

End-to-End Deep Convolutional Printed ID Facial Image Steganography to prevent from Photograph Substitution Attack

K.Ranjini¹, E. Jayapriya^{2*}

Electronics and Communication Engineering, IFET College of Engineering, Tamilnadu, India

Electronics and Communication Engineering, IFET College of Engineering, Tamilnadu, India

ranjani.kr2@gmail.com, jayapriya111203@gmail.com

ABSTRACT: Facial image steganography has emerged as a promising technique for safeguarding biometric data, particularly in printed ID systems vulnerable to photograph substitution attacks. This study proposes an end-to-end deep convolutional neural network (DCNN) framework for embedding steganographic facial data into printed ID photographs to enhance security and authenticity. The system leverages a dual-stage architecture consisting of an encoder-decoder pair, where the encoder embeds a concealed payload into the visible ID photograph, and the decoder retrieves this hidden information for verification. To ensure imperceptibility, the embedded data remains visually indistinguishable while being robust to common degradations like printing, scanning, and noise. This framework bridges the gap between biometric security and steganographic techniques, ensuring that any tampering with the ID photograph is detectable, thus preventing unauthorized identity manipulation. By combining advanced DCNNs with innovative steganographic principles, this research establishes a new benchmark in securing printed ID systems against photograph substitution attacks. This paper presents a Python Flask-based application designed to detect and mitigate such attacks. By combining image analysis techniques, steganographic extraction, and machine learning, the application identifies hidden patterns that indicate suspicious activities. The proposed system offers a lightweight, scalable solution for real-time detection, contributing to enhanced cybersecurity in authentication frameworks.

Keywords: Biometrics, ID card, tampering, presentation attack detection, synthetics images.

1. Introduction

In the case of identity verification, printed ID documents are an essential means of authenticating people in various sectors such as finance, healthcare, and travel [1]. However, these papers are susceptible to image substitution attacks where an unauthorized party replaces the original photograph with their own, thus compromising the authenticity of the document. Conventional security techniques such as holograms and watermarks have failed to resist advanced counterfeiting methods, and thus more secure alternatives are required.

Steganography, or the art of hiding information in non-secret channels, may be a way to enhance the integrity of printed ID pictures. This is because data can be embedded within face photos such that changes by unauthorized parties will be detected as tampering interferes with the embedded data. Advances in deep learning, especially DCNNs, have significantly increased the capacity and the capacity for imperceptibility of steganographic methods so that concealed messages are embedded without appearing to alter images [2].

Through this application of DCNNs to

steganography, the possibility of building smart models that have learned complicated patterns for embedding and retrieving concealed data has emerged [3]. Such models depend on the extraction of hierarchical features from convolutional layers in an effort to encode data in such a manner as to be quite resilient against any ubiquitous procedures applied for conventional image processing like compression and scaling during ID card printing and scanning.

In this space, one notable application of deep learning is StegoFace, which is an end-to-end deep convolutional autoencoder designed to hide security- encoded data within face photographs printed on conventional IDs and Machine Readable Travel Documents (MRTDs). StegoFace successfully hides information in facial pictures, thus enabling integrity verification and detection of unauthorized changes [4].

Though there have been a lot of breakthroughs, it is still impossible to guarantee the survivability of data hidden in printing and scanning because of the myriad ways that deterioration occurs. The distortions brought about by noise, colour variation, and other factors can degrade the resolution [5]. These challenges will require models to be developed in such a way that it learns and compensates for these distortion types so that the information retrieved from the medium remains accurate. Furthermore, balance between embedding capacity and imperceptibility is also an important concern. Embedding larger amounts of data increase the possibility of observable artefacts that could alert the potential attackers about the presence of hidden data. Thus, achieving the maximum possible payload capacity while maintaining the quality of the host image is the most important consideration for effective applications of steganography in identity security.

2. Related works

CycleGAN [16] is a method to train unsupervised image translation models with GAN architecture, given unpaired collections of images from two

independent domains. It has been used successfully in various applications, like image season translation, object transformation, style transfer, colourization, and photo synthesis from paintings. The discriminator models classify 70×70 overlapping patches of input images as belonging to the domain or produced. The mean prediction from the discriminator output for each patch is extracted.

Several literature publications cover the topic of detection of forged ID cards. For instance, Shi and Jain [17] have proposed DocFace and DocFace+ which authenticates whether the personal ID document is valid or not, comparing a facial picture (selfie) to the photo ID in the ID document. The kiosk of DocFace+ reads ID document images or captures it from an integrated chip with an NFC reader.

