AUTOMATING THE CUSTOMER VERIFICATION PROCESS IN A CAR SHARING SYSTEM BASED ON MACHINE LEARNING METHODS

Convenient and accurate verification of the user of a car sharing system is one of the key components of the successful functioning of the car sharing system as a whole. The machine learning-based KYC (Know your customer) process algorithm makes it possible to improve the accuracy of customer data validation and verification. This makes it possible to eliminate possible losses and reputational losses of the company in case of unforeseen situations while using the client's car sharing services. The object of this study is to find a solution to the problem of user verification in a car sharing system based on the KYC process using deep learning methods with a combination of OCR (Optical Character Recognition) methods.

The statement of the user verification problem in the car sharing system was formalized and the key parameters for the KYC process have been determined. The algorithm of the KYC process was constructed. The algorithm includes six successive stages: separating a face in the photograph, comparing faces, checking documents and their validity period, establishing and recognizing ROI (region of interest), formulating a verification decision. To separate the face in the client's photograph and compare faces, methods based on deep learning, as well as the quick HoG method (Histogram of oriented gradients), were considered and implemented. Verification of these methods on a test dataset, which includes images of documents of two thousand clients, showed that the recognition accuracy was 91 % according to Jaccard's metric. The average time of face separation using the HoG method was 0.2 seconds and when using trained models – 3.3 seconds. Using a combination of ROI and ORC separation methods makes it possible to significantly improve the accuracy of verification. The proposed client verification algorithm is implemented as an API on an ML server and integrated into the car sharing system

Keywords: know your customer, pattern recognition, optical character recognition, car sharing

Beibut Amirgaliyev
Candidate of Technical Sciences, Professor*
Gulzhan Yegemberdiyeva
Master of Computational and Cognitive Neuroscience, Lecturer*
Alexander Kuchansky
Corresponding author
Doctor of Technical Sciences, Head of Department
Department of Information Systems and Technology
Taras Shevchenko National University of Kyiv
Volodymyrska str., 60, Kyiv, Ukraine, 01033
E-mail: kuczanski@gmail.com
Yurii Andrashko
PhD, Associate Professor
Department of System Analysis and Optimization Theory**
Ihor Korol
Doctor of Physical and Mathematical Sciences, Professor,
Vice-Rector for Academic Policy and Research
Department of Algebra and Differential Equations **
*Department of Computer Engineering
Astana IT University
Mangilik Yel ave., 55/11, Nur-Sultan,
Republic of Kazakhstan, 010000
**Uzhhorod National University
Narodna sq., 3, Uzhhorod, Ukraine, 88000

1. Introduction

The increase in the number of cars leads to a wide range of environmental and economic problems [1]. Exhaust gases from cars significantly degrade air quality in a region. This is harmful to the health of the inhabitants of this region, which in turn leads to additional costs for their medical care. In general, emissions from personal cars account for almost 60.7 % of total CO_2 emissions from road transport in Europe [2].

The increase in the number of cars leads to a complication of traffic, the occurrence of congestion, which negatively affects the economic condition of the region and the state. That is why the governments of different countries are working on the implementation of concepts to reduce the number of cars in the traffic of cities [3]. One such solution is car

sharing. Some people use the same car when reaching their destination. Other people use the car until its main owner needs it. The concept of car sharing is gaining momentum, reaching a market value of more than USD 60 billion, and continues to grow. The volume of growth is projected at the level of 20 % in the coming years [4].

Some car sharing systems provide customers with their own cars, however, other car sharing companies are implementing a system in which people share their cars with other members of the community. The first approach is considered the traditional way of a car sharing company. The second approach is known as a peer-to-peer approach that requires fewer resources on the part of the company. For all these car exchange systems, one of the first and main steps is to get customers into the system so that they can legally drive a car.

For the car sharing company, the client's data must be available: name, ID number, driver's license number, photograph, etc. The main tasks of machine learning used by car sharing systems are face recognition and comparison, identification of identity cards and driver's licenses, character recognition and verification of documents for forgery, etc. This is important because companies will be able to protect themselves in the case of possible road accidents or replacing the driver with a person not registered in the system, etc. That is why the development of a user verification method in the car sharing system is a relevant task.

2. Literature review and problem statement

The development of user verification systems based on KYC (Know your customer) processes is a task for banking and exchange regulation for companies that work with the funds of individuals and financial institutions. The purpose of verification is to establish the identity of the counterparty before transactions are carried out. The principles of document verification and selection of key data from documents in KYC processes are often similar but they also have specific characteristics, in particular, depending on the data that need to be selected from the document. Paper [5] describes the advantages of using machine learning for the task of user verification.

