

Nikolaievskiy O. Yu., Skliarenko O.V. *European University, Kyiv*

Sidorchuk A.I. *Odessa National O.S. Popov Academy of Telecommunications, Odessa*

ANALYSIS AND COMPARISON OF FACE DETECTION APIS

Today, more and more information is being accumulated in digital form, including media content, which is a growing segment of the Internet, searching for such content is an important task, but it is significantly different from textual information search and involves image recognition. Images recognition is an automatic comparison of images to objects in a class. There are two major issues in this area: classification and identification. The first helps the search engine to understand what type of object is in the media resource. Only correctly having solved this fundamental problem, the computer will be able to distinguish, for example, a dog from a cat. The second allows not only to find the object category, but also to identify it. The article defines the mathematical formulation of the problem, which is important in order to formalize the accuracy of recognition and to determine the procedure for comparing the capabilities of existing algorithms. The results show that existing APIs allow you to largely solve the problem of face recognition in images, and the next important step is to recognize the face of a person on the move, that is, on video content.

Keywords: *neural network, convolutional neural network, tensorflow, computer vision, deep learning, convolutional kernel, weight of the neuron, activation function, feature map, pattern recognition.*

Ніколаєвський О.Ю., Скляренко О.В. *ПВНЗ «Європейський університет», Київ*

Сидорчук А.І. *Одеська національна академія зв'язку ім. О.С. Попова, Одеса*

АНАЛІЗ І ПОРІВНЯННЯ АРІ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ

Сьогодні, в цифровому вигляді накопичується все більше інформації, зокрема і медіа-контенту, який складає зростаючий сегмент мережі Інтернет, пошук по такому контенту є важливою задачею, але він суттєво відрізняється від пошуку текстової інформації та передбачає розпізнавання образів. Розпізнавання образів – це автоматичне зіставлення образів до об'єктів класу. У цій області є дві основні проблеми: класифікація та ідентифікація. Перша допомагає пошуковій системі зрозуміти, який тип об'єкта знаходиться у медіа-ресурсі. Тільки правильно вирішивши цю фундаментальну проблему, комп'ютер зможе відрізнити, наприклад, собаку від кішки. Другий дозволяє не тільки знайти категорію об'єкта, а й ідентифікувати її. У статті визначено математичне формулювання задачі, що є важливим для того, щоб формалізувати точність розпізнавання та визначити процедуру порівняння можливостей існуючих алгоритмів. Результати показують, що існуючі АРІ дозволяють значною мірою розв'язувати завдання з розпізнавання обличчя на зображеннях, а наступним важливим кроком є розпізнавання обличчя людини в русі, тобто на відео-контенті.

Ключові слова: *нейронна мережа, згорткова нейронна мережа, tensorflow, комп'ютерний зір, глибинне навчання, згорткове ядро, вага нейрона, функція активації, карта об'єктів, розпізнавання образів.*

Николаевский А.Ю., Скляренко Е.В. *ЧВУЗ «Европейский университет», Одесса*

Сидорчук А. И. *Одесская национальная академия связи им. О.С. Попова, Одесса*

АНАЛИЗ И СРАВНЕНИЕ АРІ РАСПОЗНАВАНИЯ ЛИЦ

Сегодня, в цифровом виде накапливается все больше информации, в том числе и медиа-контента, который составляет растущий сегмент сети Интернет, поиск по такому

© Ніколаєвський О.Ю., Скляренко О.В., Сидорчук А.І. 2019

контенту является важной задачей, но он существенно отличается от поиска текстовой информации и предусматривает распознавания образов. Распознавание образов - это автоматическое сопоставление образов к объектам класса. В этой области есть две основные проблемы: классификация и идентификация. Первая помогает поисковой системе понять, какой тип объекта находится в медиа-ресурсе. Только правильно решив эту фундаментальную проблему, компьютер сможет отличить, например, собаку от кошки. Второй позволяет не только найти категорию объекта, но и идентифицировать ее. В статье определены математическую формулировку задачи, что важно для того, чтобы формализовать точность распознавания и определить процедуру сравнения возможностей существующих алгоритмов. Результаты показывают, что существующие API позволяют в значительной мере решать задачи по распознаванию лица на изображениях, а следующим важным шагом является распознавание лица человека в движении, то есть на видео-контенте.

