UDC 004.8

**Гацан Сергій, Олійник Володимир**

МОДЕЛЬ КОМП’ЮТЕРНОГО ЗОРУ ДЛЯ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧА З ЗАХИСТОМ ВІД СПУФІНГУ

**Hatsan Serhii, Oliinyk Volodymyr**

COMPUTER VISION BASED AUTHENTICATION MODEL WITH SPOOFING PROTECTION

У статті пропонується приклад реалізації підсистеми комп’ютерного зору для ідентифікації користувача з захистом від спуфінгу. Базовими моделями для даної підсистеми виступили сіамська нейронна мережа на базі SqueezeNet для розпізнавання обличь та мережа на основі ResNet для антиспуфінгу обличь.

**Ключові слова:** розпізнавання обличь, комп’ютерний зір, нейронна мережа, захист від спуфінгу, сіамська нейронна мережа, SqueezeNet, ResNet.

Рис.: 4. Табл: 1. Бібл.: 5.

The paper offers an example of the implementation of a computer vision subsystem for user identification with protection against spoofing. The basic models for this subsystem were the Siamese neural network based on SqueezeNet for face recognition and the network based on ResNet for face anti-spoofing.

**Key words:** face recognition, computer vision, neural network, antispoofing, siamese neural network, SqueezeNet, ResNet.

Fig.: 4. Tabl.: 1. Bibl.: 5.

**Relevance of the research topic.** Today we’re coming into such future, when everything needs protection: from our smartwatches, smartphones, laptops and up to our houses, banks, etc. And everyday we’re founding new and new types of vulnerabilities, that can put in danger our personal information.

Subsystems, that are based on a face recognition are the easiest way to protect such appliances and buildings. They do not require any specific action from user, such as specific placement, like in fingerprint scanners, or physical interaction, like with most of the security systems. User just need to stand in front of a camera to get recognized and immediately granted their access.

**Problem statement.** The problem of creating and developing such system lies inside the task of training such neural networks, because of a differentiality of many different spoofing methods. Building such systems requires large count of samples of different types of spoofing, so that in case of an testing in a wild neural network won’t recognize it as a real person.

For face recognition we have different approaches, such as comparing two faces or identifying it based on a different geometrical or any other features [1] [2]. Downside of a identifying it with a geometrical method lies inside similarity of different faces. If faces are really similar, or there is something, that is covering such face – such systems can falsely recognize or not recognize such test case. One of the most basic face identification solution relies on comparing two faces or, more specifically, two images of faces, and grants us similarity score between them, so that we can place a margin, which will split similar and different faces as it should be.

**Analysis of recent research and publications.** In recent years there is visible growing interest in different approaches related to face recognition system, because of a COVID-19 pandemic [1] [3].

**Identification of unexplored aspects of the general problem.** This article focuses on studying and analyzing the proposed approach for constructing two models: for face recognition based on a Siamese network architecture and face antispoofing based on a ResNet. The research is concentrated on studying working with Siamese networks and differentiating different types of spoofing.

**Target setting.** The task is to train base models to perform face recognition task and in the same time perform recognition of a possible spoofing methods, such as printed attacks or video attacks.

**The statement of basic materials.** Solving this task can be split into three different stages. First stage requires creation two new models from the ground up for two separate tasks. By the second stage there will be prepared datasets for training those neural networks and retrieving weights for both networks. The final stage includes processing obtained results by showing neural network different types of spoofing and real face of same and different person, which will serve as the final result.

**Model training.** To solve the target problem of face recognition, it was decided to choose the Siamese neural network architecture. The SqueezeNet neural network architecture, which is an improved version of the AlexNet architecture, is chosen as the convolutional network that will process the input images.

AlexNet is the name of the convolutional neural network (CNN) architecture developed by Alex Kryzhevsky in collaboration with Ilya Sutzkever and Jeffrey Ginton. Developed in 2012, this network was the best solution according to the speed and quality of the received object classifications on images up to 1000 classes.

In computer terms, SqueezeNet [4] is a deep neural network for image classification that was developed in 2016 by DeepScale researchers from the University of California, Berkeley, at Stanford University. When developing the SqueezeNet model, the goal of the developers was to create a smaller neural network tha the existing AlexNet, with fewer parameters, but at the same time achieve the same high accuracy. As a result, according to the developers themselves, it was possible to reduce the number of parameters by 50 times and achieve a model size of less than 0.5 megabytes.

The SqueezeNet network itself is only responsible for the classification of images, therefore, in order to compare them, we will pass both images to the neural network, and then compare their outputs using a neural network based on the Siamese network model, which is somewhat similar to ensemble model, described in [5].

