[1.] Adnan RAMAKIĆ, [2.] Zlatko BUNDALO, [3.] Dušanka BUNDALO, [4.] Mirko SAJIĆ

DEEP LEARNING BASED INTELLIGENT SECURITY SYSTEM FOR HOMES

[1.]Technical Faculty, University of Bihać, Bihać, BOSNIA & HERZEGOVINA
[2.]Faculty of Electrical Engineering, University of Banja Luka, Banja Luka, BOSNIA & HERZEGOVINA
[3.]Faculty of Philosophy, University of Banja Luka, Banja Luka, BOSNIA & HERZEGOVINA
[4.]Independent University of Banja Luka, Banja Luka, BOSNIA & HERZEGOVINA

Abstract: In today's world, technology is omnipresent. In many areas of human life, people use some kind of new technology. From smartphones that people use every day to intelligent systems built into the cars. Accordingly, many technologies are also used in homes. Some of these include smart thermostats, appliances, lighting and various voice assistants. In some cases, all of that is implemented as one system. In this paper, some methods that can be used for security purposes in homes and other buildings have been analyzed and studied. The methods mentioned in this case form an intelligent security system based on deep learning. The described example of a security system can be used for both access to the house and access to the outside of the house. In this case, the methods of gait recognition, face recognition and vehicle recognition, i.e., cars recognition, were used and analyzed. Accordingly, three deep learning models were developed and described for the gait, face and vehicle recognition. The analysis, the defined settings and the results obtained in relation to the developed models were also described.
Keywords: Intelligent security system, deep learning, gait recognition, face recognition, vehicle recognition

INTRODUCTION

The use of new technology is inevitable in many areas of human life. Technology is advancing every day, so its use will be even more present in the future. If we take a period of just the last 30 years as a reference, we should notice many advances in technology. One example of this is the smartphone, what people use on a daily basis for many tasks. From basic texting in e–mail correspondence to the use of advanced artificial intelligence tools for translations, image creating and editing, etc. Another interesting example is modern car that is equipped with different types of technologies that support the driver while driving, such as modern advanced driver assistance systems *(ADAS)* or systems for autonomous driving.

Based on the above, even modern houses are not immune to new technology. Modern homes contain various elements such as smart thermostats, heating, ventilation and air conditioning systems (HVAC), appliances, lighting, various voice assistants, etc. In some cases, this is all combined in one system. The security aspect is also crucial for any modern home. Different identification methods can be used to achieve different levels of security in home access. For example, to enter the house, people can enter different combinations of numbers (Personal Identification Number, PIN) or use more advanced methods such as voice or speech recognition, face recognition, etc. The methods mentioned are usually based on certain characteristics, i.e., extracted features, of the human body such as fingerprint, palm, face, voice or speech, eye elements (retina or iris), gait, etc.

This paper is primarily concerned with the aspect of security required for access and entering the house and house yard. In other words, for access and entering the interior and exterior of the property, house or building. Three methods are proposed and used for this purpose. Two of the methods are used for person identification, gait recognition and face recognition. Gait and face recognition are intended for access and entering all parts of the property, both indoors and outdoors, for the persons. The third method is used to recognize vehicles, in this case cars. The mentioned vehicle recognition is intended for access and entering the yard from the street and access and entering the garage. Accordingly, three deep learning models were developed and described in this paper. In connection with the developed models, the analysis, the defined settings and the obtained results were described.

The paper is structured as follows. After the introductory section, the second section describes an example of an intelligent security system for houses and buildings. The architecture of the models, the defined settings and the datasets used are described in the third section. The fourth section describes the results obtained in relation to the models and datasets used. Concluding remarks are made at the end.

EXAMPLE OF AN INTELLIGENT SECURITY SYSTEM FOR HOUSES

An intelligent security system can be realized with different methods and technologies. In this paper, the intelligent security system is based on the deep learning approach. The system is based on two known methods for person identification, face and gait recognition, as well as method for vehicle recognition. The two mentioned methods for people recognition are intended for the identification of persons entering all parts of the property, both indoors and outdoors. In other words, gait and face recognition are used when entering the house, garage or yard from the street. The vehicle recognition is used when a car enters the yard from the street or enters the garage. An example of scheme of a house and adjoining yard with an intelligent security system is shown in Figure 1.

