IRSTI 28.23.15

M. Yerlan[1]
[1]Suleyman Demirel University, Kaskelen, Kazakhstan

FACE-RECOGNITION TO AUTHENTICATE STUDENTS

Abstract. Within the framework of this project, a face recognition system is being de- veloped, which will be used in educational institutions to identify students who are taking exams. To achieve both high quality and fast results, I focused on deep learning approaches to face and object detection and recognition. This research is mainly aimed at providing neural networks and other models with enough data to achieve the desired results. Starting with the basics of neural networks, in which I described and explored a neuron, the smallest unit of deep learning, I brought my research to the point where I could detect a person's face or an object in a photograph. This research began with the development of neural networks and went on to train them on both the CPU and GPU. In a technique called matrix backpropagation, multiple GPUs were used in conjunction with the CUDA core and the cuBLAS library. Face identification was performed using a pretrained Facenet model combined with deep convolutional neural networks. Numerous ap- proaches to deep learning have been developed from the study of neural networks and their application to face recognition.

Keywords: Face-Recognition, OpenCV, Facenet, open-source library, Deep Learning, YOLO**.**

***

Аннотация. В рамках этого проекта разрабатывается система распознавания лиц, которая будет использоваться в учебных заведениях для идентификации студентов, сдающих экзамены. Чтобы добиться как высокого качества, так и быстрых результатов, я сосредоточился на подходах глубокого обучения к обнаружению и распознаванию лиц и объектов. Это исследование в основном направлено на обеспечение нейронных сетей и других моделей достаточным количеством данных для достижения желаемых результатов. Начав с основ нейронных сетей, в которых я описал и исследовал нейрон, наименьшую единицу глубокого обучения, я довел свое исследование до точки, где я мог обнаружить лицо человека или объект на фотографии. Это исследование началось с разработки нейронных сетей и продолжилось обучением их как на CPU, так и на GPU. В методе, называемом матричным обратным распространением, несколько графических процессоров использовались вместе с ядром CUDA и библиотекой cuBLAS. Идентификация лиц выполнялась с использованием предварительно обученной модели Facenet

в сочетании с глубокими сверточными нейронными сетями. На основе изучения нейронных сетей и их применения для распознавания лиц были разработаны многочисленные подходы к глубокому обучению.

Ключевые слова: распознавание лиц, OpenCV, Facenet, общедоступная библиотека, глубокое обучение, YOLO.

***

Аңдатпа. Осы жоба аясында білім беру ұйымдарында емтихан тапсыратын студенттерді анықтау үшін қолданылатын тұлғаны тану жүйесі жасалынды. Жоғары сапалы және жылдам нәтижелерге қол жеткізу үшін мен бет-әлпет пен затты анықтау мен тану бойынша "терең оқыту" тәсілдеріне назар аудардым. Бұл зерттеу негізінен нейрондық желілерді және басқа модельдерді қажетті нәтижеге жету үшін жеткілікті мәліметтермен қамтамасыз етуге бағытталған. Нейрондық желілердің және нейронды терең оқытудың негіздерінен бастап, кейіннен мен өзімнің зерттеулерімнің арқасында жүйені адамның бет-әлпетін немесе фотосуреттегі затты анықтай алатын деңгейге жеткіздім. Бұл зерттеу нейрондық желілерді дамытудан басталып, оларды орталық процессорға да, графикалық процессорға да үйретумен жалғасты. Матрицалық артқа көшіру деп аталатын әдісте CUDA ядросымен және cuBLAS кітапханасымен бірге бірнеше графикалық процессорлар қолданылды. Face ID терең конволюциялық нейрондық желілермен біріктірілген алдын ала зерттелген Facenet моделін қолдану арқылы жүзеге асырылды. Нейрондық желілерді зерттеу және оларды тұлғаны тануға қолдану негізінде терең оқытудың көптеген тәсілдері жасалды.

Түйін сөздер: тұлғаны тану, OpenCV, Facenet, ашық жүйе, терең оқыту, YOLO.

Introduction

Artificial intelligence is the creation of software or a computer system that can execute a variety of tasks utilizing human intelligence. To put it another way, artificial intelligence is the process of simulating intelligent machine behavior. Visual perception, verbal task identification, decision-making, and other behaviors are examples of this behavior. The most difficult aspect of artificial intelligence is duplicating the human brain's precision and efficiency in its operation.