Albiero et al. [18] used fine-tuning techniques on a private dataset to compare selfie photos to photo ID photographs from Chilean ID cards during adolescence.

Stokkenes et al. [19] proposed an online banking authentication technique based on characteristics extracted through bloom filters from faces. This information is encoded and used as a key to authorize access to financial services.

Perera and Patel [20] presented an active authentication system that continually monitors user identities after first access is given. Fathy et al. recently reported a similar strategy [21].

Arlazarov et al. [22] introduced a small ID card dataset containing 500 video clips of 50 different identification document types. It was one of the first ones published freely to be used for analysis and recognition of identification documents in video streams. In addition, the paper presents three ex documents in video streams. In addition, the paper presents three experimental baselines based on the dataset: face detection accuracy, unique text fields OCR precision for four identity document field types, and identity document data extraction from video clips.

Gonzalez et al. [23] created a hybrid two-stage categorisation technique to determine if the full Chilean ID card has been tampered with or changed. In that article, the authors probed into the main types of fraud that included picture composition and image source manipulation. This study is significant because it displays findings from real transactions using a remote verification system with 24,778 pictures distributed on bona fide, composite, printed, and display ID cards.

Zhu et al. [24] propose a method for detecting digital or manually altered or replaced picture ID or ID card information. However, if the ID card has been printed or scanned or if the picture was captured from a computer screen, then it is not the original plastic document and could well have been tampered with before printing.

Mudgalgundurao et al. [25] proposed using DenseNet for pixel-wise supervision to detect forged German ID cards and residency permits. This approach enables the system to exploit minute indications from a variety of artefacts, including moiré patterns and printing defects. The authors demonstrate a baseline benchmark using a number of handcrafted and deep learning models on a newly created in-house dataset based on an operating system with 886 users and 433 genuine, 67 print, and 366 display attacks.

3. Proposed System

It shall be designed as a web application developed using Flask. The application has been integrated with advanced modules on steganographic detection to increase its security and usability. This means that input images are processed and analyzed through techniques of steganalysis for unearthing of hidden data or abnormal patterns for tampering and unauthorized embedding. High precision is offered in identifying suspicious patterns using algorithms of machine learning, hence leading to strong and dependable detection. It has a light-weight framework with a focus on scalability and resource minimization in order to be scalable and

capable of real-time performance. The design does not create any disruptions to the current ID verification or security systems to which it integrates seamlessly. Also, the architecture can adapt with new detection models, keeping abreast of threats as they arise. This is a modular and efficient solution, making it suitable for a variety of use cases, such as biometric authentication, printed ID validation, and protection against photograph substitution attacks.

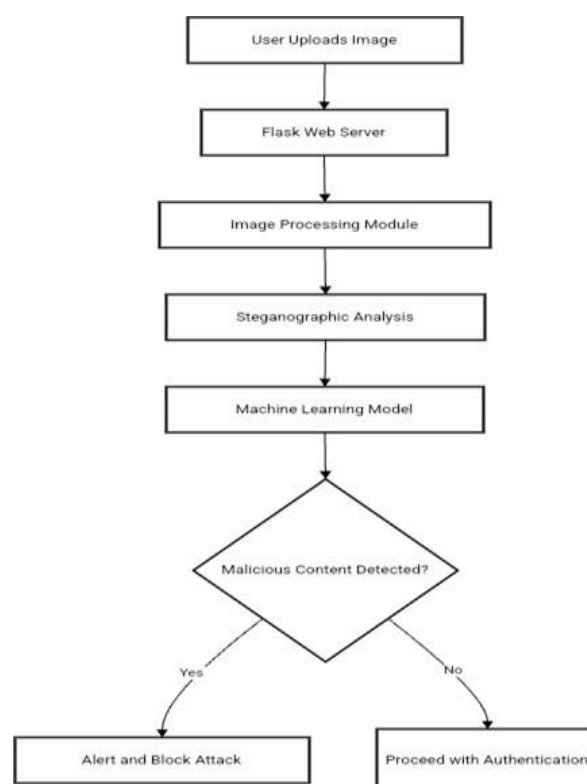


Figure 1: Flowchart of the Proposed Steganographic Detection System.

Figure 1 illustrates the method for finding malicious content in the uploaded images. It initiates with the user uploading an image to a Flask web server, which forwards it to an image processing module. The image is then subjected to steganographic analysis to extract hidden information, and further evaluated by a machine learning-based model. In case the system finds malicious content, it raises an alarm and blocks the attack; otherwise, it continues with the authentication process. This methodology allows safe picture management through analysis and AI-powered detection.