Figure 1 shows an example of a plot of land (property) on which the house (H), garage (G) and yard (Y) are located with associated security elements (cameras). In this example, there are five RGB (Red, Green, Blue) cameras to monitor the entire property. Different cameras with different ranges can be used in this context. RGB–D sensors (Red, Green, Blue – Depth) can also be used. The use of a particular type of camera depends on the methods to be implemented. In other words, if a method requires depth images, for example, it is necessary to install an RGB–D sensor. In this example, methods intended for identification use RGB images, so it is sufficient to install RGB cameras.

The cameras are located at the main entrances to the house, garage and yard. The cameras are labeled *C1*, *C2*, *C3*, *C4* and *C5*. The cameras are used to detect persons and cars and capture images of anyone attempting to enter the house, garage and yard. To gain access to the yard, a person or vehicle must be detected by the camera *C1* located at the main entrance. That means that images should be taken for a person or vehicle. The taken images are used in all used methods, gait, face and vehicle recognition. In this part, all mentioned methods should be implemented. The person should be identified by gait or face recognition methods, while the vehicle should be identified based on vehicle recognition method. If a person or vehicle is positively identified, they can enter the yard from the street.

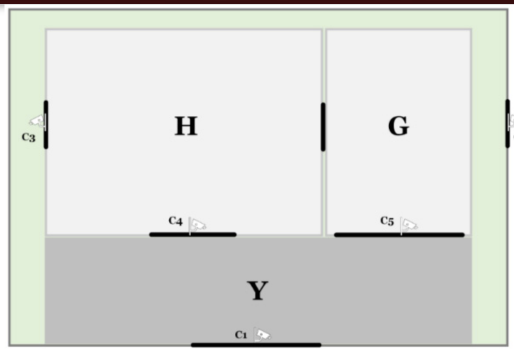

Figure 1. An example of scheme of a house and adjoining yard with an intelligent security system

In addition, persons can enter the property via the side entrance (auxiliary entrance), which is monitored by camera *C2*. This means that persons who want to gain access to the yard from that side must be detected by camera *C2*. In this case, images should also be taken for a person, as in the case of the main entrance and camera *C1*. The only difference in this case is that the side entrance is only intended for people. This means that gait and face recognition should be implemented so that persons can be detected and identified, to enable access and entrance.

To enter the house, persons should use one of two entrants monitored by cameras *C3* and *C4* and must be detected with cameras *C3* and *C4*. In this example, camera *C4* is located at the main entrance of the house, while camera *C3* is located at the side entrance. Gait and face recognition methods should be implemented in this case.

Camera *C5* is used to access and enter the garage. To enter the garage, all methods should be implemented as in the case of access to the main entrance to the property.

If, for example, a certain person wants to enter the house through the main entrance, this person should be detected by the camera *C4*. When the person is detected, the identification process begins. The persons can be identified by their gait or their face. Gait recognition is suitable for identification purposes at a greater distance. If the person approaches the main entrance from a greater distance, it will be identified by their gait, as this method requires more steps to identify the person. In the event of a positive identification, the door at the main entrance is unlocked. If for some reason the person is not identified by the gait, identification is done by the face. Over 90 percent should be the confidence in the identification of a certain person or vehicle to gain access to the house or yard. The default identification method for longer distances is gait recognition, while face recognition is used for shorter distances. If, for any reason, gait recognition is unable to identify a certain person, face recognition begins.

A *Gait Energy Image (GEI)* [5] was used as a method for gait recognition. GEI is an image containing the silhouettes of a person during a gait cycle, where mentioned silhouettes are normalized, aligned and temporally averaged [5]. Examples of the GEI images from *Casia Dataset B* [22][23][12] are shown in Figure 2.