Deep learning is another type of machine learning that allows a machine to learn, teach, understand, and experience the real world through a hierarchy of concepts, allowing a machine or computer to internalize complex concepts by building them from simpler concepts. There is no need for human assistance to manually operate the computer in order to realize all of the knowledge required by the computer, because the computer itself unites the field of experience. Deep learning is, first and foremost, a self-directed, self-learning system that uses

existing data to train algorithms to find various patterns [2].

Neural networks are typically computer hardware or software systems that are designed to model the structure and functions of neurons in the human brain. Artificial neural networks, or neural networks, are used for deep learning.

Aim and objectives of research

The primary technique can be applied to any modal quality, however, due to the currently available research, the results are limited to face recognition using Siamese Networks and OpenCV. I used the Siamese Convolutional Neural Networks in this study, which are capable of determining the similarity between two objects. In addition, these networks may be trained using traditional optimization approaches on randomly sampled pair sets, and they present a viable method that doesn't require any domain knowledge to use deep learning techniques. To help me improve my knowledge of the Siamese convolutional neural networks, I first studied how to construct a neural network consists one hidden layer and then how to construct with several hidden layers.. The Multi-Layer Perceptron was trained using the back-propagation technique, and then the matrix form of the back-propagation technique was created, which made use of many GPUs to do a better job. The results of this research found that GPUs could be used not just for rendering graphics, but also for speeding up the processing of a large number of algorithms when applied correctly. Once the network has been trained using the picture dataset, it is fed new photos to see if the network is able to identify the face in the images. Additionally, we utilized the same trained network for face detection and recognition, as well as object detection and recognition. In addition to setting the standard for facial recognition, object identification, and text detection, this research produced the most accurate system for each of these. In order to train the network, this research selected the predictor and set standards for the parameterization of the network. In light of the fact that I am not planning to learn the measure, my methodology will strictly concern itself with the deep learning elements. The aim is to begin with an accurate image of the neural network, and then use that image to perform face and object recognition and recognition of different types of text, such as an address or phone number, as well as identification and recognition of individual faces.

Background of Literature Review

Additionally, The research is centered on object identification and recognition, as well as word recognition in photos and videos. These are the topics that have received considerable attention in recent years. By defining numerous methods in the literature, this work focuses primarily on deep architectures for face and object recognition. All of these methods are defined by their usage of Neural Networks, that are trainable functions constructed from a collection of linear and non-linear operators. Face Aligner is a representative system of this class of methods. This approach uses the Siamese Network to train and aligns the faces once they are detected using facial landmarks and the image

is resized and cropped to remove the background. Furthermore, when performing real-time face identification, it draws a bounding box around the face in the video or live camera, displaying the system's accuracy as well as the individual's name if the network was trained for that individual. The system reads and captures images and videos using OpenCV, and a web service for image registration using images or a live camera is also included. Currently, this approach gets the highest performance on pictures and live cams.

Methods

In the 1960s and 1970s, when scientists from various fields such as psychology, computer graphics, biology, and computer vision were just starting to use and come up with new methods of facial recognition, it was logical to calculate the distance between important parts of the face, such as the distance between the eyes or the angle between the eyes and the lip for face recognition.

But to automate the face recognition system, this data is not enough. Therefore, the following methods are also used:

- Holistic Matching Methods ;
- Feature-based Methods;
- Hybrid Methods.

Holistic Matching Methods

With a holistic approach, a person's face will be used as input data into the system.

Examples of such a system are: one's own person, linear discriminant analysis. In the first block of the diagram (Figure 1), you can see the stages of face recognition using the Eigenface method.

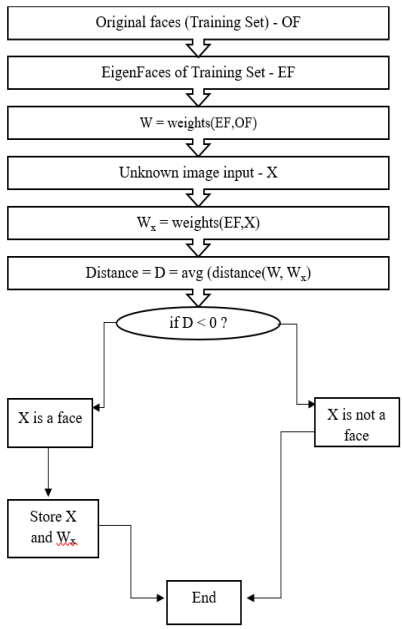

Figure 1. Eigenface-based algorithm

Feature-based Methods

A feature-based method is to find faces by extracting the structural features of the face. Therefore, firstly facial elements like eyes, lips and nose are extracted. And their location is taken as input to the classifier. But the disadvantage of this method is that it becomes almost impossible to determine the object when even the head position changes. Therefore, in my opinion, this method will be very difficult to use for recognizing students' faces in real time.

