Input Data Requirements for Person Identification Information Technology

Kateryna Merkulova[a] and Yelyzaveta Zhabska[a]

[a] *Taras Shevchenko National University of Kyiv, Volodymyrs'ka str. 64/13, Kyiv, 01601, Ukraine*

Abstract

This paper describes the research of the information technology for person identification, that is based on mathematical models of local-texture descriptors in order to increase efficiency and reduce the probability of identification errors.

Analysis of previous studies has shown that the effectiveness of the face recognition and identification algorithm can vary significantly after its appliance to images from different datasets. Therefore, the main purpose of this research is to determine the requirements for face images that will be used in the identification process.

During the research, experiments were conducted on several of the most common face image databases, as a result of which the efficiency of the algorithm reached an identification accuracy rate of 95% on images captured under controlled conditions and adjusted within single database. Based on the results of the experiments, the conditions of capturing and parameters of face images were determined, under which the accuracy of identification is the highest.

Keywords[1]

Face recognition, person identification, local-texture descriptors, information technology

1. Introduction

As of today, the ability of software to interpret and analyze images is one of the key areas of interest and innovation in the field of information technology. At the beginning of 2023, face recognition and identification technologies continue to develop rapidly, and the use of these technologies is becoming more and more common. Face recognition is one of the best biometric technologies because face images can be taken without interacting with the person being identified, and these images are instantly captured and verified through existing databases. The purpose of face recognition and identification technology is to achieve successful and accurate identification on available hardware and software, such as CCTV channels and standard computer equipment.

From the rapid progress in face recognition and identification technology over the past 60 years several main branches of industry have taken advantages, including mobile, retail, banking, finance, border control. Biometrics take an important part of real-time police security in increasingly crowded and diverse locations around the world. New technologies such as augmented reality, virtual reality and computer vision applications are based on image recognition by artificial intelligence.

The authors of the paper [1] distinguish the following most common areas of application of face recognition technology:

- Access control – a person's face can be seen as a biometric signature, so face recognition can be used to confirm an identity of a person.
- Criminal investigations – face recognition can be used to locate and confirm the identity of a suspect at a crime scene using surveillance images or sketches described by witnesses.

CITI'2023: 1st International Workshop on Computer Information Technologies in Industry 4.0, June 14–16, 2023, Ternopil, Ukraine
EMAIL: kate.don11@gmail.com (K. Merkulova); y.zhabska@gmail.com (Y. Zhabska)
ORCID: 0000-0001-6347-5191 (K. Merkulova); 0000-0002-9917-3723 (Y. Zhabska)

- Support for identification of wanted persons – face recognition technologies can be used to identify wanted persons in real time using surveillance cameras, that allows to quickly neutralize suspects and increase the level of security in public places.

Improvement of face recognition and biometric technologies have led to increasing of accuracy, availability, and, consequently, to widespread usage of automated face recognition. It means that face recognition can be used on an even larger scale and in more complex environments. Particularly, the Intelligence Advanced Research Projects Activity, the high-level research agency of the US intelligence community, created a program in collaboration with Federal Bureau of Investigation scientists and some of the national leading experts in computer vision to develop and test software that would was able to quickly and accurately identify faces at partially obscuring angles captured by surveillance cameras in public places, such as subway stations and crowded streets. The advanced face recognition system eventually became part of the searching tool called Horus and became available to the Combating Terrorism Technical Support Office of Pentagon, that helps provide military technology to civilian police forces [2].

Face recognition technologies are also being implemented for use by special operations forces on drones to gather intelligence and assist in other missions. RealNetworks, a drone software manufacturer, claims that with their software, unmanned aerial vehicles can already be used for rescue missions, perimeter protection and indoor search operations [3]. In Dubai, police use drones equipped with face recognition to track reckless drivers [4].

Currently, the Ukrainian government widely uses face recognition and identification technologies during the Russian-Ukrainian war. Ukrainian investigators and independent organizations use Clearview AI software, a web platform based on face recognition technology that contains more than 30 billion face images obtained from publicly available web sources. During the war, this program was used to search and identify refugees for the purpose of family reunification; exposing false war-related posts on social networks; increasing security at checkpoints; identification and confirmation of dead soldiers; identification and detention of Russian intruders [5].

However, the ways in which face recognition and identification technologies are used for military purposes, are subject of public criticism, because any errors of the software or those who using it can have irreversible consequences. Therefore, the study of recognition and identification algorithms is relevant in order to increase their efficiency and reduce the probability of false identification.

2. Task definition and solution methods

In paper [6], the person identification information technology based on local-texture descriptors was proposed, in which the algorithm is underlying that consists of the following methods: Haar features for detecting and localizing the face region in images, Gabor wavelet transform for image processing, histograms of oriented gradients (HOG) and local binary patterns in one-dimensional space (1DLBP) for extracting image feature vectors.

Let's consider the proposed algorithm in more detail. As input data for the operation of the algorithm, images of faces on which it is necessary to identify a person are used, presented in the form of a pixel intensity matrix.

