National Research University Higher School of Economics

as a manuscript

Pavlova Anastasiya Dmitrievna

Effective face recognition methods and models based on sequential analysis of neural-net features and automatic detection of minority classes

PhD Dissertation Summary

for the purpose of obtaining academic degree

Doctor of Philosophy in Computer Science

Academic Supervisor:

Doctor of Technical Sciences

Savchenko Andrey Vladimirovich

Nizhny Novgorod - 2025

DISSERTATION TOPIC

Relevance of the topic. Face recognition in photographs and videos is a relevant applied and scientific task. Despite long-term research on face recognition and identification, there are still a number of cases in which existing algorithms demonstrate low accuracy results. For example, low quality of input data (different scale, low brightness, low resolution), age-related changes, changes in appearance and etc. Moreover, in the modern world there is a tendency to use various filters and masks, especially in social networks, which greatly change a person's appearance. If the system receives a wide variety of images of the same person at the input, which are very different from the reference image, in other words, anomalous images, then there is a high probability of a false decision of the algorithm.

The problem of anomalous (rare) data becomes especially acute in the context of face recognition, where the accuracy of detecting faces from all classes is critical, such as in the areas of security and prosecution [1]. Insufficient training of models on rare classes can lead to serious consequences, such as missing faces of dangerous criminals or wrongful accusation of innocent individuals.

Currently, convolutional neural networks (CNNs) have shown significant progress in face recognition. They represent an image as a descriptor (a vector of features) to identify the most important characteristics of the image [2]. However, when working with descriptors, their size must be taken into account. Some CNNs have high-dimensional feature vectors at the output, which raises a new problem related to the linear complexity of decision-making algorithms depending on the number of people to be recognized [3].

There are ways to improve computational efficiency, such as reducing the dimensionality of features or excluding some data from the analysis, but both involve data loss which can lead to a decrease in the accuracy of the algorithm. To solve these problems, the presented dissertation proposes a sequential decision-making method.

It is also worth noting the problem of computational efficiency when working with video sequences, especially when building modern biometric identification and video surveillance systems [1, 3, 4, 5]. Video files can be large and contain lots of frames. Faces can appear in frames, disappear, move, change their viewing angles and lighting. Processing and analyzing such data require large computing resources and efficient algorithms.

The aim of the work: to improve the efficiency of face recognition from images by identifying rare data in the input flow and sequentially analyzing neural network descriptors.

The tasks of the work:

1. Develop a neural network model for detecting anomalous images to identify faces.
2. Implement a computationally efficient image recognition algorithm through sequential analysis of neural network descriptors.
3. Develop an efficient method to organize images of people for video surveillance systems in the “unsupervised” mode based on the selection of homogeneous tracks and aggregation of video frames.
4. Implement the proposed methods for recognizing faces in images and video sequences as a software suite.

KEY RESULTS

The scientific novelty of the work is as follows:

1. For image classification (open-set classification), a neural network model for detecting anomalous images is proposed based on a fully connected neural network trained on a specially created set of anomalous photographs of faces, including people of different ages, different races, and images with filters applied. Unlike the existing approaches to identifying anomalies, the proposed model allows for high-precision detection of not only certain types of anomalies (presence of noise, poor quality), but also more diverse types of rare images (age change, appearance change, corrupted image, etc.).
2. A new efficient computational algorithm based on sequential analysis of image features sorted by significance has been developed.
3. For a collection of face video data, an efficient method of organizing images of people is proposed using the aggregation of descriptors of similar video frames.

The main provisions submitted for defense:

1. A neural network model for detecting anomalous images for face recognition is proposed. An automatic method for generating rare images has been developed, with which a sample of face images with various anomalies was collected and labeled, used to train a neural network model.
2. A new algorithm based on the sequential analysis of neural network descriptors is proposed, which allows increasing the computational efficiency of nearest neighbor search methods up to 5 times.
3. An efficient method for organizing frames from video sequence using aggregation of frames is proposed, which ensures a low rate of incorrect assignment of a frame to a group of images of the same face.

The personal contribution of the author consists in carrying out the main theoretical and practical studies presented in the dissertation. The scientific supervisor A.V. Savchenko is responsible for setting the tasks, and A.D. Pavlova is responsible for all the results obtained.

PUBLICATIONS AND APPROBATION OF RESEARCH

The results of the work are presented in 10 printed publications.