Hybrid Methods

Hybrid methods use these two methods together, which gives it high efficiency. The hybrid method takes a person's face in 3D format. Therefore, with its help, you can not only read the distance or size of the face elements but the depth or, for example, the shape of the forehead or chin. There are 5 steps for this: Detection, Position, Measurement, Representation, Matching.

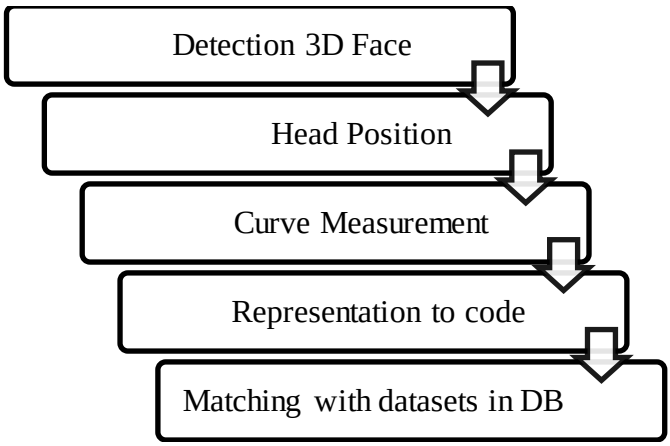

Figure 2. Steps of Hybrid Method

Results

From these studies, we realized that the most effective method is the hybrid method. And now the task remains to implement a program that recognizes the student's face. For this, we have chosen the open-source library OpenCV as the basis. But improving it with deep learning algorithms and YOLO algorithm.

The system We created, like most face recognition systems, consists of 3 stages:

1. *Data Collection*

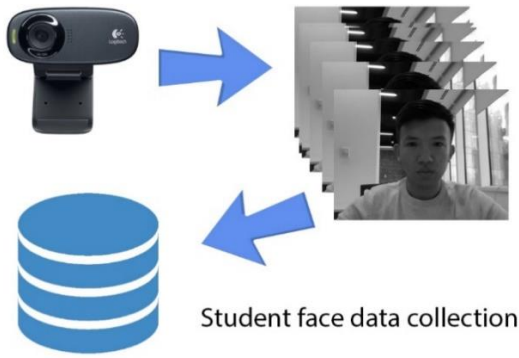

Figure 3. Data collection stage

Before a recognition system can identify a face, it must first be taught to a collection of images known as a training kit. The set allows you to compare it content with a new image to identify the test user.

To more or less identify a student, you need at least 25-30 images. To simplify the task of collecting images, We wrote a script that takes picture on the webcam, the student's face 30 times.

1. *Train the System*

After the system has detected a person's face, you can add to training set. So that this data can be used the recognition algorithm is correct, several factors should be considered. First, the size the test image and the trained image must be the same. If the test image is larger than the image of the training set, it can be reduced to a smaller size while maintaining the aspect ratio of the large image is the same as that of the smaller image. Without aspect ratio adjustments, resized images may be stretched vertically, horizontally, or both, negatively impacting algorithm recognition accuracy. If the test image is much smaller than the image in training set, you can enlarge it to the same size as the image in the training set. At this stage, we give the previously collected images to the system to train. To train my system, We used the OpenCV Recognizer, which uses a special OpenCV function. We transfer all data for training to OpenCV Recognizer, the result of which will be a file with the extension .yml which we will use for the next step For student identification, We saved them by student ID on the SDU portal.

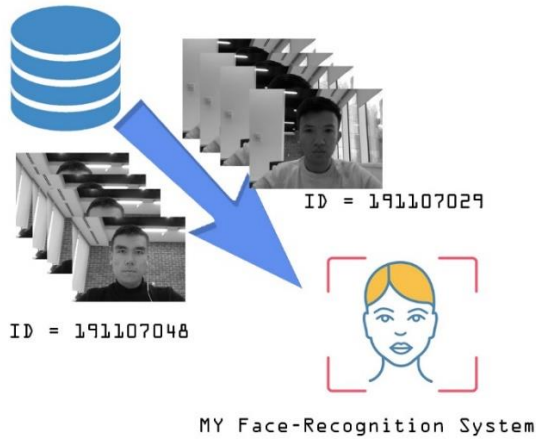

Figure 4. Training stage

2. *Recognition*

And the last stage is recognition. At this point, the system recognizes the student ID. And with this id gets student's name and surname from database and displays all the information on the screen.