In particular, machine learning methods allow for priority of calls when searching for matches, they eliminate errors and reduce efforts for further control and unload the personnel responsible for compliance. According to [5], learning with a trainer using decision trees makes it possible to get satisfactory recognition accuracy. Paper [6] describes technologies that can be used for the user verification process, in particular convoluted neural network (CNN), OpenCV, OCR and MRZ packets. The solution proposed in [6] involves scanning the document in order to obtain the necessary personal data from the passport, the validity of the passport, etc.

Then this information is compared with the information that is available in the user database in order to verify the submitted document. Based on the verification, the model proposed in [6] can classify the document as verified, overdue, canceled, or indicate that the data do not match. According to the calculations given in [6], the machine learning-based KYC solution reduces manual effort to test KYC by 70 %. However, in addition to choosing the technology for implementing the user verification system, it is necessary to develop a methodological apparatus for this task, in particular to ensure maximum accuracy of verification. This is due to the fact that in the

case of a false operation of the system during user verification in the car sharing system, extremely negative cases may arise, which will harm the stable development of the company. In particular, this may be due to insurance payments in the event of a traffic accident, etc. Some companies, including financial institutions, do not implement user verification systems or implement them at a low level. If a case of violation is detected, then such companies resort to some procedures to resolve the situation, which costs less than ensuring the full operation of the user verification system. This is described in [7]. However, if a verification error can cost people's lives and health, as is the case when implementing a car sharing system, user verification rules cannot be neglected.

In addition to the mathematical and technical problem of introducing user verification systems, there is an issue of ensuring the security of such systems. Therefore, the system should provide means of combating traditional attacks and a relatively new type of attack: face substitution (digital spoofing). Paper [8] describes the method of protecting identity documents from these attacks. Article [9] describes the algorithm for detecting deepfake in KYC systems, which achieves the Logloss result of 0.5189, which is one of the best results. The development of the direction of face identification to increase productivity and accuracy is described in [10]. In [10], an architectural solution using YOLO, MTCNN, FaceNet technologies was created, and noise reduction methods are described. That made it possible to increase the accuracy and performance of facial verification, however, for car sharing systems, the accuracy of identity verification is extremely critical, which should be as high as possible, even if it requires more system operating time. In [11], a new facial recognition scheme using deep learning is proposed and high detection performance is achieved on the FDDB (Face Detection Data Set and Benchmark) test.

Paper [12] describes the method of facial recognition, which uses an LBP-based classifier. An average of 3.2 seconds is required for image processing, in contrast to the classifier based on Haar functions, which requires an average of 6.3 seconds to process the image. The use of the Haar function for facial recognition is described in [13]. In general, combined car sharing systems that allow for security and confidentiality have now gained popularity, as well as authenticity, traceability and reliability, scalability and compatibility. In particular, blockchain and the Internet of Things technologies are used to ensure these tasks. Features of the implementation of such systems are described in [14]. Another task that can help verify a person is the use of OCR technology to recognize text from customer information in a photograph document. However, in [15] it is investigated that computer vision methods, which are an integral part of OCR, have relatively low recognition accuracy.

Automation of the verification process of clients of the car sharing system under conditions of limited time and information is an unresolved task. The problem can be solved on the basis of the KYC process using deep learning methods with a combination of OCR methods.

3. The aim and objectives of the study

The aim of this study is to develop a user verification algorithm in a car sharing system based on the KYC process. This will make it possible to increase the efficiency of car sharing system management.

To accomplish the aim, the following tasks have been set:

– to formalize the verification task of the user of the car sharing system and build an algorithm for the KYC process for this system;

– to choose methods for implementing the stages of the user verification algorithm in the car sharing system and check its operation on real data.

4. The study materials and methods

The object of this study is the method of user verification in the car sharing system based on the KYC process. To develop a method of user verification in the car sharing system, two groups of methods were used: machine learning methods and image processing methods. This is necessary for processing a photograph of a document with the identification of areas with the necessary information for verification. It is also important to provide system training to improve the accuracy of facial recognition.

A research hypothesis can be formulated: customer verification of a car sharing system can be automated with the required accuracy using machine learning approaches, in particular by deep learning with a combination of OCR methods. Confirming the hypothesis would improve the accuracy and speed of customer service. A car sharing management system was built to verify the user in which the data were previously prepared. The tasks of verification of the client by his/her ID card, driver's license, identification number, etc. have been implemented. For 2000 real persons, images of their documents have been prepared, as well as photographs of these persons with a document for training. For these images, bounding boxes have been selected that indicate objects in the images. Since the images contain personal data, only generalized results for their recognition are given. Accordingly, the dataset cannot be accessed publicly.

