UDC: 004.8; 004.9; 004.93 https://doi.org/10.15407/jai2023.01.049

INFORMATION-ANALYTICAL TECHNOLOGY FOR MONITORING THE FLOW OF VISITORS TO THE UNIVERSITY

V. Hrytsyk[1], O. Babii[2]

[1,2]Lviv Polytechnic National University
28a Bandery St., Building 5, Room 801, 79012
volodymyrhrytsyk@gmail.com
[1] https://orcid.org/0000-0002-9696-5805
[2] https://orcid.org/0000-0003-4155-3923

Annotation. In the 21st century, one of the most widespread problems in developed countries is the unraveling of complex tasks related to the security of citizens. An example is the need to conduct a security check at universities, when at one checkpoint there may be a need to let a thousand people pass within 5 minutes. Inspection of each (even a formal presentation of the document) will lead to the disruption of 4 classes; automated turnstiles will not ensure quality inspection + queues will be created (or will require many turnstiles that will actually be used for a short time). The Covid'19 pandemic only transfers the problem to another plane - a distance of one and a half meters + the risk of infecting the guard, who will turn into a source of infection. Military and, especially, terrorist events (when civil infrastructure objects with a large concentration of civilians become the targets of attacks) in Ukraine show the need to simultaneously ensure high throughput and for people and the safety of the object itself. The paper considers the concept of impersonal monitoring of the number of visitors. A safe approach is considered, when a recognition system based on the use of artificial neural networks allows checking and accompanying a large number of people impersonally at the same time. The system is implemented as a pattern recognition technology with statistical analysis. The system (visualization in the figures in the text) was tested on the video streams of the security cameras of the main building of the Lviv Polytechnic. The purpose of the work is the first phase of testing the hypothesis of the possibility of impersonal verification by using several impersonal classifiers. In the work, people are recognized not by their faces, but by a large set of parameters that allow classifying a person, but not identifying them.

Keywords: digital, statistical data, analysis, monitoring of university attendance.

Introduction

The purpose of the work is to increase the accuracy of statistical data collection in a way that does not interfere with work, does not create traffic jams, and avoids problems of leakage/use of personal data. This is especially important during a terrorist threat and potential air attacks on civilian objects, where there are many people with a need to respond quickly to the threat and quickly return to work. The data describe events in the visible range of the spectrum. Data analysis makes it possible to obtain real-time information about attendance and occupancy of certain areas of the university.

Analysis of recent research and publications

Object recognition is one of the most important areas of computer science and cybernetics, as well as the most important result of deep and machine learning algorithms. People easily recognize people, objects, scenes, and visual details when viewing a photo or video. The goal is to teach computers to do what humans can do naturally – to understand what an image contains in order to further enable reactions such as classification [8,9] or hardware reactions [7].

In Ukraine today, there are two serious problems that are related to the need to analyze the movement of people's flows – a terrorist threat (the war for independence) and a biological threat (the COVID-19 virus [18]). Both threats involve the prevention of mass gatherings and computer monitoring of the state of gathering places and nodal/key points. Both threats can be monitored using automatic recognition and classification systems in the visible spectrum. And, these two problems cannot be solved by the usual biometric methods [17, 20-21] due to the need to stop the movement – hence the additional threat of crowding.

The used methods largely simulate the cognitive, diagnostic and classification abilities of a person [14-16]. A set of various tasks for identification in digital images and video streams is usually called the process of human recognition. Typically, the process is that the system takes an image from a camera, then uses an algorithm to change the color gamut (e.g. [19]), brightness, alignment, scaling, etc., thereby transforming the input stream into a final given shape that allows the desired operations to be performed. . After that, the methods of searching for moving objects, their cracking and further processing are used. People recognition can be used in the context of institutions that receive large flows of people, including higher education institutions. Monitoring and further processing of this data can reveal attendance statistics and patterns that depend on external factors, one of which today is not only the COVID-19 virus but also the serious challenge of the terrorist threat, which has undoubtedly affected all areas of life, and especially those , which are closely related to visiting places of large gatherings of people. Therefore, the development of a method and the creation of a corresponding software solution that can be used to analyze the flow of people is a promising and expedient task for today.

Problem formulation

The task of the work is research and optimization of a technique that can process video streams, classify people, determine the direction of their movement and automate the collection of university statistics. The task of recognition and classification involves working with the visible range of the spectrum.