First-tier publications:

1. Sokolova, A. Organizing Multimedia Data in Video Surveillance Systems Based on Face Verification with Convolutional Neural Networks / A. Sokolova, A. Savchenko, A. Kharchevnikova // in van der Aalst, W., et al. Analysis of Images, Social Networks and Texts. AIST 2017. Lecture Notes in Computer Science. Springer, Cham. – T. 10716. – 2018. – pp. 223–230. (Scopus, Q2). – https://doi.org/10.1007/978-3-319-73013-4_20.
2. Sokolova, A. Fast Nearest-Neighbor Classifier based on Sequential Analysis of Principal Components / A. Sokolova, A. Savchenko // in van der Aalst, W., et al. Analysis of Images, Social Networks and Texts. AIST 2019. Lecture Notes in Computer Science. Springer, Cham. – T. 11832. – 2019. – pp. 73–80. (Scopus, Q2). – https://doi.org/10.1007/978-3-030-37334-4_7. – The main co-author.

Second-tier publications:

1. Sokolova, A. Open-Set Face Identification with Sequential Analysis and Out-of-Distribution Data Detection / A. Sokolova, A. Savchenko // 2022 International Joint Conference on Neural Networks (IJCNN). IEEE Xplore. – 2022. – pp. 1–8. (Core B). – The main co-author.
2. Sokolova, A. Computation-Efficient Face Recognition Algorithm Using a Sequential Analysis of High Dimensional Neural-Net Features / A. Sokolova, A. Savchenko // Optical Memory and Neural Networks (Information Optics). – 2020. – T. 29(1). – pp. 19–29. (Scopus, Q3).
3. Sokolova, A. Search for rare data in the problem of face recognition in images / A. Sokolova, A. Savchenko, S. Nikolenko // Computer Optics. – 2022. – T. 46(5). – pp. 801–807. (Scopus, Q3). – The main co-author.

Other publications:

1. Sokolova A. Clustering of video sequences in video surveillance systems based on convolutional neural networks / A. Sokolova, A. Savchenko // Proceedings of the XXIII International Scientific and Technical Conference “Information Systems and Technologies-2017”. – 2017. – pp. 870–875.
2. Sokolova, A. Data organization in video surveillance systems using deep learning / A. Sokolova, A. Savchenko // Proceedings of the IV International Conference on “Information Technology and Nanotechnology” (ITNT-2018). CEUR Workshop Proceedings. – T. 2210. – 2018. – pp. 243–250. (Scopus). – The main co-author.
3. Sokolova A. Data organization in video surveillance systems based on deep learning technologies / A. Sokolova, A. Savchenko // Collection of works of the IV International Conference and Youth School “Information Technologies and Nanotechnologies” (ITNT 2018), Enterprise “New Technology”, Samara. – 2018. – pp. 946–952.
4. Sokolova, A. Cluster Analysis of Facial Video Data in Video Surveillance Systems Using Deep Learning / A. Sokolova, A. Savchenko // in Kalyagin, V., et al. Computational Aspects and Applications in Large-Scale Networks. NET 2017. Springer Proceedings in Mathematics and Statistics. – T. 247. – 2018. – pp. 113–120. (Scopus). – https://doi.org/10.1007/978-3-319-96247-4_7.
5. Sokolova, A. Effective face recognition based on anomaly image detection and sequential analysis of neural descriptors / A. Sokolova, A. Savchenko // Proceedings 9th IEEE International Conference on Information Technology and Nanotechnology (ITNT 2023). IEEE Xplore. – 2023. – pp. 1–5. (Scopus). – The main co-author.

Approbation of the work. The results of the work were reported at the following events:

- International scientific and technical conference “Information systems and technologies-2017”, Nizhny Novgorod, 2017.

- IV International Conference and Youth School “Information Technologies and Nanotechnologies” (ITNT 2018), Samara, 2018.
- International summit “Computer Vision and Deep Learning summit”, Moscow, 2018.
- International Conference on Image, Social Media and Text Analysis AIST19, Kazan, 2019.
- Conferences with international participation “Mathematical Methods of Pattern Recognition” (MMPR), Moscow, 2019.
- 4th visiting seminar on machine learning of the Faculty of Computer Science of the National Research University Higher School of Economics, Moscow, 2019.
- International summit “Computer Vision and Deep Learning summit”, Moscow, 2019.
- Machine Learning Workshop 2019 (Huawei), Sochi, 2019.
- 10th International Conference on Network Analysis NET2020, Online, 2020.
- IJCNN IEEE World Congress on Computational Intelligence, Padua, Italy, 2022.
- IX International Conference “Information Technologies and Nanotechnologies” ITNT-2023, Samara, Russia, 2023.