5. User verification method in the car sharing system

5. 1. Formalization of the user verification task and the Know your customer process algorithm

Convenient and reliable verification of the user of the car sharing system is one of the key components of the successful functioning of the car sharing system as a whole. The user verification task can be formalized as follows: using the available data, it is necessary to establish that the client sending the request corresponds to the profile in the database of the car sharing system. Significant limitations of this problem are incomplete input data. It is also significant to limit the time for which the task must be solved because the long wait for customers negatively affects their satisfaction with the operation of the system.

In the course of the study, it was established that the key data for user verification in the car sharing system are:

1) a photograph of the client's face;

2) a photograph of the identity document and/or a photograph of the driver's license;

3) personal data of the user: last name and first name, age, etc.;

4) data from public services.

The specified data can be relatively easily obtained and analyzed. The necessary and sufficient data for user verification are a photograph of the client at the moment when s/he forms a request to use the car sharing system, and data for comparison. An identity document is a reliable source that contains photographs and information about the client.

The solution to the verification problem of the user of the car sharing system is a certain Boolean value. This value is true if the obtained data confirm the client's profile, respectively, s/he is provided with a car sharing service. Another value of false means that the client has not passed verification. In this case, it is possible to provide the service personally by an employee of the company, after verification of the client based on the results of personal communication, or otherwise. Otherwise, the client will be denied the service.

To solve the problem of user verification in the car sharing system, the KYC process algorithm is proposed, which is shown in Fig. 1 and includes 6 main stages:

1) identification of the face in the photograph. That is, in the image that is obtained from the camera of the client's smartphone, it is necessary to select a rectangle that limits the face from all sides;

2) face comparison. It is necessary to establish a comparison of the face, which is separated from the client's photograph with the face in the photograph, which is retracted from the document certifying his identity;

3) verification of documents, in particular their validity period;

4) determining parts of the document that contain important information, which are called ROI (region of interest);

5) recognition of information in the ROI;

6) making the final decision on user verification of the car sharing system. In the case of verification, the client is provided with a car sharing service. Otherwise, the client is invited to contact the company employee personally to confirm the identity of the client and complete the verification.
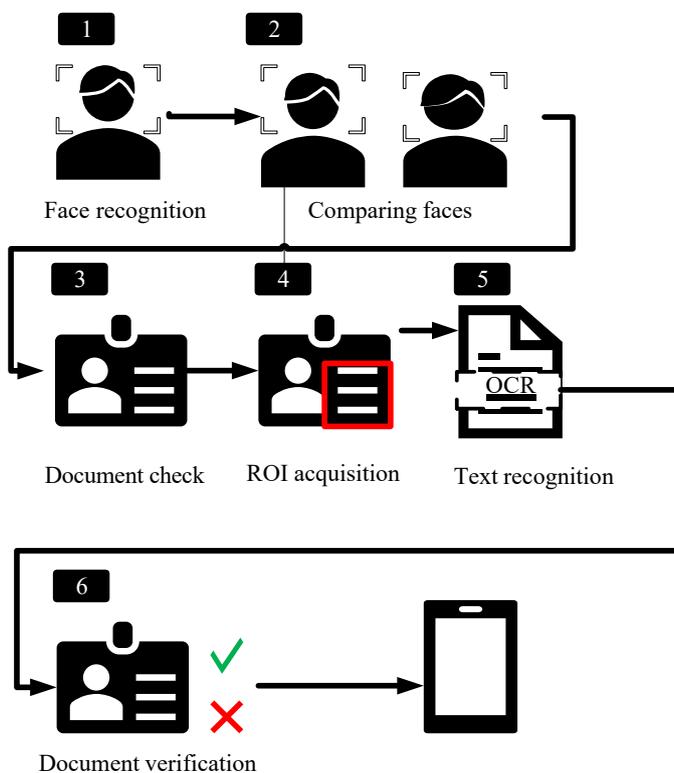
Fig. 1. Know your customer process algorithm for user verification in the car sharing system

5. 2. Selection of methods for implementing the stages of the user verification algorithm in the car sharing system

Separating the face in the client's photograph and comparing faces.

The identity document contains a photograph of the client. A prerequisite for user verification in the car sharing system is to download his/her selfie. Moreover, the client must hold an identity card in his hand. It is necessary to separate the face in the client's photograph and compare it with the face on the identity card.

