

СТВОРЕННЯ ТА ПІДГОТОВКА ДАНИХ ДЛЯ РОЗПІЗНАВАННЯ БІОМЕТРИЧНИХ ПАРАМЕТРІВ ЛЮДИНИ ЗА ДОПОМОГОЮ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ

Вінницький національний технічний університет

Анотація

Запропоновано підходи до розпізнавання біометричних параметрів людини, а саме різних емоцій, віку й статі. Велику увагу приділено автоматизації збору даних для тренування алгоритму розпізнавання емоцій та лиць, виборі архітектури використаних алгоритмів машинного навчання, процедурі тренування та оптимізації моделей й імплементації бізнес логіки додатку. Більший акцент зосереджено на створенні додатку для персональних комп'ютерів.

Ключові слова: розпізнавання біометричних параметрів людини, тренування розпізнавання емоцій, машинне навчання.

Abstract

The recognition of human biometric parameters, namely different emotions, age and gender, are proposed in this paper. Much attention is paid to the automated data collection for training the algorithm of recognizing emotions and faces, choosing the machine learning algorithms architecture, procedures for preparing and optimizing models and implementing the application business logic. Main focus is set on creation of application for personal computers.

Keywords: recognition of human biometric parameters, emotion recognition training, machine learning.

Вступ

У наш час розвиток алгоритмів машинного навчання [1] відкриває безліч можливостей для створення додатків, що будуть оптимізувати певні процеси. Такий швидкий розвиток сфери штучного інтелекту пов'язаний з двома факторами:

1) Зростання кількості даних, таких як текстові дані, зображення, відео й дані про активність користувачів, а також покращення якості цих даних.

2) Покращення методів машинного навчання, оптимізація архітектур й алгоритмів.

У роботі описано повний життєвий цикл проекту з машинного навчання, починаючи зі збору даних й тренування моделей та закінчуючи оптимізацією моделей, створенням додаткової логіки й самого додатку.

Результати дослідження

Суть проекту - розпізнавання біометричних параметрів людини у реальному часі. Аби досягти цієї мети, додаток повинен розпізнавати біометричні параметри з певного відеопотоку, що представляє собою окремі зображення. Для того, щоб розпізнати емоцію, вік та стать людини, спочатку потрібно вирішити проблему локалізації людського обличчя. Таким чином було використано наступний підхід:

1) знаходження області на зображенні, що відповідає людському лицю й її вирізання для подальшого використання;

2) використання моделей для розпізнавання емоцій, статі й віку на даних з 1-го пункту.

Оскільки виникла потреба розпізнавати не лише лица людей, а й їх тіла, з'явилася необхідність створення нового датасету. У якості початкового було використано WIDER FACE датасет [2], що містить анотації лиць людей (їх локацію на зображеннях). Так як датасет мав велику кількість зображень, а його анотації були у форматі, що не підходив для тренування обраних моделей, датасет було проанатовано з нуля. Для автоматизації цього процесу було використано моделі для розпізнавання об'єктів із сімейства Yolo. Yolo - аббревіатура назви You Look Only Once, що є нейронною мережею [3] для розпізнавання об'єктів й потребує лише одного проходу через усі з'єднання аби розпізнати й

локалізувати всі об'єкти. Для створення анотацій лиць було використано YoloV3 натреновану на WIDER FACE даних та для анотації людських тіл - YoloV3 натреновану на COCO [4] даних (використовуючи лише клас людини у якості вихідних даних з моделі) (рис. 1).

В результаті було отримано анотації у форматі tf-record, що були придатними для використання.

