Daniel DROZD[1], **Zenon JABŁOŃSKI**[2], **Grzegorz LITAWA**[3]

[1] University of Applied Sciences in Nowy Sącz, Faculty of Engineering Sciences, Zamenhofa 1a, 33-300 Nowy Sącz, e-mail:ddrozd@ans-ns.edu.pl
[2] University of Applied Sciences in Nowy Sącz, Faculty of Engineering Sciences, Zamenhofa 1a, 33-300 Nowy Sącz, email: zjablonski@ans-ns.edu.pl
[3] University of Applied Sciences in Nowy Sącz, Faculty of Engineering Sciences, Zamenhofa 1a, 33-300 Nowy Sącz, email: glitawa@ans-ns.edu.pl

Recognition of people in an IoT system with limited resources

Abstract

The Internet of Things (IoT) is currently a rapidly growing field of science. It is often the case that devices working in IoT cannot exceed predetermined dimensions. The small size does not allow the use of large computing power. For this reason, we are dealing with the use of Internet of Things in resource-limited environments. In this paper, a study was conducted to test the accuracy of face recognition. By analyzing the results obtained in this paper, the reliability of the solutions ob-tained with the library used can be checked.

Keywords: face recognition; IoT; microcontroller; IoT system with limited resources; OpenCV; recognition accuracy; cybersecurity.

Rozpoznanie osób w systemie IoT z ograniczonymi zasobami

Streszczenie

Internet rzeczy (IoT) to obecnie dynamicznie rozwijająca się dziedzina nauki. Często zdarza się, że urządzenia działające w ramach IoT nie mogą przekraczać wcześniej określonych rozmiarów. Niewielkie rozmiary nie pozwalają na użycie dużych zasobów obliczeniowych. Z tego powodu mamy do czynienia z wykorzystaniem Internetu rzeczy w środowiskach o ograniczonych zasobach. W niniejszym artykule przeprowadzono badanie w celu sprawdzenia dokładności rozpoznawania twarzy. Analizując wyniki uzyskane w tym artykule, można zweryfikować niezawodność rozwiązań uzyskanych przy użyciu używanej biblioteki.

Słowa kluczowe: rozpoznawanie twarzy; Internet rzeczy (IoT); mikrokontroler; system IoT z ograniczonymi zasobami; OpenCV; dokładność rozpoznawania; cyberbezpieczeństwo.

1. Introduction

Devices that are part of the "Internet of Things" are increasingly becoming an integral part of our live. They allow us to save our time, automate processes or increase our safety. Smart homes are becoming something normal and are more and more often taken into account when planning investments. The growing number of companies producing IoT equipment, and the increasing affordability of this type of solutions, resulted in the dynamic growth of the industry. The aim of the article was to check the operation of the automated facial recognition system and recording using devices with limited performance. Checking the operation of the system on the most popular Raspberry Pi microcomputers. Integration with the Xbee module and the STM32 microcontroller, through which you can integrate the proposed system with IoT devices operating in sensor networks. Additionally, a database of 15 actors was prepared, 4 people were selected from them, and the behavior of the system was examined in more detail.

2. Materials about IoT – Internet of Things

But what IoT really is? IoT, or Internet of Things, is a network of interconnected things. Put simply, it will be a network of interconnected "smart" devices. They can be equipped with many sensors used for monitoring the environment, which permit for in-dependent operation but their full potential is revealed when they are connected in a net-work. It is estimated that the current number of IoT devices is around 20 billion world-wide. This is a great amount that gives us to understand, how commonly IoT is used (Arora et al., 2021; Alam et al., 2020).

Sample and the most common places of IoT application will be devices from the "smart home" family. Everything, starting of intelligent lighting, automatic heating system, air conditioning control, and so on, adds up to so-called smart home. The possibility of creating "scenes" gives almost unlimited possibilities of personalizing all scenarios.