To separate the face in the photograph, it is proposed to use a combined method that includes a histogram of oriented gradients (HoG) [12] and a convoluted neural network (CNN) [18]. The main difference between HoG and CNN methods is that HoG can only separate the face at small angles. This means that HoG is a good "frontal" face detector and cannot separate the face from different angles. CNN can separate faces from a wide range of angles. As a result of testing on the test dataset, it was found that HoG works 16 times faster than CNN. The average time of separation of the face using the HoG method is 0.2 seconds, and according to the CNN method – 3.3 seconds. Testing was carried out on the basis of the Intel Core I5-3570 CPU.

When separating the face, one needs to make sure that there are key landmarks on the face such as eyes, eyebrows, and lips. If one of these guidelines is missing, then the effectiveness of the method is significantly reduced. There are several cases where these facial landmarks may not be defined. Common cases are associated with the inability to identify the lips. In particular, given the fact that due to the spread of the COVID-19 pandemic most people began to wear face masks. Therefore, a necessary condition for user verification in the car sharing system is photographing without a protective mask on the face. Another case that prevents the establishment of a guideline on the face, in particular the eyes, is the wearing of glasses, a beard, or long hair. Given these cases, a user manual should be added to the system that will show one «what to do» and «what not to do» when taking selfies. Photographs that do not have one of the landmarks require separate manual verification, which makes it impossible to automatically verify the user.

The result of separating the face in the photograph are two rectangles. One limits the client's face to a selfie, the other limits the face on the document. It is necessary to compare these faces and determine whether they belong to the same person. For each face, a code is built in the form of a one-dimensional array. When using convoluted neural networks, facial encoding is obtained automatically as a value at the outputs of the previous layer of neurons.

To find the face code based on the HoG method, it is necessary to build HoG functions. Consider the image as a function of two variables $f(x, y)$. The value of this function corresponds to the color of the pixel. To find the HoG function, one must first calculate the centered horizontal and vertical gradients:

$$g_x = \frac{\partial f}{\partial x}, \ \ g_x = \frac{\partial f}{\partial y}, \tag{1}$$

where $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are derivatives in directions x and y, respectively. To calculate derivatives based on the direction, we use their approximation using a difference scheme.

Applying the values of derivatives for the direction, we calculate the direction and magnitude of the gradient. The direction of the gradient is determined by the angle of inclination θ:

$$\theta = \arctan \frac{g_y}{g_x}. \tag{2}$$

The length of the gradient is calculated from the formula:

$$\sqrt{g_x^{\ 2} + g_x^{\ 2}}. \tag{3}$$

The original face image is divided into 16 blocks horizontally and 16 blocks vertically. To obtain a vector of facial signs, the gradient of each of the blocks is calculated. In this implementation, the face is encoded by a one-dimensional array $k=[k_0, k_1, ..., k_{127}]$ with a length of 128. The values of each element of the array belong to the segment [−1; 1].

To establish the similarity of faces, we use the Euclidean distance between the vectors encoding these faces. Let k and p be the vectors encoding the face. The distance between them is calculated from the formula:

$$\rho(k, p) = \sqrt{\sum_{i=0}^{127} (p_i - k_i)^2}. \tag{4}$$

If the distance is close to 0, then this means that the faces are similar. Otherwise, this means that the vectors are not similar, and the probability that the faces belong to one person, respectively, is low. To make a decision on the similarity of faces, a threshold value is used. Empirically, it has been established that for the problem of establishing the similarity of faces, the best is the threshold value, which is equal to 0.6. Therefore, if the distance between the vectors is less than 0.6, then this means that two faces belong to one person, otherwise – to different people.

The solution to establish facial similarity can be implemented using a wide range of algorithms [17–19]. In [20], a comparison of five methods for identifying a face was made: Haarcascade, OpenCV DNN, Dlib, MTCNN (Multi-task Cascaded Convolutional Networks), Facenet. In general, these algorithms show similar results regarding the accuracy of identification. However, according to the time criterion, the fastest is the Dlib algorithm – 30 seconds. Next are the Facenet and OpenCV DNN algorithms – about 50 seconds, Haarcascade approx. 100 s, and MTCNN approx. 300 s. The comparison diagram is shown in Fig. 2.
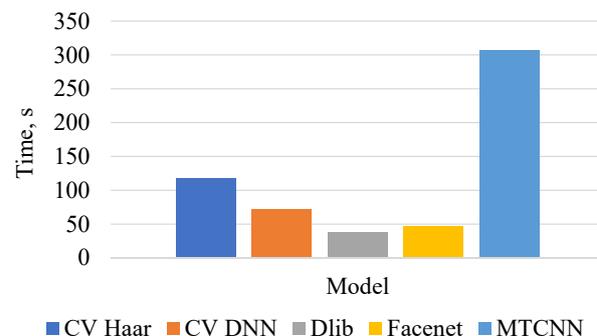
Fig. 2. Comparison of facial recognition methods by processing time criterion

To implement the KYC process in the car sharing system, the Dlib model was used [18].