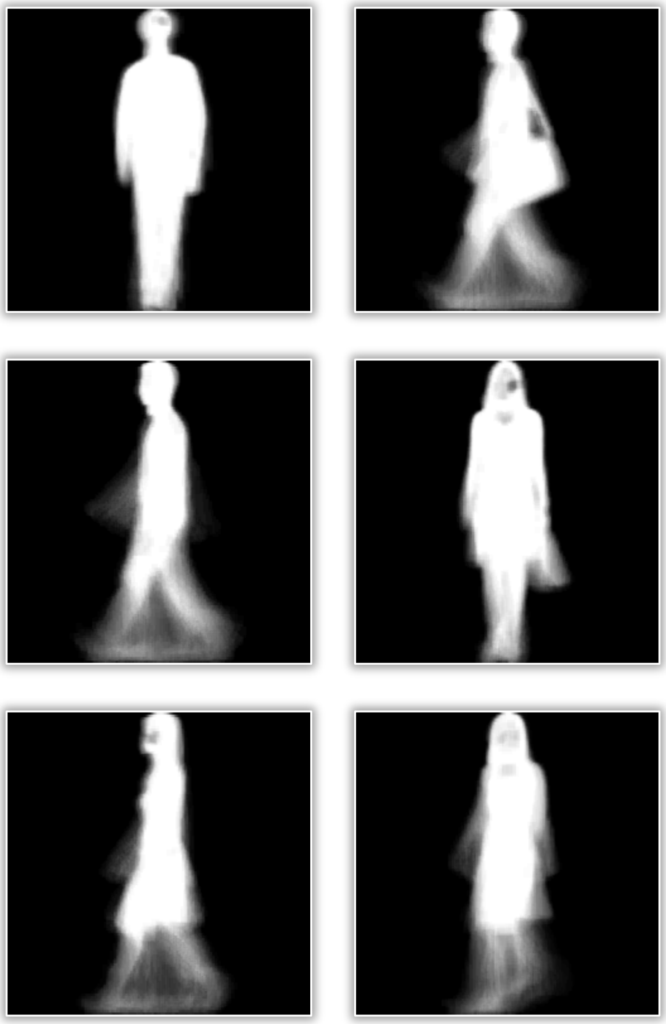

Figure 2. Examples of GEI images (*Casia Dataset B*) [22][23][12]

Approaches to gait recognition [14] can generally be divided into two categories, those based on appearance and those based on models. In appearance–based approaches, silhouettes of the individuals are usually the basis for the methods presented. In model–based approaches, a model serves as the basis for the methods presented. The mentioned model is based on some human body characteristics such as the length of the legs or arms etc. Some interesting works on gait recognition can be found in [1,3,4,6,7,10,13,15–20].

In addition to gait recognition, face recognition is also an interesting method for identifying people. It is a method that is widely used today. This type of method is usually used for smaller distances, but can also be used at a greater distance similar as for gait recognition. This is possible because today there are various cameras with a long range, so that face images can be taken from a greater distance and people can be identified. Compared to gait recognition, face recognition does not require as many steps, making it quicker and easier to implement in this context. Some interesting elements related to face recognition can be found in [2].

In general, an identification system can be roughly divided into two parts. One part is *the identification part* and the other is *the database (dataset) creation part*. In the database creation part, the images for each person and vehicle are captured and stored (image acquisition). The identification part involves capturing a new image of a certain person or vehicle and comparing it with the images stored in the database. In this context, features can be extracted from the images and then the features form the database.

With the rapid development of artificial intelligence and easier access to the tools and platforms for developing machine or deep learning models, the identification process is now usually carried out using machine and deep learning approaches. In this context, a model should be developed to identify individuals and vehicles and a database should be created, i.e., images should be captured for each person and vehicle (image acquisition). The database is used to train and validate the model. To identify a specific person or vehicle, a new image of the person or vehicle must be captured and transferred to the developed model. For the methods used in this paper, this is shown in the Figure 3.

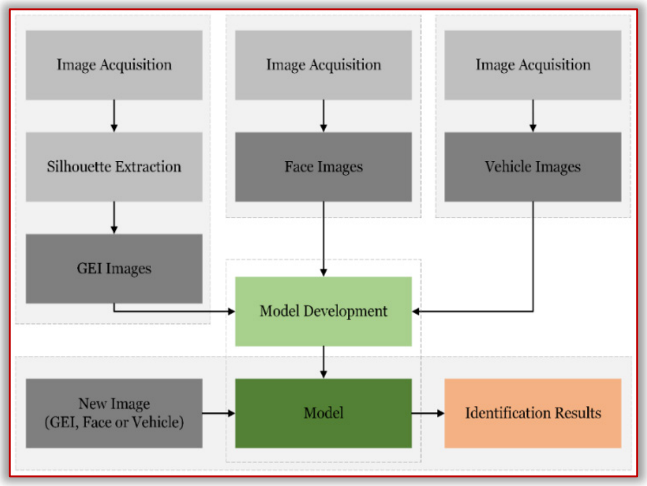

Figure 3. Identification methods based on deep learning

Figure 3 shows how each of the models was developed. In other words, Figure 3 shows the steps required to develop each model. In the case of gait recognition, it is necessary to capture images for each person while walking, i.e. in gait, and then perform image processing. After that, the silhouettes of a person should be extracted from the images, processed, and then GEI images should be created. When GEI images have been created for each person, the database was created. Then, a model for gait recognition should be developed, trained and validated on GEI images. The same process takes place for face and vehicle recognition. The only difference is that in the case of gait recognition, the silhouettes must be extracted from the images and then the GEI images have to be created. In the case of face and vehicle recognition, it is sufficient to capture images of faces and vehicles, respectively. The above text describes the process of database creation and model development for each of the methods. During identification, a new image of a person must be taken and passed to the model so that the person can be identified.