There are many off-the-shelf platforms that allow us to easily build IoT hardware. The Arduino platform is undoubtedly the most popular for its ease of use. Numerous libraries and community support mean that when we encounter a problem, we can solve it quickly with the help of the Internet. Another popular platform is ESP32, equipped with Wi-Fi and Bluetooth, which is perfect for creating Internet of Things devices. Other plat-form Raspberry Pi with high computing power finds its application in places requiring a large number of resources. It is no different in the case of facial recognition – it is a complicated process that requires a camera and a strong computing unit. The next of the most popular platforms is the STM32. These are 32-bit processors on the ARM core. Depending on the needs, we meet high-power microcontrollers, with a large number of peripherals and energy-saving. Producer shares configurator in which everyone will surely find something that will meet his expectations.

As in the case of microcontrollers, there are also many solutions to facilitate communication between individual devices. The obvious ones will be Wi-Fi and Bluetooth, but in the case of IoT, other communication interfaces have gained the greatest popularity. The most common communication standards in the case of the Internet of Things are: ZigBee, Z-Wave, Threed, standard cellular network, Sigfox or LoRa (Saravanan et al., 2021; Hancke et al., 2016; Ziegler, 2019).

ZigBee is the most widespread standard in the area of smart home devices. Based upon the IEEE 802.15.4 standard communication on the 2.4 GHz band is characterized by a mesh topology that guarantees stability and a long connection distance (up to 100 m). Thanks to the securing with a 128-bit cryptographic key, communication is adequately protected against unauthorized access and error detection allows for possible correction. ZigBee Green Power is an energy-saving version, designed for battery-powered devices (Yasuura et al., 2017; Ahmed, 2018; Suganthi, 2021).

Z-Wave is a standard developed in 1999 by the Danish company Zensys. It operates at frequencies below 1 GHz, which makes it resistant to interference from other devices communicating via Bluetooth or Wi-Fi. One network can be composed of at most 232 devices. The mesh topology used in this solution significantly increases the range, which in communication between two devices is up to 100 meters in an open space (Veneri et al., 2018).

3. Research methodology

The main objective of this research is to recognize people in IoT systems with limited resources. IoT devices are typically characterized by low computational power, which can repeatedly generate limitations. By using wireless connectivity and additional hardware to support higher computation, this can be prevented.

4. Recognition of people in an IoT system with limited resources

In order to run on limited resources, a team of devices consisting of a raspberry pi 4 responsible for image capture and STM32 responsible for data transmission and storage was created. In our research, we used Python language to implement the face recognition system. It is a high-level programming language and an open source project managed by the Python Software Foundation.

OpenCV is one of the most popular libraries for real-time image analysis. It was initiated by Intel, based on open source. The big advantage is multiplatform, which allows easy implementation regardless of the operating system. OpenCV can be used in:

- face recognition;
- objects detection;
- set of 2D and 3D tools;
- evaluation of emotions;
- understanding of movement;
- stereo vision from two cameras, taking perception into account;
- motion tracking;
- segmentation and recognition;
- augmented reality;
- gesture recognition;
- human-computer interaction (HCI);
- mobile robotics.

Face recognition is a library that allows us to quickly implement face recognition. It includes many useful functions such as tagging faces in images, recognizing faces and describing who a person is – based on previously learned patterns (Chen et al., 2020; Li et al., 2011).

Facial recognition is an involuntary activity performed by a human. To our brain, this seems like a simple and effortless activity. In the case of computer and electronics, it looks much more complicated. The program needs to be provided with images for analysis, for which a camera is necessary. All of them must be analyzed and then properly interpreted. The project used the prototype kit STM32, RaspberryPi and XBee modules, which helped in communication between individual devices. It is presented in the diagram below (Figure 1).