At the first stage of the identification process, it is necessary to localize the region of the face in the image. For this, Haar features are used, which is a quite effective method of detecting objects in images [7]. This method is based on machine learning, where the cascade function is trained on a large number of images that contain human faces and those that do not contain faces. Haar features are a set of primitives (white and black blocks) which correspond for certain face features and organized into a cascade structure. Each feature is the difference value between the sum of pixels under the black block and the sum of pixels under the white block. As a result of features learning f_j, it is possible to obtain the value of comparability by modulo p_j and the limit value θ_j, that allows to describe the classifier as:

$$h_j(x) = \begin{cases} 1, if\ p_j\theta_j > p_j f_j(x), \\ 0, else. \end{cases} \tag{1}$$

After the face region is localized, the image that was given to the input of the algorithm is reduced and only the part that represents the region of the image in which the human face is localized remains.

At the second stage, the obtained image region is processed by Gabor wavelets. This method is biologically significant due to the fact that wavelets have a shape similar to the receptive fields of simple cells of the primary visual cortex. Accordingly, image representation is based on the principles of image representation in the human brain, and computer vision modeling becomes a more effective and efficient process. The Gabor function is defined as follows [8]:

$$g(x,y) = e\big(j(2\pi(u_0x + v_0y) + P)\big) \times Ke\big(-\pi(a^2(x - x_0)^2_\tau + b^2(y - y_0)^2_\tau)\big). \qquad (2)$$

Function parameters: (u_0, v_0) are spatial frequencies of the sinusoidal carrier in Cartesian coordinates; P is the phase of the sinusoidal carrier; K is scaling parameter of the value of the Gaussian envelope; (a, b) are scaling parameters of two axes of the Gaussian envelope; (x_0, y_0) are the coordinates of the peak of the Gaussian envelope.

The image processed by the Gabor wavelet transform is given to the input of feature extraction methods, which are local texture image descriptors. The first method is the histogram of oriented gradients (HOG), that allows to distinguish features of the image shape. To create a histogram of local gradients, orientation gradients are initially calculated for each area of the image, which are then normalized. The gradient is calculated using a one-dimensional horizontal discrete derivative mask D_x by first filtration and a one-dimensional vertical discrete derivative mask D_y by convolution. The resulting value is the sum of the adjacent pixels with account of the weight of the mask. This process can be described by the following formulas [9]:

$$I(x) = I \cdot D_x, where\ D_x = [-1 \quad 0 \quad 1], \qquad (3)$$

$$I(y) = i \cdot D_y, where\ D_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}. \qquad (4)$$

The next method for image feature extraction is local binary patterns in one-dimensional space (1DLBP). LBP operators are used for texture distinction and demonstrate effective performance under conditions of changing rotation angles and lighting. Most LBP operators characterize the texture distribution of the face image for each pixel with only its edges. However, differences between two face images can be demonstrated by the relative relationship to other pixels. For this purpose, the original image is decomposed into several sub-images of different sizes to better characterize the details and relationships between all the pixels in the image. Next, the received image descriptors are combined into one global vector. This technique allows to obtain small details and the relative relationships between all the pixels of a full image. 1DLBP-vector is formed by a binary code and is determined by the formula [10]:

$$1DLBP = \sum_{n=0}^{N-1} 2^n \cdot S(g_n - g_0), \qquad (5)$$

where g_0 and g_n are the values of the central pixel element and its one-dimensional neighbors.

The HOG and 1DLBP vectors are concatenated to form a global feature vector of the face image, that is used in the further classification process.

According to the results of previous studies [11, 12], it was found that the effectiveness of the algorithm used for person recognition and identification on face images can significantly vary on different datasets. Such problem can become a significant obstacle to the correct identification of a person and cause errors in the identification process, if the parameters of the images stored in the database and the images that are given to the input of the algorithm vary essentially. In such cases, it is appropriate to formulate requirements for the images that will be used in the process of recognition and identification.

This paper is devoted to the research of the information technology for person identification proposed by the authors, which is based on the developed mathematical model containing the combination of such methods, as Haar features, Gabor wavelet transform, histograms of oriented gradients (HOG) and local binary patterns in one-dimensional space (1DLBP). The purpose of the research is to determine the requirements for the formation of test and etalon image samples, the use of which would allow to obtain the highest accuracy rate results of the person identification after applying this combination of methods.

3. Experimental research

To conduct the experiments databases were selected, that contain images captured in constrained, as well as in unconstrained conditions close to the natural environment in which human faces can be observed. The research was carried out on images from the following databases: AgeDB, CFP (Celebrities in Frontal-Profile data set), Database of Faces, FERET (Face Recognition Technology database), LFW (Labeled Faces in the Wild), SCface and TinyFace.

The selected databases contain different numbers of face images for different numbers of individuals. In order for the results obtained during the experimental study to be more objective, it was decided to choose a fixed number of research subjects based on the smallest number of persons for whom images were captured within one database. Thus, images of 40 people were selected from each database to form test and etalon samples of images to conduct experiments.

Since face image databases usually have variability in image properties such as format, resolution, and the size of the image region containing the person's face, it was decided to conduct experiments using different listed properties to form requirements in regard to the technical characteristics of face images to which the developed algorithm will be applied.

Thus, the experiments were conducted in three stages. The first set of experiments was performed on face images that were converted from the original format in which the images are stored in the aforementioned databases to the image formats BMP and PNG, which are compressed image formats, and JPG, which is an uncompressed image format. If the initial images were already stored in any of the studied formats, they were still converted to the same format using software-implemented image format conversion methods for a more objective evaluation of the experimental results.