A Siamese neural network takes input from another neural network that has processed two images and produced results, and outputs the similarity of these two images according to a feature. For the purpose of facial recognition with a neural network, we just need to teach it to find similarities between faces, so that it can then be checked how similar the face that is currently in the frame is to the face that should be.

After defining the architecture, the model was trained with the following hyperparameters: optimizer – SGD; loss – “contrastive loss” for Siamese network and “binary crossentropy” for ResNet; metrics – ‘binary accuracy’, ‘loss’, ‘f1 score’, ‘precision’ and ‘recall’; and number of epochs – 100-150.

**General model structure.** In our case, the proposed solution for face recognition looks like the one, that shown on figure 1.

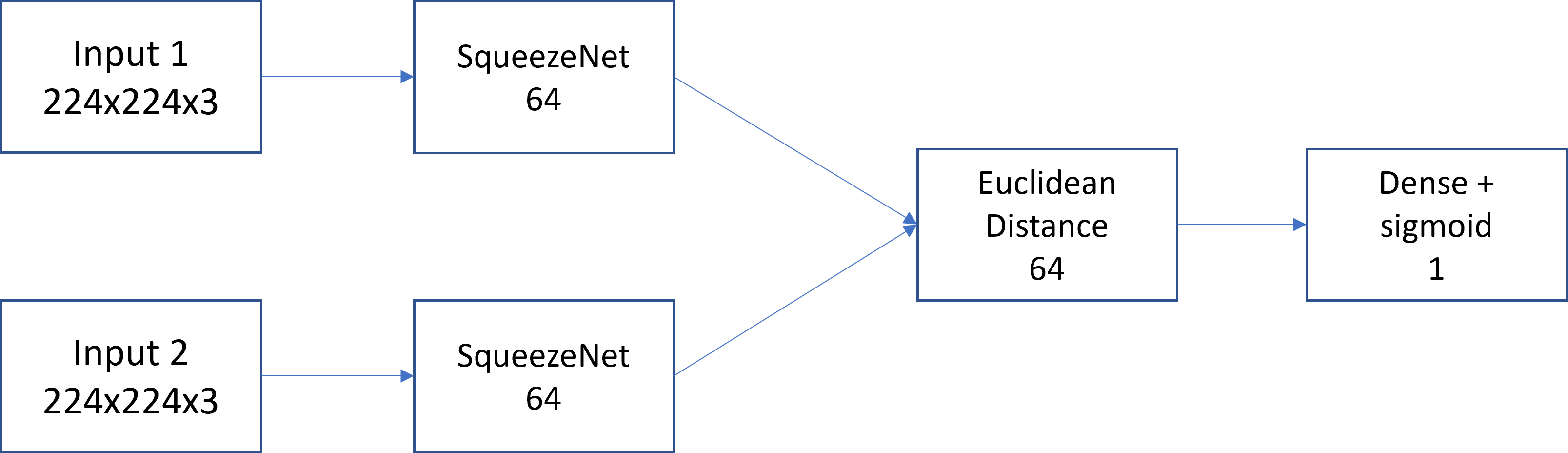


Fig. 1. General model structure for face recognition

It is working as following: we’re feeding two images to two similar neural networks that will extract feature vectors from those images. One image will be the one, that was saved inside the system and another one is the real one, that is currently captured from camera. Based on an extracted feature vectors we’re calculating Euclidean distance to get final difference value between images. The higher result means more similar images, and lower result means dissimilar images.

For a face antispoofing it is a regular ResNet architecture with slightly changed numbers of filters on some of the layers, that will minimize overall network size but won’t loss it’s accuracy. It’s architecture is shown on figure 2.