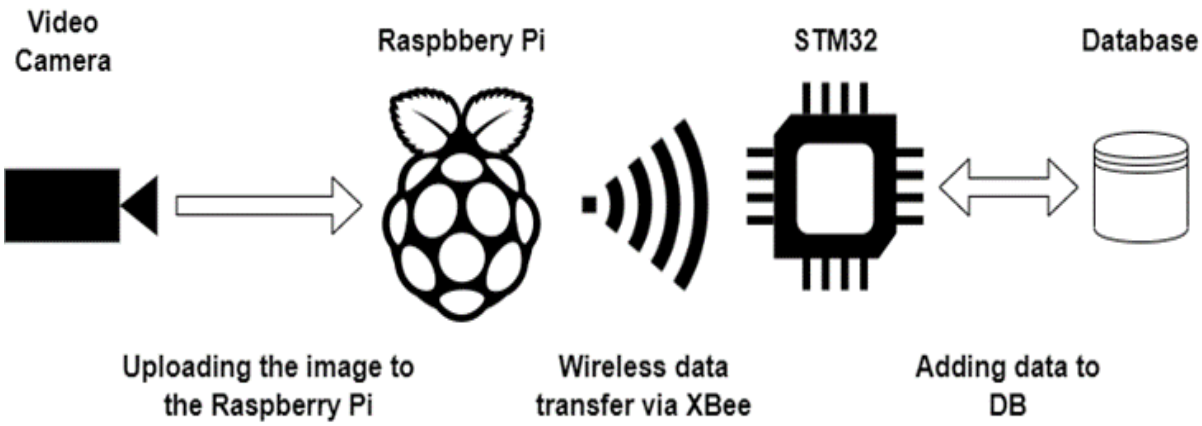

Figure 1. Diagram showing communication

The first step in the facial recognition is finding the face in the image. For this purpose, a histogram of oriented gradients, called HOG in brief, is created. The illustration below shows the traditional photo and the visualization of the histogram (Figure 2).

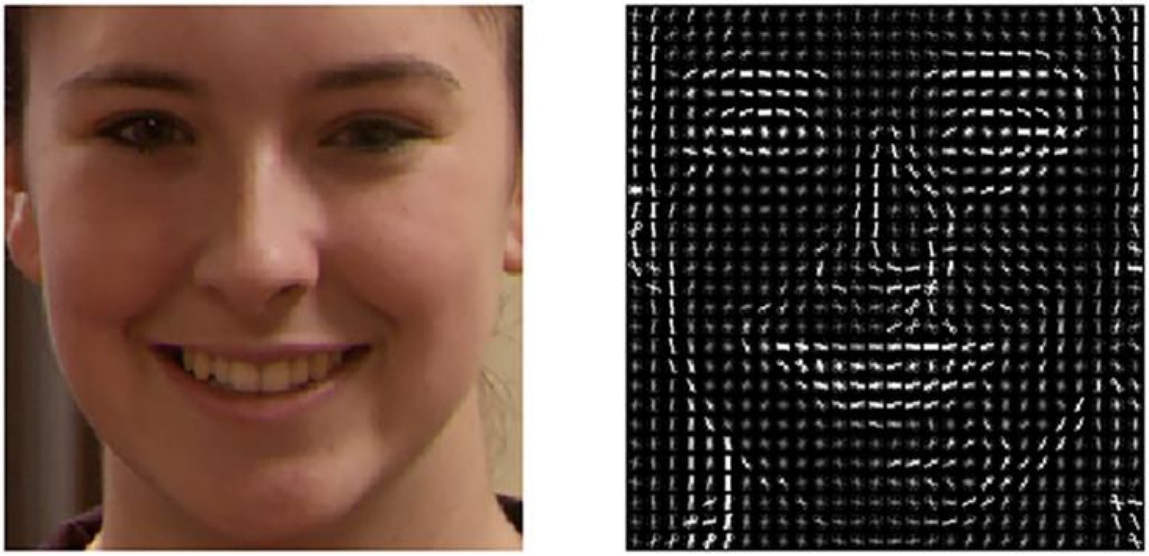

Figure 2. Photography and histogram representation of oriented gradients
(source: https://peerj.com/articles/cs-236/)

Each photo is a set of pixels. When a HOG is created, each pixel is analyzed in the following way:

1. We "look" at each pixel together with the adjacent pixels.
2. "We draw an arrow" to the darkest adjacent pixel.

The biggest advantage of a gradient is that regardless of the brightness of the image, it does not change. Depending on the system performance, the image is analyzed in larger or smaller blocks. The image processed in this way will be used in the face recognition and dlib libraries in order to find the face.

