

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені І.І.МЕЧНИКОВА

(повне найменування вищого навчального закладу)

Факультет математики, фізики та інформаційних технологій

(повне найменування інституту, назва факультету (відділення))

Кафедра математичного забезпечення комп'ютерних систем

(повна назва кафедри (предметної, циклової комісії))

Кваліфікаційна робота

на здобуття ступеня вищої освіти «магістр»

(освітньо-кваліфікаційний рівень)

на тему «Інформаційна технологія пошуку об'єктів відеопослідовності.

Пошукова модель об'єкта відеопослідовності»

«Information technology of searching objects in video sequence.

Video sequence object search model»

Виконав: студент денної форми навчання

напряму підготовки 126 – Інформаційні системи та технології

(шифр і назва напряму підготовки, спеціальності)

Освітня програма «Інформаційні системи та технології»

(назва освітньої програми)

Шлемко Олександр Віталійович

(прізвище, ім'я, по-батькові)

Керівник д-р тех.н., проф. Малахов Є. В

(науковий ступінь, вчене звання, прізвище та ініціали, підпис)

Рецензент канд. фіз.-матем. н., доц. Петрушина Т.І.

(науковий ступінь, вчене звання, прізвище та ініціали)

Рецензент канд. тех. н., доц. Глава М.Г.

(науковий ступінь, вчене звання, прізвище та ініціали)

Рекомендовано до захисту:

Протокол засідання кафедри

№ __ від «__» ____ 2022 р.

Завідувач кафедри

Євгеній МАЛАХОВ

(підпис)

(прізвище, ініціали)

Захищено на засіданні ЕК № ____

протокол № __ від «__» ____ 2022 р.

Оцінка ____ / ____ / ____

(за національною шкалою, шкалою ECTS, бали)

Голова ЕК

Володимир ВИЧУЖАНІН

(підпис)

(прізвище, ініціали)

Одеса – 2022

АНОТАЦІЯ

У дипломній роботі розроблено систему пошуку та класифікації облич у відеопослідовності з використанням методів машинного навчання.

Сьогодні задачі розпізнавання та класифікації облич – один з найпріоритетніших напрямків у сфері машинного навчання. Компанії із сфер автомобілебудування, розробки програмного забезпечення, індустрії безпеки та електронної комерції, застосовують ці технології для автоматизації процесів, побудови аналітики та покращення надаваних послуг, тим самим збільшуючи свої прибутки.

Мета даного проекту — підвищення ефективності пошуку та класифікації облич у відеопослідовності, шляхом розробки моделі класифікації для вилучених кадрів з відеоряду.

Об'єкт дослідження — технологічний процес класифікації та пошуку облич у відеопослідовності.

Предметом дослідження є методи та алгоритми пошуку та класифікації облич у відеопослідовності

Модель складається з модуля отримання даних, модуля препроцесингу даних, модуля класифікації об'єктів та модуля збереження зображень.

Модулі препроцесингу та класифікації примітні тим, що використана у них згортова нейронна мережа родини FaceNet не займається порівнянням наданих зображень з вже відомими, а генерує 128 метрик обличчя. Ці метрики після відпрацювання нейронної мережі передаються до класифікатора на основі SVM. Також результати зберігаються до бази даних, що створює додаткові можливості для аналізу та пошуку, наприклад, за іменами розпізнаних людей, часом появи на відео, тощо.

Розроблена модель має кілька важливих переваг. По-перше, завдяки використанню метрик обличчя вдалося зменшити час, необхідний для класифікації окремого зображення. Так, використання окремого статистичного дозволило знизити витрати часу на обробку одного кадру на 20мс. Також шляхом модифікації обраних методів класифікації вдалося підвищити точність класифікації облич на зображеннях високої якості на 1-2%.

ABSTRACT

In this thesis a system for searching and classifying faces in video sequences using machine learning methods is developed.

Today, the tasks of face recognition and classification are one of the most important directions in the field of machine learning. Companies from the automotive, software development, security and e-commerce industries are using these technologies to automate processes, build analytics and improve services, thereby increasing their profits.

The purpose of this project is to increase the efficiency of face search and classification in video sequences by developing a classification model for extracted frames from the video sequence.

The object of the study is the technological process of classifying and searching for faces in a video sequence.

The subject of research is methods and algorithms for searching and classifying faces in a video sequence

The model consists of a data acquisition module, a data preprocessing module, an object classification module, and an image storage module.

