

УДК 004.056.523:159.942]:004.032.26

DOI: 10.31673/2412-9070.2023.065457

О. Г. ВАРФОЛОМЕЄВА¹, канд. техн. наук, доцент;С. І. ОТРОХ², доктор техн. наук, професор;К. Д. СУШИЛЬНИКОВ², студент;М. О. ШАЛИГІН², студент,¹ Державний університет інформаційно-комунікаційних технологій, Київ² Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ

РОЗПІЗНАВАННЯ ЕМОЦІЙ ЛЮДИНИ В РЕАЛЬНОМУ ЧАСІ

Висвітлено значущість та актуальність застосування нейронних мереж для розпізнавання емоцій людини в реальному часі. Досліджено можливості використання згорткових нейронних мереж для ефективного аналізу та класифікації емоцій на обличчях. Описано процес розроблення програми з використанням JavaScript та TensorFlow, а також розглянуто потенційні можливості оптимізації цієї програми для вдосконалення швидкості та точності розпізнавання емоцій. Запропоновано високоефективне вирішення для розпізнавання емоцій, яке може бути застосовано в різних сферах: від медицини та соціальних досліджень до маркетингу та розваг. Особливу увагу приділено розвитку технологій, спрямованих на покращення взаємодії між людьми та комп'ютерами, та відображає важливість використання інноваційних методів для поліпшення якості життя людей.

Ключові слова: нейронна мережа; згортка; JavaScript; TensorFlow.

ВСТУП

У сучасному цифровому світі все помітніше зростає важливість розуміння та інтерпретації людських емоцій. Це має особливе значення в таких галузях, як соціальна взаємодія, медицина, маркетинг, реклама, розваги та багато інших. Відомості про емоції людини можуть використовуватися для покращення користувацького досвіду, персоналізації послуг, аналізу настрою та багатьох інших цілей.

Отже, розроблення програми, яке здатне точно розпізнавати емоції людини за допомогою використання нейронних мереж у реальному часі, має великий потенціал для застосування в різних сферах життя. Це сприятиме поліпшенню взаємодії між людьми та технологіями, даючи змогу системам краще розуміти та відповідати на емоційний стан користувачів. Водночас це може мати значущий вплив на розвиток емоційного інтелекту та позитивно впливати на підвищення загального добробуту людей.

Метою статті є розроблення програмного забезпечення, яке здатне точно розпізнавати та класифікувати емоційний стан людини на основі обличчя з використанням нейронних мереж у реальному часі. Основним завданням є створення ефективною системи, яка може ідентифікувати широкий спектр емоцій, таких як щастя, сум, злість, страх, здивування, з високою точністю та швидкодією. Це дасть змогу використовувати таку технологію для автоматичного аналізу емоцій у реальному часі в різних контекстах, включно з оцінюванням реакцій на відеоконтент, моніторингом емоційного стану під час медичних процедур, психологічних досліджень, оцінюван-

ням емоційних реакцій на рекламу та маркетингові кампанії, а також для покращення взаємодії людини зі штучним інтелектом та інтерфейсами користувача.

ОСНОВНА ЧАСТИНА

Розпізнавання емоцій нейронними мережами стало значущим кроком у сфері штучного інтелекту та оброблення природної мови. Перші дослідження в цьому напрямку виникли у 1990-х роках, коли вчені почали використовувати нейронні мережі для аналізу та класифікації людських емоцій на основі текстів і звукових сигналів. Згодом із розвитком технологій і збільшенням обчислювальних можливостей дослідники розширили цей підхід на оброблення зображень, зокрема обличчя людей, що відкрило шлях до точного розпізнавання емоцій за допомогою комп'ютерних систем.

Незважаючи на переваги таких мереж, існують і аналоги методів розпізнавання емоцій, такі як використання методів оброблення природної мови для аналізу текстової інформації та методів аналізу звукових сигналів для інтерпретації голосових виразів. Однак нейронні мережі є більш ефективними у розпізнаванні складних візуальних шаблонів та емоцій на обличчях людей, що робить їх перспективними для використання в різних сферах застосування.

Одним із сучасних прикладів нейронних мереж, якими послуговуються для розпізнавання емоцій на обличчях, є модель «Facial Expression Recognition using Deep Learning» (FER) на основі згорткових нейронних мереж. Модель навчається на великому наборі зображень обличчя, що охоплюють різні вирази обличчя та емоцій, зокрема щастя, сум,

злість, страх, здивування тощо. Цей набір даних може містити тисячі зображень людей з різними емоційними виразами. Розробка застосовує згорткові шари для виявлення важливих візуальних шаблонів на обличчях, зокрема очі, рот, брови та їхні конфігурації. Після згортки використовуються повністю зв'язані шари для класифікації емоцій на основі здобутих ознак. До додаткових функцій можуть належати авґментація даних для збільшення набору даних для навчання та використання методів оптимізації, зокрема стохастичний градієнтний спуск, для вдосконалення точності моделі.