Winning competitions:

- Machine Learning Workshop 2019 (Huawei) for the best poster – 1st place.
- NRU HSE NRWS-2019 in the nomination “Best research paper in computer science for master's students and graduates of 2019” – 2nd place.
- NRU HSE NRWS-2018 in the nomination “Best research paper on Business Informatics for master's students and graduates of 2018” – 3rd place.

The study was conducted within the framework of the scientific and educational group “Analysis of multimedia data of mobile device users” and

within the framework of a grant from the Russian Science Foundation (project № 20-71-10010).

CONTENTS

The dissertation consists of an introduction, 3 chapters and a conclusion.

In the first chapter, "Overview of the Subject Area", an analysis of existing approaches to image recognition was performed: the use of neural networks, the application of various classification and clustering methods, and the detection of anomalous images.

There are popular classifiers that perform well in distributing images into the necessary groups: K-Nearest Neighbors method (k-NN), the support vector machine (SVM), random forest, linear discriminant analysis (LDA), neural networks for image classification. However, the listed classical approaches may have the following drawbacks: with large amounts of data and high-dimensional features, calculating distances between all objects can be computationally expensive, or large amounts of labeled data may be required for training, and assumptions about the normal distribution of data must be made. But one of the main problems in the field of face image recognition is the diversity of input images that are not present in the training set [6]

In the absence of a sufficient amount of labeled data for training, clustering methods can be used. The work considered: K-Means method, hierarchical clustering, DBSCAN [7], graph-based clustering methods [8], auto associative clustering methods [9], image grouping based on imitation learning [10]. The main problem with these approaches is that a pre-defined number of clusters or sensitivity to the choice of hyperparameters is required. However, in most cases, it is unknown how many unique people are present in the sample.

It is also worth noting the computational efficiency of the listed classification and clustering approaches, since in the modern world real-time solutions are often required. And the computational efficiency of some approaches can significantly depend on the characteristics of the data, such as the number and presence of "rare" (out-of-distribution) objects.

To improve the quality of face recognition, it is possible to apply an approach to detecting rare data [11, 12] to bring the input data to a normal distribution. For example, calculating PSNR and/or SSIM and comparing with a predetermined threshold, using an autoencoder, One-Class SVM, a one-class neural network [13], data cleaning using stochastic gradient descent [14], confidence assessment based on the Mahalanobis distance, using a confidence estimate, transfer learning, using generative adversarial networks. But the disadvantage of many methods is that they help to detect only certain image distortions, whereas it would be useful to know how to detect all anomalies. Moreover, at the present time, there are no datasets that cover all types of anomalies and on which the listed algorithms could be trained.

To reduce the computational complexity of image recognition methods, it is often recommended to optimize the architecture or tune the hyperparameters, but for many traditional approaches these methods are not feasible. Therefore, it is worth paying attention to working with input data. An analysis of existing approaches to data compression was presented: Principal Component Analysis, Linear Discriminant Analysis, t-Distributed Stochastic Neighbor Embedding (t-SNE) [15], Uniform Manifold Approximation and Projection (UMAP) [16], autoencoder, nonlinear dimensionality reduction methods, factor analysis, sparse coding, feature selection. The most optimal approach is the principal component method, which can be used to highlight the most significant features of an image descriptor regardless of the data features.

In the second chapter, “Face Recognition Methods with Automatic Detection of Minority Classes”, a neural network model for detecting anomalous images for face recognition was proposed (Figure 1).

