки більшість наборів є однотипними і недостатньо великими, використання лише одного може спричинити вразливість та неякісну роботу системи на реальних даних. У таблиці 1 наведено основні характеристики проаналізованих датасетів.

Таблиця 1

Огляд наборів даних, що містять вирази обличчя людей з відповідними емоціями [1]

База даних	Кількість прикладів	Кількість моделей	Роздільна здатність	Колір	Кількість емоцій	Особливості
JAFFE	213	10	256×256	Сірий	7	Азіатська зовнішність
CK+	920	123	640×480, 640×490	RGB, сірий	8	Обличчя різних національностей
KDEF	4900	70	562×562	RGB	7	Фіксування обличчя у п'ятих ракурсах
FER2013	35887	—	48×48	Сірий	7	Зображення з інтернету

Етапи процесу розпізнавання емоцій. Автоматизований процес розпізнавання емоцій поділяють на три основні етапи: розпізнавання обличчя; вилучення ознак (орієнтирів обличчя, контурних точок обличчя); класифікація емоцій.

Розпізнавання обличчя. Виявлення обличчя на зображенні є складним завданням перш за все через можливі варіації обличч. Різні розміри, кути, пози формують безмежний спектр людських лиць. Окрім цього,

численні емоції, які переживають люди, та характеристики самих зображень, наприклад, освітлення та фон, теж мають значний вплив на вигляд обличчя.

Одним з найпоширеніших й ефективних інструментів для цього є каскади Хаара. Їх основою є набори прямокутних областей з різними характеристиками яскравостей пікселя, що перемішуються по зображенню і порівнюються з його частинами для виявлення збігів. Метод передбачає використання монохромних зображень, оскільки це надає приріст у швидкості часу розпізнавання [9].

Для застосування вищеописаного методу було використано бібліотеку OpenCV у мові Python.

Вилучення ознак. Основним викликом, на який потрібно відповісти для успішного вилучення орієнтирів обличчя, є вибір достатнього набору контурних точок, які представляють важливі характеристики людського обличчя. Зазвичай, використовуються підходи на основі геометрії та симетрії обличчя. Симетрія в межах обличчя допомагає виявляти орієнтири обличчя незалежно від відмінності у їх формі, розмірі і структурі. При роботі з відеопослідовностями, виділяють зображення, що відповідають кожному кадру, тоді виявляються початкові ознаки на основі першого кадру, які після цього зіставляються із з кожним іншим кадром.

У роботі використано згорткову нейронну мережу, яка на відміну від традиційних класифікаторів, автоматизує процес вилучення ознак, самостійно визначаючи значущі ознаки на основі навчальних даних.

Класифікація емоцій. За останні кілька десятиліть запроваджуються та розвиваються різні підходи для класифікації емоцій. Усі ці підходи зосереджені на класифікації шести універсальних емоцій за П. Екманом. Це радість, сум, здивування, відраза, злість, страх [10]. Інші, небазові емоції, не враховуються. Хороший класифікатор емоцій повинен вміти розпізнавати емоції незалежно від статі, віку, етнічної приналежності, пози, умов освітлення, фону, зачіски, наявності окулярів чи бороди тощо. При роботі з відеопослідовностями, виділяють зображення, що відповідають кожному кадру, тоді виявляються початкові ознаки на основі першого кадру, які після цього зіставляються із з кожним іншим кадром.

Етап класифікації емоцій реалізується за допомогою натренованої згорткової нейронної мережі на наборі даних з частини датасетів вказаних у табл. 1. Для побудови нейронної мережі було використано швидкісний та ефективний при роботі з алгоритмами глибинного навчання фреймворк MXNet у мові Python. Однією з найбільш його переваг є підтримка графічного процесора і відповідно швидкість, що є особливо важливим, оскільки модель нейронної мережі достатньо складна, а набір даних великий.

Навчання нейронної мережі відбувалось протягом 25 епох. Час – 22 хв. Отримана точність – 84 %.

Після програмної реалізації усіх вищеописаних процесів за допомогою графічної бібліотеки Tkinter у Python було розроблено простий інтерфейс для взаємодії користувача з системою. Приклад роботи системи з розробленим інтерфейсом наведено на рис. 1.

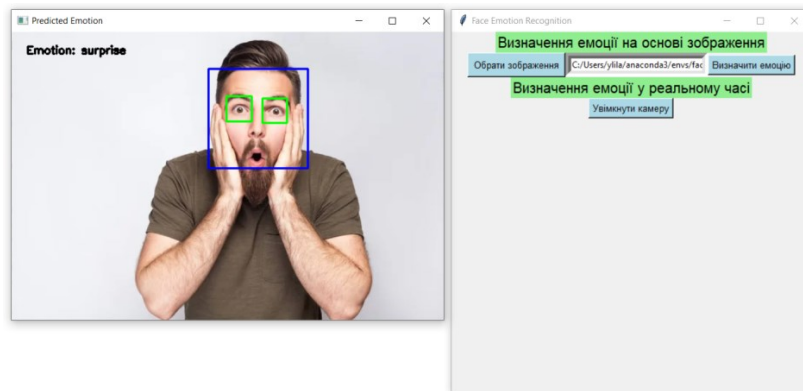


Рис. 1. Приклад роботи системи при визначенні емоції на основі зображення

Висновки

В результаті проведеної роботи створено систему розпізнавання емоційної характеристики обличчя людини. Така розробка надає можливість прогнозування емоції на завантаженому зображенні, а також підтримує режим ідентифікації емоції у реальному часі. В загальному, детальне дослідження задачі розпізнавання емоцій показало, що система розв'язання такої задачі характеризується значною затребуваністю та корисністю для різних сфер життя, а оскільки ця галузь недостатньо досліджена в Україні, не впроваджено аналогів, то така розробка є дуже доцільною та актуальною. Точність розробленої системи становить 84 %, що є хорошим показником, проте наступні дослідження можуть бути спрямовані на досягнення вищої якості завдяки збільшенню епох під час навчання та детальнішого підбору гіперпараметрів нейронної мережі.