Materials and methods

Today, "standard" specialized problems require the selection of an optimal combination of methods for an effective solution: it is necessary to take into account the properties of hardware, external factors, features of objects, etc. To test the effectiveness of each individual location, the software solution must be carefully adjusted with respect to many factors, which include lighting, the scale of the frame, the number and texture of the surrounding elements, etc. Accordingly, the mathematical model on which such technology is based must take into account all parameters. The test was carried out using Python, the OpenCV library. In addition, the source of the video stream is needed (Fig. 1): a stream from an IP surveillance camera, prepared video or connected video equipment.

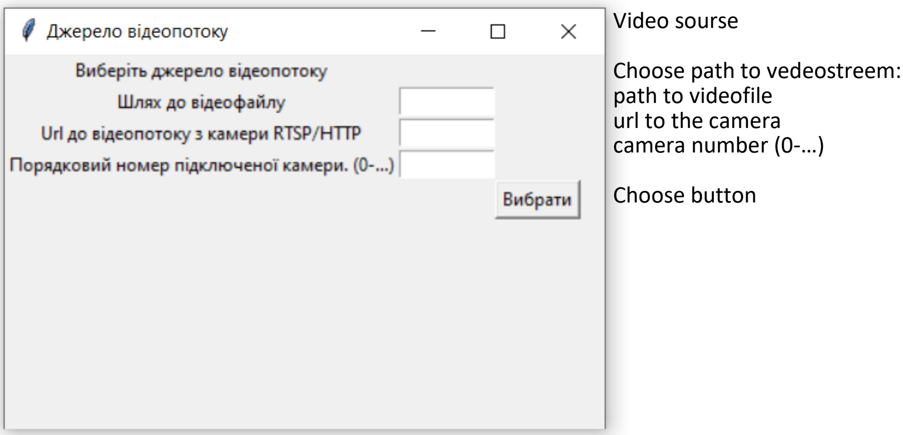

Fig. 1. Interface window for selecting the source of the video stream

The KNN algorithm is the first algorithm that is used in the system. It is a background (foreground) segmentation algorithm based on the nearest neighbor method [1,10]. It is applied by lazy learning algorithm because there is no special training phase and it uses all the data for training during classification. Also, a non-parametric learning algorithm is used because there is nothing about the input baseline data.

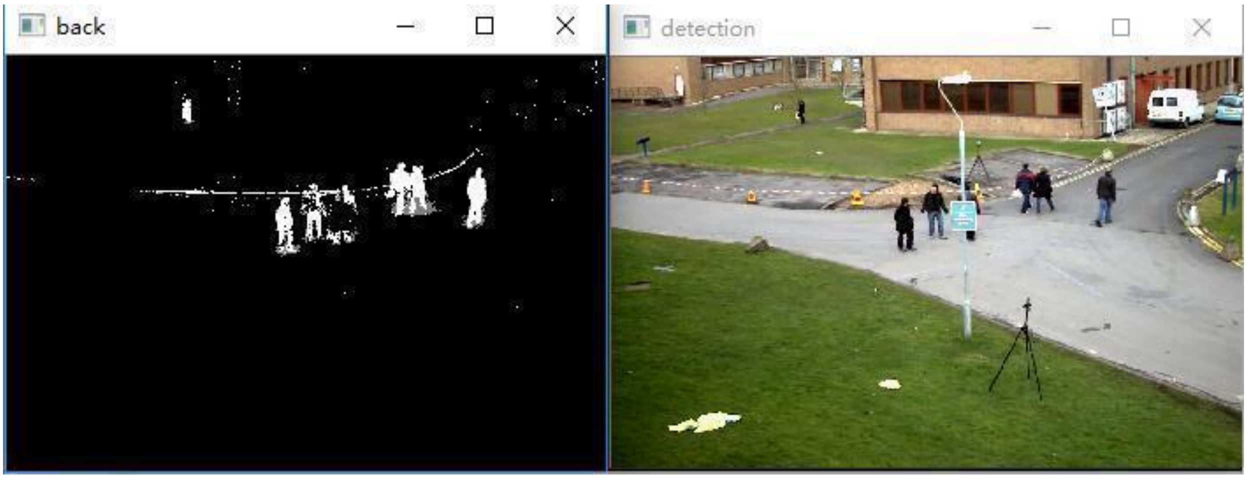

Fig. 2. An example of KNN algorithm operation on a digital video stream

The stages of KNN work are as follows:

1. Receive unclassified data.
2. Measuring the distance (Euclidean, Manhattan, Minkowski or weighted) from new data to all other data that is already classified.
3. Gets K (K is a parameter you define) smaller distances.
4. Checks the list of classes with the shortest distance and counts the number of each class that appeared.
5. Accepts as the correct class, which appeared the largest number of times.
6. Classifies the new data with the class you took in step 5.