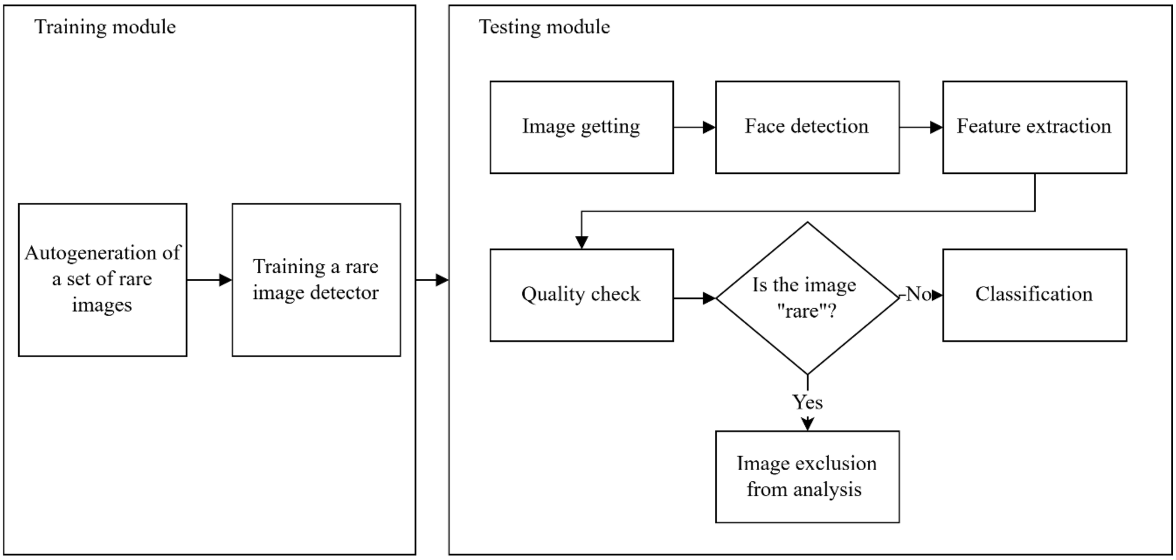

Figure 1 – Architecture of the proposed neural network model for detecting anomalous images for face recognition

To train and test the model for detecting anomalies, an approach was proposed to create a dataset with rare images, which includes 167 500 images. To create such a dataset, various methods were used (Figure 2): using the OpenCV library and deep models for image transformation, including images from existing datasets that display different age groups, including images in which the classifier demonstrates more false positives.

An experimental study of the following rare data detection methods was conducted on the collected set of anomalous images:

1) Training binary classifiers (InceptionV3, EfficientNetB3, MobileNetV1, Xception, ResNet50) - overfitting occurs: high scores on the training set and low scores on the test set.

2) Using a generative model GAN [17] – 70% accuracy in identifying rare images.

3) Data cleansing [14] – 73% accuracy in identifying rare images.

4) Using known classifiers (k-NN, SVM, Random Forest, LDA), which accept image descriptors obtained using known convolutional neural networks as input (InsightFace [18], VGGFace [19], ResNet50 [20], MobileNetV1 [21],

FaceNet [22], LCNN [23]) – accuracy varies from 64.8% to 92.4% depending on the classifier and CNN.

5) Training binary classifiers (InceptionV3, EfficientNetB3, MobileNetV1, Xception, ResNet50) - overfitting occurs: high scores on the training set and low scores on the test set.

Figure 2 – Examples of images from the rare image dataset

It was found that only a small number of existing approaches for detecting anomalous images can satisfy the accuracy requirements, so a new method for detecting anomalous images using a fully connected neural network was proposed [5, 11], consisting of three fully connected layers, a Relu activation function, and two Dropout layers.

Unlike existing approaches to detecting anomalies (binary classifier, using stochastic gradient descent and others), the proposed method allows detecting not only certain types of test distributions of anomalies, but more diverse types of rare images (age change, appearance changes, etc.) with a sufficiently high accuracy, 94.8% using MobileNetV1.

The next stage of the study considered traditional classification methods using information about anomalies detected in the input image. The analysis involved image classification algorithms from the scikit-learn library: k-NN, SVM, Random Forest, LDA. It is proposed for the experiment:

- Image classification without considering rare data information;
- Image classification based on rare data information: rare images are excluded from the analysis – the decision on them is postponed [24];
- Image classification considering rare data information: rare image descriptor transformation is applied [25]. The transformation consists of subtracting the average quality vector from the rare image feature vector. The quality vector is obtained by finding the average difference between the normal and rare images.

The neural network feature vector is used both for making the final decision in the classification methods and for detecting rare input images using proposed model. The experimental study showed that the K-Nearest Neighbors with exclusion of rare images demonstrates the highest accuracy rates using MobileNetV1 / InsightFace (Table 1) / ResNet50 for descriptor extraction, 97.9% / 97.9% / 92.4% respectively for the VGGFace2-based dataset and 90.5% / 92.5% / 91.9% for the MS1M-based dataset.

To improve the computational efficiency of the classical K-Nearest Neighbors classification approach, a method of sequential analysis of neural network descriptors was proposed (Figure 3) [5, 26, 27].

Table 1 – Classification results on VGGFace2-based dataset using InsightFace to extract feature vectors

Algorithm	Accuracy (%)	Classification time for one image (ms)
k-NN + all	93,6	13,8
k-NN + exclusion	**97,9**	5,9
k-NN + transformation	93,7	14,5
Random Forest + all	24,5	3,0
Random Forest + exclusion	63,7	2,6
Random Forest + transformation	24,5	3,7
LDA + all	89,4	**0,2**
LDA + exclusion	96,7	0,9
LDA + transformation	89,8	0,9
SVM + all	74,3	1,3
SVM + exclusion	82,7	1,2
SVM + transformation	75,2	1,4