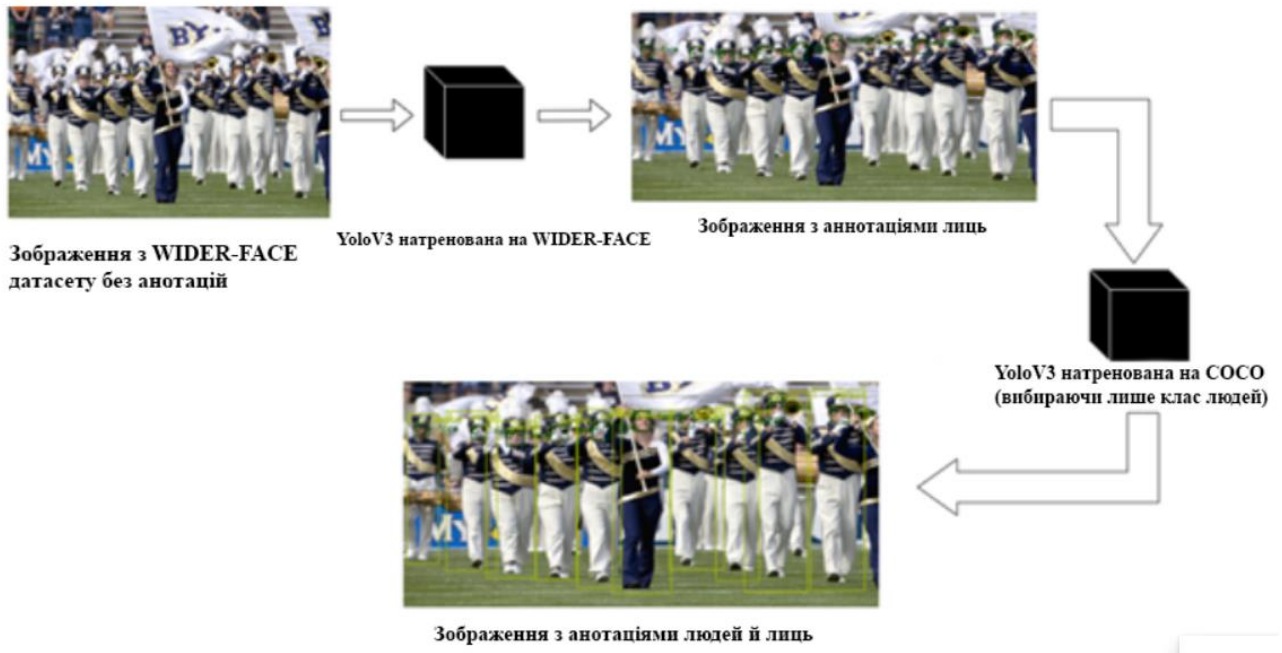


Рис. 1. Автоматизація лейблінгу даних для розпізнавання людей й облич

У якості даних для розпізнавання віку й статі було використано IMDB[5] датасет. Так як датасет з емоціями повинен був мати нову емоцію “рокетон” (ахегао), потрібно було знайти відповідні зображення й проанатувати їх. Від самого початку ідея було у тому щоб знайти “рокетон” зображення, перевести їх у чорно-білий колір й додати до FER датасету, проте зображення з FER датасету мали зовсім інший розподіл, що відрізнявся від нових зображень. Це призвело до проблеми, відомої як covariate shift [6]. Зважаючи на низьку якість зображень у FER датасеті й описану проблему, було вирішено зібрати новий датасет RGB зображень. Датасет було зібрано за рахунок парсингу інстаграму, відео у ютубі й інших джерел. Деякі зображення було узяті з IMDB та AffectNet [7] датасетів. Кожне зображення у підсумковому датасеті - обличчя людини з відповідною емоцією, що було вирізано з оригінального зображення, використовуючи YoloV3 натреновану на WIDER FACE. У результаті було отримано датасет з 6 емоцій (рис. 2).

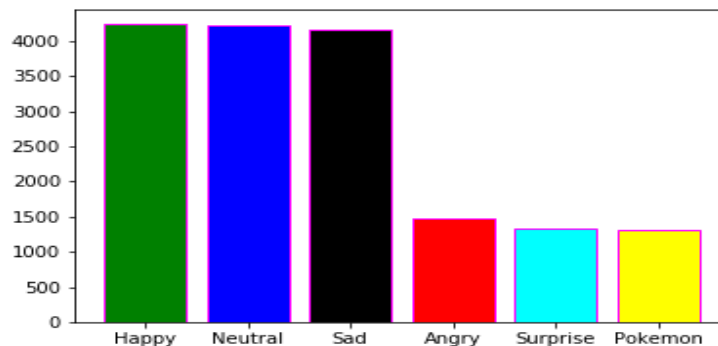


Рис. 2. Розподіл зображень у датасеті щодо кожної емоції

Висновки

У роботі було розроблено додаток на персональні комп'ютери для розпізнавання біометричних параметрів людини й мобільний додаток для розпізнавання емоцій й трекінгу людських тіл.

Практична цінність роботи полягає у описі життєвого циклу data-science продукту, використаних методів поліпшення тренування моделей та їх подальшої оптимізації й створенні кінцевого продукту.

Важливою частиною роботи є збір даних й способи оптимізації цього процесу, що зазвичай забирало велику кількість часу команди.

Основною частиною роботи є розробка й тренування нейронних мереж, використаних у якості основних компонентів додатку.

Даний додаток має велику кількість сторін для подальшого розвитку. Наприклад однією з можливих ідей є інтеграція персоналізованих графіків, щодо розподілу емоцій користувача.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Машинне навчання [Електронний ресурс]: Режим доступу: https://en.wikipedia.org/wiki/Machine_learning - Назва з екрану.
2. Shuo Yang et al; WIDER FACE: A Face Detection Benchmark [Електронний ресурс]: Режим доступу: <https://arxiv.org/pdf/1511.06523.pdf> - Назва з екрану.
3. Joseph Redmon, Ali Farhadi; YOLOv3: An Incremental Improvement [Електронний ресурс]: Режим доступу: <https://arxiv.org/pdf/1804.02767.pdf> - Назва з екрану.
4. Tsung-Yi Lin et al; Microsoft COCO: Common Objects in Context [Електронний ресурс]: Режим доступу: <https://arxiv.org/pdf/1405.0312.pdf> - Назва з екрану.
5. Rasmus Rothe and Radu Timofte and Luc Van Gool; Deep expectation of real and apparent age from a single image without facial landmarks [Електронний ресурс]: Режим доступу: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/> - Назва з екрану.
6. Covariate shift [Електронний ресурс]: Режим доступу: <https://www.analyticsvidhya.com/blog/2017/07/covariate-shift-the-hidden-problem-of-real-world-data-science/> - Назва з екрану.
7. Ali Mollahosseini et al; AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild [Електронний ресурс]: Режим доступу: <https://arxiv.org/pdf/1708.03985.pdf> - Назва з екрану.

Ковенко Володимир Андрійович, студент групи ІІСТ-18б, Факультет комп'ютерних систем та автоматизації, Вінницький національний технічний університет, urumipainblackreaper@gmail.com.

Богач Ілона Віталіївна, кандидат технічних наук, доцент кафедри автоматизації та інтелектуальних інформаційних технологій, Вінницький національний технічний університет, ilona.bogach@gmail.com.

Kovenko Volodymyr Andriyovich, the student of group IIIST-18b, the faculty of computer systems and automation, Vinnytsia National Technical University, urumipainblackreaper@gmail.com.

Bogach Ilona Vitaliyevna, PhD, Associate Professor of the department of automation and intelligent information technologies, Vinnytsia National Technical University, ilona.bogach@gmail.com.