Development of a module for identifying a person by photography to improve the quality of the machine-engineering plant in remote mode

D A Shabalina¹, S A Yurkov², T V Masharova³, E K Gerasimova⁴ and S S Pichugin⁵

¹ Department of Digital Technologies in Education, Vyatka State University, Kirov, Russia
² Department of Criminal Law, Procedure and National Security of Vyatka State University, Kirov, Russia
³ Department of Pedagogy, Institute of Pedagogy and Psychology of Education FSEI HE Moscow City Pedagogical University, Moscow, Russia
⁴ Department of Informatics, Institute of Mathematics and Information Technologies (named after Prof. Nikolay Chervyakov) North Caucasus Federal University, Stavropol, Russia
⁵ Department of General Education Academy of Public Administration, Moscow, Russia

E-mail: sh4balina1981@yandex.ru

Abstract. A machine vision performs a variety of face verification tasks in different infrastructure environments. A facial recognition system is technically implemented as an image recognition program. It serves to automatically identify a face in an image, and then identify the person by comparing and analyzing the biometric data of the human face. Cameras are able to capture an image at a distance, which is perfect for building monitoring systems and contactless biometrics. Such software systems are particularly important in the context of remote work of employees of industrial enterprises. The purpose of research is to develop and implement a facial recognition web service for the staff department of a machine-engineering enterprise. The web service should work in conjunction with the existing data bank. Tasks of research are: to consider the fundamental algorithms of facial recognition, to formulate the basic requirements for the development and implementation of the corresponding automated web service. The significant result of research is that the developed module uses library of computer vision algorithms, image processing and general-purpose numerical algorithms with open-source code. Identification system was checked by testing on the company's website.

1. Introduction
A facial recognition system is an image recognition program, the purpose of which is to automatically identify a face in an image, and then identify the person by comparing and analyzing the biometric data of the human face [1]. Such facial recognition systems provide security support, help to control access by face and to make analysis of sales. Thus, machine vision can perform face verification tasks in any infrastructure environment. At the moment the most popular method of identity verification is comparison of facial features with photo ID.
There is a difference between a facial recognition and a face comparison, though both these approaches are based on the same technology [2]. The facial recognition is a searching for a match with an image in a large database, it is widely used by law agencies (to find photos of criminals), by security service (for example, for verification of employees), at banking (to prevent fraudulent transactions), at the consumer sector (for real-time payments). And as for the face comparison just two photos of the same person are to be compared with each other [3].

The biometric facial recognition system is based on the principle of constructing and comparing mathematical models of a face, which provides a high precision of identification or identity verification. These systems hypothesize whether a face in one photo matches a face in another, regardless of its expression and other attributes, such as age, hair, accessories, and so on. Regardless of the technology used, facial recognition programs work in several stages [4].

Stage 1. Detection of the face. The system automatically identifies the face in the image (the photo scan or the frame of the video stream). The detection of the face is a very important step of recognition, it allows the further identification.

Stage 2. Analysis of the face. A human face has 80 nodal points. A facial recognition program can read the geometry of these points such as the distance between the eyes, the shape of the nose and so on. A digital facial recognition relies mainly on 2D, not 3D images, because 2D photos are easier to compare with those contained in the database.

Stage 3. Digital face print. The facial recognition system perceives the human face as data: the results of facial analysis are converted into a digital code. Such a "face signature" is unique to each person.

Stage 4. Searching for a match.

Various factors can affect the performance of Face-ID systems outside the controlled environment. Advanced facial analysis programs can cope with such changes to some extent, but under poor conditions they can produce erroneous identification results or fail to recognize at all [5].

Such software systems are particularly important in the conditions of remote work of employees of industrial enterprises. The context of remote work means the interaction of employees, representatives of legal entities and individual entrepreneurs with each other at a distance [6].

Kirov machine-engineering plant is a modern hi-tech enterprise. The production of modern and reliable products and equipment is carried out in accordance with certificates and quality policy. The company employs more than 1,500 workers, the volume of production increases annually. Modern technologies and a high level of qualification of employees enable to produce products of high quality [6]. The enterprise processes personal data of the following categories of users in remote access mode: individual employees; individual job seekers; representatives of legal entities and individual entrepreneurs and so on. It is allowed to submit reports on the enterprise website, which is a great advantage, since there is no need to visit the corresponding department personally. It is necessary to log-in with your password to be identified. But nowadays such way of identification is not safety enough. Biometric data can be used to increase reliability, because it is difficult to forge it. A human face is a type of biometric data. So, the process of identification includes the analysis of the image got from the web-camera, identification of the face on the image and its comparing with the photo of employee stored in the account card. Such a task could be solved with the help of machinery vision libraries [7].