MODEL ARCHITECTURE, SETTINGS AND DATASETS USED

A total of three models were developed, as already mentioned in the text. One model for each of the methods used, i.e., for gait, face and vehicle recognition. The *TensorFlow* [21] platform was used for model development together with *Keras* [8]. From *Keras*, the *Keras Sequential* model was used. All three models consist of a *preprocessing layer*, *convolution layers*, *pooling layers*, *reshaping layer*, *core layers* and *regularization layer* (in models for face and vehicle recognition). Three datasets were used in the development of the models. In the case of gait recognition, the *Casia Dataset B* [22][23][12] was used. *The Facial Images: Faces95* [11] dataset was used for the face recognition model. A separate own dataset with 10 classes, i.e., images for ten different cars, was used for vehicle recognition.

Casia Dataset B, a well–known gait dataset, consists of the images of 124 subjects recorded from 11 views. The subjects in *Casia Dataset B* have a normal gait or have clothing and carrying condition changes. It should be noted that silhouette images and GEI images are also available in *Casia Dataset B*. These images are especially suitable for research and methods based on appearance–based approaches. *The Facial Images: Faces95*, face dataset, contains images of faces of 72 subjects with a resolution of 180 x 200, where the background is a red curtain. Own dataset contains the images of the cars taken from different angles, taking into account all sides of the cars. The images are high–resolution.

From the *Casia Dataset B* and the *Faces95* dataset, 20 subjects (20 classes) were randomly selected and used for the training and validation process of the developed models. On the other hand, all 10 vehicle classes were used from our own dataset for vehicle recognition. All GEI images from the *Casia Dataset B* were used for each subject, taking into account all views and all conditions such as normal gait, clothing and carrying condition changes. A total of 2200 GEI images were used, 110 GEI images for each subject. In the training and validation process, the images were divided so that 80 percent were used for training and 20 percent for validation, for all three models. This means that of this number of GEI images, 1760 images were used for training and 440 images for validation.

In the case of the *Faces95* dataset, 20 face images were used for each subject. A total of 400 images, with 320 images used for training and 80 images for validation. For vehicles, 80 images per class were used, making a total of 800 images. Of this number of images, 640 images were used for training and 160 images for validation. Data augmentation was also performed for the *Faces95* and own vehicle datasets to increase the diversity of the training set. Further training options are 20 *epochs* and the *Adaptive Moment Estimation Optimizer (Adam)* [9] was used.

RESULTS AND DISCUSSION

After analyzing the developed deep learning models for gait, face and vehicle recognition, the following results were obtained. In the case of the model for gait recognition, the validation accuracy for the defined settings and the dataset used, as described in the previous section, was 99,32%. The face recognition model had a validation accuracy of 98,75%. In contrast, the model for vehicle recognition had a validation accuracy of 100%. The results obtained can be seen in Figure 4 and Table 1.

Table 1. Validation Accuracy for Developed Models

Developed Model for Method	Dataset	Validation Accuracy
Gait Recognition	*Casia Dataset B*	99,32%
Face Recognition	*The Facial Images: Faces95*	98,75%
Vehicle Recognition	*Own Dataset*	100%

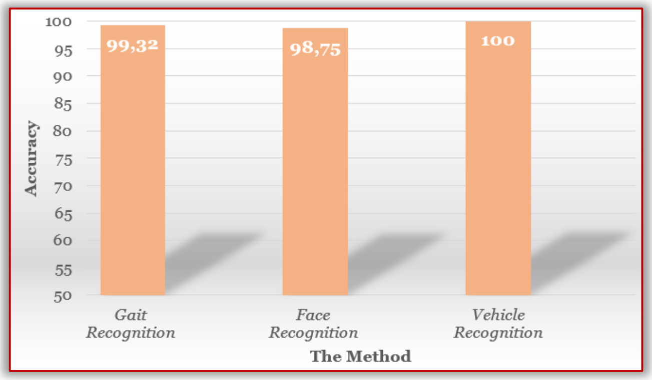

Figure 4. Validation Accuracy for Developed Models

As is shown in Table 1 and Figure 4, all three developed models for gait, face and vehicle recognition achieved high results in terms of accuracy, what is above 98% in all three cases. That means that all three methods are interesting and promising for use in various security systems. Figure 5 shows the training and validation accuracy for all three developed models. The left side of the image represents the training and validation accuracy for the model for gait recognition, the middle part of the image for the model for face recognition and the right side of the image for the model for vehicle recognition.