The second set of experiments was conducted with the resolution conversion of the original images stored in the databases. During preparation for the experiments, it was important to convert the images to the same resolution values. And since all used databases contain initial images with different resolutions, the aspect ratios of the images may have changed after applying the resolution conversion methods. Therefore, the features of the images may have changed, including the face features that these images contain, that is a crucial factor in face recognition tasks. Therefore, the image resolution was converted to a fixed height value only, and the width was automatically determined in such a way as to keep the aspect ratio constant. According to the results of the analysis of previous experimental studies described in [11], it was determined that the highest rates of identification accuracy during the appliance of the studied algorithm were obtained on images with a resolution of *width*×91, *width*×100, *width*×128 and *width*×144 of pixels, where *width* is the width value that was determined automatically for images in the way to preserve the aspect ratio of the image. The same values were used for the experiments in this paper.

In the third stage of the study, experiments were carried out on face images from databases, transformed in such a way that the images contained only the face features of the person without any other details, such as the background. The purpose of this experimental study is to establish the dependence of identification results on whether the images to be identified contain any other details besides the face itself. Considering the fact that different databases contain images of different resolutions, which were captured at different distances between the camera and the object, accordingly, the image region that will contain the face is different in size, it was decided to investigate different sizes of regions of the images containing face. To conduct the study, the image size parameters were chosen, that made it possible to obtain high results of the algorithm performance in the previous study [12], i.e. *width*×47 and *width*×78 pixels, as well as the threshold values of the size of the face region (*width*×32 and *width*×128) and average value between selected values and threshold values (*width* ×64).

3.1. AgeDB

AgeDB is a database developed by scientists of the Imperial College London for the purpose of researching the problem of age-related changes in the face for tasks of person recognition and identification on images taken at different ages of a person [13]. AgeDB contains images of multiple subjects annotated with age labels to the nearest year, making the database suitable for experiments with age-invariant face verification, age estimation, and face age change under unconstrained image fixation conditions. The use of this database allows to measure the sensitivity in the performance of face recognition algorithms with increasing age difference between instances of images of the same subject. In addition, the images of faces contained in this database are captured in completely unconstrained conditions of the real world (i.e., in different poses, with noise, with different face expressions, with occlusions, etc.). In total, AgeDB contains 16,488 JPG images for 568 individuals with an average age between image captures of 50.3 years.

The results of experiments conducted on images from the AgeDB database are shown in Table 1.

Table 1

Experimental results of algorithm performance on images from the AgeDB database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	174	174	174	174	174	174	174	174	174	174	174	174	174
Number of identified individuals	18	18	18	18	18	16	20	16	10	14	18	14	16
Number of non-identified individuals	22	22	22	22	22	24	20	24	30	26	22	26	24
Identification accuracy rate	45%	45%	45%	45%	45%	40%	**50%**	40%	25%	35%	45%	35%	40%
Identification error rate	55%	55%	55%	55%	55%	60%	50%	60%	75%	65%	55%	65%	60%

After appliance of the algorithm to images from the AgeDB database, overall identification results in range from 25% to 50% were obtained. The highest result of identification accuracy, as can be seen from Table 1, is 50% and it was obtained after converting the resolution of the original images to 128 pixels in height with preserving the aspect ratio. The following factors may have caused such average identification accuracy rates. First, the variability of the age of the persons whose faces were captured in the images. Secondly, the database mainly contains images of famous people, so many of them contain theatrical cosmetics, which in some cases distorts the features of a person's face. The last possible factor is that the images were captured in unconstrained conditions, i.e. they are not uniform in relation to head rotation, presence of noise, variability of facial expressions, presence of occlusion.

3.2. CFP

CFP (Celebrities in Frontal-Profile data set) is a data set proposed by researchers from the University of Maryland and the Rutgers University that contains images of faces collected in the open access, captured in both constrained and unconstrained environments, but processed to correspond to certain frontal and profile poses of the face in the images. The use of such a dataset allows to investigate more detailed the problem of changing poses, while other variations of image characteristics are unrestricted. According to the developers of the dataset [14], solving the problem of extreme pose variations allows to more successfully solve the general problem of unbounded pose variation, especially in the cases of video surveillance and photo tagging. The purpose of this dataset is to isolate the pose variation factor in terms of extreme poses, such as profile, in which many features of the face image are obscured from the observer's perspective. The CFP dataset contains 10 frontal and 4 profile images for 500 individuals in JPG format.

Table 2 contains the identification accuracy rates on face images from the CFP database.

For images from the CFP database, the results of correct person identification by the algorithm vary from 15% to 70%. After converting the resolution of the original images to 100 pixels in height and width that preserves the aspect ratio, the highest identification accuracy rate of 70% was obtained. It is worth mentioning that the obtained results could be affected by the variability in the poses of the persons captured in the images, since the CFP database contains both frontal images and those containing head turns up to 90 degrees, accordingly, some face features may be closed for observation by the camera. In addition, as in the case of the AgeDB database, the images in this database are mostly photographs of public figures that were captured in unconstrained conditions in terms of ambient lighting and camera flash intensity. Those factors made some areas of the image excessively illuminated, and it could cause distortion in the image of the face features of the person to be identified.