The next step is to create 128 values with the help of a convolutional neural network. This model was previously trained on a group of test faces, so that the newly "delivered" face should always return almost the same vector. The vectors encoded in this way are saved to the database and then the appropriate person is assigned to them (Kumar et al., 2020; Chantzis et al., 2021) (Rohlf, 2011; Bolle et al., 2008; Ślot, 2008).

The last step in recognition of the person's identity is comparing previously saved vectors. If the distances between them are not significant, we can determine with a high degree of confidence who the person currently being analyzed is.

The information collected in this way is encoded and then sent using the XBee module to the server, which saves it in the database. The data has been formatted in such a way to provide as much information as possible in one string. Individual parameters are separated by a previously defined separator (Li et al., 2011; Wechsler, 2009; Silhavy, 2020). The data we provide is:

1. Identifier of the recognized person.
2. The distance that separates it from the entries assigned to it.
3. Identifier of the second closest recognized person.
4. The distance that separates it from the entries assigned to it.
5. Number of entries in the database that were correctly assigned before the incorrect categorization occurred,
6. (In the case of working on photos) Name of the tested file.

All these parameters allow for even more accurate recognition and efficient testing of new entries, people, etc. The following algorithms are a very good illustration of how the program works (Figure 3).

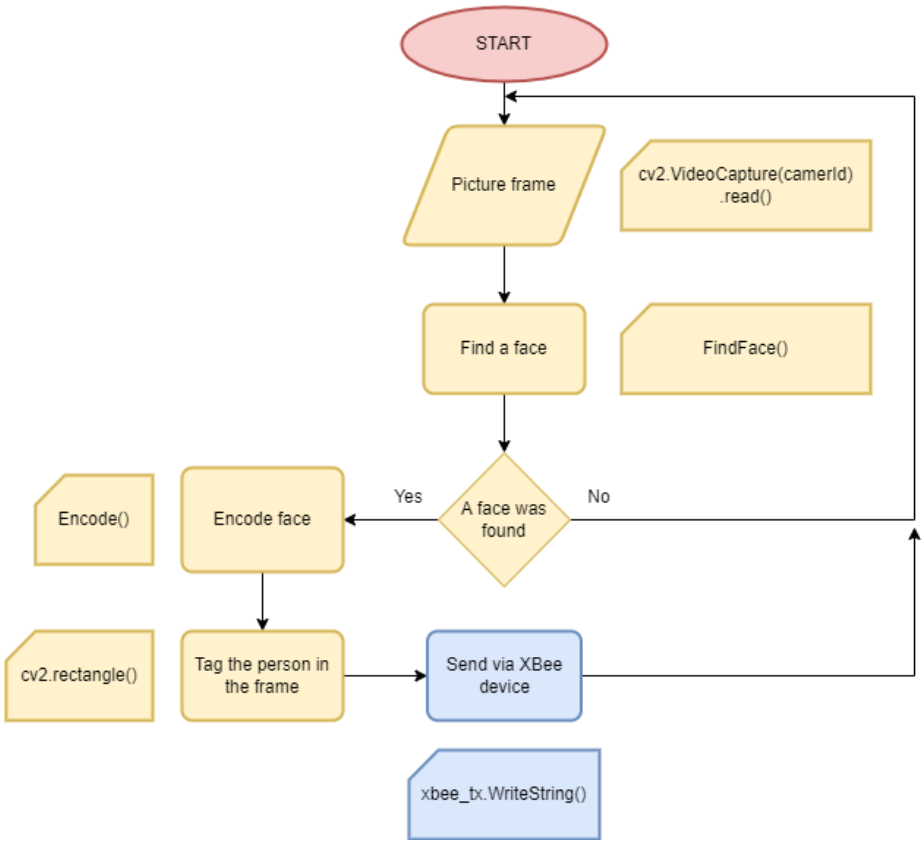

Figure 3. The general operating algorithm

At the moment of receiving the information, STM32 checks what time interval has occurred between the successive identical entries. If it is greater than a minute, information is saved to the database, otherwise it goes to the beginning of the program (Figure 4).

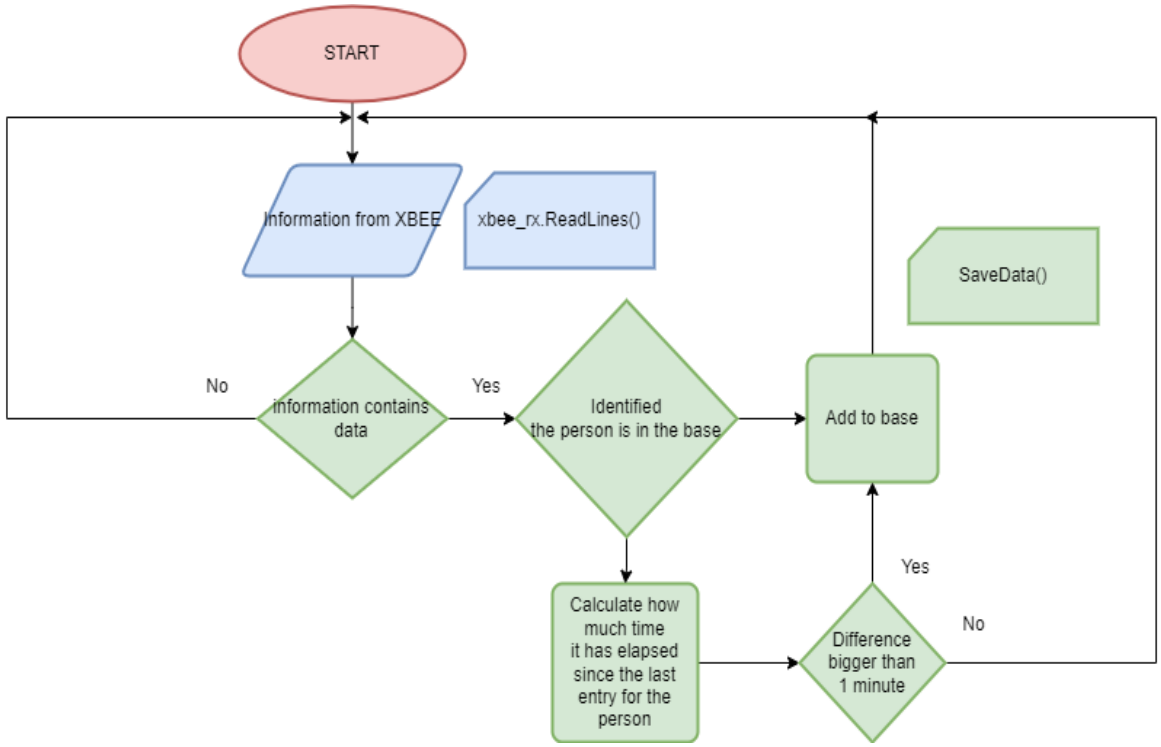

Figure 4. The algorithm of the program that puts information into the database

5. Analysis and test of the developed system

The system created for demonstration purposes provides a basis for creating interesting new functionalities. The example scenario will illustrate many possibilities of its application. For example, when a person saved in the database enters the house, we can adjust the temperature inside to individual preferences of the person entering, and when the person leaves the building, an economic heating or cooling plan will be activated. When it comes to inventing new scenarios, the only limit is the creativity of the creators.