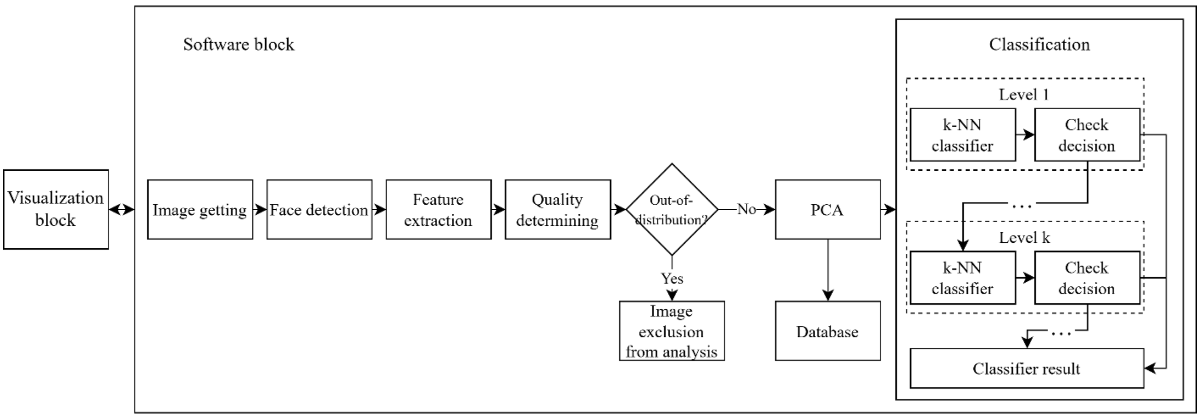

Figure 3 – Functional diagram of the proposed sequential classification algorithm

The sequential analysis of feature vectors consists of the fact that before launching the classifier, it is necessary to preprocess all descriptors using the

Principal Component Analysis without dimension reduction in order to arrange the components in order to reduce their significance. Then, the concept of levels is introduced: the resulting face descriptor's principal component vector can be divided into l parts (levels) and each part can be considered sequentially and separately. When comparing feature vectors with this approach, it is important to calculate the distance only between d_l components at each level l. Since the measure of difference (e.g., the square of the L2 distance) is often additive, the sequential analysis of descriptors by levels is fair. Accordingly, it is possible to use distances from the previous $l-1$ level to increase computational efficiency, in which case it will be sufficient to find only the distances between new components:

$$\rho\left(\tilde{x}^{(l)}, \tilde{x}_r^{(l)}\right) = \rho\left(\tilde{x}^{(l-1)}, \tilde{x}_r^{(l-1)}\right) + \sum_{d=d_{(l-1)}+1}^{d_l} \rho(\tilde{x}_d, \tilde{x}_{r;d}) \qquad (1)$$

The calculation of the distance between components will continue until the condition is met: the ratio of the distance between the component vector **x** of the input image and some class to the distance between **x** and its nearest neighbor is less than a fixed threshold:

$$C_{l+1} = \{c \in C_l | \frac{\rho\left(\tilde{x}^{(l)}, \tilde{x}_r^{(l)}\right)}{\rho\left(\tilde{x}^{(l)}, \tilde{x}_{c_l^*}^{(l)}\right)} \leq \delta\} \qquad (2)$$

It has been experimentally proven that already at the first levels it is possible to determine the object's class, so the proposed approach of sequential analysis allows to improve computational efficiency up to 5 times (depending on the dimension of the image feature vectors) with minimal losses in accuracy up to 1%. Additionally, a comparison with the Approximate Nearest Neighbors methods HNSW and FAISS was also carried out. The parameters for HNSW and FAISS were selected to be as close as possible to the computational efficiency results of the proposed sequential analysis. The results are presented in Table 2.

Table 2 – The results of sequential classification with rare image exclusion (VGGFace2 dataset)

CNN	Metrics	k-NN	HNSW	FAISS	Sequential analysis
VGGFace	Accuracy, %	87,6	84,1	81,7	84,3
	Time, ms	25,4	8,8	14,7	6,5
ResNet50	Accuracy, %	92,4	92,0	91,6	91,9
	Time, ms	20,3	5,7	8,0	4,1
MobileNetV1	Accuracy, %	**97,9**	96,9	96,3	**97,5**
	Time, ms	9,6	**3,4**	6,2	3,5
InsightFace	Accuracy, %	**97,9**	**97,5**	96,0	**97,6**
	Time, ms	5,9	2,9	5,1	**2,6**
FaceNet	Accuracy, %	**97,7**	**97,5**	94,2	**97,4**
	Time, ms	5,7	2,7	5,0	**2,4**
LCNN	Accuracy, %	91,3	88,2	88,5	89,9
	Time, ms	2,3	1,2	1,6	0,9

HNSW demonstrated similar results in accuracy in most experiments, but the time value is greater compared to sequential analysis. FAISS allows to speed up image processing only by 1,5-2 times and the loss in accuracy is more than 1%.