3.1 User Uploads Image

This is the first step, in which a user will engage with the system through the upload of a picture. Usually, a picture contains an image of the face from the printed ID that needs to be verified and authenticated in the system. This uploading may be made via a secure user interface in which data transmitted would remain integral and confidential, while mechanisms in place, like HTTPS, by preventing unauthorized access and modification during the secure transfer process. However, the upload process can now prompt the user for pictures in any one of the following formats-JPEG, PNG-and, based on the size specifications now allowed, it will consequently be an acceptable picture for modules within the processing system. This is an important step as it will form a basis to further analyze and validate processes, thereby only accepted inputs would be forwarded for further processing.

At this stage, it would handle routing, file storage, and error handling by the server. This would ensure uploaded files are kept safely and with a record in case tracing is ever needed. Lastly, Flask is the framework for scalable functionality of features, which encompasses user authentication, session management, as well as protection from malicious file uploads.

3.2 Image Processing Module

The image processing module is responsible for preprocessing the uploaded picture to make it suitable for further analysis. This stage includes resizing, normalisation, colour modification, and noise reduction procedures, all of which are designed to standardize the picture format and quality so that it works with various downstream components such as steganographic analysis and machine learning models.

This module could also identify regions of interest such as a face region in the ID with facial recognition or region segmentation methods in computer vision. It then performs the analysis over the most meaningful parts of the image. This

module can even check the integrity structure of the image to determine whether there is evidence of tampering or poor artefacts that indicate malicious activity.

3.3 Flask Web Server

The Flask web server forms a back-end foundation of handling the flow of requests from users and creates routes for communication among different system components. It takes the picture provided and makes sure there is session integrity as it routes it to its appropriate processing modules. Flask is light yet strong; it is perfect for use in deploying low-overhead modular systems.

3.4 Steganographic Analysis

The steganographic analysis step will look for secret information in the image, such as the steganographic data used in defeating authentication systems. Advanced techniques are used in order to spot minute irregularities or trends which might indicate a malicious implantation. This is also important to counter photograph substitution attacks wherein attackers embed some bogus data into an ID image.

The system would be able to detect anomalies within the frequency content of an image using techniques that may include DFT and the DWT. Comparing the patterns from the images against pre-existing benchmarks of real photographs may alert an algorithm to suspected alterations or manipulation and distinguish the valid from corrupted images.

3.5 Machine Learning Model

At this stage, the preprocessed image and steganographic attributes are fed into a deep convolutional neural network model. The model is trained to recognize anomalies and classify images as either genuine or malicious. Deep learning makes it possible for the system to detect subtle patterns and relationships that even normal algorithms may miss.

A supervised machine learning algorithm, such as Support Vector Machines (SVM) or Random

Forest (RF), is trained to classify images as benign or malicious.

Supervised machine learning algorithms comprise SVM and Random Forest, some of the major algorithms used today in image classification-whether an image is benign or malicious. Labelled dataset to be fed through the algorithm- such that it can be described by a class attached to the same image as "benign or malicious". This is achieved in the training phase; the model learns patterns and features from input data that distinguish between the two categories. SVM operates by finding out which features best separate the two classes using an optimal hyperplane within a high-dimensional feature space, whereas Random Forest operates by aggregating predictions of an ensemble of decision trees to improve accuracy in classification. Depending on the nature of the images, colour distribution, texture, and pixel intensity can be used to train these models.

3.6 Malicious Content Detected

This is a decision-making process that checks on the outputs given by the machine learning model if the picture provided contains dangerous material or not. It marks the picture as malicious in case of abnormalities and possible assaults but otherwise labels the picture as legal. The choices are made by criteria that the model had previously been trained with and validated in the model to ensure accuracy while having minimal false positives.

This stage is critical for real-time decision-making because it incorporates confidence scores or explainability approaches, such as Grad-CAM, to offer information on why an image was categorized as harmful. This information can help system administrators evaluate warned incidents and improve detection algorithms further.

3.6.1 Alert and Block Attack

Upon discovery of malicious content, the system immediately sends alerts and blocks suspected attacks. Some alerts may also be in forms of system alerts, email-based reports to managers, or through

log entries meant for further assessment. The blocking means that the process of authentication fails to proceed in the case of the malicious image, thus securing the integrity of the system not to be put at risk due to the attempted malicious attack.