Verification of identity card and driver's license.

The next step is to highlight key data from photographs of documents that make it possible to identify the client. It is usually enough to highlight the identification number since the rest of the additional information about the client can be obtained by this number.

To highlight key data from photographs of documents, it is necessary to train the appropriate ML (Machine Learning) model. For training, one needs to prepare a data set. Fig. 3 shows an example of images used for training and testing. The set contains images of ID and driver's license. The images show two points each as the corresponding documents are separated.

In this system, the MaskRCNN (region-based convolutional neural network) method was used to verify the user [21]. The general architecture of the MaskRCNN method is illustrated in Fig. 4.

MaskRCNN helps segment the identifiable object. After identifying the object, a bounding frame is constructed by setting the upper left and lower right coordinates. The use of MaskRCNN showed 91 % accuracy for the user verification task in a KYC-based car sharing system. For evaluation in this case, the Jaccard metric was used [22], which is calculated as the ratio of the intersectional area of rectangles limiting the isolated object and its true boundaries to the area of combination of these rectangles. The value of the metric belongs to the segment [0, 1]. In Fig. 5, the intersection is marked with a green rectangle, and the combination is marked with a red polygon.

Fig. 4. MaskRCNN (region-based convolutional neural network) algorithm scheme
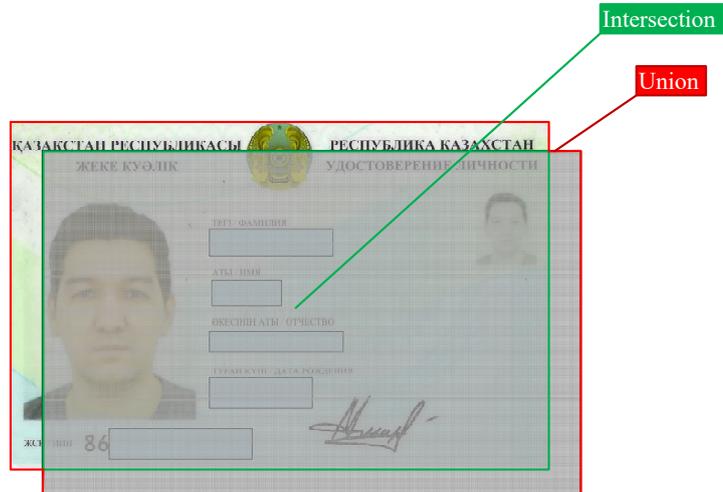
Fig. 5. Calculation of Jaccard's metric to highlight an identity card

Fig. 3. Sample data from a test set

The identification of a driver's license is generally similar to the task of identifying an ID. For the KYC user verification process in the developed car sharing management system, the MaskRCNN method is also used. The accuracy of determining the driver's license using this method is equal to 91 %. About 2000 sample data were used, of which 80 % of the data were used for training, and the remaining 20 % of the data were used for testing.

ROI acquisition, text recognition, and document validation by the system architecture.

The necessary requirements for customer verification in the car sharing system are to obtain an identification number and the expiration date of the driver's license. In the case of kazakhstani documents, the guideline for highlighting these data is a chip for a driver's license or face. MaskRCNN was used to detect the chip. According to this concept: 80 % of the data were used for training, and the remaining 20 % of the data were used for testing.

Once the ROI is highlighted and cut, the OCR procedure is activated. OCR consists of three components: text detection, character segmentation, character recognition. For the segmentation of characters in the developed system, the open-source library tesseract is used. However, before recognizing the text, one needs to clear the data from noise. To this end, one needs to use a Gaussian filter, alignment histograms, and combinations of morphological operations such as erosion and expansion, etc. Acquiring ROI based on the fixed location of the document's landmarks is shown in Fig. 6.

The methods proposed in this study are implemented on the server. Access to functions occurs through the corresponding API (Application Programming Interface) (Fig. 7).