The purpose of research is to develop and implement a facial recognition web service for the staff department of a machine-engineering enterprise. The web service should work in conjunction with an existing web site and data bank.

The main tasks are determined by the requirements for software development: the recognition function must be separated from the information system and implemented as a separate web service; web service must be implemented on the NodeJS platform; the information system should interact with the developed web service using the http protocol; information should be exchanged using JSON requests; the web service input data must be a user photo from a webcam encoded into a string format Base64; in case of successful recognition the output data must be the identification number of the user on photo; in case of fail the web service must show the description of error; all necessary information
for recognition should be stored on the side of the web service; document-oriented database must be used to store information.

2. Materials and methods
Currently the scientific fields related to computer vision are actively developing, computing power of modern software allows processing large data arrays. As a result, various libraries of face detection and recognition algorithms began to be actively created.

OpenCV (Open-Source Computer Vision) is a library of computer vision algorithms, image processing and general-purpose numerical algorithms with open-source code. It is implemented on C/C++. OpenCV contains facial detection methods based on Haar cascades [1] and several methods of face recognition.

The Eigenfaces method is based on principal components analysis (PCA) method. Let a matrix be given $X = \{x_1, x_2, ..., x_n \}$, which is a set of faces. Each element $x_i$ of $X$ is a vector, obtained as follows. Let an image of a face ($w \times h$ pixels) be given. Let us consider it as a matrix $Y$, where $y_{ij}$ is a brightness value of pixel with coordinates $(i, j)$. Then the matrix is transforming into a vector $x$ with the length $w \times h$. So the main scheme of PCA method is as follows: projecting all sample images into subspace; projecting input image into subspace; searching for the nearest neighbor between images from sample and input image. As a result we get the distance and index of the object from the database that the input image is closest to. This method is susceptible to lighting changes. Also it is necessary to build new eigenspace each time we add new face in database. The main advantage of the method is quick recognition [8].

Fisherfaces method is based on linear discriminant analysis. This method selects such projection of image space onto feature space that the ratio of the between-class scatter is maximized and the ratio of the withinclass scatter is minimized. This algorithm is better in face recognition in different illumination conditions but the training sample must contain images of faces with different lightening.

LBPH method is based on the Local Binary Pattern method. LBPH extracts features from pixel surroundings. At the beginning an image is dividing into blocks of a grid. Next, a code histogram is constructed for each block. A vector used for comparison is made by concatenation of these histograms. The codes are calculated as follows: each pixel is to be compared to its neighbor, if intensity of central pixel is more than or is equal to neighbor’s intensity then it gets code 1, otherwise code 0. So each pixel is represented as a binary code got from comparison.

Dlib is a modern tool of C++, which contains machine learning algorithm and lots of other auxiliary means. The library is free; therefore, it can be used for academic purposes. Dlib uses HOG (Histogram of Oriented Gradients) for face detection. Also it contains a recognition model, designed as neural network of 29 layers. This net is an analogue of the neural network ResNet-34. It was trained on a database of more than 3 million photos, compiled from several well-known image databases, as well as many photos from the Internet. About 7 thousand people were represented on those photos [9]. With the help of this model Dlib represents human face as a 128-dimensional vector space so that images of the same person are next to each other, and images of different people are far from each other. Thus, we can use Euclidean distance to count the distance between the vectors. If the distance threshold is 0.6 then Dlib model recognizes faces with the accuracy of 99.38% [10].

Georgia Tech Face Database is a database of images which was created at the Center for Signal and Image Processing at Georgia Institute of Technology in 1999 [11]. Database contains 750 frontal photos of 50 people. Each person is represented by 15 color JPEG images at resolution 640x480 pixels. The pictures show faces with different lighting conditions, background and distance to a person.