Let's consider an example of the algorithm.

Below, we have an image (Fig.3) with the whole process, the input is unclassified data (in red) and all your other data that is already classified (yellow and purple), each with its own class (A or B). So you calculate the distances of your new data with all the others to know which ones have the smallest distance, so you get the 3 (or 6) closest data and see which class appears the most, in the case of the image below, the closest data to the new data are those inside the first circle (inner circle) and inside this circle there are 3 more data (already classified in yellow), we will check which class is predominant there, see it is purple because there are 2 purple balls and only 1 is yellow, so this is new data that was not classified before, it will now be classified as purple.

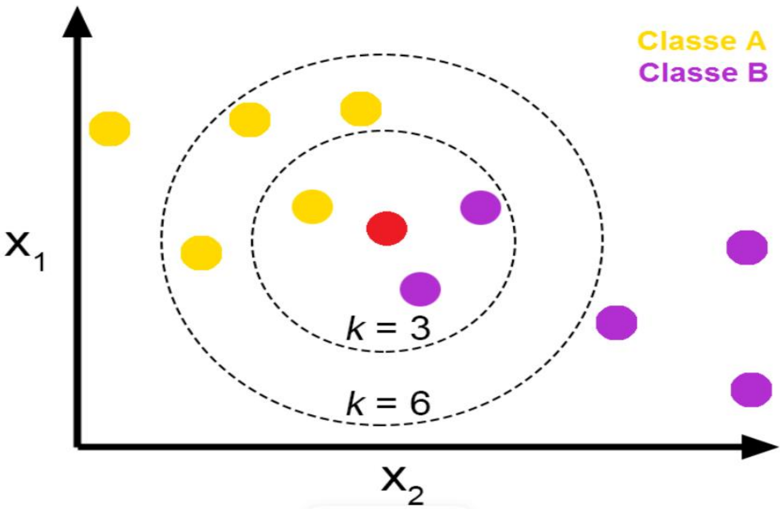

Fig. 3. Classified/unclassified data

Distance calculation.

The authors review additional, frequently applied techniques [12,13] for estimating quality and distance in image processing. There are various ways used to determine the distance between two points, but the most popular is the mapping associated with metric space is the distance function

$X \times X \to R$, which assigns each pair of points x and y. (distance from x to y). he most common distance functions encountered in image processing are Euclidean distance, city block or diamond distance and chessboard distance, which are defined as

$$d(x, y) = [\sum_{k=1}^{n}(x_k - y_k)^2]^{\frac{1}{2}},$$

$$p(x, y) = \sum_{k=1}^{n}|x_k - y_k|,$$

$$\delta(x, y) = mdx\{|x_k - y_k|: 1 \leq k \leq n\}$$

Point rules make it simple to compute the distance (norm). From the normative L^p rules, we propose the following

$$\|x\|_p = \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$$

The rule (norm) L^∞ is given (defined through):

$$\|x\|_\infty = \bigvee_{i=1}^{n}|x_i|,$$

Where

$$\bigvee_{i=1}^{n}|x_i| = max\{|x_1|, \dots, |x_n|\}$$

For a particular assignment, the Euclidean distance rule (also known as the Euclidean norm) is given as

$$\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

In summary,

$$d(x, y) = \|x - y\|_2.$$

Similar to determining the distance using the city block method, you can calculate the distance using the formula

$$p(x, y) = \|x - y\|_1$$

And the distance according to the chessboard distance rule using RRR

$$\delta(x, y) = \|x - y\|_\infty$$

Keep in mind that the $\| \quad \|_\infty: X \to \mathbb{R}$. function serves as the p-norm of the point x, which is a unary (single, one-component) operation. For the positions P(2, 3, 1) and Q(-3, 2, 2), the formula for distan ce in n-dimension, or 3-dimensional space, will be as follows:

$$\sqrt{(-3 - 2)^2 + (2 - 3)^2 + (2 - 1)^2} = 5.196$$

Following calculations, you must choose k sample items with the shortest distances. The class of the object that has to be identified is the one that manifests itself most frequently among its k nearest neighbors. As a background reader, this method was created to find moving items in the mask that could be people. As seen in the image below (Fig.4), it is intended to put the camera primarily above the steps, so further processing is not required to identify items.