The server is divided into 2 parts: ML server and backend server. ML server contains trained models. All customer requests that send them using the mobile application are processed by the backend server. The backend server makes

appropriate requests to the ML server. ML server separates and compares the face in the photograph, detects ROI, and recognizes the information in the ROI. Based on these data, the backend server verifies the identity of the client and sends the response to the mobile application.
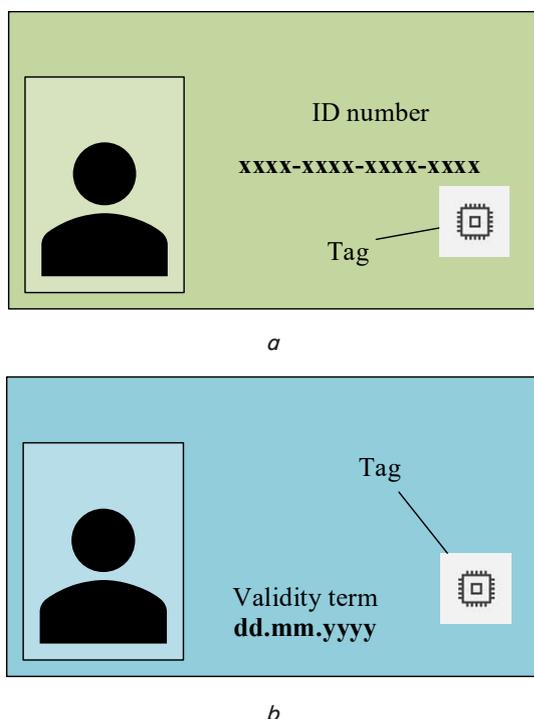
Fig. 6. Acquiring ROI (region of interest) based on a fixed location of the document's landmarks:
a — identification number; *b* — validity period
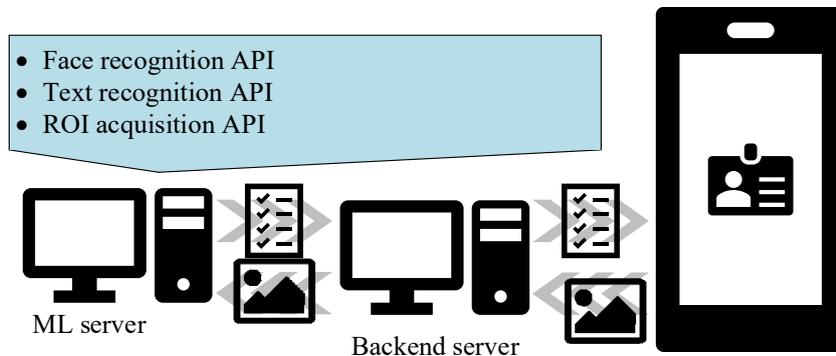
Fig. 7. Information system architecture

6. Discussion of the development of a user verification method in the car sharing system

A key feature of the system is the speed of verification. Initially, the system verifies the user using the HoG method (accuracy of 91 % of clients); in the case of unsuccessful verification, the system applies a neural network approach. If, in this case, it was not possible to verify the user, then a request for manual verification by connecting to the operator occurs. This approach, implemented on the server (Fig. 7), allows for fast customer service despite the low accuracy of verification. That is, the goal of the study has been achieved: an algorithm for verifying clients of the car sharing system

has been developed. That has made it possible to improve the management of this system. Our results are explained by the use of a combined verification method, which, in the case of unsuccessful verification, makes it possible to verify the client according to the neural network approach. Unsuccessful verification may be the result of a poor-quality selfie of the client, or a significant change in the appearance of the client, in particular the appearance of a beard or glasses.

The disadvantage of the study is the relatively small volume of the client base on which the system was trained: 2000 photographs. Also, a disadvantage is that the study does not analyze separately photographs of clients of different races. It is implied that in this case there may be a difference in verification.

A procedure based on the KYC process makes it possible to solve the problem of verifying the user of the car sharing system. But a number of related problems arise. Some additional features can be added to the KYC verification process in the car sharing system. In particular, gender recognition, customer age recognition, etc. These features help one learn more about your customers so that this information can then be used for other purposes, including marketing. Identifying the specified characteristics is called establishing the demographic characteristics of customers. Demographic attributes describe the location, gender, and age preferences of customers. If the true location of the client can be easily determined using GPS, then determining the age and gender of customers is considered a more challenging task. The age forecast is considered as a regression problem, where at the output one obtains a numerical indicator for assessing the client's age. To determine the sex, the method of classifying photograph is used, where the input photograph must be assigned to one of the limited classes.

Another important function is to establish that a real person is verified in the system. To this end, the client is given commands regarding face movements, turning the head and it is monitored whether the client performs these commands correctly. If the client executes these commands well, then the verification is successful. Otherwise, the request is rejected, and a KYC process request is sent to validate the client manually. The list of commands for the client may be as follows: close your eyes, close your right/left eye, turn your head left/right, look up/down, etc. If at least three commands are executed by the client, then it is considered that the client is a real person.