Література

1. Ryumina E. V. Analytical review of methods for emotion recognition by human face expressions / E. V. Ryumina, A. A. Karpov // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. – 2020. – Vol. 2. – P. 163–176.

2. Zhang H. A face emotion recognition method using convolutional neural network and image edge computing / H. Zhang, A. Jolfaei, M. Alazab // IEEE Access. – 2019. – Vol. 7. – P. 159081–159089.
3. Bodavarapu P. N. R. An optimized neural network model for facial expression recognition over traditional deep neural networks / P. N. R. Bodavarapu, P. V. V. S. Srinivas // International Journal of Advanced Computer Science and Applications. – 2021. – Vol. 12, No. 7. – P. 443–451.
4. Deng J. CGAN based facial expression recognition for human-robot interaction / J. Deng, G. Pang, Z. Zhang, [et al.] // IEEE Access. – 2019. – Vol. 7. – P. 9848–9859.
5. Kim J.-H. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure / J.-H. Kim, B.-G. Kim, P. P. Roy, D.-M. Jeong // IEEE Access. – 2019. – Vol. 7. – P. 41273–41285.
6. Riaz M. N. EXnet: an efficient approach for emotion recognition in the wild / M. N. Riaz, Y. Shen, M. Sohail, M. Guo // Sensors (Switzerland). – 2020. – Vol. 20, No. 4.
7. Shukla R. Facial emotion recognition by deep cnn and haar cascade / R. Shukla, L. Agilandeewari, M. Prabukumar // International Journal of Innovative Technology and Exploring Engineering. – 2019. – Vol. 8, No. 12. – P. 3433–3441.
8. Luna-Jiménez C. Multimodal emotion recognition on raveds dataset using transfer learning / C. Luna-Jiménez, D. Griol, Z. Callejas, [et al.] // Sensors. – 2021. – Vol. 21, No. 22.
9. Туманов В. И. Использование каскадов хаара для распознавания образов / В. И. Туманов, Т. С. Катермина ; Нижневарт. гос. ун-т, 2021. – С. 188–191. – ISBN 978-5-00047-589-8.
10. Коваленко А. Система розпізнавання мімічних проявів емоцій людини з використанням багатошарового перцептрону / А. Коваленко. – 2011.

References

1. Ryumina E. V. Analytical review of methods for emotion recognition by human face expressions / E. V. Ryumina, A. A. Karpov // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. – 2020. – Vol. 2. – P. 163–176.
2. Zhang H. A face emotion recognition method using convolutional neural network and image edge computing / H. Zhang, A. Jolfaei, M. Alazab // IEEE Access. – 2019. – Vol. 7. – P. 159081–159089.
3. Bodavarapu P. N. R. An optimized neural network model for facial expression recognition over traditional deep neural networks / P. N. R. Bodavarapu, P. V. V. S. Srinivas // International Journal of Advanced Computer Science and Applications. – 2021. – Vol. 12, No. 7. – P. 443–451.
4. Deng J. CGAN based facial expression recognition for human-robot interaction / J. Deng, G. Pang, Z. Zhang, [et al.] // IEEE Access. – 2019. – Vol. 7. – P. 9848–9859.
5. Kim J.-H. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure / J.-H. Kim, B.-G. Kim, P. P. Roy, D.-M. Jeong // IEEE Access. – 2019. – Vol. 7. — P. 41273–41285.
6. Riaz M. N. EXnet: an efficient approach for emotion recognition in the wild / M. N. Riaz, Y. Shen, M. Sohail, M. Guo // Sensors (Switzerland). – 2020. – Vol. 20, No. 4.
7. Shukla R. Facial emotion recognition by deep cnn and haar cascade / R. Shukla, L. Agilandeewari, M. Prabukumar // International Journal of Innovative Technology and Exploring Engineering. – 2019. – Vol. 8, No. 12. – P. 3433–3441.
8. Luna-Jiménez C. Multimodal emotion recognition on raveds dataset using transfer learning / C. Luna-Jiménez, D. Griol, Z. Callejas, [et al.] // Sensors. – 2021. – Vol. 21, No. 22.
9. Туманов В. И. Ispolzovanie kaskadov haara dlya raspoznavaniya obrazov / V. I. Tumanov, T. S. Katernina. — Nizhnevartovskij gosudarstvennyj universitet, 2021. – S. 188–191. – ISBN 978-5-00047-589-8.
10. Kovalenko A. Systema rozpoznavannia mimichnykh proiaviv emotsii liudyny z vykorystanniam bahatosharovoho perseptronu / A. Kovalenko. – 2011.

Рецензія/Peer review : 14.11.2021

Надрукована/Printed :30.12.2021