Table 2
Experimental results of algorithm performance on images from the CFP database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	202	202	202	202	202	202	202	202	202	202	202	202	202
Number of identified individuals	24	24	24	24	26	28	24	26	6	8	26	20	22
Number of non-identified individuals	16	16	16	16	14	12	16	14	34	32	14	20	18
Identification accuracy rate	60%	60%	60%	60%	65%	**70%**	60%	65%	15%	20%	65%	50%	55%
Identification error rate	40%	40%	40%	40%	35%	30%	40%	35%	85%	80%	35%	50%	45%

3.3. Database of Faces

Database of Faces is a set of face images of 40 people, which were captured with changes in lighting, facial expressions (open/closed eyes, smiling/unsmiling) and facial details (for example, the presence of glasses). All images were taken on the dark uniform background with subjects in an vertical, frontal position (some sideways movements were allowed). The set contains images in PGM format, size 92×112 pixels, with 256 gray levels per pixel. For each of the 40 subjects, 10 images were stored. The dataset was generated by researchers at AT&T Laboratories Cambridge between 1992 and 1994 for use in face recognition research conducted in collaboration with the Speech, Vision and Robotics Group of the Cambridge University Engineering Department [15].

The results of the algorithm performance after its appliance to face images from the Database of Faces are contained in Table 3.

Table 3
Experimental results of algorithm performance on images from the Database of Faces

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	120	120	120	120	120	120	120	120	120	120	120	120	120
Number of identified individuals	29	31	29	27	32	27	29	29	2	2	19	19	29
Number of non-identified individuals	11	9	11	15	8	13	11	11	38	38	21	21	11
Identification accuracy rate	72.5%	77.5 %	72.5 %	67.5 %	**80%**	67.5 %	72.5 %	72.5 %	5%	5%	47.5 %	47.5 %	72.5 %
Identification error rate	27.5%	22.5 %	27.5 %	32.5 %	20%	32.5 %	27.5 %	27.5 %	95%	95%	52.5 %	52.5 %	27.5 %

The overall results of experiments on images from the Database of Faces range from 5% to 80%. At the same time, the lowest rate of identification accuracy was obtained during experiments on images that contain only the face region and have the smallest resolution among the entire set of experiments. However, the experiments with changing the image format and resolution look more promising, as they vary between the values of 67.5% and 80%.

A rather interesting result is that after performing the conversion between the image formats, there was an increase in the identification accuracy rate compared to experiments on the original images stored in the PGM format. The highest rate in the set of experiments with the format is 77.5% of correctly identified images, that was obtained after converting the image to JPG format. Thus, it can be concluded that the format of the image given to the input of the algorithm is an important requirement for obtaining more accurate identification results.

The improvement of the aforementioned result by 2.5% was also facilitated by the transformation of the resolution from 92×112 pixels to 75×91 pixels, which made it possible to obtain a result of identification accuracy of 80%, which is the highest rate among all of the experiments conducted on images from the Database of Faces.

3.4. FERET

FERET (Face Recognition Technology database) is a database that contains 1564 sets of images of 1199 people with a total number of 14,126 images with a resolution of 256×384 pixels. The database, distributed by the National Institute of Standards and Technology (NIST), was created by researchers at George Mason University as part of a program dedicated to developing an automatic face recognition algorithm that could be used to assist security, intelligence and law enforcement personnel in carrying out their duties. Database images were captured during 15 sessions between 1993 and 1996 in a semi-constrained environment. To maintain some consistency across the database, the same physical settings were used for each photo session [16].

Rates of the correctness of the person identification in images from the FERET database are shown in Table 4.

Table 4
Experimental results of algorithm performance on images from the FERET database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	99	99	99	99	99	99	99	99	99	99	99	99	99
Number of identified individuals	30	29	30	30	22	25	30	30	2	14	28	29	29
Number of non-identified individuals	10	11	10	10	18	15	10	10	38	26	12	11	11
Identification accuracy rate	**75%**	72.5%	**75%**	**75%**	55%	62.5%	**75%**	**75%**	5%	35%	70%	72.5%	72.5%
Identification error rate	25%	27.5%	25%	25%	45%	37.5%	25%	25%	95%	65%	30%	27.5%	27.5%

Identification accuracy rates on images from the FERET database vary from 5% to 72.5% in experiments on images of different resolutions containing only the face of a person without other details, from 72.5% to 75% in experiments with image format conversion, and from 55 % to 75% in image resolution conversion experiments. The highest accuracy rates were obtained on the original images contained in the database after converting them from TIFF to PNG and BMP formats, and reducing the resolution from 256×384 pixels to 85×128 pixels and 96×144 pixels.

3.5. LFW

LFW (Labeled Faces in the Wild) is a database presented by scientists from the University of Massachusetts, that contains 13,233 images of 5,749 different people, of which 1,680 people have two or more images in the database, the remaining 4,069 people have only one image. Images are available in JPG format with a size of 250×250 pixels. Most of the images are in color, but some are only in grayscale. The primary purpose of the database was to provide a large dataset of face images with a wide range of variations in pose, lighting, facial expression, race, ethnicity, age, gender, background, camera quality, color saturation, focus, and other characteristics. That is, the presented dataset does not have any limitations for face detection on the image and it is as complete as possible to the natural environment in which human faces can be observed. The motivation for creating this database, as stated by the developers [17], was to investigate the problem of matching pairs of face images captured in unconstrained conditions, as this is the most general and fundamental problem in the field of face recognition.

Table 5 contains the results of experiments on images from the LFW database.