The limitation in the system is the quality of the photograph. The image quality may vary depending on the type of smartphone and the characteristics of the camera. Low camera parameters and poor lighting conditions create low-quality images. This, in turn, will lead to poor performance of the KYC ML system where end customers will have problems during verification. Therefore, one can first improve the quality of the images, so that the text on the documents is more readable, and there is less noise. The first case can be solved using deblurring algorithms, where the main function is to improve image quality. To eliminate blur, one can use generative networks (GAN) [23] and various types of noise removal algorithms [24, 25]. The results of verification in combination with the use of a dynamic model of profit maximization [26] are important components of the successful functioning of the car sharing

system. To predict the profit of the car sharing system, one can use combined methods of forecasting and analyzing time series [27–30]. It is in this direction that the development of this study may proceed.

7. Conclusions

1. We have formalized the verification task of the user of the car sharing system and the KYC process algorithm for this system was constructed. The algorithm consists of the following stages: identifying the face in the photograph, comparing faces, checking documents and their validity period, establishing and recognizing ROI, formulating the final verification solution. The algorithm contains components that are necessary for user verification in the car sharing system. This is required to prevent losses of the company in the event of road accidents or other unforeseen situations when using car sharing services by the client.

2. For the KYC client verification process, methods for each stage of the verification algorithm have been proposed and a system architecture that is tested on real data was suggested. The system with the proposed architecture made it possible to verify 91 % of customers. To train the system, 2000 photographs of clients and their photographs with documents were introduced. The key feature of the system is the speed of verification, which is lower in the case of using neural network methods. The average time of face identification using the HoG method is 0.2 seconds and using trained models – 3.3 seconds. The proposed architecture can be applied to various car sharing systems. The proposed approaches are based solely on computer vision and machine learning solutions. The machine learning server can be expanded by adding additional features such as gender classifier and user age prediction.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Acknowledgments

References

1. Cerovsky, Z., Mindl, P. (2008). Hybrid electric cars, combustion engine driven cars and their impact on environment. 2008 International Symposium on Power Electronics, Electrical Drives, Automation and Motion. doi: https://doi.org/10.1109/speedham.2008.4581321

2. CO_2 emissions from cars: facts and figures (infographics) (2019). EU Parliament News. Available at: https://www.europarl.europa.eu/news/en/headlines/society/20190313STO31218/co2-emissions-from-cars-facts-and-figures-infographics

3. Kagawa, S., Hubacek, K., Nansai, K., Kataoka, M., Managi, S., Suh, S., Kudoh, Y. (2013). Better cars or older cars?: Assessing CO_2 emission reduction potential of passenger vehicle replacement programs. Global Environmental Change, 23 (6), 1807–1818. doi: https://doi.org/10.1016/j.gloenvcha.2013.07.023

4. Shared mobility (2020). McKinsey Center for Future Mobility. Available at: https://www.mckinsey.com/features/mckinsey-center-for-future-mobility/overview/shared-mobility

5. Why Machine Learning brings up to 57 percent savings in the KYC process. Actico. Available at: https://www.actico.com/blog-en/why-machine-learning-brings-up-to-57-savings-in-the-kyc-process/#:~:text=With%20their%20feedback%2C%20they%20improve,to%2057%20percent%20of%20clarifications

6. Sitaraman, S. (2019). How we built an intelligent automation solution for kyc validation. Available at: https://www.srijan.net/resources/blog/intelligent-automation-solution

7. Charoenwong, B. (2019). The one reason why ai/ml for aml/kyc has failed (so far). Available at: https://medium.com/@BenCharoenwong/the-one-reason-why-machine-learning-for-aml-kyc-has-failed-so-far-38e3388ad85c

8. Pic, M., Mahfoudi, G., Trabelsi, A. (2019). Remote KYC: Attacks and Counter-Measures. 2019 European Intelligence and Security Informatics Conference (EISIC). doi: https://doi.org/10.1109/eisic49498.2019.9108787

9. Do, T.-L., Tran, M.-K., Nguyen, H. H., Tran, M.-T. (2021). Potential Threat of Face Swapping to eKYC with Face Registration and Augmented Solution with Deepfake Detection. Lecture Notes in Computer Science, 293–307. doi: https://doi.org/10.1007/978-3-030-91387-8_19

10. Darapaneni, N., Evoori, A. K., Vemuri, V. B., Arichandrapandian, T., Karthikeyan, G., Paduri, A. R. et. al. (2020). Automatic Face Detection and Recognition for Attendance Maintenance. 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS). doi: https://doi.org/10.1109/iciis51140.2020.9342670