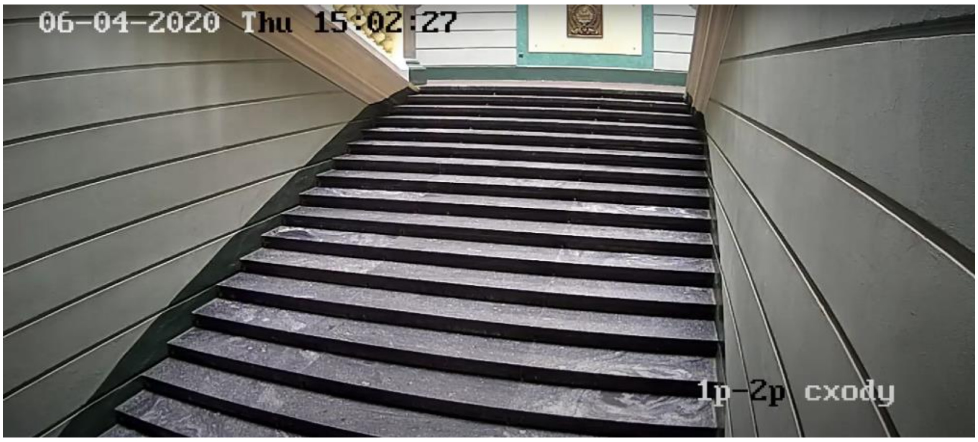

Fig. 4. An example of a video stream source installed in the system

Image processing

Setting threshold values

For every pixel, the same threshold value is used. The pixel value is set to 0 if it is less than the threshold; otherwise, it is set to the maximum value. Use the cv.threshold function to apply threshold values. The original image, which must be in grayscale, is the first defense. A threshold value is used as the second argument to categorize pixel values. The maximum value that is assigned to pixel values that are higher than the threshold is the third argument. The fourth parameter of the function specifies the various threshold settings that OpenCV offers. The cv.THRESH BINARY [2] type is used to implement the fundamental threshold value as it was previously stated.

Morphological transformations

We send the background reader object to the OpenCV morphologyEx() library's standard methods to process the image using morphological transformations.

Simple procedures that are based on the shape of the image are known as morphological transformations. Typically, they are carried out on binary images[3]. It requires two inputs: our source image as the first, and the structuring element or kernel as the second, which establishes the purpose of the operation.

A multidimensional vector with the indices "0" and "1" with a size of (n, n) is the structuring element or kernel. The size of the element is determined by the dimension of this vector, and its linear shape is determined by the positions of "0" and "1" in the vector. Although it is possible for the element's starting point to be outside the element, it is recommended that it be in the matrix's center. The goal form that we need to extract from the processed image determines the kernel's size and shape. [11] These components can be produced utilizing the OpenCV or Numpy libraries (Fig. 5).

11111 11111 11111 11111 11111	00100 00100 11111 00100 00100	0001000 0011100 0111110 1111111 0111110 0011100 0001000
Square, dimension 5	Cross, dimension 5	Rhombus, dimension 5

Fig. 5. Types of structural elements

Erosion and dilatation are two essential procedures in mathematical morphology. These operators are coupled linearly in various techniques. The kernel corresponds to the image if each pixel in the original image has a value of 1. The element is considered to be in the image if there is at least one pixel that meets the aforementioned criteria.

Assuming that L is a complete set, supremum and infinimum are denoted by ∧ and ∨, respectively, and that universal set and empty set are defined as U and ∅, respectively.

In this case, the erosion operation can be defined:

$$\varepsilon : L \to L \mid \bigwedge_i \varepsilon(X_i) = \varepsilon\left(\bigwedge_i X_i\right), \varepsilon(U) = U$$

Dilatation (expansion) operation is defined as follows:

$$\delta : L \to L \mid \bigvee_i \delta(X_i) = \delta\left(\bigwedge_i X_i\right), \delta(\emptyset) = \emptyset$$

The Open operation, which comprises first calling the erosion operator and then the dilation operator on the generated image, is used in the project's initial step (Fig. 5). In this case, the MORPH OPEN argument is used with the cv.morphologyEx() function.

```
opening = cv.morphologyEx(img, cv.MORPH_OPEN, kernel)
```

Fig. 6. The result of the disclosure

The system then executes the Closure function, which entails first invoking the dilation operation and then the erosion operation on the output image.

```
closing = cv.morphologyEx(img, cv.MORPH_CLOSE, kernel)
```

Closing is the opposite of opening, expanding followed by blurring. It is useful for closing small holes inside foreground objects, or small black dots on the object. In our case, this method also connects white foreground areas (Fig.7).