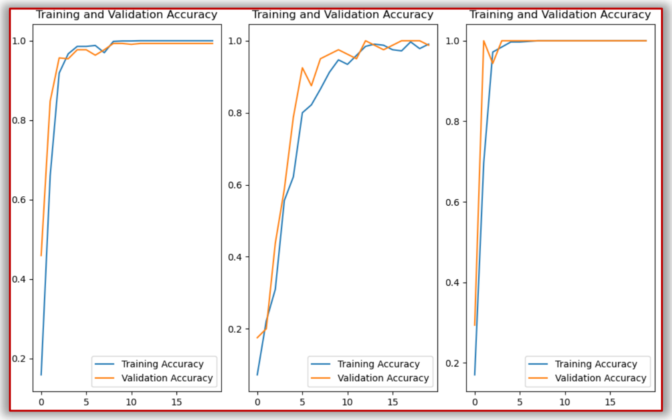

Figure 5. Training and validation accuracy for developed models (left – gait recognition, middle – face recognition, right – vehicle recognition)

As can be seen from the previous text on the models developed and the results presented, all three methods are promising for use in various security systems and applications. The models developed for the methods provided good results in all three cases. Although a relatively small number of classes were used in all three cases (20 for gait recognition, 20 for face recognition and 10 for vehicle recognition), the obtained results are very promising. To support this, the models for gait and face recognition were also trained and validated with a larger number of classes from datasets used, using 50 randomly selected subjects in both cases. The obtained results in terms of accuracy were also high, exceeding 95% in both cases. In the case of gait recognition, different GEI images were also used. This means that the model was trained and validated only with GEI images taken at 0 degrees and with images taken at 90 degrees.

It should also be noted that gait recognition methods such as GEI require more steps to identify a person. This means that RGB images of a person must first be captured during a gait cycle, after which the person's silhouettes must be extracted and then GEI images are created. This is a challenging task that must be performed in real–time, as the person can quickly disappear from the camera's field of view. In the case of face recognition, this is easier as only the face images should be captured. The advantage of gait recognition methods is that a person can be identified over a greater distance and without the person knowing that the identification process is underway. The face recognition is used for smaller distances and is faster. But, a potential problem and disadvantage of face recognition can be changes to a person's face, such as a beard, glasses, wearing a hat, etc. All this should be taken into account when preparing a dataset for the development of a face recognition model.

In the case of vehicle recognition, only the images of the cars are taken into account during model development. This is potentially disadvantageous, as the same car with the same color can enter a house yard. For this reason, the vehicle recognition model should be coupled with a license plate recognition system, i.e. *automatic number plate recognition (ANPR)*, to achieve the highest level of security. In view of the above, it would be interesting to analyze another method of gait recognition instead of the GEI method in future research. It would also be useful to compare the performance of the appearance–based and the model–based method of gait recognition. In particular, it would be useful to obtain results in terms of speed of identification and reliability of identification.

CONCLUSION

An example of an intelligent home security system based on deep learning was analyzed and described in this paper. The intelligent system is based on three methods to identify people or vehicles. These methods are gait, face and vehicle recognition. Gait and face recognition is for identifying people who want to access and enter certain part of the property such as house, garage or yard, i.e., both indoors and outdoors. Vehicle recognition is intended for entering the yard from the street and entering the garage. A well–known method called Gait Energy Image or GEI was used as the method for gait recognition.

In light of the above, three deep learning models were developed for each of the methods. The *TensorFlow* platform with *Keras* was used to develop the models. The *Keras Sequential* model was used. Three datasets were used in the development of the models for the training and validation process. The well–known *Casia Dataset B* was used as the gait dataset. On the other hand, *The Facial Images: Faces95* was used as the face dataset. A separate own dataset with 10 classes was used for the vehicles. The results obtained in terms of validation accuracy were very promising for all three models developed. The results achieved are above 98% for all three models.

In the future, it would be interesting to analyze another method of gait recognition instead of GEI. It would also be useful to compare the performance of the appearance–based and the model–based method of gait recognition, taking into account the speed of identification and the reliability of identification.