Складова програми

Для реалізації цієї роботи було вибрано два основних інструменти: JavaScript для фронтенду та TensorFlow для створення та тренування нейронних мереж.

JavaScript є однією з найпоширеніших мов програмування, особливо у сфері веброзроблення. Вона забезпечує можливість легкої інтеграції із вебзастосунками та дає змогу запускати програми прямо у веббраузері. Це уможливило створення динамічних та інтерактивних інтерфейсів для користувачів. За допомогою цієї мови програмування відображаємо вхідні дані від користувача та відеопотік для подальшого оброблення.

TensorFlow зі свого боку є однією з найпопулярніших відкритих бібліотек машинного та глибокого навчання. Вона надає широкий спектр інструментів для створення, тренування та розгортання різноманітних типів нейронних мереж. З використання бібліотеки було реалізовано згорткову нейронну мережу, здатну аналізувати та класифікувати емоційні стани на основі вхідних зображень облич.

Взаємодія між цими двома інструментами полягає в тому, що відеопотік обробляється за допомогою JavaScript, а потім передається до моделі TensorFlow для аналізу емоцій. Результати аналізу можуть бути відтворені на вебсторінці за допомогою JavaScript для подальшого відображення користувачеві. Така архітектура дає змогу створювати ефективні та інтерактивні інтерфейси для розпізнавання емоцій у реальному часі.

Алгоритм написання програми

Для написання програми було вибрано згорткову нейронну мережу — це такий тип нейронної мережі, яка спеціалізується на обробленні зображень та виявленні просторових патернів у вхідних даних. Вона використовує фільтри (які також називаються ядрами) для згортання зображення та отримання карт ознак, що дає можливість виявляти різноманітні візуальні особливості, зокрема контури, краї та текстури.

Основні компоненти алгоритму згорткової нейронної мережі охоплюють такі етапи:

- **згортка (Convolution)**. Це перший крок у згортковій нейронній мережі. Згортання здійснюється за допомогою фільтрів, які ковзають по вхідному зображенню, виконуючи операцію згортання для створення функцій карт ознак. Ці функції карт ознак виокремлюють різні візуальні особливості, такі як краї, форми та текстури;

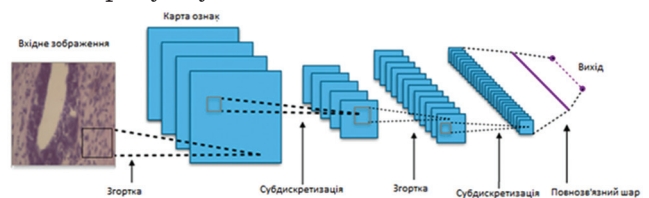
- **неалінійність (Activation)**. Після згортки застосовується неалінійна функція активації ReLU (Rectified Linear Unit) для додавання нелінійності до моделі;

- **пулінг (Pooling)**. Далі виконується операція пулінгу для зменшення розмірності карт ознак і зменшення обсягу обчислень. Це допомагає знизити кількість параметрів і уникнути перенавчання моделі;

- **повторення шарів**. Згортка, неалінійність і пулінг можуть повторюватися кілька разів для отримання більш складних функцій карт ознак та виявлення вищих рівнів абстракції;

- **повністю зв'язані шари (Fully Connected Layers)**. Після кількох повторень згортки та пулінгу дані проходять через повністю зв'язані шари для кінцевої класифікації або прогнозування.

Алгоритм згорткової нейронної мережі зображено на рисунку.



Класифікація зображень за допомогою згорткової нейронної мережі

Субдискретизація — це процес зменшення розмірності сигналу або зображення через видалення деякої інформації або зведення даних до меншої кількості точок чи пікселів. Це також називають децимацією сигналів.

У контексті зображень субдискретизація часто використовується в процесі пулінгу в згорткових нейронних мережах. Під час пулінгу розмір ділянки зменшується завдяки об'єднанню інформації з деякого регіону зображення (наприклад, у разі максимального пулінгу вибираються лише максимальні значення пікселів у кожному регіоні). Це дає змогу зменшити розмірність зображення і скоротити кількість параметрів, що зменшує обчислювальні витрати і ризик перенавчання.

Після тренування нейронної мережі вона стає здатною класифікувати емоції або характеристики на зображеннях із високою точністю. Процес тренування включає в себе ваги оптимізації фільтрів у кожному шарі так, щоб мережа була здатна

виявляти та класифікувати емоції на обличчях людей у реальному часі з високою точністю.

Основну формулу для згорткової операції можна подати в такому вигляді:

$S(i,j) = (I \cdot K)(i,j) = \sum_m \sum_n I(i+m, j+n) \cdot K(m,n)$
де $S(i,j)$ — значення пікселя у вихідному зображенні; I — вхідне зображення; K — ядро або фільтр; (i,j) — координати пікселя на вихідному зображенні; (m,n) — координати елементів ядра.

Згорткові шари можуть використовуватися для виявлення різних функцій та особливостей на зображенні, які потім подаються наступним шаром мережі для подальшого аналізу та класифікації. Це дає можливість згортковим нейронним мережам ефективно працювати з великими обсягами зображень та виявляти складні патерни в даних із високою точністю.