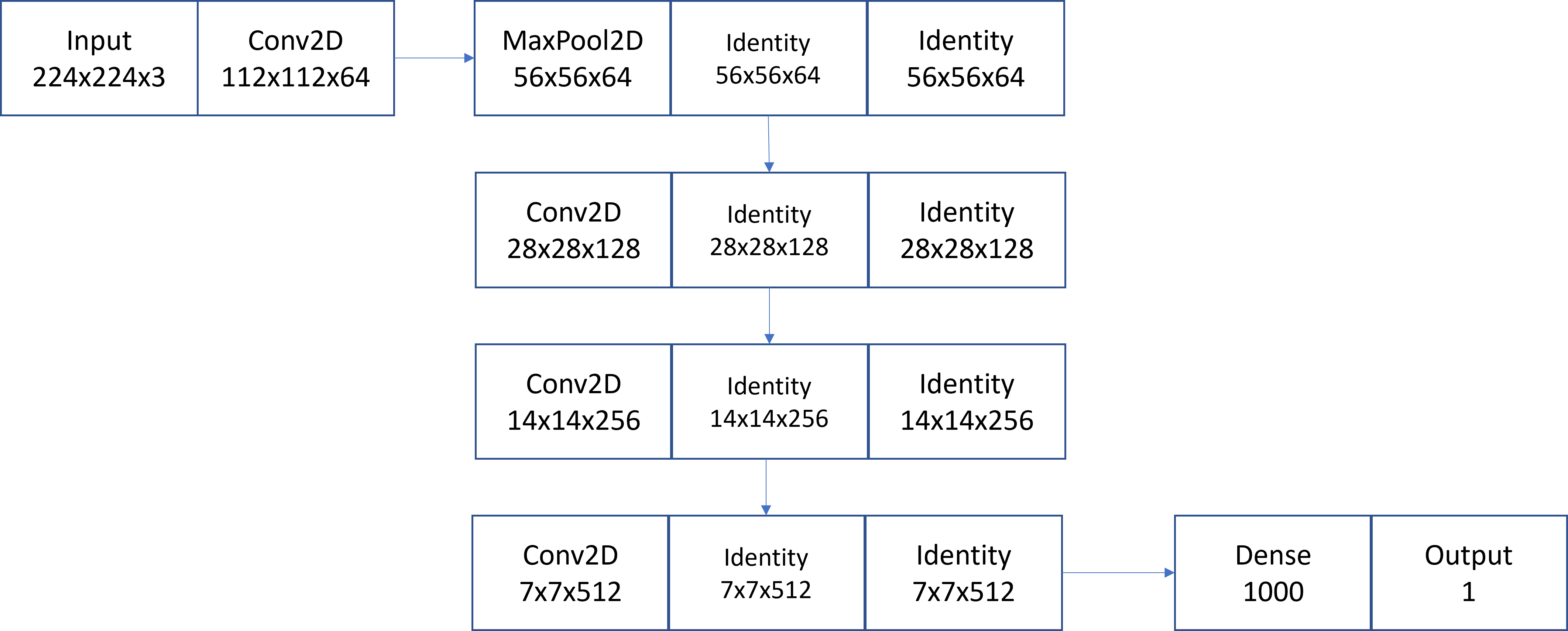


Fig. 2. General model structure for face antispoofing

General architecture of combining results of both networks will be done in sequential way. That means that images will be sequentially processed through antispoofing network and face recognition network. If antispoofing network is rejecting the input image – then there is no need to process it through face recognition, because we already have the required result – spoofed image. Whole processing order can be seen on a Fig. 3.