Table 5

Experimental results of algorithm performance on images from the LFW database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	125	125	125	125	125	125	125	125	125	125	125	125	125
Number of identified individuals	22	22	22	22	16	14	14	18	2	10	18	20	24
Number of non-identified individuals	18	18	18	18	24	26	26	22	38	30	22	20	16
Identification accuracy rate	55%	55%	55%	55%	40%	35%	35%	45%	5%	25%	45%	50%	**60%**
Identification error rate	45%	45%	45%	45%	60%	65%	65%	55%	95%	75%	55%	50%	40%

The accuracy of identification obtained as a result of conducting experiments on images from the LFW database ranges from 5% to 60%. On the original images from the database, the identification accuracy is 55% and it is constant after image format conversion. After converting the resolution, the rates of correct identification decreased on 10-20%.

However, the highest identification accuracy rate was achieved after converting the image to one that contains only the human face region and having a resolution of 128 pixels in height and width that preserves the aspect ratio. In addition, this is the only database on the images from which the algorithm was more effective in this set of experiments. This result can be explained by the fact that the LFW database includes images where several faces of different people are recorded. Therefore, situations could arise in which the face detection method implemented in the algorithm detected the face of the wrong person whose image was in the test set.

Since the LFW database is a set of images that approximate to the conditions of the natural environment in which human faces can be observed, and therefore has no restrictions on the variability of the conditions under which the images were created, this could significantly affect the performance of the algorithm during its appliance to the images set from this database.

3.6. SCface

SCface is a face image database developed by researchers from the University of Zagreb. It contains 4,160 static images of 130 people. The images were captured under unconstrained lighting conditions in the same accommodation, from different fixed distances and angles of view of the camera on the subject, with head positions typical for commercial surveillance systems, i.e. the camera is placed slightly above the subject's head, during the recording by surveillance cameras subjects did not look at a fixed point. During the capture of images, five video surveillance cameras of different quality and resolution were used, installed and fixed in the same positions, that allows to simulate real conditions and test face recognition algorithms for different scenarios of use by law enforcement agencies and national security services. The resulting images were processed in such a way to remove as much background as possible, and as a result of processing have the following resolution: 75×100 pixels for a distance between the camera and the subject of 4.20 m, 108x144 pixels for a distance of 2.60 m and 168×224 for a distance of 1 m [18].

Table 6 presents the results obtained after appliance of the algorithm to images that the SCface database contain.

Table 6
Experimental results of algorithm performance on images from the SCface database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	160	160	160	160	160	160	160	160	160	160	160	160	160
Number of identified individuals	38	38	38	38	26	28	34	38	2	27	35	34	35
Number of non-identified individuals	2	2	2	2	14	12	6	2	38	13	5	6	5
Identification accuracy rate	**95%**	**95%**	**95%**	**95%**	65%	70%	85%	**95%**	5%	67.5%	87.5%	85%	87.5%
Identification error rate	5%	5%	5%	5%	35%	30%	15%	5%	95%	32.5%	12.5%	15%	12.5%

As a result of the algorithm appliance to the original face images from the SCface database, an identification accuracy rate of 95% was obtained, that is the highest result among all experiments conducted on various databases. Conversion of image formats did not affect the efficiency of the algorithm. During experiments with image resolution conversion, results ranging from 65% to 95% were obtained. The largest number of persons was correctly identified by converting the resolution to the parameters of 108×144 pixels.

3.7. TinyFace

TinyFace is a large-scale set of face images created by researchers from Queen Mary University of London to test deep learning algorithms in face recognition tasks, specifically investigating natural face recognition on low-resolution images. The TinyFace dataset consists of face images of 5,139 individuals labeled with identifiers, a total of 169,403 original low-resolution face images (average size is 20×16 pixels) in JPG format, captured for testing 1:N recognition algorithms. All low-resolution images in TinyFace were collected from publicly available web data containing a variety of image capture scenarios under unconstrained conditions regarding face positioning, lighting, presence of occlusion, and background. According to the authors, this is the first systematic study devoted to face recognition on low-resolution images in natural conditions, excluding works based on artificially reduced samples of high-resolution face images for testing recognition performance [19].

The accuracy rates of the person identification on the images from the TinyFace database are presented in Table 7.

The results of the experiments on the images from the TinyFace database are identification accuracy rates ranging from 10% to 45%, that is the lowest result among all sets of experiments. Such results of the algorithm can be explained by the specificity of the parameters of the images contained in this database and the conditions under which they were captured. The resolution of the images contained in the TinyFace database is critically low for recognition by most standard algorithms. To apply the algorithm, studied in this paper, to these images, it was necessary to preprocess the images, enlarging them for the possibility of detecting the face region. Such transformations could affect the correctness of the facial features representation, blurring them and making it impossible to further extract the feature vector in a form that is suitable for further classification. Moreover, the original images were captured in unconstrained conditions of the natural environment, that could also significantly affect the efficiency of the algorithm.