Ключевые слова: нейронная сеть, сверточная нейронная сеть, tensorflow, компьютерное зрение, глубокое обучение, сверточное ядро, вес нейрона, функция активации, карта объектов, распознавание образов.

Introduction.

The human face plays an important role in our social interaction, conveying people's identity. Using the human face as a key to security, biometric face recognition technology has received significant attention in the past several years due to its potential for a wide variety of applications in both law enforcement and non-law enforcement.

As compared with other biometrics systems using fingerprint/palmprint and iris, face recognition has distinct advantages because of its non-contact process. Face images can be captured from a distance without touching the person being identified, and the identification does not require interacting with the person. In addition, face recognition serves the crime deterrent purpose because face images that have been recorded and archived can later help identify a person.

Thus, face recognition is a method of identifying or verifying the identity of an individual using their face. Face recognition systems can be used to identify people in photos, video, or in real-time. Law enforcement may also use mobile devices to identify people during police stops. Face recognition systems use computer algorithms to pick out specific, distinctive details about a person's face. These details, such as distance between the eyes or shape of the chin, are then converted into a mathematical representation and compared to data on other faces collected in a face recognition database. The data about a particular face is often called a face template and is distinct from a photograph because it's designed to only include certain details that can be used to distinguish one face from another.

Goal.

The goal is to show all modern analogues of the face recognition API. Analyze the weak and strong sides of them, make a final comparison and figured out the most powerfull.

Pattern recognition problem

Nowadays, all powerful analogues used neural network approach, as it showed itself as most accurate and the quickest solution. The mathematical problem statement is following: there is a set M of objects ω . The objects are defined by the values of some features: x_i , $i = 1, \dots, N$. kits of which are the same for all objects. The aggregate of the features of the object ω defines in some way its description it is shown at the 1.

$$I(\omega) = (x_1(\omega), x_2(\omega), \dots, x_N(\omega)) \quad (1)$$

Features can be expressed in terms of yes/no, yes/no/unknown, numeric values, values from the set of possible variants, etc.

On the whole set of M there is a division into subsets (object classes), that is shown in the 2.

$$M = \bigcup_{i=0}^M E_i \quad (2)$$

The division into classes can be set completely or determined by some a priori information I_0 about class E_i - like the characteristic description of the objects they contain.

The task of recognition is shown in the 3 and it is to ensure that for each object ω by its $I(\omega)$ and for prior(learning) information calculate the predicate value:

$$P_i = (\omega \in E_i), \quad i = 1, \dots, m \quad (3)$$

To describe the impossibility to recognize objects, the P_i predicates are replaced by the following values that shows in the 4:

$$a_i \in \{0(\omega \notin E_i), 1(\omega \in E_i), \Delta(\text{unknown})\} \quad (4)$$

Thus, for the object under consideration ω it is necessary to calculate the information vector that shows in the 5.

$$a(\omega) = (a_1(\omega), a_2(\omega), \dots, a_m(\omega)) \quad (5)$$

In this case, the procedure of building the information vector $a(\omega)$ expresses the algorithm of making a decision on attributing an object ω to one or another class and it is called the "decisive function".

AWS face recognition API.

The first analogue – Amazon Recognition API. It is API that provide AWS users find a faces in their applications simply and quickly. It is not free – most of AWS services are well-known optimized and are priced accordingly. The figure 1 shows searching face in the collection.



Fig.1. AWS Recognition searching process

But there is one limitation – users can process only documents which contains in AWS S3. It is container for storing document from Amazon. Also AWS Recognition Api supports dynamic recognition function. Thus it makes it possible to track movement. The figure 2 shows AWS movement tracker.

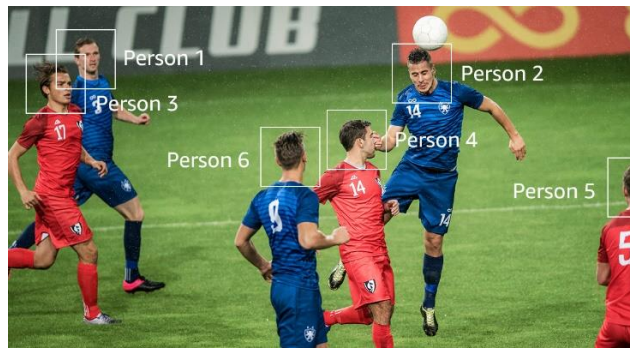


Fig.2. AWS API – movement track

Microsoft Azure Facial Recognition API.