In order to create a database of people to be used for research purposes, photographs of famous people downloaded from the Internet were used. The data of ten people were collected. Each entry in the database was generated from the ten photographs.

The tests were conducted with the help of two to five photos. To make it easier to test the application's capabilities, it was adapted to work with photos. This allows both the creation of a larger base of people and faster testing. The program has several modes and depending on the needs, one can run it with different parameters. Two prepared web interfaces present the test results. The first view shows the records from the database. The information is stored in five columns. The first one gives the date and time when a person was spotted. If someone is in the field of view of the camera for less than 60 seconds, the program will register his/her presence only once. Otherwise, their presence will be registered every minute. During this time the program can easily record the presence of other, non-repeating people. The next column stores the file name in case the application is testing with photo files. This facilitates later analysis of the collected information. The third column stores information about who was recognized, in this case in the photo. The next value is the distance. As described earlier, the distance tells us how similar the person is to the person to whom it was assigned. The last column

is the rating. It displays the percentage of certainty in a given analysis and the previously described ratio of photos identified before the appearance of the most similar person (Table 1).

Table 1
Web interface showing entries from the database

File	Who	Distance	Rate
target1.jpeg	Jessica Alba	0.1654938	93.333% (14/15)
target2.jpeg	Angelina Jolie	0.3052380	90.0% (9/10)
target3. jpeg	Jcssica Alba	0.32S9543	100.0% (15/15)
target11.jpg	Jessica Alba	0.3727370	93.333% (14/15)
target12.jpeg	Jessica Alba	0.3812532	93.333% (14/15)
target13.jpeg	Jessica Alba	0.4219757	93.333% (14/15)
target14.jpeg	Jessica Alba	0.3664828	93.333% (14/15)
target20.jpeg	Unknow Person	0.6071871	0% (0/n)
target31.jpeg	Angelina Jole	0.3454402	90.0% (9/10)
target32.jpeg	Angelina Jole	0.4112590	100.0% (10/10)
target4l.png	Unknow Person	0.5779084	0% (0/n)
target42.jpeg	Unknow Person	0.5631111	0% (0/n)
target43.jpeg	Ellie Sattler	0.1669919	100.0% (10/10)
target44.jpeg	Ellie Sattler	0.3519967	100.0% (10/10)
target5l.jpeg	Claire Dearing	0.3827268	100.0% (9/9)
target52.jpeg	Claire Dearing	0.4382213	77.778% (7/9)

The second feature of the web application is the recognition statistics view. This view presents the minimum, average, and maximum distance for each person's entries in the form of a bar graph. The range of values is from 0 to 1, where 0 means the distance calculated from the photo or image frame equal to the one stored in the database, which may suggest that a photo that is already in the database was used. The given value indicates the distance between the analyzed image and the nearest corresponding record. Based on it, several conclusions can be made.

The first one is the discrepancy between the minimum and maximum value. This value suggests that a wide variety of photos/frames were given to the test. The system operated on relatively easy (low minimum value) and difficult (high maximum value) images. The average itself allows to determine the specific coefficient of difficulty level. The closer its value is to zero, the easier the data were to recognize.

The second one is as follows. If a person has maximum values close to the threshold level indicated by the yellow line in the graph, it means that there are entries in the database that are on the verge of being marked as unknown. The recognition threshold is fixed and based on testing, but it is certainly not perfect. These two facts suggest that an entry should not be trusted 100 percent. On the other hand, low minimum values suggest that there are cases that are easy to classify. One can assess whether the database

is well learned by checking whether the entry "Unknown Person" appears in the graph and how large the bars are. The smaller the minimum, maximum, and average values, the better the learned base is because the distances to known images are small (Figure 5).