3. Results
A face recognition process can be reduced to two main steps: a face detection and localization; the recognition process itself. In the first step there is a search of a face on image and its selection. The process of face detection and its localization has already been described in detail [1]. In the second step the face image is aligned if necessary and then the necessary characteristics are calculated. After that
they are to be compared with reference features from database. Currently there are many facial recognition algorithms. For example, comparing of templates.

Website of Kirov machine-engineering plant is designed on Drupal [7] and deployed on a server with Ubuntu Server 16.04. Drupal is a content management system with open-source code. It makes process of developing and supporting websites easier. The content management system has a module structure. Drupal is a free software, so there are a lot of repositories with various modules. Thus, it is enough to find and install the appropriate module to expand the functionality of the site. Besides it is possible to develop new modules or change existing ones. Quiz module operations test forms, which contain a parameter that is responsible for the method of employee identification.

An identification function is implemented with the help of user interface module Webcamproctor. In case of choosing a face identification method the module uses the functions of an external web service for facial recognition. At the beginning of testing module Webcamproctor gets an image from webcam, converts it into a string format Base64 and sends the recognition demand to web service.

The web service checks whether the user is in the database and returns a response about the possibility of the operation. On the basis of response module Webcamproctor will grant access or display the appropriate message. In addition to the identification function, Webcamproctor performs the function of collecting images to create a recognition model. Basic components of the web service are: a control unit that interacts with the OS, processes the received requests and writes information to the database; an interface module that controls the reception and transmission of requests; facial recognition library; a database for storing the necessary information.

The main idea of PCA method is to find a linear combination of features that maximizes the overall variance of the data. This method does not take into account object classes, therefore, some of the distinctive information may be lost due to discarding components. This problem could be solved by the linear discriminant analysis (LDA) algorithm [5]. The linear discriminant analysis selects such a projection of image space onto feature space that the ratio of the between-class scatter is maximized and the ratio of the within-class scatter is minimized. This algorithm is better in face recognition in different illumination conditions but the training sample must contain images of faces with different lightening. The PCA and LDA methods have mostly general approach to face recognition. These algorithms treat the image as a vector of large dimension. It is difficult to work with such vectors, therefore, a subspace of smaller dimension is used, in which significant information is stored.

The Local Binary Pattern algorithm (LBPH) extracts features from image surroundings, such a set of features has an undersized structure. This algorithm considers the neighborhood of each pixel and gives a value to each pixel with the help of a function. After that the image is divided into sub-areas, for each of which a histogram is calculated. Then histograms are concatenated and are compared using machine learning methods. This method is simple to implement, has a sufficiently high speed of operation, which can still be increased due to various modifications of the algorithm. The algorithm is also resistant to changing lighting. Thus, facial recognition algorithms combined with the performance of modern computers allow to implement effective systems for identifying people in real time.

Labeled Faces in the Wild is a database of face pictures intended for investigation of the problem of unconditional facial recognition. It contains more than 13000 JPEG pictures of size 250x250 pixels, collected from the Internet [10]. 1680 people have two or more photos of themselves, the other ones – only one photo. The Database Labeled Faces in the Wild is more difficult to work with then the Georgia Tech Face Database, so the testing results appeared to be worse. But Dlib library again showed almost 100% accuracy of recognition. For LBPH and Fisherfaces methods, at least 7 photos per person were needed to achieve 80% accuracy, and for the Eigenfaces method – at least 10.

Next the application architecture was designed. Computer vision libraries OpenCV and Dlib were reviewed and tested on special image bases. Based on the test results, the Dlib library was selected to solve the face recognition problem. During the design process, two main parts were highlighted - a web service for facial recognition and the Webcamproctor interface interaction module. The process of developing of these modules is described below.
To develop a web service, the Node.js platform is chosen, which extends the capabilities of the JavaScript by adding the functions of connecting external libraries written in different languages and accessing them from JavaScript code. Module face-recognition.js is selected to implement recognition functions. This module uses the Dlib computer vision library, which contains effective facial recognition tools. Data storage and processing are carried out by means of document-oriented (NoSQL) database management system MongoDB. MongoDB uses JSON-like documents and database schema and does not require a table schema description when designing.