11. Sun, X., Wu, P., Hoi, S. C. H. (2018). Face detection using deep learning: An improved faster RCNN approach. Neurocomputing, 299, 42–50. doi: https://doi.org/10.1016/j.neucom.2018.03.030

12. Chang-Yeon, J. (2008). Face Detection using LBP features. Available at: http://cs229.stanford.edu/proj2008/Jo-FaceDetectionUsingLBPfeatures.pdf

13. Mita, T., Kaneko, T., Hori, O. (2005). Joint Haar-like features for face detection. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. doi: https://doi.org/10.1109/iccv.2005.129

14. Malhotra, D., Saini, P., Singh, A. K. (2021). How Blockchain Can Automate KYC: Systematic Review. Wireless Personal Communications, 122 (2), 1987–2021. doi: https://doi.org/10.1007/s11277-021-08977-0

15. Liu, Y., James, H., Gupta, O., Raviv, D. (2021). MRZ code extraction from visa and passport documents using convolutional neural networks. International Journal on Document Analysis and Recognition (IJDAR), 25 (1), 29–39. doi: https://doi.org/10.1007/s10032-021-00384-2

16. Dalal, N., Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). doi: https://doi.org/10.1109/cvpr.2005.177

17. Face recognition. Available at: https://pypi.org/project/face-recognition/

18. Cnn based face detector. Dlib. Available at: http://dlib.net/cnn_face_detector.py.html

19. Face Recognition. URL: https://github.com/ageitgey/face_recognition/

20. Stefanovic, S. (2020). Face detection algorithms comparison. Available at: https://datahacker.rs/017-face-detection-algorithms-comparison/

21. He, K., Gkioxari, G., Dollar, P., Girshick, R. (2017). Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV). doi: https://doi.org/10.1109/iccv.2017.322

22. Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's Cosineversus the Jaccard index. Journal of the American Society for Information Science and Technology, 59 (1), 77–85. doi: https://doi.org/10.1002/asi.20732

23. Lu, B., Chen, J.-C., Chellappa, R. (2020). UID-GAN: Unsupervised Image Deblurring via Disentangled Representations. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2 (1), 26–39. doi: https://doi.org/10.1109/tbiom.2019.2959133

24. Verma, R., Ali, J. (2013). A comparative study of various types of image noise and efficient noise removal techniques. International Journal of advanced research in computer science and software engineering, 3 (10), 617–622. Available at: https://www.researchgate.net/publication/307545428_A_comparative_study_of_various_types_of_image_noise_and_efficient_noise_removal_techniques

25. Yegemberdiyeva, G., Amirgaliyev, B. (2021). Study Of AI Generated And Real Face Perception. 2021 IEEE International Conference on Smart Information Systems and Technologies (SIST). doi: https://doi.org/10.1109/sist50301.2021.9465908

26. Amirgaliyev, B., Andrashko, Y., Kuchansky, A. (2022). Building a dynamic model of profit maximization for a carsharing system accounting for the region's geographical and economic features. Eastern-European Journal of Enterprise Technologies, 2 (4 (116)), 22–29. doi: https://doi.org/10.15587/1729-4061.2022.254718

27. Kuchansky, A., Biloshchytskyi, A., Andrashko, Y., Biloshchytska, S., Honcharenko, T., Nikolenko, V. (2019). Fractal Time Series Analysis in Non-Stationary Environment. 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T). doi: https://doi.org/10.1109/picst47496.2019.9061554

28. Kuchansky, A., Biloshchytskyi, A., Andrashko, Y., Biloshchytska, S., Shabala, Y., Myronov, O. (2018). Development of adaptive combined models for predicting time series based on similarity identification. Eastern-European Journal of Enterprise Technologies, 1 (4 (91)), 32–42. doi: https://doi.org/10.15587/1729-4061.2018.121620

29. Kuchansky, A., Biloshchytskyi, A., Bronin, S., Biloshchytska, S., Andrashko, Y. (2020). Use of the Fractal Analysis of Non-stationary Time Series in Mobile Foreign Exchange Trading for M-Learning. Advances in Intelligent Systems and Computing, 950–961. doi: https://doi.org/10.1007/978-3-030-49932-7_88

30. Kuchansky, A., Biloshchytskyi, A. (2015). Selective pattern matching method for time-series forecasting. Eastern-European Journal of Enterprise Technologies, 6 (4 (78)), 13. doi: https://doi.org/10.15587/1729-4061.2015.54812