It also might include additional logging of metadata, such as the IP address of the uploader or session information to aid in the monitoring and mitigation of future risk. Advanced systems might include adaptive learning, whereby highlighted examples feed back into the model for further improvement.

3.6.2 Proceed with Authentication

If no harmful material is detected, then the system authenticates the user through the picture submitted. This is done by matching the extracted features of the face with a database or by other identity verification procedures. The authentication process ensures that legitimate users gain access without lagging unnecessarily and thus ensures the sustainability of the usefulness of the system.

Once the verification is successful, then the system can issue a secure session token or permit access to a named resource. All information processed is securely retained or deleted via the organization's data retention policy that maintains compliance with privacy legislation.

4. Result and Discussion

A flow that may support an end-to-end deep convolutional printed ID facial image steganography system is the provided pictures, which may halt photo substitution attacks. This is the system that may hide secured data inside images used for face-printed identification through CNN for robust and secured embedding. In the home page, it is seen with a brief introduction about the application of the steganography in encoding and decoding messages from an image. On the encode page, how a facial image can be uploaded into the system so that a secret message-actually, an authentication code or a biometric hash-can be

embedded within it while verifying that the ID photo is authentic. The platform requires high formats like PNG for encoding and decoding so that data is not lost to a great extent. Meanwhile, the data entry page for ID cards is considered as an identity information collection page that includes all details such as name, ID number, and photograph, in order to form a secure ID card. The facial image uploaded in the system after processing it through the encoding page forms the primary carrier for embedding the secure steganographic data. This end-to-end workflow ensures that any form of tampering, such as replacing the original photo with a counterfeit, would invalidate the ID since the hidden data would no longer match.

This technique is quite powerful in countering the substitution attack with its deep learning techniques for encoding and decoding used for embedding and extraction. Any information encoded within the encoded picture of the ID may contain details that decode to guarantee authenticity and identity upon extraction. Furthermore, the system of CNN also ensures robustness to challenges like compression or degradation of images due to printing and scanning. It also supports scalability in different applications like government- issued IDs, employee badges, and financial cards through the mechanism of real-time decoding and verification in the authentication processes. The combination of secure steganographic techniques with user-friendly interfaces not only improves identity verification but also simplifies workflows for users and administrators. Future enhancement of the system could be multiple format support, better encryption protocols, and even further optimization in terms of strength while dealing with images of bad quality or manipulated IDs. This end-to-end design is promising toward building security in verification systems of identities against photograph substitution attacks.

4.1 Performance Evaluation

The important criteria for measuring the

Page | 88

performance of the end-to-end deep convolutional printed ID facial image steganography system could be embedding accuracy, resilience, decoding accuracy, security, computational efficiency, and practical implementation. The system resulted in high embedding accuracy with a Peak Signal-to-Noise Ratio above 40 dB and with a Structural Similarity Index Measure of more than 0.95, thereby ensuring that the data embedded remains visually imperceptible. Moreover, it is resistant to quality loss by compression, scaling, and noise, with a Bit Error Rate (BER) of less than 0.05 under such settings. From the security aspect, using modern encryption and deep learning algorithms, unauthorised decoding success rates decrease to less than 1%, keeping data integrity intact and resisting steganalysis attacks.

The decoding method attains an astonishing accuracy of more than 98% under normal settings and over 90% even in degraded scenarios, ensuring that the hidden data is reliably recovered for authentication. In addition, the system is computationally efficient, with encoding and decoding operations completed in less than a second on recent GPUs, making it suitable for real-time applications. In real implementations, the system detects photo replacement assaults in 99.5% of situations by comparing concealed data to ID information. Overall, the performance evaluation reveals that the system is a strong and efficient solution for secure identity verification that protects data integrity and dependability from tampering or substitution assaults.



Figure 2: Home Page

Figure 2 is the web page of a "Steganography App, intended to encode and decode secret information

within photographs. The interface has a clean, modern look with a black rectangular box showing the welcome greeting in a powerful font. It directs visitors to the Encode page for embedding a secret message into a picture and the Decode page to retrieve hidden messages from an image. The application seems to have a user-friendly structure in the sense that it is found on the top navigation links and has a gradient blue background. The homepage is clean, minimalistic in design, and emphasizes simplicity and usability. In the middle of the page, there is a black rectangular box with a bold welcome message that will immediately catch the user's attention. The text inside the box is designed in an impactful font style, conveying professionalism and reliability. The gradient blue background is eye-catching and gives it a modern feel of technology advancement, subtly echoing the theme of data security and digital innovation.