This analogue provide next possibilities: detect and compare human faces, organize images into groups based on similarities, identify previously tagged people in images, run locally on-premises or in the cloud.

Check the likelihood that two faces belong to the same person. The API will return a confidence score about how likely it is that the two faces belong to one person.

The figure 3 shows Microsoft Azure API compare process.



Fig.3. Microsoft API – face identification

The Face API now integrates emotion recognition, returning the confidence across a set of emotions for each face in the image such as anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. These emotions are understood to be cross-culturally and universally communicated with particular facial expressions. The figure 4 shows emotion recognition process.

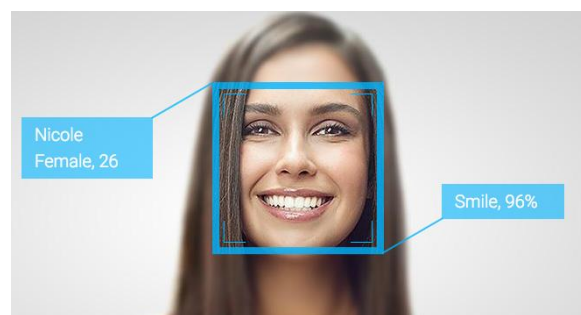


Fig. 4. Microsoft API – emotions detection

FindFace Recognition Services.

The next analogue – FINDFACE service. It is a high certified resource with a lot of extra features as dynamic recognition, person identification, verification and so on.

Unfortunately, this resource should be downloaded and paid before using. The figure 5 shows FindFace main interface.

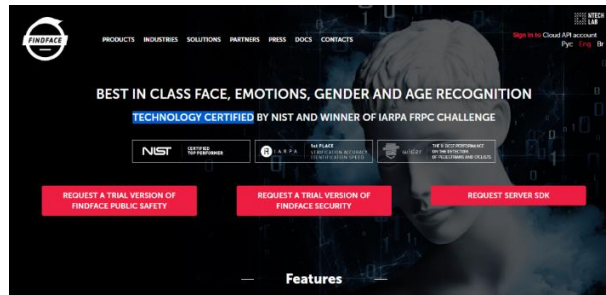


Fig.5. FindFace: main interface

FindFace provide a full documentation of features usage and has 250 million photos in database. The time of processing request is approximately 1 second. The figure 6 shows available features of FindFace.

This app uses facial recognition to identify strangers on social is very popular in Russia. A new facial recognition app has racked up 500,000 users in only two months. Face recognition system allowed successful detaining of almost 180 people from wanted lists.



Fig.6. FindFace's available features

Face X.

This API also has an offline SDK for iOS & Android for you to use. The offline SDK does not provide face recognition, but it can perform face detection, comparing, tracking and landmarks, all while the phone does not have cell service. The figure 7 shows Face X main interface.

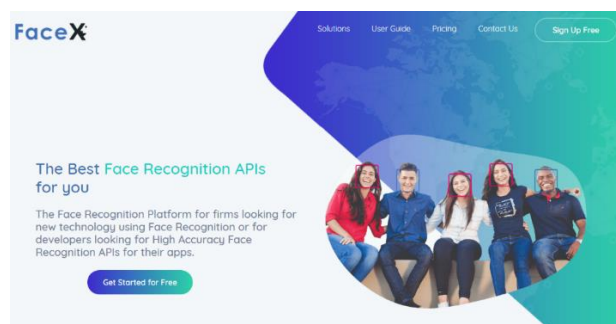


Fig. 7. Face X – main interface page

This API is very useful and not an expensive. FaceX API also provide a step by step user guide. The figure 8 shows Face X documentation information.

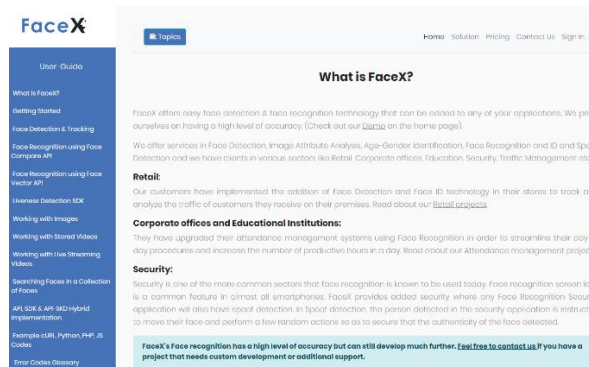


Fig. 8. Face X – user guide

Comparison of API analogues.