Table 7
Experimental results of algorithm performance on images from the TinyFace database

	Original	Format converted			Resolution converted				Face region converted				
		JPG	PNG	BMP	width x 91	width x 100	width x 128	width x 144	width x 32	width x 47	width x 64	width x 78	width x 128
Number of individuals	40	40	40	40	40	40	40	40	40	40	40	40	40
Number of images	89	89	89	89	89	89	89	89	89	89	89	89	89
Number of identified individuals	4	4	4	4	18	18	14	14	10	10	10	10	10
Number of non-identified individuals	36	36	36	36	22	22	26	26	30	30	30	30	30
Identification accuracy rate	10%	10%	10%	10%	**45%**	**45%**	35%	35%	25%	25%	25%	25%	25%
Identification error rate	90%	90%	90%	90%	55%	55%	65%	65%	75%	75%	75%	75%	75%

However, it is worth mentioning that, just as in the case of the AgeDB images, when performing experiments on images containing only the face of a person at the lowest resolution, the highest person identification rate was obtained among all databases. Such results can be useful for solving one of the main tasks of the field of face recognition, that is recognition on small-sized images.

4. Analysis of results and discussion

Comparative diagram of the results of experiments conducted on original images and with format conversion is presented in Figure 1. As can be seen from the diagram, the highest rates of identification accuracy were obtained on images from the Database of Faces databases (77.5% after image conversion to JPG format), FERET (75% on original TIFF format images and after format conversion to PNG and BMP) and SCface (95% - on all investigated image formats). The results of experiments on images from other databases range from 10% to 60% identification accuracy.

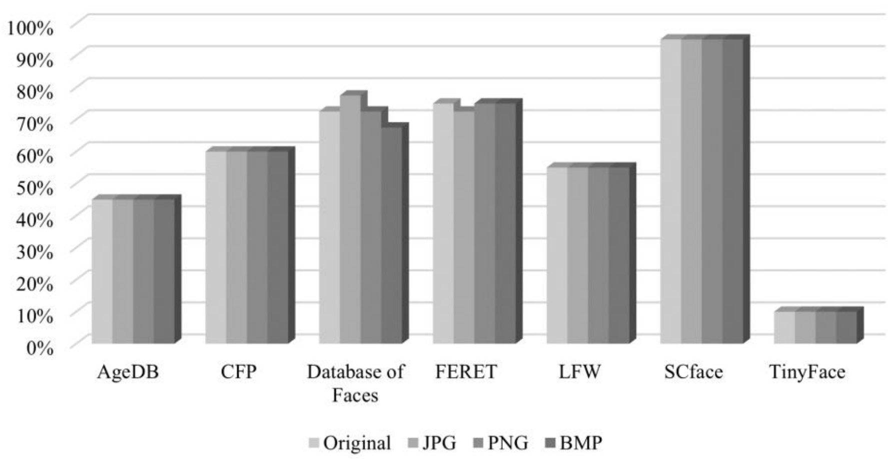

Figure 1: Experimental results on images with format conversion

Analyzing the results that obtained on the Database of Faces and FERET database, it can be concluded that the format of the images, that are given to the input of the algorithm, in some cases affects the efficiency of the algorithm. On all other databases, the results are constant, regardless of changing the image format. It is worth noting that all these databases contain images in JPG format, so this format is the most suitable for images to which the studied algorithm will be applied in the future.

Figure 2 presents a comparative diagram of the results of experiments conducted to investigate the resolution of input images at which the algorithm is most effective. The highest identification accuracy rate among all sets of experiments, that is 95%, was obtained on images from the SCface database, the resolution of which was converted to 108×144 pixels. However, quite high rates on images from various databases were also obtained after conversion of resolution parameters to 75×91 pixels and 96×128.

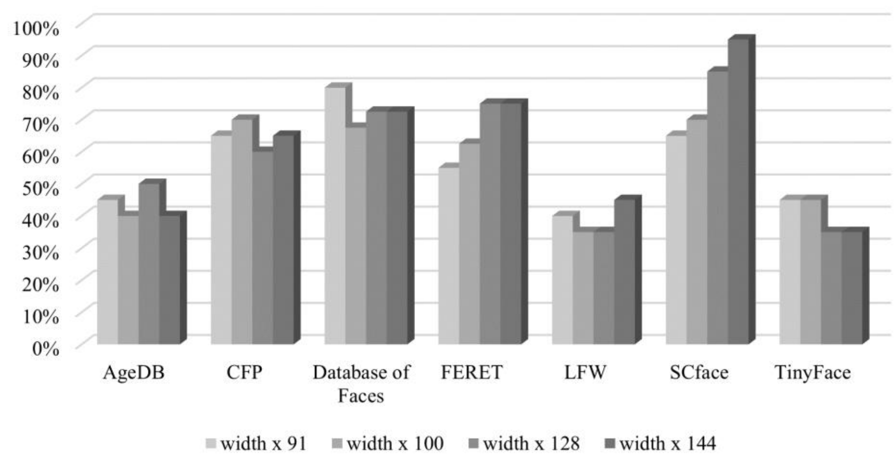

Figure 2: Experimental results on images with resolution conversion

A diagram of the results of experiments on images that contain only the face region without any other details is shown in Figure 3. Analyzing the obtained results, it can be concluded that the identification accuracy rate is significantly reduced when this type of transformation is applied to images. This may indicate that the face detection and region resizing methods available in the studied algorithm successfully process the input images, regardless of whether the input data contains any details other than human facial features.