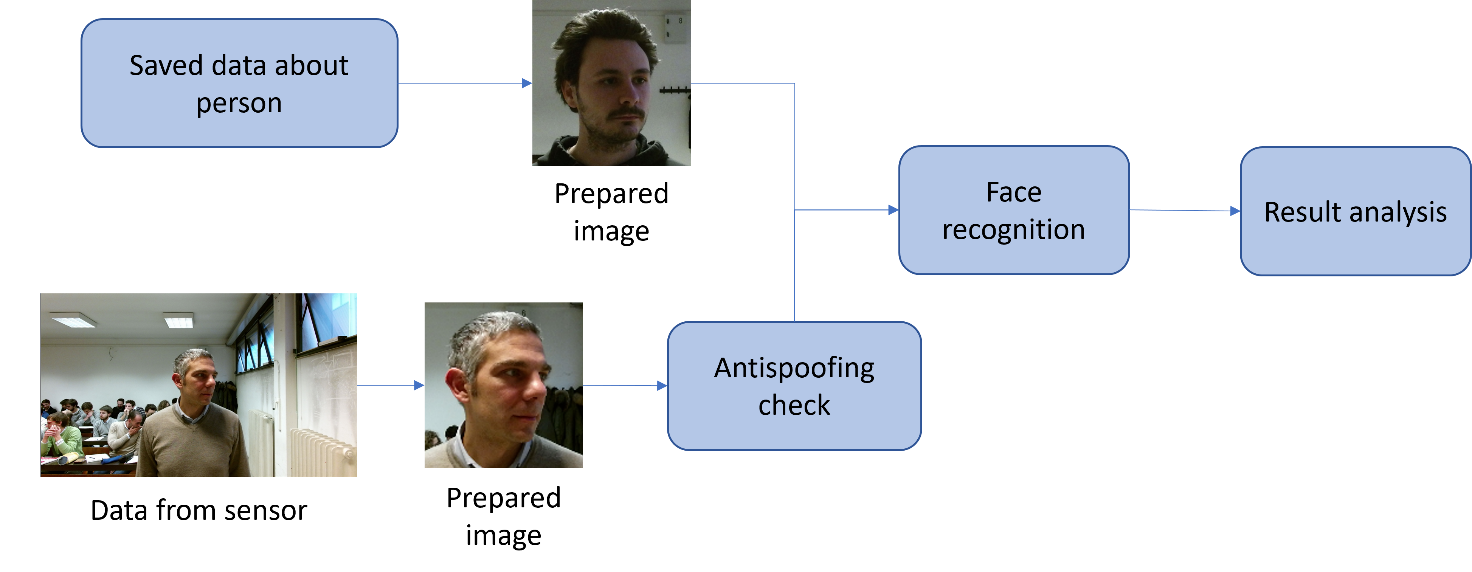


Fig. 3. Image processing order

**Results.** The obtained results of the face recognition and face antispoofing are presented in Table 1. This table displays the accuracy of each individual model, ranging from 97% to 98%, loss values and timings for each model an a summary time, elapsed for whole process – 34 ms.

Table 1

Model accuracy results

|  |  |  |
| --- | --- | --- |
| Criteria | Antispoofing | Face recognition |
| Accuracy | 97.3% | 97.5% |
| Loss | 9.06% | 4.75% |
| Average frame processing time | 16 ms | 17 ms |
| Video frame processing | 2 ms | 2 ms |
| Summary time | 34 ms | |

**Experiments.** A simple graphical user interface was created for testing purpose of the developed neural networks. Series of experiments were conducted to analyze model performance. For instance, two images of different people were shown on a phone screen under different angles. Also, printed image on a plastic card was shown to a neural network to perform spoofing protection.

Based from retrieved results, users’ face is shown without direct sunlight, that could interrupt face recognition task. Confidence of face recognition network was 98% and more, which means, that is is confident in users’ face. On the second example antispoofing network immediately kicks in and stops image from processing on a face recognition network. Due to low confidence of image reality, which was 13.8% and 7.1% on a third test, those images were considered as spoofed and because of that immediately rejected.

The test results are displayed in Fig. 4.



Fig. 4. Test results

**Conclusions.** The paper proposes an approach to solve the task of face recognition with spoofing protection. This approach involves the application of modern machine learning methods, such as ResNet architecture and Siamese networks. The utilization of a neural networks in this task allowed for receiving low processing time – from 2 ms on a video to 17 ms on a single photo with a high accuracy of 97.5%. This is high performance and accuracy, considering amount of data, that is being fed to a neural networks. For an antispoofing task, dataset volume and quality is crucial for achieving state-of-the-art result. Also, for further multi factor authentication purpose it is possible to utilize security keys or other identification systems, such as fingerprint scanner, to ensure security of system. In our case, using the proposed model is looking promising and we plan to experiment with it in future work.

# References

|  |  |
| --- | --- |
| 1. | Deng J. Masked face recognition challenge: The insightface track report, / Deng J., Guo J., An X., Zhu Z., Zafeiriou S., // Proceedings of the IEEE/CVF International Conference on Computer Vision – 2021. – P. 1437-1444. |
| 2. | Пантелеев А. С. Метод визуального мультитрекинга в реальном времени на основе корреляционных фильтров, / Пантелеев А. С., Олейник В. В., // Міжвідомчий науково-технічний збірник "Адаптивні системи Автоматичного Управління", К: Політехніка – 2018. – т. 1, № 32, – С. 97-106. |
| 3. | Oliinyk V. An efficient face mask detection model for real-time applications, / Oliinyk V., Ryzhiy A., // Adaptive systems of automatic control – 2022. – vol. 1, no. 40, – P. 54-64. |
| 4. | Iandola F. N. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size., / Iandola F. N., Han S., Moskewicz M. W., Ashraf K., Dally W. J., Keutzer K., // arXiv preprint – 2016. – no. arXiv:1602.07360. |
| 5. | Ponochovnyy P. Classification model of military aviation based on neural network ensemble, / Ponochovnyy P., Oliinyk V., // Proceedings of the international conference ICSFTI 2023, Kyiv, Ukraine, 2023. |

# ДОВІДКА ПРО АВТОРІВ

Олійник Володимир Валентинович – к.т.н., доцент, доцент кафедри інформаційних систем та технологій, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Oliinyk Volodymyr – Ph.D., associate professor, Department of Information Systems and Technologies, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: [oliinyk.volodymyr@gmail.com](mailto:oliinyk.volodymyr@gmail.com)