The main features from API analogues were figured out and compared (table 1).

Table 1

The most important characteristic for each API analogue

	AWS Face Recognition API	Face X API	Microsoft Azure Face Rec API
Additional info: age, sex, emotions	+	+	+
Recognize face in motion	+		
Face identification	+		+
Well-structured documentation	+		+
Authentication		+	+

The most powerful API analogue is AWS Face Recognition API. It supports:

- 1) additional information: age, sex, emotions;
- 2) recognize face in motion;
- 3) face identification;
- 4) well-structured documentation.

Unfortunately, it does not provide Authentication function. Nevertheless, it is the most powerful analog of considered in APIs.

Conclusions.

Till now, no intelligent system has been developing which gives hundred percent correct outputs. Fortunately current percent of recognition is enough to person’s needs. It leads to creation new technical problems in different sphere like medicine, chemistry, math and other. In theory, neural network can execute all tasks that can be done by human’s brain.

Our main plan is to detect and recognize a facial image in motion. It is almost difficult to detect face from a moving picture. But we wish to perform the work in future. We wish to continue research other neural network architectures to develop ability in deep learning problem solving. However modern AI technologic progress so rapidly that modern developers should follow certain rules to avoid such a thing as "Rise of the Machines."

References

1. Antonio Gulli, Sujit Pal “Deep Learning with Keras” (2017): 25-36. Print
2. Richard Burton, Iffat Zafar, Giounona Tzanidou “Hands- On Convolutional Neural Networks with TensorFlow” (2018): 35-85. Print
3. Santanu Pattanayak “Pro Deep Learning with TensorFlow: A Mathematical Approach to Advanced Artificial Intelligence” (2017):10-65. Print
4. Tariq Rashid “Make Your Own Neural Network” (2016):14-37. Print
5. *CNN in TensorFlow course – official site. [Electronic resource]. – Access mode: <https://www.coursera.org/learn/convolutional-neural-networks-tensorflow/>*
6. *Face Recognition article.– [Electronic resource]. – Access mode: <https://towardsdatascience.com/cutting-edge-face-recognition-is-complicated-these-spreadsheets-make-it-easier-e7864dbf0e1a/>*
7. *Image augmentation tutorial – [Electronic resource] – Access mode: <https://machinelearningmastery.com/image-augmentation-deep-learning-keras/>*
8. *Math method recognition patterns – [Electronic resource] – Access mode: http://window.edu.ru/resource/800/73800/files/lect_Lepskiy_Bronevich_pass.pdf*
9. *Keras documentation – official site. [Electronic resource]. – Access mode: <https://keras.io/documentation/>*
10. *Python Manual – official site. [Electronic resource]. – Access mode: <https://www.python.org/>*
11. *TensorFlow documentation – official site. [Electronic resource]. – Access mode: <https://www.tensorflow.org/>*
12. *Transfer learning documentation – [Electronic resource] – Access mode: <https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a/>*
13. *VGG16 documentation – [Electronic resource] – Access mode: <https://neurohive.io/ru/vidy-nejrosetej/vgg16-model/>*

Автори статті (Authors of the article)

Ніколаєвський Олександр Юрійович – к.т.н., старший викладач кафедри інформаційних технологій, кібербезпеки та математичних дисциплін (Nikolaievskiy Oleksandr – PhD in Technics, Senior Lecturer of the Department of Information Technology, Cybersecurity and Mathematical Sciences). Phone +38044 424-13-86. E-mail: lingvokot@gmail.com.

Склярєнко Олена Вікторівна – к. ф.-м.н., доцент, завідувач кафедри інформаційних технологій, кібербезпеки та математичних дисциплін (Skliarenko Olena – PhD in Physical and Mathematical Sciences, Associated Professor, Head of the Department of Information Technologies, Cyber Security and Mathematical Disciplines). Phone +38044 424-13-86. E-mail: sigma.inet@gmail.com.

Сидорчук Артем Ігорович – магістрант Одеської національної академії зв’язку ім.О.С.Попова (Sidorchuk Artem – master’s degree student in Odessa National O.S. Popov Academy of Telecommunications). Phone: +380 48 7050224. E-mail: artee.sid98@gmail.com.