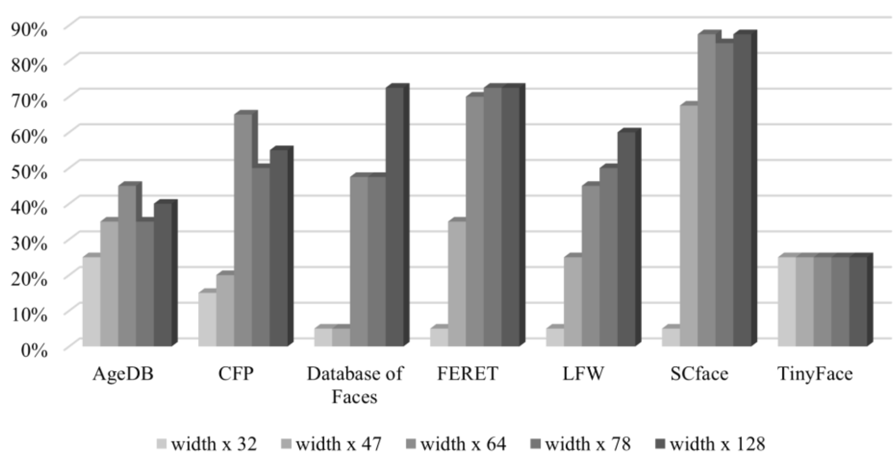

Figure 3: Experimental results on images that contain only face region

It is important to note that the images contained in all image databases, on which it was possible to achieve higher accuracy results after applying the studied algorithm, were captured in constrained or semi-constrained conditions: Database of Faces – images on a uniform background with frontal positions of faces; FERET – frontal images, aligned and captured under the same physical settings; SCface – images captured under the same lighting conditions using cameras that had an unchanged position relative to the subject, i.e. uniform by head position.

Instead, the images from the AgeDB, CFP, LFW and TinyFace databases, after applying the studied algorithm to which the obtained identification accuracy rates are lower, were captured in unconstrained conditions and have the following properties: large variability of the age of the person whose face was captured in the images at different time intervals; the presence of cosmetics that distort a person's facial features; occlusive conditions and head postures within those limits of the angle of rotation, under which the features of the face are partially or completely closed for observation by the camera; variability of human facial expressions; excessive amount and intensity of lighting, which affects the possibility of extracting the image feature vector; the presence of several regions containing faces in the image; low resolution images contained in the database. Thus, after analyzing the results of the experiments, it can be concluded that the algorithm is not effective for the tasks of identifying a person on images captured in unconstrained conditions.

One of the mechanisms for increasing the accuracy of recognition and identification is a clear formulation of requirements for images. Considering all the studies carried out in this work, it is possible to clearly formulate and describe in detail the following input images requirements:

- Image format – JPG.
- Image resolution in the range from *width*×91 pixels to *width*×144 pixels, where *width* is the width of the image, that is calculated automatically so that the aspect ratio of the image and, accordingly, facial features do not change.
- The subject's head position should be frontal. Head turns up to 45 degrees are allowed, as all facial features of the person must be visible in the image.
- Images must be captured under standard lighting conditions, with adequate use of flash to avoid shadows or excessive illumination of certain areas of the image.
- The distance between the subject and the camera can be between 1 m and 4.20 m.
- An average time interval of 2 years between the captures of the images forming the sample for one individual is acceptable.
- The faces included in the images should not have cosmetics, theatrical make-up, occlusive elements that distort facial features or make them invisible in whole or in part from the point of view.
- The image must contain the face of only one person.
- Slight variability in facial expressions is allowed, e.g. open/closed/squinted eyes, smile/no smile.

5. Conclusion

This paper is devoted to the research of information technology for person identification based on mathematical models of local-texture descriptors to determine the requirements for the formation of samples of face images that the algorithm can process with the highest efficiency.

The studied mathematical model consists of the combination of the following methods: Haar features for detecting and localizing the face region in images, Gabor wavelet transform for image processing, histograms of oriented gradients (HOG) and local binary patterns in one-dimensional space (1DLBP) for extracting image feature vectors.

Experimental studies of the algorithm were conducted on the AgeDB, CFP, Database of Faces, FERET, LFW, SCface and TinyFace databases. According to the results of the experiments, the following rates of identification accuracy were obtained in all sets of experiments on face images from the databases: AgeDB – from 25% to 50%, CFP – from 15% to 70%, Database of Faces – from 5% to 80%, FERET – from 5% to 75%, LFW – from 5% to 60%, SCface – from 5% to 95%, TinyFace – from 25 to 45%. At the same time, the lower limit of the accuracy of identification on images from all datasets was obtained during experiments with transformation of images in such a way that they did not contain other details, except for the face region. It means that the face detection and localized region resizing methods contained in the algorithm are sufficient for further processing and successful classification of the image feature vector. Also, after analyzing the results, the impact of the format and resolution of the images given to the algorithm input on the subsequent efficiency of the identification process was found, which indicated the need to set requirements for such image parameters.

In particular, the analysis of the results of experimental studies showed that the highest accuracy of identification by the algorithm was obtained after its appliance to the Database of Faces, FERET and SCface databases, that contain face images captured under constrained conditions and uniformed within the database. The reasons for the low efficiency of the algorithm on all other databases may be the following factors that indicate the unconstrainability of the image capturing conditions: the variability of the age of the person in the images in the same dataset; the presence of cosmetics or makeup; presence of occlusion; head postures and facial expressions that make it impossible to fully recognize features, etc.