The JSON format is used to exchange data between the information portal and the web service. JSON is a textual format of data exchange based on JavaScript. This format is independent of the implementation language and can be used with almost any programming language. Basically, JSON is used to transmit data over the network. JSON-text could be organized in one of the following structures: set of key/value pairs (different programming languages, this concept is implemented as an object, record or structure); an ordered list of values (in most languages it is implemented as an array, vector, list, or sequence).

Web service implements the following functions: storing and processing of information intended for identification; identification of the user by the recognition model; creation of a recognition model; function of the external interface M2M.

Machine-to-machine interaction (M2M) is the general name of technologies that allow machines to exchange information among themselves. Interaction between the enterprise site and the web service for facial recognition is carried out using http-requests. A POST-request with the title Content-Type: application/json is sending to the address of the web service, and a string in JSON format will be sent as a parameter. After performing all the operations, the web service must send a JSON response, which is to be processed on the site. To implement the photo identification function it is necessary to create a model for recognition.

The model is built on the basis of processing several photographs depicting a human face. The more photos, the fuller the model and therefore higher the accuracy of recognition. The model is stored in the database along with the rest of the user's characteristics in a format of a JSON string.

To create a recognition model, the function train is developed. To initialize this function it is necessary to send a POST-request to an address http://<web service address>/train. The number of photos in the query may vary. Function train have the following input parameters: identification key; usrID; file extensions; array of images. The query with these parameters is generated on the side of enterprise web portal. The array of images consists of photos coded into string format Base64. The generated request is sent to the web service.

The received request is being processed as follows. All images are decoded and stored on the side of web service as temporary files of the original format. The model for recognition is created using the received files and the employee ID. If the photos do not meet the requirements of the model (there is no any person or more than one person on the photo), then the process of creating of the model is terminated and a message about the inability to create a model is sending to the site. When the model is successfully created, the information in the database is updated, and the employee ID is used as the key. And a JSON response is sent that contains information about the results of the execution. When the model modification process is complete, resources are cleared by deleting temporary files and clearing memory.

The developed recognition function identifies the user by recognizing the face from the model stored in the database. To initialize this function, it is necessary to send a POST-request to an address http://<web service address>/recognition with the the following input parameters: identification key; user ID; file extension; an image in the format Base64.

The function recognition searches for the user with the specified ID in the database. If there is no such employee then the corresponding JSON response must be sent to a web site. If a record exists the recognition model is loaded from it. The image in the Base64 format is decoded and stored as a temporary file in the format that was specified in the request. The face is detected in the image and the Euclidean distance is calculated using the recognition model. If the distance is more than 0.6 it is assumed that the photo shows another person, and as a result, the function sends a JSON response with
the value «unknown» in the usrID field. If the distance is less than 0.6, the recognition is regarded as successful. In this case the employee ID is written in the usrID field, and the Euclidean distance is written in the «distance» field of the JSON response. When the recognition function is complete, resources are cleared by deleting temporary files and clearing memory.

The Webcamproctor module is implemented in PHP, it performs the interaction between the enterprise website and the web service for face recognition. The module implements the following functions: interaction with the user through a graphical interface via a web browser; collection of photos obtained using a webcam; encoding of photos in the Base64 string format; M2M interface. The Webcamproctor module contains a description of the following forms: a form for identity identification before the operation begins; form for creating a recognition model.

When you click the «Take a photo» button, the photo is uploaded to the enterprise server and the «Start Operation» button appears. After pressing this button, an identification request is generated and sent. The message received in response is processed and, depending on the identification results, the corresponding message is displayed. This form could be called from the employee's personal account or automatically when you receive a response containing information about the absence of the necessary model in the database.

Form modules allow to take three photos using a webcam. Photos are created consistently one after another when click on «Take a photo» button. After that the button «Send» appears, that generates and sends a request to create a model. The response received from the web service is processed and a message about the results of the recognition model creation is displayed on the form.

Putting into operation is an important stage in the life cycle of a software product. At this stage the effective and long-term use of the implemented product is laid. Purposes of implementation are: adaptation of the program to the working environment; identification of errors in program code and debugging; training of accompanying personnel. At the same time the following tasks are solving: to deploy a web-service; to make changes to existing software; to test; to solve detected problems; to perform acceptance tests; to conduct a training workshop. To put the developed application into operation a web service for facial recognition was deployed in a Docker container on the enterprise server, and the Webcamproctor module was integrated into the site software. Docker is software that implements application virtualization technology that isolates an application from operating system resources.