In the third chapter, “Methods of face analysis for video surveillance systems based on face recognition”, a method for aggregating frames of a homogeneous segment with subsequent clustering was proposed (Figure 4) [28].

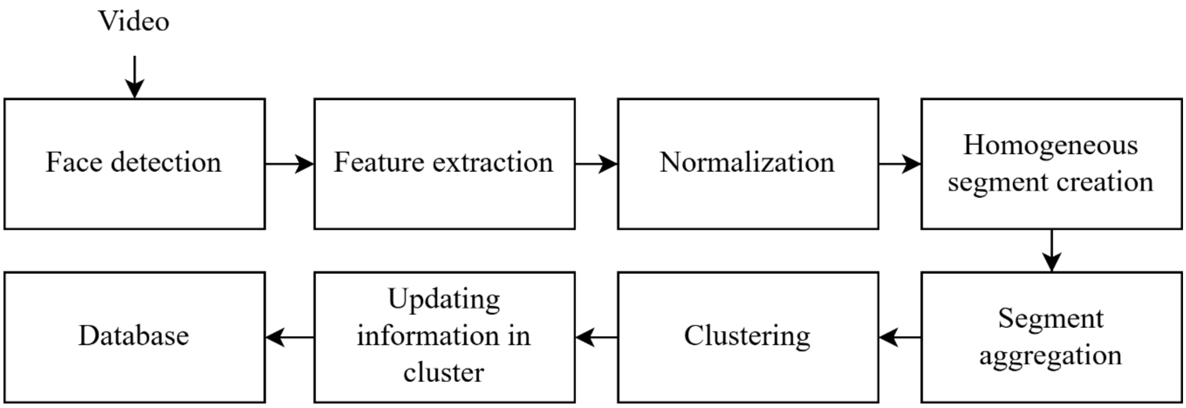

Figure 4 – Functional diagram of the proposed algorithm for organizing images of people from a video data collection using clustering

To aggregate close frames, it is proposed to use two-level periodization [28-30]. The first level involves identifying a set of consecutive frames with the image of the same person, the second – in combining the selected homogeneous tracks into a cluster.

For the first-level periodization, the threshold ρ_0 = 0,52 was experimentally selected, which was compared with the Euclidean distance between frame descriptors. If the distance between consecutive frames is less than the threshold, then these frames belong to the same segment, otherwise a segment change occurred. For the second-level periodization, a test was performed to see how normalizing descriptors by the Euclidean metric would affect the accuracy of identifying consecutive homogeneous frames. The YouTubeFaces dataset [31] was taken, which contains frames from different video sequences. The following methods were considered: average distance and variance between elements of two tracks, medoid, average distance among two tracks, average track vector, median, aggregate vector. The following CNNs were considered to obtain descriptors: ResNet50 [20], MobileNetV1 [21], InsightFace [18]. It has been experimentally proven (Table 3) that it is not necessary to process each consecutive frame of one segment, and the most efficient algorithm is to normalize all descriptors and then find the average track vector (AUC = 99.1% / 98.5% / 99.3% for each CNN, respectively).

Table 3 – The results of frame verification using InsightFace for feature vector extraction

Algorithm	AUC	Time, sec
average distance and variance between elements of two tracks	92,1	0,014
medoid + normalization	98,1	0,23
normalization + average distance among two tracks	99,0	0,016
normalization + average track vector	**99,3**	**0,015**
average track vector + normalization	98,1	**0,014**
median	95,4	0,018
normalization + median	95,6	0,019
normalization + aggregate vector	92,2	0,017

After aggregation of close frames, clustering is performed (Figure 4, blocks 6-7). Clustering allows solving the problem of dividing images of people into unique groups without preliminary information about the object's belonging to any class. This approach is useful when working with video surveillance systems, where hundreds/thousands of people are captured in the frame. Within clustering, an experimental study of several algorithms was conducted: agglomerative hierarchical clustering, Rank-Order [32], DominantSet [33]. The most accurate results of image distribution into groups were demonstrated by the hierarchical clustering approach based on Rank-Order using InsightFace convolutional neural network to extract image descriptors (Table 4). The obtained clusters turned out to be the closest to the groups from the validation sample, which contained images of 1595 people. The percentage of wrong clusters is 1.8%.