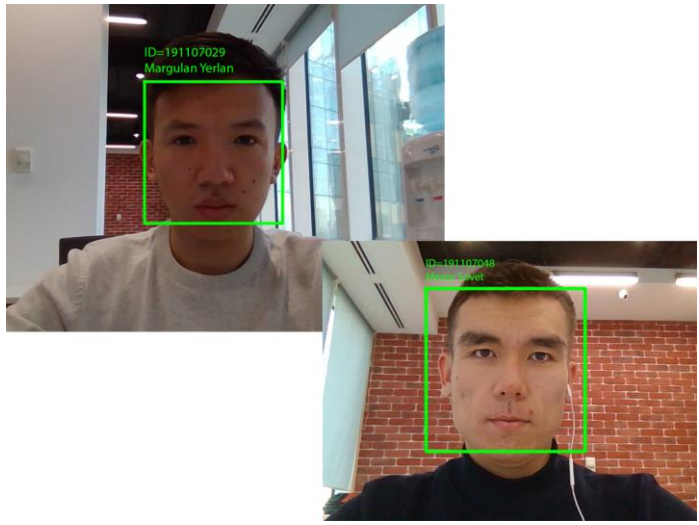

Figure 5. Recognition stage

Conclusion

While deep learning is still emerging as a game-changing technology that will aid humans in not just recognizing faces in images, videos, speech, or audio translations, but also in creating new animals with human-like brains, the majority of people still don't recognize its true potential. I have extensively discussed and re- searched this fundamental and critical subject. It's clear that most existing neural network models lay the groundwork for what will surely be a tremendous growth of deep learning in the future. Following the work of

various well-known learning algorithms and their use of noisy or noise-canceling encoders and decoders, models such as CNN, Deep CNN, and other such models, as well as various noisy or noise-canceling encoders and decoders, demonstrate the usefulness of deep learning in numerous areas and show that deep learning is capable of delivering positive results in a wide range of disciplines. Several goals remain in deep learning, including higher accuracy in face recognition, speech detection, and translation, as well as learning in networks, and an easy method of machine learning, all of which will lead to more efficient and theoretical approaches to deep learning. We must do feature extraction on every level of the neural network to the point where the underlying backpropagation algorithm, due to a practical limitation, is incapable of performing the same tasks performed by the matrix-based form of the backpropagation algorithm. It has also been found that it is not necessary to completely retrain the model or network, but instead it is sufficient to merely update it with fresh data and knowledge. This is essential, as well, because it is essential to develop and explore more complex algorithms that employ a selective optimization strategy. At this time, existing deep learning algorithms seem to work well on many tasks in a lab setting, but when put to the test, they have failed. For ex- ample, a robot was patrolling a pool and drowned due to a lack of knowledge and training. When training models or implementing deep learning algorithms, their size can be reduced by extracting features from each layer of the network or even from each input. Additionally, we will need to design more scalable models capable of parallelizing the training of deep models or networks, which could result in the optimization of several approaches. Taking everything into consideration, we intend to further improve our product and make it as efficient and quick as possible.

References

1 Sombir Singh Bisht. Image Registration Concept and Techniques: A Review. *Journal of Engineering Research and Applications.* 4 (4), (2014): pp. 30-35.

2 James Philbin, Florian Schroff , Dmitry Kalenichenko. *FaceNet: A Unified Em- bedding for Face Recognition and Clustering.* (2015): pp.1-10.

3 Graupe, D. Principle of artificial Neural network. *In: World Scientific Publishing Co Pte Ltd.* 384 p.

4 Ujjwal Karn. *A Quick Introduction to Neural Networks.* URL: https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/.

5 Adrian Rosebrock. *Face Alignment with OpenCV and Python.* URL: https://www.pyimagesearch.com/2017/05/22/face-alignment-with-opencv-and-python/

6 Zhang, S., Gunupudi, P., Zhang, Q. Parallel back-propagation neural network training technique using CUDA on multiple GPUs. *2015 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO),* 2015, pp. 1-3

7 Y. Bengio and A. Courville.*Deep Learning. The MIT Press, Cambridge, MA,USA.* URL: https://www.deeplearningbook.org/

8 Felzenszwalb, P.F., Girshick, R.B., McAllester, D., Ramanan, D. Object Detection with Discriminatively Trained Part-Based Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (9), (2010): pp. 1627-1645.

9 Girshick, R., Redmon, J., Divvala, S. *You Only Look Once, Uni-fied, Real time Object Detection. GPUs.* (2015): pp. 1-10.