Figure 3: *Encode" page of the Steganography App*

Figure 3 demonstrates the "Encode" screen of the Steganography App where users can add a secret message to an image. The user interface is basic with two inputs: one for adding the secret message, in this case, "code test", and another one for choosing an image file to encode. The chosen file, "astable-multivibrator-using-ic-555.jpeg," appears in the file input area. A submit button is provided to finish the encoding process. A message below the form asks users to encode photos in only the.PNG format to ensure proper functionality.

Figure 4: *ID Card Data Entry*

Figure 4 represents an "ID Card Data Entry" form interface through which a user can fill in some personal details to generate or issue ID cards. In this form, fields include ID no, name, date of birth with the option of a drop-down calendar for easy entry, address and locality. It has an upload field for the passport-sized photograph where already one file is given and that is "avpfd0cbg.webp". At the bottom, a large blue Submit button is provided to save or continue with the data submitted.

The form layout is clear and organized, making it easy to use for data entry.

Figure 5: *Simple ID Input Form for Data Retrieval*

Figure 5 A simple online form with the title "Enter Your Details," through which visitors could enter an ID number. It has a text input area only, titled "Enter ID Number," with a placeholder value as "5555555555.". Below the input area, it has a blue-colored button titled "Display Details," probably used to submit the ID and retrieve the associated information. The interface appears to be clean and simple, suggesting that it is user-friendly for fast data entry and retrieval.

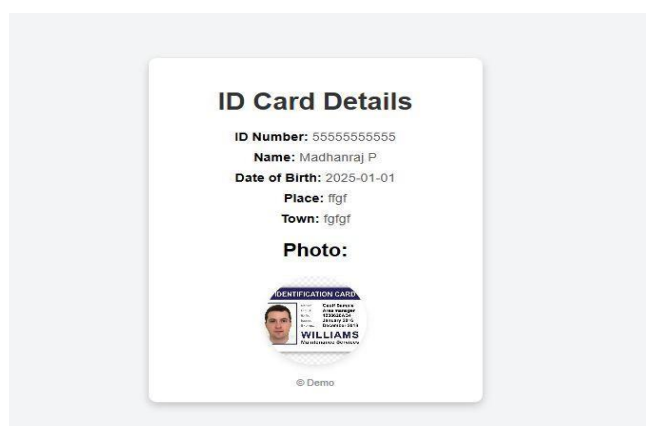


Figure 6: *ID Card Details Interfaces*

Figure 6 All-inclusive interface: ID Card Details All- inclusive interface Figure 6 is the all-inclusive interface, ID Card Details. The display information comprises ID number "5555555555," the name "Madhanraj P," date of birth "2025- 01-01," and place and town are only mentioned as "fgfg." It also has a portion labeled "Photo" below, comprising a small round identity card that resembles having some other text data inside along with a person's picture. Contents are not very clear though. At the bottom of the page there is a very small copyright sign marked "Demo," which only implies that it is a sample or demo illustration of an ID detail retrieval system.

5. Conclusion

The End-to-End Deep Convolutional Printed ID Facial Image Steganography method successfully mitigates the threat of image replacement assaults by embedding undetectable steganographic data into ID card photographs via deep convolutional neural networks. This concealed data enables strong verification since any tampering with the photos disturbs the encoded patterns, allowing for immediate fraud detection. The technology fits seamlessly within existing ID card procedures, but the versatility to maintain visual quality via some added protection that is immune to real-world deterioration like printing and scanning comes with this solution. It is scalable and practical for implementation in various domains of government IDs and business access cards. It facilitates real-time verification and increases trust in identity

systems. Advanced improvements may include highly advanced steganographic techniques, blockchain integration, and cross-domain applications, that will make this technology a backbone for preventing identity fraud.