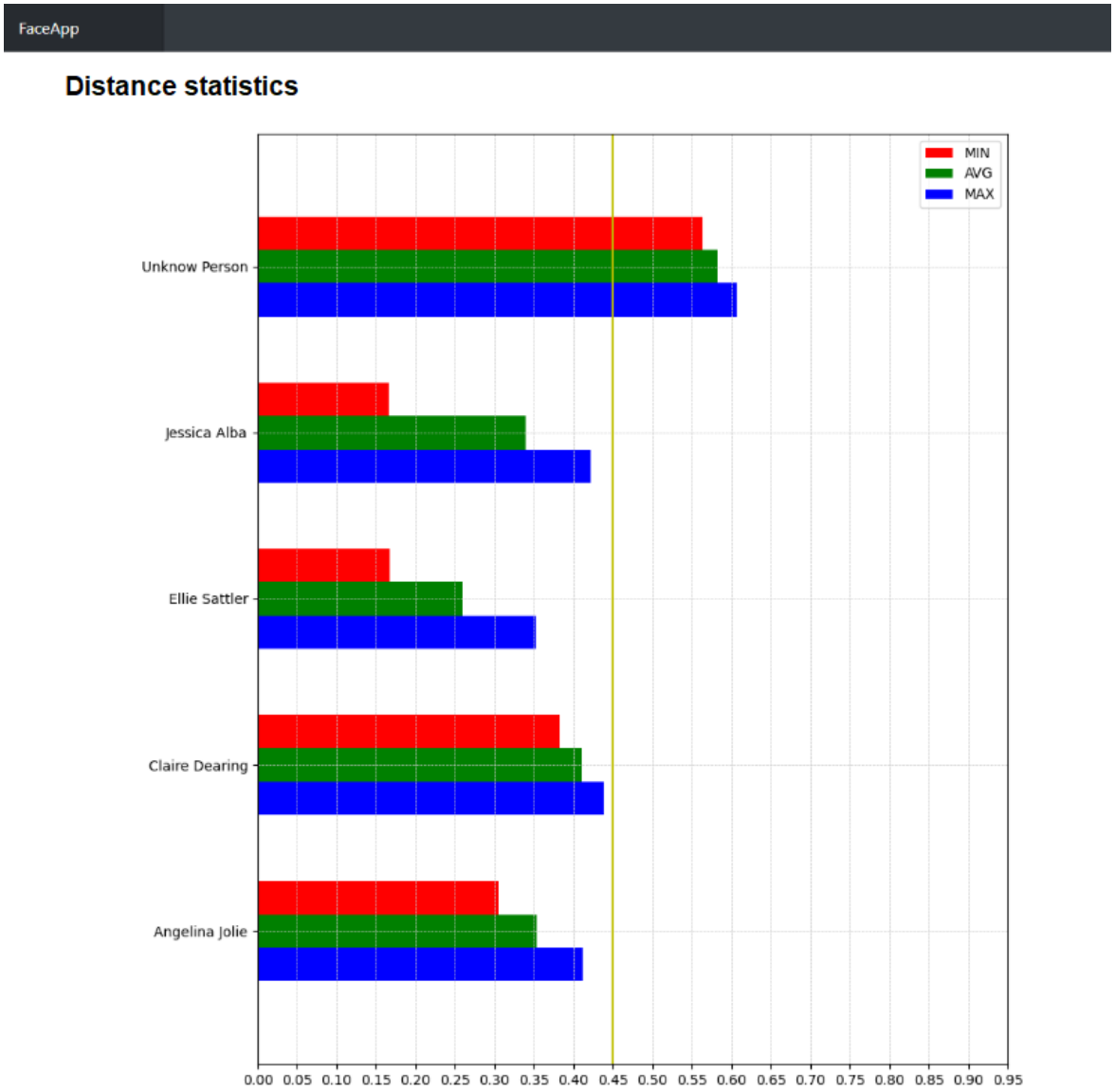

Figure 5. Distance statistics

6. Conclusions

Due to the conducted research, we can conclude that the effectiveness of the program is quite high. The created database of famous people allows to recognize them easily. It could be observed that in some samples the evaluation coefficient had the value of 14/15, which gave the fact that one photo from the group of 15 known people allowed for wrong assignment of a given person.