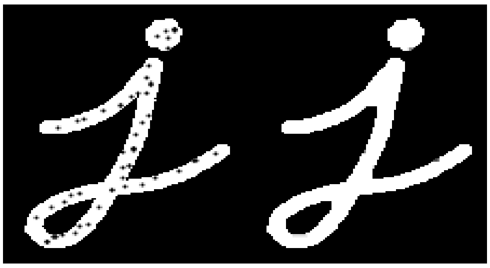

Fig. 7. The result of the closure execution

Tracking of objects

After image processing using the

KNN algorithm, it is important to identify the contours in order to track foreground objects. For this, the method findContours() with the parameter RETR_EXTERNAL is utilized. The function merely identifies the pixels that have a noticeable variation in color when it determines the borders of the input image [10]. The majority of contours are an abstract collection of segments and points that represent the reflected forms of objects found in the photos the system has analysed. As a result, the programs [2–6] that allow access to contours allow for their manipulation.

There are various ways to accomplish this, including counting the number of contours in an image and using that information to categorize the forms of objects, segment photos, and perform a variety of other similar tasks.

With the RETR EXTERNAL argument, the findContours() method only looks for contours at the third level, the highest level of search. Below is a description of the function in action. [4]

Fig. 8. Found contours in the hierarchy with the parameter RETR_EXTERNAL

Analysis of the method of determining the action of entering or leaving the visitor

Every time a person is recognized in the frame, the system generates a special identifying number for that person. The frame is divided into 4 sections that are connected by 4 lines. Red and blue lines serve as the real confirmation of a person entering or leaving the focus area (which in our case is responsible for entering and exiting the territory of the second floor). Additionally, there are two auxiliary lines that serve as a "start" to establish the motion of the recognized object. For instance, if object A crossed the lower auxiliary line, the program will start counting the number of people who could potentially enter at that point. Additionally, it counts as a "visit" if a person who has already been localized and is being followed by the program (tracking) crosses the upper main line, adding another entry unit to the counter.

To graphically display the logical component of this algorithm, we will use the UML activity diagram [5].

Activity diagrams are visual representations of organizational processes (workflows), iterations, stepwise or choice-supported activities, and data flows that cross over with the actions connected to them. Activity diagrams may include features that simulate the movement of data between activities through one or more data stores, even if their primary purpose is to depict the overall flow of control.

The activity diagram to determine the action of entering or leaving the visitor will look like this:

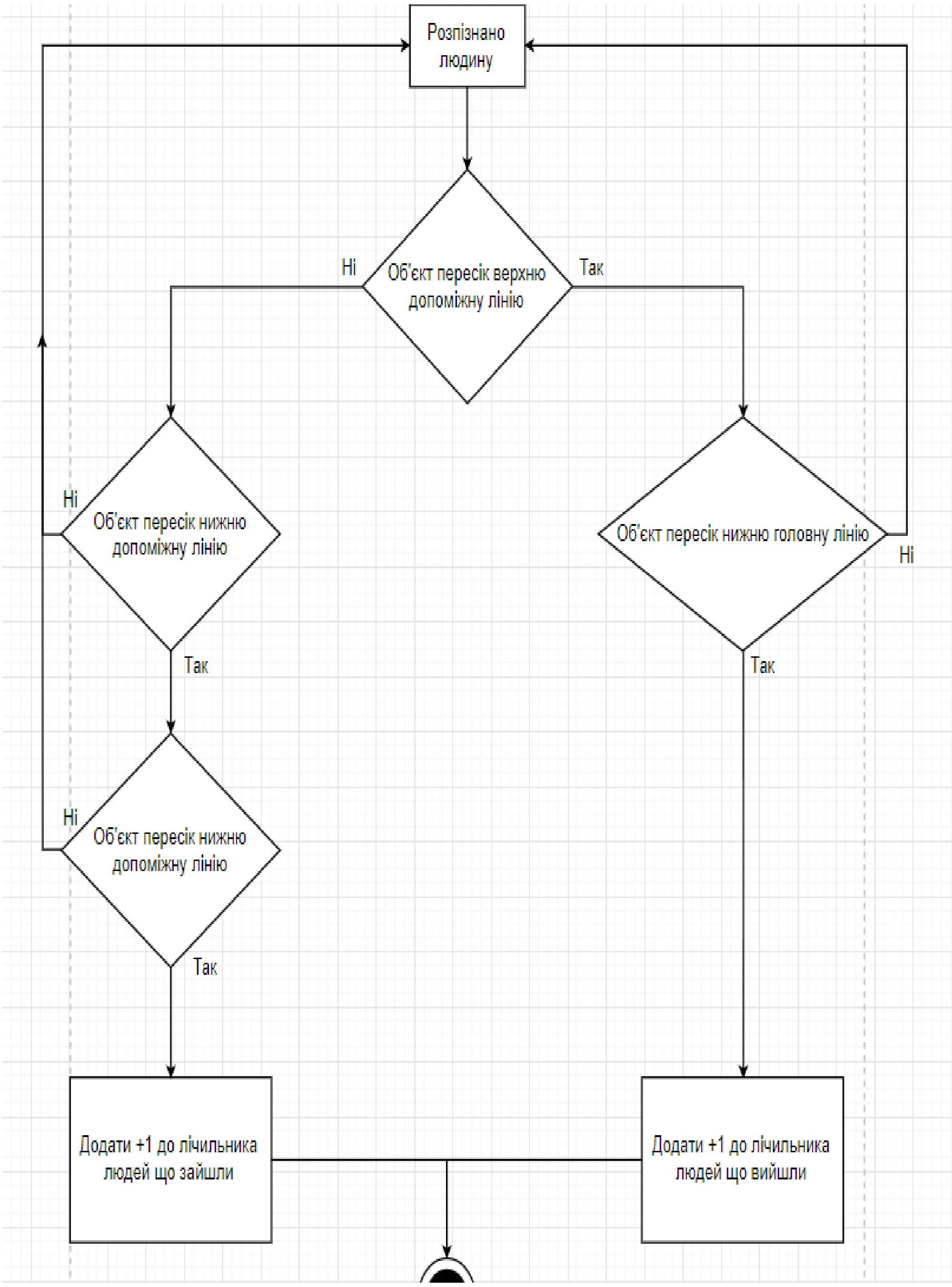

Fig. 9. Algorithm for determining the action of entering or leaving the visitor

A Person class object that is stored in the database is created for each moving object. Each object has coordinates that can be used to build a track in the frame, find the intersection of the same auxiliary lines, and determine the object's direction of movement.

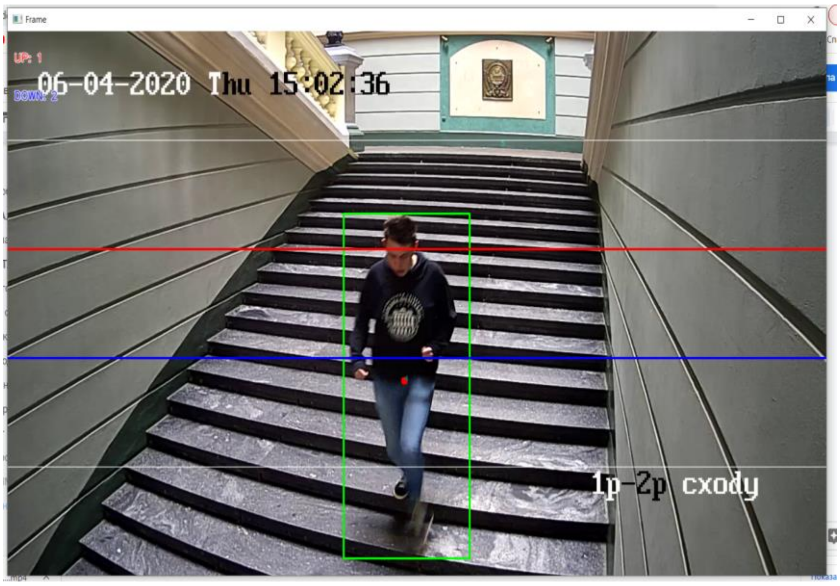

Fig. 10. An example of human recognition and tracking down the stairs

Also, the system is able to combine several objects that go side by side into a group, this data can also be processed further. An example of merging can be seen below:

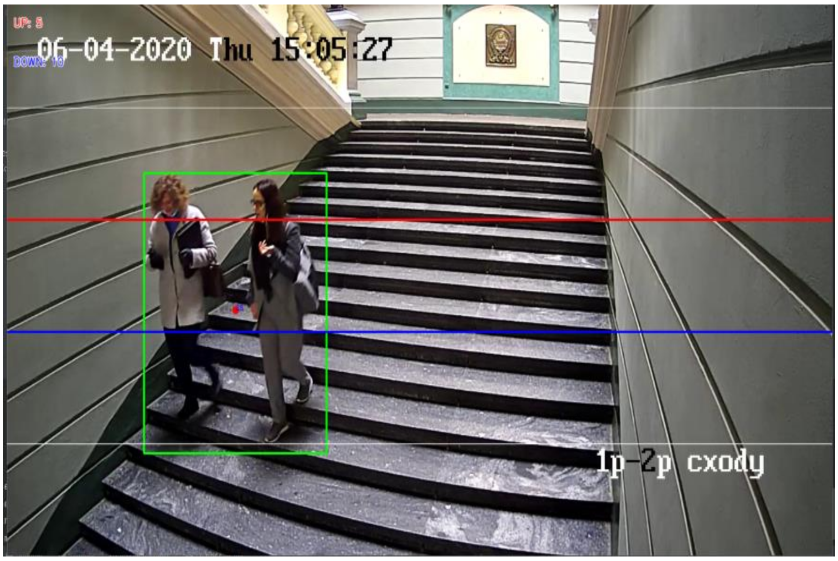

Fig. 11. Two women grouped together by the program

Logging

Additionally, by writing to a text file in the txt format, the system records every movement of every object in the frame. The file is created at the beginning of the process and may be processed whenever is necessary. [6] The file has the form:

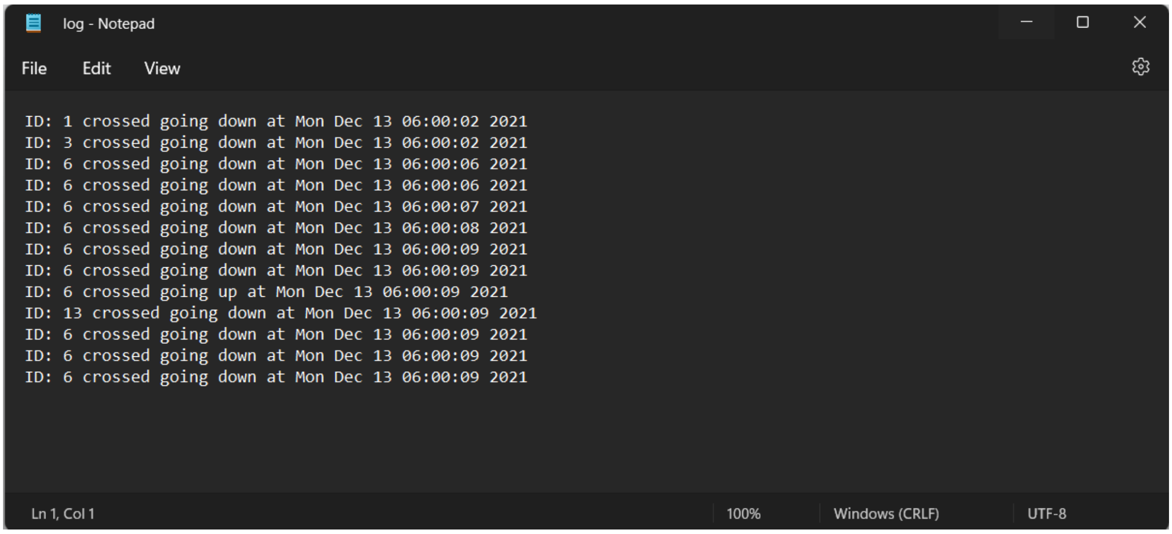

Fig. 12. File with system operation logs

Testing

It was decided to run the face recognition of 10 people in order to test the system, analyze the data, and assess how accurate the software was. The surveillance camera's video file was selected as the testing source. File containing system operation logs, Fig. 12.

Results:

- The first person is correctly recognized and correctly enrolled in the entry/exit counter.
- The second person is correctly recognized and correctly enrolled in the entry/exit counter.
- The third person is correctly recognized and correctly enrolled in the entry/exit counter.
- The fifth person is correctly recognized and correctly enrolled in the entry/exit counter.
- The sixth person - was recognized incorrectly, the program detected the second person nearby, reacting to the shadow, 2 output values were recorded in the counter, that is, false data.
- The seventh person is correctly recognized and correctly enrolled in the entry/exit counter.
- The eighth person is correctly recognized and correctly enrolled in the entry/exit counter.
- The ninth person is correctly recognized and correctly enrolled in the entry/exit counter.
- The tenth person is correctly recognized and correctly enrolled in the entry/exit counter.

Conclusion

As a result of the work, an information and analytical system was implemented for the selection of statistical data from the visible range of the spectrum. The system, in addition to the statistical method itself, includes procedures for recognition, classification and localization with object management (tracking). An information and analytical system for monitoring the flow of visitors to the second and third floors of the university based on machine vision was created and tested.

The project exists in its first version, which is an experimental base with the possibility of expanding tasks. For example, the technique was tested at the main entrance, which allows you to make only a general statistical picture. All the entrances to the second floor can be monitoring by installing such monitoring systems at each entrance and synchronize the cameras, and then we will get a complete information picture. Another example of improvement is the addition of the ability to identify objects – in this case, it will be possible to keep not only general statistics, but also statistics on specific objects. Of course, there are many opportunities to expand the capabilities of the system that works on the core of recognition.