The important concepts are image and container. An image is the basic element of each container. The image includes everything you need to run the application on a Docker machine: OS, program execution environment and the application ready to deploy. A container is an executable instance that encapsulates the required software. The Ubuntu 16.04 image was taken as the initial image. Further, the developed software, the Dlib computer vision library and the Mongodb database management system were installed in this image. The PM2 program was also installed, which allows to maintain applications written in NodeJS in working condition and to operational and reboot them without downtime. The container was started after the installation of all required programs. It was tested by sending special requests to the web service.

4. Conclusion

The company's web site is implemented on the Drupal 6 platform, which controls the work of Webcamproctor module. To install this module the folder with the program code was copied to the path sites/all/modules, and then a module in the appropriate platform section was connected. The installation of the Webcamproctor module was completed in this way.

To check the functionality of the identification system a testing was conducted. There was a model of recognition created and then tested on the web site. Different images of users on the base of which the recognition model was made were downloaded. The process of the model creation was successfully completed and confirmed by a corresponding message.
After that some incorrect images were given for the creation of the recognition model, such as an image without a human face on it. The program diagnosed an error in the input data so the model was not created.

After that the following scenarios were modeled: an identification of a verified user, an identification of an unverified user, a processing of an image with two or more human faces on it, a processing of an image without a human face on it. All the scenarios were processed correctly with the corresponding message shown.

After testing a demonstration combined with a training seminar was held for customer representatives. During the demonstration a complex of actions of an ordinary employee was performed, also some issues related to the operation of this system were discussed.

Thus the tasks of the research were completely solved: algorithms of facial recognition were considered; the most popular facial recognition libraries were analyzed and tested; the subject area of work of the stuff department in terms of identification was considered; the architecture of the software providing a function of identification by photo was designed; the software was developed and integrated to the system.

At the moment the developed software meets the requirements of identification of enterprise's employees. However the implemented method has some disadvantages because the process of recognition is initializing once at the beginning of operation so it could be falsified. The further research assumes the on-line processing of video stream from a web-camera in order to monitor identified users in real time.

References

[1] Devitsyna S N, Eletskaia T A, Balabanova, T N and Gakhova N N 2019 Development of an intelligent system of biometric user identification Belgorod State University Scientific Bulletin. Economics, Computer Science. 46(1) 148-160

[2] Amandeep K 2015 Face Recognition using Template Matching. International J. of Comp. Applications. 115(8) 10-13

[3] Vizilter IU V, Gorbatshevich V S and Moiseenko A S 2020 Single-shot detector of faces and landmarks on digital images Computer Optics. 44(4) 589-595

[4] Ezrokh I U 2020 Problems of establishing the practice of remote identification of clients in the Russian banking system (economic aspects) St Petersburg University Journal of Economics Studies. 36(2) 266-286

[5] Awad A 2016 From classical methods to animal biometrics: a review on cattle identification and tracking Computers and Electronics in Agriculture. 123 423-435

[6] Soboleva E V, Karavaev N L, Shalaginova N V and Perevozchikova M S 2018 Euro. J. Contemp. Edu. 7(4) 845

[7] Sulavko A E, Lozhnikov P S, Choban A G, Stadnikov D G, Nigrey A A and Inivatov D P 2020 Evaluation of EEG identification potential using statistical approach and convolutional neural networks Control and Systems Engineering. 6 37-49

[8] Iskhakov A IU, Iskhakova A O, Meshcheriakov R V, Bendrau R and Melekhova O 2018 Application of user behavior thermal maps for identification of information security incident. Informatics and Automation (SPIIRAS Proceedings). 6(61) 147-171

[9] Zinoveva E S and Voitenko IU 2015 Development of information and bio-technologies: impact on personal security MGIMO Review of International Relations. 5(44) 154-161

[10] Spitsyn V G, Bolotova IU A, Shabaldina N V, P N Hoang and B T T Trang 2016 Facial recognition based on PCA using the Haar and Daubechies wavelets Scientific Visualization. 8(5) 103–112

[11] Cantelon M, Harter M, Holowachuk T J, Rajlich N 2014 Node.js in action Manning Publications Co 2014