Table 4 – The results of clustering using InsightFace for feature vector extraction

Algorithm	**Total number of clusters**	**Invalid number of clusters**
DominantSet	1799	160
Rank - Order	1987	29
Hierarchical clustering, FAR=1%	2056	21
Hierarchical clustering, FAR=10%	1918	119

All the proposed methods were implemented as a software package for organizing face images in video surveillance systems: https://github.com/SokolovaNastya/FaceOrganization. The program of the effective face recognition method consists of the system graphical interface (Qt framework), face detection program (TensorFlow Models), image feature extraction program (CNN), anomalous image detection program (Figure 1), database (Apache Cassandra), and sequential image classification program (Figure 3). The face grouping program for video data (Figure 4) consists of the system graphical interface (Qt framework), face detection program (TensorFlow Models), image feature extraction program (CNN), homogeneous segment detection program, aggregated vector finding program, clustering program, and database (Apache Cassandra).

CONCLUSION

The following results were obtained during the study:

1. The problem of image classification with automatic detection of minority classes is considered. A classification approach on an open set is proposed using a new neural network model for detecting anomalous data trained on a specially created dataset that includes people of different ages, different races, and images with filters. According to the results of the experimental study, the developed detection method demonstrates up to 20% more accurate result compared to existing approaches to detecting anomalies. Deferring the decision on detected anomalous data allows for an increase in the accuracy of image classification up to 50%, depending on the selected classification method.
2. The problem of improving the computational efficiency of face classification is considered. A method of sequential image classification is developed, which allows for determining the class membership based on the first elements of the descriptor. An experimental study demonstrated that sequential classification allows for improving computational efficiency up to 5 times with accuracy losses of no more than 1%.
3. The problem of organizing images of people when working with a video sequence is considered. A two-level periodization approach is proposed, where instead of processing each frame, normalization of all descriptors is performed with subsequent finding of the average track vector. Experimental results show that the implemented face clustering allows for distributing video frames of faces into groups with a share of incorrectly identified clusters not exceeding 2%.
4. The proposed models and methods are implemented as a set of programs for recognizing faces in photos and videos with open source code. The program set includes, in addition to the proposed methods for recognizing faces,

obtaining images through the system's graphical interface, detecting faces and extracting image features, and working with a database.

Further work on the presented topic can be directed towards studying more complex classifiers with the aim of increasing recognition accuracy. Moreover, it is important to develop a method for applying information about rare data without deferring decisions on them. For example, using certain weights or transforming descriptors in a special way. Additionally, it is necessary to consider other modern methods of data aggregation instead of simple averaging of individual features.

REFERENCES

1. Mi Y., Zhong Z., Huang Y., Ji J., Xu J., Wang J., Wang S., Ding S., Zhou S. Privacy-preserving face recognition using trainable feature subtraction // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – pp. 297-307. – 2024.
2. Li Y. Research and application of deep learning in image recognition // 2022 IEEE 2nd international conference on power, electronics and computer applications (ICPECA), IEEE. – pp. 994-999. – 2022.
3. Savchenko A. V. Sequential three-way decisions in multi-category image recognition with deep features based on distance factor // Information Sciences. – T. 489. – pp. 18-36. – 2019.
4. Terhörst P., Ihlefeld M., Huber M., Damer N., Kirchbuchner F., Raja K., Kuijper A. Qmagface: Simple and accurate quality-aware face recognition // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. – pp. 3484-3494. – 2023.
5. Sokolova A., Savchenko A. Effective face recognition based on anomaly image detection and sequential analysis of neural descriptors // 2023 IX International Conference on Information Technology and Nanotechnology (ITNT), IEEE. – pp. 1–5. – 2023.
6. Song Y., Wang F. Coreface: Sample-guided contrastive regularization for deep face recognition // Pattern Recognition. – T. 152. – pp. 110483. – 2024.
7. Singh H. V., Girdhar A., Dahiya S. A Literature survey based on DBSCAN algorithms // 6th International Conference on Intelligent Computing and Control Systems (ICICCS). – IEEE. – pp. 751-758. – 2022.
8. Bai Y., Chu L. A Graph-based Approach to Estimating the Number of Clusters // arXiv preprint arXiv:2402.15600. – 2024.