As can be easily seen, the real-time face recognition requires adequate performance. Raspberry Pi 4 in the version with 2 GB of RAM at the most optimal settings, provides the ability to analyze approximately one frame per second. Although this value is not very high, however, it proves to be sufficient in many applications. Attempts to run the

system on other Raspberry microcomputers were made, but their recognition performance was unacceptable. The same system was also tested on Raspberry Pi Zero W, Raspberry pi 3, Raspberry Pi 4. In the case of the Raspberry Pi Zero W, the frame refresh rate was significantly below 1 frames per second [FPS]. For the Raspberry Pi 3, the image refresh rate with the recognition script was around 2-3 FPS, which was still unsatisfactory. The Raspberry Pi 4 approached the value of 5 FPS, which can be considered the absolute minimum if we want to properly carry out the face recognition process and we don't want to miss any information. In this case, due to the use of a popular single-board computer, the simplicity of the solution in relation to the achieved effect is very satisfactory. IoT is something that was once portrayed as science fiction in literature. Full automation of cyclical processes (and more) is slowly becoming our everyday life. The electronics we interact with almost all the time allow us to have constant control over our other devices, homes, cars, etc. The Internet of Things is with us in our homes, workplaces and other places. The ability to control lighting and heating with a smartphone is becoming the norm.

References

Ahmed, B. (2018). *Secure and Smart Internet of Things (IoT)*. Denmark: River Publishers, Gistrup.

Alam, M., Ara, K., Khan, S. (2020). *Internet of Things (IoT)*. Springer Nature.

Arora, S. Bhatia, M.P.S. (2021). Challenges and opportunities in biometric security: A survey. *Information Security Journal: A Global Perspective, 30,* 1-21.

Bolle, R.M., Connell, J.H., Pankanti, S. (2008). *Biometria*. Warszawa: WNT Scientific and Technical Publishers (in Polish).

Chantzis, F., Stais, I., Calderon, P.; Deirmentzoglou, E., Woods, B. (2021). *Practical IoT Hacking*. San Francisco: No Starch Press.

Chen, X., Yan, H., Yan, Q., Zhang, X. (2020). *Machine Learning for Cyber Security*. Springer Nature.

Hancke, P.G., Markantonakis, K. (2016). *Radio Frequency Identification and IoT Security*. 12th International Workshop, RFIDSec Hong Kong, China, November 30 -- December 2, 2016.

Kumar, S., Bhushan, B., Narayan, C.N. (2020). *IoT Security Paradigms and Applications*. Abingdon: Taylor & Francis Group.

Li, Z.S., Jain, K.A. (2011). *Handbook of Face Recognition*. Springer Science & Business Media.

Rohlf, J.F. (2011). *Biometry*. W.H. Freeman, Gordonsville.

Saravanan, V., Anpalagan, A., Poongodi, T., Khan, F. (2021). *Securing IoT and Big Data*. Abingdon: Taylor & Francis Group.

Silhavy, R. (2020). *Intelligent Algorithms in Software Engineering*. Springer Nature.

Suganthi, K. Machine (2021). *Learning and Deep Learning Techniques in Wireless and Mobile Networking Systems*. Abingdon: Taylor & Francis Group.

Ślot, K. (2008). *Wybrane zagadnienia biometrii*. Warszawa: WKŁ Transport and Communication Publishers (in Polish).

Veneri, G., Capasso, A. (2018). *Hands-On Industrial Internet of Things*. Birmingham: Packt Publishing.

Wechsler, H. (2009). *Reliable Face Recognition Methods*. Springer Science & Business Media.

Yasuura, H., Kyung, C., Liu, Y., Lin, Y. (2017). *Smart Sensors at the IoT Frontier*. Springer International Publishing.

Ziegler, S. (2019). *Internet of Things Security and Data Protection*. Springer, Cham.