Отже, програма для розпізнавання емоцій працює в такий спосіб: коли людина вмикає камеру, програма починає оброблення вхідного відеопотоку за допомогою згорткової нейронної мережі. Ця мережа аналізує кожен кадр відео, виявляючи обличчя людей та визначаючи емоційні вирази на цих обличчях. Вона встановлює різні ключові елементи, які допомагають визначити емоційний стан, зокрема розташування очей, рота та їхні вирази.

На окремій панелі, яка відображається для користувача, відтворюється інформація про виявлені емоції. Ця інформація може бути подана у вигляді тексту або графічних піктограм, що відбивають найвиразніші емоційні стани, виявлені на обличчях.

У майбутньому, для подальшого вдосконалення цієї програми, можна розглянути такі напрямки оптимізації. Застосування більш ефективних алгоритмів навчання, зокрема адаптивні методи оптимізації чи методи зменшення розмірності даних, може допомогти підвищити швидкість та точність моделі. Також можна використати більш складні архітектури нейронних мереж, зокрема глибокі згорткові нейронні мережі зі зворотним зв'язком, що сприятиме покращенню здатності програми розпізнавати складніші шаблони та вдосконалювати результати. Використання спеціалізованих апаратних рішень, таких як графічні прискорювачі або тензорні прискорювачі, в нашому разі значно поліпшить швидкодію програми та знизить час оброблення даних.

ВИСНОВКИ

У статті було запропоновано розроблення програми, яка використовує згорткові нейронні мережі для розпізнавання емоцій на обличчях людей у реальному часі. Використання JavaScript та TensorFlow дало змогу створити потужний та ефективний інструмент, здатний точно класифікувати широкий спектр емоцій із високою швидкістю та точністю.

Застосування цієї технології може мати значущий вплив на багато сфер, включно з соціальними дослідженнями, медициною, медіа, рекламою тощо. Відомості про емоції можуть стати важливим інструментом для поліпшення взаємодії між людьми та технологіями, сприяючи покращенню загального добробуту та розвитку суспільства.

Список використаної літератури

1. **Основи згорткових нейронних мереж для аналізу зображень** [Електронний ресурс]. URL:

<https://towardsdatascience.com/basics-of-image-classification-with-convolutional-neural-networks-cnn-cbeba760abaе>

2. **Як працюють алгоритми глибокого навчання?** [Електронний ресурс]. URL:

<https://www.digitalocean.com/community/tutorials/how-do-deep-learning-algorithms-work>

3. **Глибокі нейронні мережі для вирішення завдань розпізнавання і класифікації зображення?** [Електронний ресурс]. URL:

<https://itcm.comp-sc.if.ua/2017/Sineglazov.pdf>

4. **Abrahamsen T. Hands-On Computer Vision with TensorFlow 2** [Електронний ресурс]. URL:

<https://www.packtpub.com/product/hands-on-computer-vision-with-tensorflow-2/9781838827069>

5. **D'Mello S. K., Kory J., Smtih B. A Review of Facial Expression Recognition in Learning Environments** [Електронний ресурс]. URL:

<https://dl.acm.org/doi/abs/10.1145/2556288.2557114>

6. **Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks** [Електронний ресурс]. URL:

<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>

O. G. Varfolomeieva, S. I. Otrakh, K. D. Sushylnykov, M. O. Shalyhin

RECOGNITION OF HUMAN EMOTIONS IN REAL TIME

This article emphasizes the importance and relevance of using neural networks to recognize human emotions in real time. It explores the possibilities of using convolutional neural networks for efficient analysis and classification of emotions based on facial expressions. The article describes the process of developing an application using JavaScript and TensorFlow and highlights potential optimization opportunities to improve the speed and accuracy of emotion recognition. The solution proposed in this paper offers a highly effective approach to emotion recognition, applicable in a variety of fields ranging from medicine and social research to marketing and entertain-

ment. This highlights the importance of advanced technologies that improve human-computer interaction and reflects the importance of using innovative methods to improve the quality of human life. The convolutional neural network (CNN) algorithm used in this study uses a deep learning approach for facial expression recognition. The architecture of the CNN model includes several convolutional layers, each of which is responsible for identifying specific features of face images, such as edges, textures, and patterns. These layers are followed by pooling layers to reduce dimensionality, allowing the network to focus on the most meaningful features. In addition, the use of rectified linear units (ReLU) as activation functions improves the model's ability to capture complex relationships in the data. In addition, the paper examines a deep learning facial expression recognition (FER) model based on convolutional neural networks, demonstrating its effectiveness in accurately identifying and classifying different emotional states based on facial features. The robust architecture of the FER model and the learning process allow it to distinguish the nuances of facial expressions, contributing to a more detailed understanding of human emotions. Consequently, such a system could be useful for medicine, social research, marketing, and entertainment, helping to analyze emotional responses and improve user interaction strategies.

Keywords: neural network; convolution; JavaScript; TensorFlow.