References

[1] Arora, P. and Srivastava, S. (2015). Gait Recognition Using Gait Gaussian Image. In: 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 791–794. IEEE.
[2] Barnouti, N. H., Al–Dabbagh, S. S. M. and Matti, W. E. (2016). Face Recognition: A Literature Review. International Journal of Applied Information Systems, 11(4), 21–31.
[3] Bashir, K., Xiang, T. and Gong, S. (2009). Gait Recognition Using Gait Entropy Image.
[4] Chattopadhyay, P., Roy, A., Sural, S. and Mukhopadhyay, J. (2014). Pose Depth Volume Extraction from RGB–D Streams for Frontal Gait Recognition. Journal of Visual Communication and Image Representation, 25(1), 53–63. Elsevier.
[5] Han, J. and Bhanu, B. (2005). Individual Recognition Using Gait Energy Image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(2), 316–322. IEEE.
[6] Hofmann, M., Bachmann, S. and Rigoll, G. (2012). 2.5D Gait Biometrics Using the Depth Gradient Histogram Energy Image. In: 5th International Conference on Biometrics: Theory, Applications and Systems (BTAS), 399–403. IEEE.
[7] Iwashita, Y., Uchino, K. and Kurazume, R. (2013). Gait–based Person Identification Robust to Changes in Appearance. Sensors, 13(6), 7884–7901. MDPI.
[8] Keras. Link: https://keras.io [Accessed 25/03/2024]
[9] Kingma, D. P. and Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
[10] Kumar, M. N. and Babu, R. V. (2012). Human Gait Recognition Using Depth Camera: A Covariance Based Approach. In: Proceedings of the 8th Indian Conference on Computer Vision, Graphics and Image Processing, 1–6.
[11] Libor Spacek's Facial Images Databases. Link: https://cmp.felk.cvut.cz/
[12] Official Web Page of the Institute of Automation, Chinese Academy of Sciences. Link: http://www.cbsr.ia.ac.cn/
[13] Preis, J., Kessel, M., Werner, M. and Linnhoff–Popien, C. (2012). Gait Recognition with Kinect. In: 1st International Workshop on Kinect in Pervasive Computing, 1–4. New Castle, UK.
[14] Ramakić, A. and Bundalo, Z. (2022). Gait Recognition as an Approach for People Identification. In: International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies, 717–726. Springer.
[15] Ramakić, A., Bundalo, D. and Bundalo, Z. (2023). An Approach to Gait Recognition Using Deep Neural Network. Acta Technica Corviniensis–Bulletin of Engineering, 16(2), 1–6.
[16] Ramakić, A., Bundalo, Z. and Bundalo, D. (2020). A Method for Human Gait Recognition from Video Streams Using Silhouette, Height and Step Length. Journal of Circuits, Systems and Computers, 29(7), 2050101. World Scientific.
[17] Ramakić, A., Bundalo, Z. and Bundalo, D. (2023). An Example of Solution for Data Preparation Required for Some Purposes of People Identification or Re–Identification. Journal of Circuits, Systems and Computers, 32(10), 2350164. World Scientific.
[18] Ramakić, A., Bundalo, Z., Bundalo, D. and Vidović, Ž. (2023). Frontal Gait Recognition Based on Deep Learning Approach. Annals of the Faculty of Engineering Hunedoara–International Journal of Engineering, 21(3).
[19] Sivapalan, S., Chen, D., Denman, S., Sridharan, S. and Fookes, C. (2011). Gait Energy Volumes and Frontal Gait Recognition Using Depth Images. In: International Joint Conference on Biometrics (IJCB), 1–6. IEEE.
[20] Sivapalan, S., Chen, D., Denman, S., Sridharan, S. and Fookes, C. (2012). The Backfilled GEI–A Cross–capture Modality Gait Feature for Frontal and Side–view Gait Recognition. In: International Conference on Digital Image Computing Techniques and Applications (DICTA), 1–8. IEEE.
[21] TensorFlow. Link: https://www.tensorflow.org [Accessed 25/03/2024]
[22] Yu, S., Tan, D. and Tan, T. (2006). A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. 18th International Conference on Pattern Recognition (ICPR), 441–444. IEEE.
[23] Zheng, S., Zhang, J., Huang, K., He, R. and Tan, T. (2011). Robust View Transformation Model for Gait Recognition. In: 18th International Conference on Image Processing, 2073–2076. IEEE.

ISSN: 2067–3809