Гацан Сергій Юрійович – бакалавр, кафедра інформаційних систем та технологій, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Hatsan Serhii – BS, Department of Information Systems and Technologies, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: [serhiihatcan@gmail.com](mailto:serhiihatcan@gmail.com)

# РОЗШИРЕНА АНОТАЦІЯ

**С. Ю. Гацан, В. В. Олійник**

МОДЕЛЬ КОМП’ЮТЕРНОГО ЗОРУ ДЛЯ АВТЕНТИФІКАЦІЇ КОРИСТУВАЧА З ЗАХИСТОМ ВІД СПУФІНГУ

**Актуальність теми дослідження.** Сьогодні ми входимо в таке майбутнє, коли все потребує захисту: від наших розумних годинників, смартфонів, ноутбуків і до наших будинків, банків тощо. І щодня ми знаходимо нові й нові типи вразливостей, які можуть поставити під загрозу наші персональна інформація.

Підсистеми, які базуються на розпізнаванні осіб, є найпростішим способом захисту таких пристроїв і будівель. Вони не вимагають жодних особливих дій від користувача, наприклад певного розміщення, як у сканерах відбитків пальців, або фізичної взаємодії, як у більшості систем безпеки. Користувачеві достатньо стати перед камерою, щоб його впізнали та негайно надали доступ.

**Постановка проблеми.** Проблема створення та розвитку такої системи лежить всередині завдання навчання таких нейронних мереж через відмінність багатьох різних методів спуфінгу. Побудова таких систем вимагає великої кількості зразків різних типів спуфінгу, щоб у разі тестування в реальному середовищі нейронна мережа не помилилась розпізнати спуфінг як реальну людину.

Для розпізнавання облич у нас є різні підходи, такі як порівняння двох облич або їх ідентифікація на основі різних геометричних чи будь-яких інших ознак. Негативною стороною ідентифікації за геометричним методом є подібність різних ознак між обличчями. Якщо обличчя дійсно схожі, або є щось, що прикриває таке обличчя – такі системи можуть помилково розпізнати або не розпізнати такий тест. Одне з найпростіших рішень для ідентифікації обличчя ґрунтується на порівнянні двох облич або, точніше, двох зображень облич, і надає нам оцінку подібності між ними, щоб ми могли розмістити межу, яка розділить схожі та різні обличчя, як це має бути.

**Аналіз останніх досліджень та публікацій.** В останні роки через пандемію COVID-19 помітно зростає інтерес до різних підходів, пов’язаних із системою розпізнавання облич.

**Виділення недосліджених частин загальної проблеми.** Ця стаття присвячена вивченню та аналізу запропонованого підходу для побудови двох моделей: розпізнавання облич на основі архітектури сіамської мережі та антиспуфінгу облич на основі ResNet. Дослідження зосереджено на вивченні роботи з сіамськими мережами та диференціації різних типів спуфінгу.

**Постановка завдання.** Завдання полягає в тому, щоб навчити базові моделі виконувати задачу розпізнавання обличчя та в той же час виконувати розпізнавання можливих методів спуфінгу, таких як друковані атаки або відео атаки.

**Викладення основного матеріалу.** Вирішення цього завдання можна розбити на три різні етапи. Перший етап вимагає створення двох нових моделей з нуля для двох окремих завдань. На другому етапі будуть підготовлені набори даних для навчання цих нейронних мереж і отримання вагових коефіцієнтів для обох мереж. Завершальний етап включає обробку отриманих результатів шляхом показу нейронної мережі різних типів спуфінгу та реального обличчя тієї самої та іншої людини, що буде служити кінцевим результатом.

**Висновки.** У статті запропоновано підхід до вирішення задачі розпізнавання обличчя із захистом від спуфінгу. Цей підхід передбачає застосування сучасних методів машинного навчання, таких як архітектура ResNet і сіамські мережі. Використання нейронної мережі в цьому завданні дозволило отримати низький час обробки – від 2 мс на відео до 17 мс на одній фотографії з високою точністю 97,5%. Це висока швидкість і точність, враховуючи кількість даних, які передаються в нейронні мережі. Для завдання боротьби зі спуфінгом обсяг і якість набору даних є вирішальними для досягнення найсучаснішого результату. Для подальшого підвищення захисту також можна використати такі засоби багатофакторної автентифікації, як фізичні ключі безпеки або сканери відбитку пальців, щоб гарантувати безпеку системи. У нашому випадку використання запропонованої моделі виглядає перспективним і ми плануємо експериментувати з нею в подальшій роботі.

**Ключові слова:** розпізнавання обличь, комп’ютерний зір, нейронна мережа, захист від спуфінгу, сіамська нейронна мережа, SqueezeNet, ResNet.