

ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ В ЗАДАЧАХ РОЗПІЗНАВАННЯ ОБЛИЧ ТА КЛАСИФІКАЦІЇ ЕМОЦІЙ

Чугунов В.Ю., Поворознюк О.А.

Національний технічний університет

«Харківський політехнічний інститут», м. Харків

Методи машинного навчання активно застосовуються в задачах комп'ютерного зору, зокрема в розпізнаванні облич та класифікації емоцій [1]. Найпоширенішими підходами до розпізнавання емоцій є:

1. Класичні методи з використанням ознак: такі як HOG (Histogram of Oriented Gradients), LBP (Local Binary Patterns), SVM (Support Vector Machine), які використовуються для виділення ознак з обличчя та подальшої класифікації[2].

2. Глибоке навчання: моделі на основі згорткових нейронних мереж (CNN), які здатні автоматично навчатися релевантних ознак з зображення без ручної інженерії ознак[3].

3. Transfer learning: застосування попередньо натренованих моделей (наприклад, VGGFace, ResNet, або MobileNet), які були навчені на великих обличчєвих датасетах, і можуть бути донавчені на задачі розпізнавання емоцій.

4. До основних відкритих датасетів, що використовуються для задачі розпізнавання емоцій, належать FER2013, AffectNet, RAF-DB. Ці набори містять тисячі зображень облич, маркованих відповідно до основних емоційних категорій: радість, сум, страх, злість, здивування, відраза, нейтральний стан.

Для реалізації емоційного тесту в межах мобільного застосунку було вирішено використати модель згорткової нейронної мережі (CNN), зокрема на основі MobileNetV2, оскільки ця архітектура є оптимальною за швидкістю та розміром для розгортання на мобільних пристроях. MobileNetV2 дозволяє досягти балансу між точністю класифікації та ресурсами пристрою. Модель буде донавчатися на датасеті FER2013, який є відкритим та добре відомим у сфері емоційного розпізнавання. Кінцевим результатом стане модель, яка зможе класифікувати емоційний стан користувача за фото (наприклад, "радість", "сум", "нейтральність") у реальному часі під час проходження тесту. Інші методи, зокрема класичні SVM або LBP+SVM, були відхилені через нижчу точність та недостатню адаптивність до складних варіацій зображень в реальних умовах (освітлення, ракурси, тощо).

Література

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
2. Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *IEEE Transactions on Affective Computing*, 10(1), 18–31.
3. Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.