So, as a result of the research, it was determined that the proposed algorithm allows to obtain the identification accuracy rate of up to 95% during its appliance to images that were captured under constrained conditions and uniformed within the same database. As a result of the analysis of experimental results, the requirements for the images, to which the application of the algorithm is the most effective, were determined.

6. References

[1] Y. Zennayi, F. Bourzeix and Z. Guennoun, "Analyzing the Scientific Evolution of Face Recognition Research and Its Prominent Subfields," in IEEE Access, vol. 10, pp. 68175-68201, 2022. doi: 10.1109/ACCESS.2022.3185137.

[2] D. Harwell, "FBI, Pentagon helped research facial recognition for street cameras, drones," The Washington Post, March 7, 2023. URL: https://www.washingtonpost.com/technology/2023/03/07/facial-recognition-fbi-dod-research-aclu/

[3] S. Brodsky, "The Air Force's Drones Can Now Recognize Faces. Uh-Oh," Popular Mechanics, February 24, 2023. URL: https://www.popularmechanics.com/military/a43064899/air-force-drones-facial-recognition/

[4] S. Ahmed, "Dubai Police Will Now Track Careless Drivers With AI Drones," ProPakistani, February 17, 2023. URL: https://propakistani.pk/2023/02/17/dubai-police-will-now-track-careless-drivers-with-ai-drones/

[5] Clearview AI. War in Ukraine. URL: https://www.clearview.ai/ukraine

[6] O. Bychkov, K. Merkulova, Y. Zhabska, A. Shatyrko, "Development of information technology for person identification in video stream," Proceedings of the II International Scientific Symposium "Intelligent Solutions" (IntSol-2021), CEUR Workshop Proceedings, 3018, pp. 70-80, Kyiv - Uzhhorod, Ukraine, September 28-30, 2021. URL: http://ceur-ws.org/Vol-3018/Paper_7.pdf

[7] T. Mantoro, M. A. Ayu and Suhendi, "Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods," 2018 6th International Conference on Multimedia Computing and Systems (ICMCS), Rabat, Morocco, 2018, pp. 1-5. doi: 10.1109/ICMCS.2018.8525935.

[8] R. R. Isnanto, A. A. Zahra, A. L. Kurniawan and I. P. Windasari, "Face Recognition System Using Feature Extraction Method of 2-D Gabor Wavelet Filter Bank and Distance-Based Similarity Measures," 2022 Seventh International Conference on Informatics and Computing (ICIC), Denpasar, Bali, Indonesia, 2022, pp. 1-4. doi: 10.1109/ICIC56845.2022.10007016.

[9] B. Attallah, A. Serir, Y. Chahir and A. Boudjelal. "Histogram of gradient and binarized statistical image features of wavelet subband-based palmprint features extraction", J. Electron. Imag. 26(6) 063006, November 8, 2017. doi: 10.1117/1.JEI.26.6.063006.

[10] A. Benzaoui, A. Boukrouche, H. Doghmane and H. Bourouba, "Face recognition using 1DLBP, DWT and SVM," 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), Tlemcen, Algeria, 2015, pp. 1-6. doi: 10.1109/CEIT.2015.7233002.

[11] K. Merkulova and Y. Zhabska, "Investigating Methods of Input Data Preparation for Person Identification Information Technology," 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 2022, pp. 495-498, doi: 10.1109/CSIT56902.2022.10000539.

[12] O. Bychkov, K. Merkulova and Y. Zhabska, “Preprocessing Methods Study to Improve Information Technology for Person Identification by Occluded Image," Proceedings of the 2nd International Workshop on Information Technologies: Theoretical and Applied Problems (ITTAP 2022), Ternopil, Ukraine, November 22-24, 2022, pp. 77-82. URL: https://ceur-ws.org/Vol-3309/paper6.pdf

[13] S. Moschoglou, A. Papaioannou, C. Sagonas, J. Deng, I. Kotsia and S. Zafeiriou, “AgeDB: The First Manually Collected, In-the-Wild Age Database,” 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1997-2005. doi: 10.1109/CVPRW.2017.250.

[14] S. Sengupta, J.-C. Chen, C. Castillo, V. M. Patel, R. Chellappa and D.W. Jacobs, “Frontal to profile face verification in the wild,” 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), 2016. doi:10.1109/wacv.2016.7477558

[15] F. S. Samaria and A. C. Harter, “Parameterisation of a stochastic model for human face identification,” Proceedings of 1994 IEEE Workshop on Applications of Computer Vision, Sarasota, FL, USA, 1994, pp. 138-142. doi: 10.1109/ACV.1994.341300.

[16] P.J. Phillips, H. Moon, S.A. Rizvi and P.J. Rauss, “The FERET Evaluation Methodology for Face Recognition Algorithms,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, pp. 1090-1104, 2000.

[17] G. B. Huang, M. Ramesh, T. Berg and E. Learned-Mille, “Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments,” University of Massachusetts, Amherst, Technical Report 07-49, October, 2007.

[18] M. Grgic, K. Delac and S. Grgic, “SCface - surveillance cameras face database,” Multimedia Tools and Applications Journal, Vol. 51, No. 3, February 2011, pp. 863-879.

[19] Z. Cheng, X. Zhu, S. Gong. TinyFace: Face Recognition in Native Low-resolution Imagery. URL: https://qmul-tinyface.github.io/index.html