

ВИБІР АРХІТЕКТУРИ ГЛИБИННОГО НАВЧАННЯ ДЛЯ СИСТЕМИ РОЗПІЗНАВАННЯ ЖЕСТИВ РУК

Сіряк Р.В., Скарга-Бандурова І.С.

*Східноукраїнський національний університет ім. В. Даля,
м. Сєвєродонецьк*

За останні роки, завдяки розвитку обчислювальних технологій і появі нових алгоритмів в області штучних нейронних мереж, з'явилися можливості для реалізації раніше важкоздійснюваного розпізнавання жестів.

Технологія розпізнавання жестів складається з кількох етапів: сегментація руки, витяг ознак з отриманого регіону і визначення того, до якого класу належить жест. Залежно від підходів і цілей ці завдання можуть вирішуватися різними методами. На даний момент не існує універсального підходу, який працює з високою швидкістю і точністю розпізнавання при будь-яких умовах. Візуальні засоби розпізнавання, що використовують відеокамери дають високу точність розпізнавання, проте сильно залежать від зміни освітлення. Застосування кольорових міток допомагає усунути цю проблему, але даний підхід не є природним і достатньо незручний в повсякденному житті, особливо, якщо йде мова про спеціальні додатки конвертації жестів у звуки. Використання спеціально обладнаних рукавичок допомагає з високою точністю визначити ознаки руки, але також є незручним засобом взаємодії з комп'ютером, і, до того ж, досить дорогим. Тривимірні сенсори, будучи абсолютно незалежними від зміни освітлення, дозволяють легко знаходити ключові ознаки руки, але при цьому дають досить низькі показники розпізнавання.

Метою роботи є створення основи для безконтактного інтерфейсу шляхом розпізнавання жестів рук на основі нейронної мережі. Робота виконується з застосуванням Jupiter Notebook, що є частиною пакету Anaconda для програмування на Python.

До завдань проекту, для досягнення поставлених цілей, входить дослідження існуючих методів відстеження та розпізнавання жестів руки, і на основі їх аналізу, розробка вдосконалених або нових методів. На даний момент розроблена працююча модель згорткової нейронної мережі, здатної до розпізнавання жестів руки, що в подальшому може послужити основою для побудови повноцінної системи людино-машинного взаємодії.

В доповіді надано результати аналізу архітектур глибокого навчання, в якості моделі для реалізації нейронної мережі обрано згорткову нейронну мережу (Convolutional Neural Network) з чотирнадцятьма шарів. У трьох згорткових шарах з функцією активації ReLU застосовано відповідно 16, 32 і 64 згортальних ядра з встановленим режимом valid. За кожним згортковим шаром слідує шар MaxPooling. Далі, дані в шарі flatten перетворюються з 2D в одновимірний вектор, і в кінці проходять через два повнозв'язних шари dense. Для роботи мережі не потрібно ніякого додаткового обладнання, крім веб-камери. Наразі модель здатна до розпізнавання трьох жестів в режимі реального часу.