Preprocessing and classification modules are notable for the fact that the convolutional neural network of the FaceNet family used in them does not compare the provided images with already known ones, but generates 128 face metrics. These metrics, after training the neural network, are transferred to the SVM-based classifier. Also, the results are saved to the database, which creates additional opportunities for analysis and search, for example, by the names of recognized people, time of appearance in the video, etc.

The developed model has several important advantages. First, thanks to the use of face metrics, it was possible to reduce the time required to classify a single image. Thus, the use of a separate statistic classifier made it possible to reduce the time spent on processing one frame by 20ms. Also, by modifying the selected classification methods, the accuracy of face classification on high-quality images increase by 1-2%.

ЗМІСТ

ВСТУП	5
1 МЕТОДИ КЛАСИФІКАЦІЇ ОБ’ЄКТІВ ВІДЕОПОСЛІДОВНОСТІ. 6	
1.1 Класифікація з вчителем	7
1.2 Класифікація без вчителя	7
1.3 Згорткові нейронні мережі	8
1.4 Метод опорних векторів	9
1.5 Метод К найближчих сусідів	10
1.6 Дерева рішень	11
1.7 Порівняння розглянутих методів. Висновки	13
2 ПРОЕКТУВАННЯ СИСТЕМИ КЛАСИФІКАЦІЇ ОБ’ЄКТІВ	15
2.1 Архітектура системи	15
2.2 Модуль отримання даних	16
2.3 Модуль препроцесингу даних	16
2.4 Модуль класифікації об’єктів	23
2.5 Модуль збереження класифікованих зображень	25
3 РЕАЛІЗАЦІЯ СИСТЕМИ ВИЯВЛЕННЯ ОБЛИЧ	26
3.1 Набір даних для навчання нейронної мережі	26
3.2 Модуль препроцесингу зображень	29
3.3 Модуль класифікації зображень	30
3.4 Модуль збереження класифікованих зображень	32
3.5 Тестування моделі	33
4 АНАЛІЗ ОТРИМАНИХ РЕЗУЛЬТАТІВ	36
ВИСНОВКИ	37
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	38

ВСТУП

Широке розповсюдження смартфонів, здешевлення камер і поява більш точних методів та алгоритмів відкрили нову еру для розпізнавання та класифікації зображень. Компанії із сфер, таких як автомобілебудування, розробка програмного забезпечення, індустрія безпеки та електронна комерція, застосовують ці технології для автоматизації процесів, побудови аналітики та покращення надаваних послуг, тим самим збільшуючи свої прибутки.

Розпізнавання зображень дозволяє витягувати значущі дані із зображення, і тому має багато прикладних застосувань. Однак точність роботи таких систем кратно покращилася протягом останніх 10 років завдяки появі підходів на основі глибокого навчання, що зробило задачі розпізнавання і класифікації зображень актуальними.

Зменшення розміру та вартості камери в поєднанні зі збільшенням поширення смартфонів і соціальних медіа на основі зображень зумовили ріст кількості можливих прикладних використань класифікації зображень, наприклад, у відеоспостереженні. Використання цієї технології має великий потенціал у значній кількості місць – в аеропортах, стадіонах, вокзалах, тощо для попередження небезпечних ситуацій. Розпізнавання обличчя можна застосовувати майже скрізь, включаючи навіть домашню систему безпеки.

Об'єктом дослідження є технологічний процес класифікації та пошуку облич у відеопослідовності, предметом, відповідно — методи та алгоритми пошуку та класифікації облич у відеопослідовності.

Мета даного дослідження — підвищення ефективності пошуку та класифікації облич у відеопослідовності, шляхом розробки моделі класифікації для вилучених кадрів з відеоряду.

ВИСНОВКИ

Під час виконання даної роботи спроектована та реалізована система, що знаходить та класифікує об'єкти у відеопослідовності. Для цього був проведений аналіз існуючих методів класифікації та пошуку об'єктів у відеопослідовності, та виявлені недоліки, які можуть бути виправлені в рамках даної роботи.

Модель складається з модуля отримання даних, модуля препроцесингу даних, модуля класифікації об'єктів та модуля збереження зображень.

Модулі препроцесингу та класифікації примітні тим, що використана у них згортова нейронна мережа родини FaceNet не займається порівнянням наданих зображень з вже відомими, а генерує 128 метрик обличчя. Ці метрики після відпрацювання нейронної мережі передаються до класифікатора на основі SVM. В той самий час традиційні підходи до класифікації обличчя потребують більше вхідних даних та є більш громіздкими. Також результати зберігаються до бази даних, що створює додаткові можливості для аналізу та пошуку, наприклад, за іменами розпізнаних людей, часом появи на відео, тощо.