References

1. https://www.sciencedirect.com/science/article/abs/pii/S0167865505003521 (Zoran Zivkovic. Efficient adaptive density estimation per image pixel for the task of background subtraction // 2005.
2. https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html
3. https://docs.opencv.org/4.x/d9/d61/tutorial_py_morphological_ops.html
4. https://docs.opencv.org/4.x/d9/d8b/tutorial_py_contours_hierarchy.html
5. https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-activity-diagram/
6. https://www.studytonight.com/python/python-logging-in-file
7. Yu.V. Krak. An approach to analytical construction of the equations of motion of robot manipulators// January 1994, Springer, Journal of Mathematical Sciences 71(4):2606-2610. DOI: 10.1007/BF02111966
8. KRAK, Yu.V., et al. Podkhod k opredelenyiu эffektyvnыkh pryznakov y syntezu optymalnoho polosno-razdeliaiushcheho klassyfykatora dlia эlementov daktylno-zhestovoi rechy. Kybernetyka y systemnыi analyz, 2016, 52,№ 2: 3-10.
9. V.Hrytsyk, Andrii Grondzal, Andrij Bilenkyj. Augmented reality for people with disabilities // CSIT'2015. – IEEE. – Lviv. – 14-17 Sept. – p. 188-191.
10. Volodymyr Hrytsyk, Mariia Nazarkevych. Real-Time Sensing, Reasoning and Adaptation for Computer Vision Systems. // ISDMCI 2022: Lecture Notes in Computational Intelligence and Decision Making, Springer, Cham pp 573-585. https://www.mathworks.com/help/images/morphological-dilation-and-erosion.html
11. Hrytsyk V.V, Hrytsyk V.V. Osnovni otsinky yakosti zobrazhennia pry rozviazuvanni zadach avtomatychnoho opratsiuvannia obraziv // NU «LP». – 2014. – №783: Informatsiini systemy ta merezhi. – s. 82-92.
12. Hrytsyk V.V, Hrytsyk V.V. Osnovni otsinky yakosti zobrazhennia pry rozviazuvanni zadach avtomatychnoho opratsiuvannia obraziv // Visnyk Natsionalnoho universytetu «Lvivska politekhnika». – 2014. – № 805 : Informatsiini systemy ta merezhi. – s. 106–113.
13. V.Hrytsyk, A.Hrondzal. Model audio-vizualnoho spryiniattia// MK – ISDMCI2015. - Conference Proceedings. - Kherson: KNTU, 2015. – 51-53 p.
14. V.Hrytsyk, S.Krivtsov. Doslidzhennia kohnityvnykh vlastyvostei shtuchnykh neironnykh merezh // NAN U. In-t problem modeliuvannia v enerhetytsi. – Modeliuvannia ta informatsiini tekhnolohii. – Vyp.№72. – 2014. – s.124-195.
15. Krak, I., Kuznetsov, V., Kondratiuk, S., ... Barmak, O., Padiuk, P. Analysis of Deep Learning Methods in Adaptation to the Small Data Problem Solving // Lecture Notes on Data Engineering and Communications Technologiesthis link is disabled, 2023, 149, pp. 333–352 (book chapter).
16. Nazarkevych, M., Hrytsyk, V., Kuza, A., Shevchuk, O., Kostiak, M. The Impact of the Covid-19 Pandemic on Students Studying in High Education Institutions // CEUR Workshop Proceedings, 2021, 3188, pp. 187–196.
17. Hrytsyk, V.; Medykovskyy, M.; Nazarkevych, M. Estimation of Symmetry in the Recognition System with Adaptive Application of Filters. *Symmetry* 2022, *14*, 903. https://doi.org/10.3390/sym14050903
18. Nazarkevych, M., Hrytsyk, V., Kostiak, M., Parkhuts, L., Nazarkevych, H. Biometric Protection Information System with Extension of Segmentation Methods // CEUR Workshop Proceedings, 2021, 3188, pp. 43–52.
19. Nazarkevych, M., Voznyi, Y., Hrytsyk, V., ...Havrysh, B., Lotoshynska, N. Identification of biometric images by machine learning // 2021 IEEE 12th International Conference on Electronics and Information Technologies, ELIT 2021 - Proceedings, 2021, pp. 95–98.

The article has been sent to the editors 22.10.22.
After processing 29.11.22.