9. Hastie T. The elements of statistical learning: data mining, inference, and prediction. – New York: springer. – T. 2. – pp. 1-758. – 2009.
10. He Y., Cao K., Li C., Loy C. Merge or not? Learning to group faces via imitation learning. // Proceedings of the AAAI Conference on Artificial Intelligence. – T. 32(1). – 2018.
11. Sokolova A., Savchenko A. Open-Set Face Identification with Sequential Analysis and Out-of-Distribution Data Detection // International Joint Conference on Neural Networks (IJCNN). – pp. 1-8. – 2022.
12. Sokolova A., Savchenko A., Nikolenko S. A. Search for rare data in the problem of face recognition in images // Computer Optics. – T. 46(5). – pp. 801-807. – 2022.
13. Chalapathy R., Menon A. K., Chawla S. Anomaly detection using one-class neural networks // arXiv preprint arXiv:1802.06360. – 2018.
14. Hara S., Nitanda A., Maehara T. Data cleansing for models trained with SGD // Advances in Neural Information Processing Systems. – T. 32. – 2019.
15. Pareek J., Jacob J. Data compression and visualization using PCA and T-SNE // Advances in Information Communication Technology and Computing: Proceedings of AICTC 2019. – Springer Singapore. – pp. 327-337. – 2021.
16. McInnes L., Healy J., Melville J. Umap: Uniform manifold approximation and projection for dimension reduction // arXiv preprint arXiv:1802.03426. – 2018.
17. Lee K., Lee H., Lee K., Shin J. Training confidence-calibrated classifiers for detecting out-of-distribution samples // arXiv preprint arXiv:1711.09325. – 2017.
18. An X., Deng J., Guo J., Feng Z., Zhu XH., Yang J., Liu T. Killing two birds with one stone: Efficient and robust training of face recognition

CNNs by partial FC // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – pp. 4042-4051. – 2022.

19. Parkhi O.M., Vedaldi A., Zisserman A. Deep face recognition // Proceedings of the British Machine Vision Conference. – T. 1(3). – pp. 6. – 2015.
20. Cao Q., Shen L., Xie W., Parkhi O.M., Zisserman A. VGGface2: A dataset for recognising faces across pose and age // In: 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). – pp. 67-74. – 2018.
21. Savchenko A. V. Efficient facial representations for age, gender and identity recognition in organizing photo albums using multi-output ConvNet // PeerJ Computer Science. – e197. – 2019.
22. Schroff F., Kalenichenko D., Philbin J. FaceNet: A unified embedding for face recognition and clustering // In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – pp. 815-823. – 2015.
23. Wu X., He R., Sun Z., Tan T. A light CNN for deep face representation with noisy labels // IEEE Transactions on Information Forensics and Security. – T. 13(11). – pp. 2884-2896. – 2018.
24. Vareto R. H., Linghu Y., Boult T.E., Schwartz W.R., Günther M. Open-set face recognition with maximal entropy and Objectosphere loss // Image and Vision Computing. – T. 141. – pp. 104862. – 2024.
25. Kim M., Su Y., Liu F., Jain A., Liu X. KeyPoint Relative Position Encoding for Face Recognition // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – pp. 244-255. – 2024.
26. Sokolova A.D, Savchenko A.V. Computation-efficient face recognition algorithm using a sequential analysis of high dimensional neural-net

features // Optical Memory and Neural Networks. – T. 29(1). – pp. 19-29. – 2020.

27. Sokolova A., Savchenko A. Fast Nearest-Neighbor Classifier based on Sequential Analysis of Principal Components // Analysis of Images, Social Networks and Texts: 8th International Conference, AIST 2019, Springer. – pp. 73-80. – 2019.
28. Sokolova A., Savchenko A. Data organization in video surveillance systems using deep learning // CEUR Workshop Proceedings. – pp. 243-250. – 2018.
29. Sokolova A., Savchenko A. Cluster Analysis of Facial Video Data in Video Surveillance Systems Using Deep Learning // Computational Aspects and Applications in Large-Scale Networks. Springer Proceedings in Mathematics & Statistics, Springer. – pp. 113-120. – 2018.
30. Sokolova A.D., Kharchevnikova A.S., Savchenko A.V. Organizing multimedia data in video surveillance systems based on face verification with convolutional neural networks // In: International Conference on Analysis of Images, Social Networks and Texts (AIST), Springer, Cham. – pp. 223-230. – 2017.
31. Huo J., van Zyl T. L. Unique faces recognition in videos // 23rd International Conference on Information Fusion (FUSION). – IEEE. – pp. 1-7. – 2020.
32. Zhu C., Wen F., Sun J. A rank-order distance based clustering algorithm for face tagging // Computer Vision and Pattern Recognition (CVPR). – IEEE. – pp. 481-488. – 2011.
33. Denitto M., Bicego M., Farinelli A., Vascon S., Pelillo M. Biclustering with dominant sets // Pattern Recognition. – T. 104. – pp. 107318. – 2020.