Розроблена модель має кілька важливих переваг. По-перше, завдяки використанню метрик обличчя вдалося зменшити час, необхідний для класифікації окремого зображення. Так, використання окремого статистичного класифікатора який помітно пришвидшує процес зіставлення із вже відомими зразками обличчя дозволило знизити витрати часу на обробку одного кадру на 20мс. Також шляхом модифікації обраних методів класифікації вдалося підвищити точність класифікації обличчя на зображеннях високої якості на 1-2%.

Майбутній розвиток даної роботи полягає у розширенні та покращенні аналітичних методів пошуку зображень, додавання можливості роботи у реальному часі, а також можливості створення інтерфейсу користувача з системою нотифікацій, що дозволить спростити відстеження появи нових об'єктів.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Liu, Qiong & Wu, Ying. (2012). Supervised Learning. 10.1007/978-1-4419-1428-6_451.
2. Dridi, Salim. (2021). Unsupervised Learning - A Systematic Literature Review. 10.13140/RG.2.2.16963.12323.
3. Zhuang Liu, Hanzi Mao, Chao-yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie. A ConvNet for the 2020s.
4. Lam M. Nguyen, Yilan Chen, Wei Huang, Tsui-Wei Weng. On the Equivalence between Neural Network and Support Vector Machine. 2021. p 7-10.
5. Srivastava, Durgesh, Bhambhu, Lekha. (2010). Data classification using support vector machine. Journal of Theoretical and Applied Information Technology. 12. 1-7.
6. Guo, Gongde & Wang, Hui & Bell, David & Bi, Yaxin. (2004). KNN Model-Based Approach in Classification.
7. Rokach, Lior & Maimon, Oded. (2005). Decision Trees. 10.1007/0-387-25465-X_9.
8. Liu, Wenfei & Wei, Jingcheng & Meng, Qingmin. (2020). Comparisons on KNN, SVM, BP and the CNN for Handwritten Digit Recognition. 587-590.
9. Zaghdoudi, Rachid & Seridi, Hamid. (2017). Combination of Multiple Classifiers for Off-Line Handwritten Arabic Word Recognition. International Journal of Information Technology. 14. 713-720.
10. Hayder Hasan. A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Hyperspectral Image Classification IOP Conf. Ser.: Earth Environ. Sci. 2019. 357.
11. Шлемко О. В., Малахов Є. В. Інформаційна технологія пошуку об'єктів відеопослідовності. Пошукова модель об'єкта відеопослідовності. Інформатика, інформаційні системи та технології: тези доповідей

дев'ятнадцятої всеукраїнської конференції студентів і молодих науковців. Одеса, 29 квітня 2022 р. – Одеса, 2022. – с. 84-86.

12. Shlemko O., Yakushyna A. Information technologies for object recognition and classification in video sequence / Conference proceedings of the 1st Student Scientific Conference of Joint Research Cooperation between Odesa I.I. Mechnikov National University and Huaiyin Institute of Technology (March 28-29, May 16, 2022) / MOES of Ukraine; Odesa I.I. Mechnikov National University; – Dnipro: Serednyak TK, 2022. – p.136-139.
13. Adam Geitgey. Modern Face Recognition with Deep Learning. 2017.
14. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005, pp. 886-893
15. Ouanan, Hamid & Ouanan, Mohammed & Aksasse, B.. Facial landmark localization: Past, present and future. 2016. 487-493.
16. Kazemi, Vahid & Sullivan, Josephine. One Millisecond Face Alignment with an Ensemble of Regression Trees. 2014.
17. P. Dollar, P. Welinder, and P. Perona. Cascaded pose regression. 2010. 1078–1085.
18. M. Unser, M. A. Neimark and C. Lee, "Affine transformations of images: a least squares formulation," Proceedings of 1st International Conference on Image Processing, 1994, pp. 558-561.
19. William, Ivan & Setiadi, De Rosal Ignatius Moses & Rachmawanto, Eko & Santoso, Heru & Sari, Atika. Face Recognition using FaceNet (Survey, Performance Test, and Comparison). 2019. 1-6.
20. Schroff, Florian & Kalenichenko, Dmitry & Philbin, James. FaceNet: A unified embedding for face recognition and clustering. 2015. 815-823.
21. Mahamkali, Naveenkumar & Ayyasamy, Vadivel. OpenCV for Computer Vision Applications. 2015.