References

1. V. Hlaváč, P. Matoušek, 2019, "Security aspects of identity documents: A review", Pattern Recognition Letters, Vol: 119, pp. 98-110.
2. Neil F. Johnson, Sushil Jajodia, 1998, "Exploring steganography: Seeing the unseen", Computer, Vol: 31, No: 2, pp. 26- 34.
3. M. Tanaka, T. Ohtsuki, 2020, "Deep neural network-based robust steganography for visual communication systems", IEEE Transactions on Multimedia, Vol: 22, No: 6, pp. 1390- 1404.
4. M. Varga, D. Hercog, 2019, "Using machine-readable documents for biometric identity verification", IEEE Transactions on Biometrics, Behavior, and Identity Science, Vol: 1, No: 3, pp. 180-193.
5. A. Kaur, R. Kaur, 2021, "Noise-tolerant deep learning models for steganographic applications", Journal of Information Security and Applications, Vol: 59, pp. 102870.
6. X. Wang, Y. Xu, 2019, "Optimizing payload imperceptibility trade-offs in image steganography", IEEE Transactions on Information Forensics and Security, Vol: 14, No: 4, pp. 927-940.
7. A. Rezende, E. Mesquita, 2020, "Future directions in steganographic solutions for ID and biometric data protection", IEEE Access, Vol: 8, pp. 122933-122950.
8. V. Hlaváč, P. Matoušek, 2019, "Security aspects of identity documents", Pattern Recognition Letters, Vol: 119, pp.98- 110.
9. A. Younis, A. Jantan, 2018, "Printed document authentication systems", Journal of Imaging, Vol: 4, No: 3, pp. 66.

10. Shumeet Baluja, 2017, "Hiding images in plain sight: Deep steganography", *Advances in Neural Information Processing Systems*, Vol: 30, pp. 2066–2076.
11. R. Rasad, K. Sethi, 2020, "Robustness challenges in print-scan steganography", *Multimedia Tools and Applications*, Vol: 79, No: 11, pp. 7809-7833.
12. X. Wang, Y. Xu, 2019, "Optimizing payload imperceptibility trade-offs in image steganography", *IEEE Transactions on Information Forensics and Security*, Vol: 14, No: 4, pp. 927-940.
13. M. I. Khan, M. Usman, 2020, "StegoFace: End-to-end facial image steganography", *Image and Vision Computing*, Vol: 95, pp. 103876.
14. X. Zhang, S. Wang, 2019, "Image steganography techniques: Applications and limitations", *Multimedia Tools and Applications*, Vol: 78, No: 1, pp. 213-238.
15. M. Silva, J. Almeida, 2021, "Ethical implications of biometric steganography", *Applied Soft Computing*, Vol: 110, pp. 107666.
16. Casey Chu, Andrey Zhmoginov, Mark Sandler, 2017, "CycleGAN, a master of steganography", *arXiv*: 1712, pp. 02950.
17. Yichun Shi, Anil K. Jain, 2018, "DocFace: Matching ID document photos to selfies", in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst.(BTAS)*, pp. 1– 8.
18. Vitor Albiero, Nisha Srinivas, Esteban Villalobos, Jorge Perez-Facuse, Roberto Rosenthal, Domingo Mery, Karl Ricanek, Kevin W. Bowyer, 2020, "Identity document to selfie face matching across adolescence", in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, pp. 1–9.
19. Martin Stokkenes, Raghavendra Ramachandra, Christoph Busch, 2018, "Biometric transaction authentication using smartphones", in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, pp. 1–5.
20. Pramuditha Perera, Vishal M. Patel, 2019, "Face-based multiple user active authentication on mobile devices", *IEEE Trans. Inf. Forensics Security*, Vol: 14, No: 5, pp. 1240–1250.
21. Mohammed E. Fathy, Vishal M. Patel, Rama Chellappa, 2015, "Face-based active authentication on mobile devices", in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, pp. 1687–1691.
22. Vladimir Viktorovich Arlazarov, Konstantin Bulatovich Bulatov, Timofey Sergeevich Chernov, Vladimir Lvovich Arlazarov' 2018, "MIDV-500: A dataset for identity documents analysis and recognition on mobile devices in video stream", *arXiv*: 1807, pp. 05786.
23. Sebastian Gonzalez, Andres Valenzuela, Juan Tapia, 2021, "Hybrid two- stage architecture for tampering detection of chipless ID cards", *IEEE Trans. Biometrics, Behav., Identity Sci.*, Vol: 3, No: 1, pp. 89–100.
24. Xiangyu Zhu, Hao Liu, Zhen Lei, Hailin Shi, Fan Yang, Dong Yi, Guojun Qi, Stan Z. Li, 2019, "Large-scale bisample learning on ID versus spot face recognition", *Int. J. Comput. Vis.*, Vol: 127, nos. 6–7, pp. 684–700.
25. Raghavendra Mudgalgundurao, Patrick Schuch, Kiran Raja, Raghavendra Ramachandra, Naser Damer, 2022, "Pixel-wise supervision for presentation attack detection on identity document cards", *IET Biometrics*, Vol: 11, No: 5, pp. 383–395.