

СИСТЕМА РОЗПІЗНАВАННЯ ЕМОЦІЙ ЛЮДИНИ З ВИКОРИСТАННЯМ АЛГОРИТМІВ ГЛИБОКОГО НАВЧАННЯ

Ювченко К.С.

Науковий керівник – канд. техн. наук., доц. Єсілевський В.С.
Харківський національний університет радіоелектроніки, каф. ПМ
м. Харків, Україна

тел. (050) 042-73-80, email: kateryna.iuvchenko@nure.ua

In our research, we introduce and implement a flexible framework for constructing convolutional neural networks (CNNs) that are capable of real-time processing. We validate the effectiveness of our models by creating a real-time vision system that combines face detection and emotion classification tasks using our proposed CNN architecture. We provide details of the training procedure setup and evaluate our models on standard benchmark datasets. Our results show an accuracy of 66% on the FER-2013 emotion dataset.

Емоційний стан демонструє внутрішній стан людини та її ставлення до навколишнього світу. Технології розпізнавання емоцій використовуються у різних галузях, таких як покращення обслуговування клієнтів, співбесіди з кандидатами та оптимізація реклами.

Мета цієї роботи – розробка системи, що визначає емоційний стан людей на відео на основі міміки.

Було запропоновано створити структуру згорткової нейронної мережі (CNN), що працює в режимі реального часу, виконує завдання розпізнавання обличчя та класифікації емоцій одночасно. Модель буде оцінена на тестовому наборі даних, демонструючи точність 66% на наборі даних FER-2013, що включає 7 класів емоцій.

Платформи роботів, що вирішують подібні завдання, потребують надійних та ефективних систем аналізу особи. У цій статті пропонуємо реалізацію загальної структури CNN для використання в реальному часі. Сучасні методи в області зображень, такі як класифікація та виявлення об'єктів, засновані на згорткових нейронних мережах з мільйонами параметрів, що ускладнює їх розгортання на платформах роботів та системах реального часу.

CNN – це версії багатoshарових персептронів з регуляризацією. Багатoshарові персептрони зазвичай означають мережі з повною зв'язністю, де кожен нейрон з'єднаний з усіма наступними нейронами. Проте ця повна зв'язність може призводити до перенавчання. CNN використовують інший підхід до регуляризації, використовуючи ієрархічний шаблон в даних та збираючи шаблони зростання складності з менших та простіших візерунків у їх фільтрах. Таким чином, CNN забезпечують меншу зв'язність та складність, порівняно з іншими моделями.

Було запропоновано дві моделі, які були оцінені з точки зору їх точ-

ності відносно кількості параметрів. Моделі були розроблені з метою досягнення найкращої точності при мінімальній кількості параметрів. Зменшення кількості параметрів допомагає нам подолати проблеми, такі як обмеження обчислювальних ресурсів, наприклад на робототехнічних платформах.

Наша перша модель базується на ідеї видалення повністю пов'язаних шарів, тоді як інша модель комбінує видалення повністю пов'язаних шарів з роздільними згортками в глибину та залишковими модулями. Обидві архітектури були навчені за допомогою оптимізатора ADAM [1].

Результати наших експериментів показали, що обидві запропоновані моделі досягають високої точності при мінімальній кількості параметрів. Модель, яка базується на видаленні повністю пов'язаних шарів, демонструє високу швидкість розпізнавання та ефективне використання ресурсів обчислювальної системи.

Друга модель, яка комбінує видалення повністю пов'язаних шарів з роздільними згортками в глибину та залишковими модулями, показала високу точність в розпізнаванні обличчя, зокрема в умовах змінної освітленості та різних ракурсів. Це може бути корисно для розпізнавання обличчя в реальних ситуаціях, таких як системи безпеки, автоматичне розпізнавання особи на вході в приміщення, або в системах розпізнавання емоцій та виразів обличчя в режимі реального часу.

Ми запропонували та випробували власну архітектуру CNN, яка може працювати в режимі реального часу. Архітектури були розроблені з урахуванням зменшення кількості параметрів. Було продемонстровано, що моделі можуть бути успішно використані для класифікації багатокласових задач навіть в режимі реального часу. Конкретно, було розроблено комп'ютерну систему зору, яка здійснює розпізнавання обличчя та класифікацію емоцій в одному інтегрованому модулі.

Показники результатів збігаються з показниками людського рівня у наших задачах класифікації за допомогою єдиної CNN, яка використовує сучасну архітектуру.

Список використаних джерел:

1. Diederik Kingma, Jimmy Ba. Adam. (2014). *A method for stochastic optimization*. <https://arxiv.org/pdf/1412.6980.pdf>.
2. Francois Chollet. (2016). *Xception: Deep learning with depthwise separable convolutions*.
3. Valueva M. V., Nagornov N. N., Lyakhov P. A., Valuev G. V., Chervyakov N. I. (2020). *Application of the residue number system to reduce hardware costs of the convolutional neural network implementation*.