

Automated Compressed Image Recovery Implementation for Video Surveillance Systems as Information Security Improvement

Iryna Moiseienko^a, Ivanna Dronyuk^b, Olena Gorina^b, Halina Mykhailyshyn^c and Kateryna Dubska^b

^a Lviv State University of Internal Affairs, Gorodotska 26, Lviv, 79001, Ukraine

^b Lviv Polytechnic National University, Bandery 12, Lviv, 79013, Ukraine

^c Vasyl Stefanyk Precarpathian National University, Shevchenko57, Ivano-Frankivsk, 76018, Ukraine

Abstract

The article is devoted to researching the architecture of generative adversarial networks capable of restoring the resolution of images compressed with great resolution losses. Particular attention is paid to the comparison of the resulting networks, the training of which was carried out using different loss functions, including perceptual loss functions. The created software is prepared for implementation into video surveillance systems.

In the field of information security for financial transactions, with the help of the proposed generative adversarial networks neural network software increases the level of details for video images of banking operations customers in video surveillance systems.

Keywords 1

Compressed image recovery, automated image recovery, generative adversarial networks, video surveillance systems, information security financial operations, video verification

1. Introduction

Constant innovation in deep learning and neural networks, much progress has been made in the last few years in restoring and improving image resolution. Image resolution enhancement or restoration can be defined as increasing the size of small or artificially compressed images while maintaining minimal quality degradation or restoring high-resolution images from saturated details derived from low-resolution images.

Enhancing image resolution with neural networks is a topic with a wide range of applications for specific areas, including: video surveillance security; compressed image / video enhancement; speed up the transfer of images or videos over the network. In the field of surveillance for information security of financial transactions and customer verification, with the help of this technology you can get a clearer image that contains more important detailed information. Moreover, in the field of visualization of the images obtained in the observations, the increased resolution of the image can help to obtain more information, which is hidden in the images with low resolution. This increases the level of visibility and verification of customers, which determines the information security of financial transactions.

The brief survey in information security in banking is presented in [1]. In works [2,3] it was emphasized the importance of implementation innovative video and software systems for improving financial operation security. Images analysis as a main software element for video surveillance systems is a basic of technological solution in markets and banks [4, 5]. Neural networks for recovery

IntelITSIS'2021: 2nd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 24–26, 2021, Khmelnytskyi, Ukraine

EMAIL: ko@lvduvs.edu.ua (I. Moysenyko); ivanna.m.droniuk@lpnu.ua (I. Dronyuk); lplilon@ukr.net (O. Gorina); galmuh60@gmail.com (H. Mykhailyshyn); kateryna.dubska.mknm.2019@lpnu.ua (K. Dubska)

ORCID: 0000-0002-3140-461X (I. Moysenyko); 0000-0003-1667-2584 (I. Dronyuk); 0000-0003-1589-9242 (H. Mykhailyshyn)



© 2021 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

missing data in IoT is presented in [6]. Some aspects for video traffic quality increasing is discuss in [7,8]. The most important for achievement super-resolution is using neural network and deep learning implementation [9]. But for the specific of image resolution improving task not all neural network is suitable. The best of all are convolutional neural network [10]. In future video surveillance systems have to be more smarter [11,12]. The article [13] introduces a technique for face super-resolution based on a deep convolutional neural network (Deep CNN). For realization, these tasks the video surveillance systems and its software have to recognize small objects detections [14, 15]. For supporting person identification in surveillance video, it is used a deep learning approach [16].

Although the above-mentioned models achieve high accuracy and speed in image recovery, there is one problem that is not solved. The problem is how to restore the finer details of the texture from a low-resolution image so that the image is not distorted. For restoring small elements the Generative Adversarial Network (GAN) was proposed in 2014 [17]. GAN is a method of machine learning without a teacher, capable of generating new data, with the same statistical parameters as the training set. This result is achieved as a result of a game with zero sum of two networks - a generator and a discriminator, each of which adjusts the weight of the other. With this architecture, it is possible to restore the original high-resolution image from a low-resolution compressed image. Our work is devoted to GAN investigation for purpose of recognition small details in video images for surveillance systems.

2. Enhancement images resolution with GAN

Most compressed image resolution recovery algorithms rely on the entropy encoding of frequently repeated image patterns. The two most commonly used criteria are image similarity and natural image statistics. Currently, several methods of restoring the resolution of compressed or damaged images are popular, such as the use of arithmetic coding, autoaggressive models, domain transformation statistics, and others. Most basic techniques are described in [18].

For the correct operation of such a generative adversarial network, you need a correct assessment of the results of training the generator network, which is carried out using the loss function. In this paper, we will consider two types of loss function for the created network - the loss function that preserves pixel-by-pixel correspondence between objects, the root mean square reconstruction error and the loss function that correctly estimates the perceptual differences between generated and original image. This loss function was described in [19].

In this work it was creating and training models of generative adversarial neural networks with the same topology, but trained by variation of the loss function, which preserve pixel-by-pixel or perceptual similarity of objects on the restored and original images to increase image resolution and use in video surveillance systems. Developing new learning methods can help improve the network training process. The object of this study is generative adversarial networks - a class of artificial intelligence algorithms used in teaching without a teacher, implemented by a system of two artificial neural networks that compete with each other in a zero-sum game. The aim of the study is to improve the resolution of images for use in video surveillance systems for improving security of financial operations.

In the research the topology of the neural GAN was developed, the software for the neural network training process was created, the training was performed using various loss functions to assess the accuracy of the model during training, and conclusions were drawn from the model training results.

3. Software for restoring resolutions of compressed images

3.1. Components of creating software

As you can see in Fig. 1, the compressed images restoring resolutions software for video surveillance systems consists of three main components:

“General” - a software component that creates a common interface and stores auxiliary functions that are responsible for downloading data, processing it, drawing graphs and the resulting illustrations;

“Modeling” is a software component that is responsible for the design of generator and discriminator networks. The basis of this component has been implemented using the third-party Keras library, which will be described below.

“Training” is a software component that is responsible for the process of creating, training and saving a model. This process includes the stage of loading and processing data, the design of the model and, accordingly, the process of training the network and saving the resulting weights. The detail training process is shown in Fig.2.

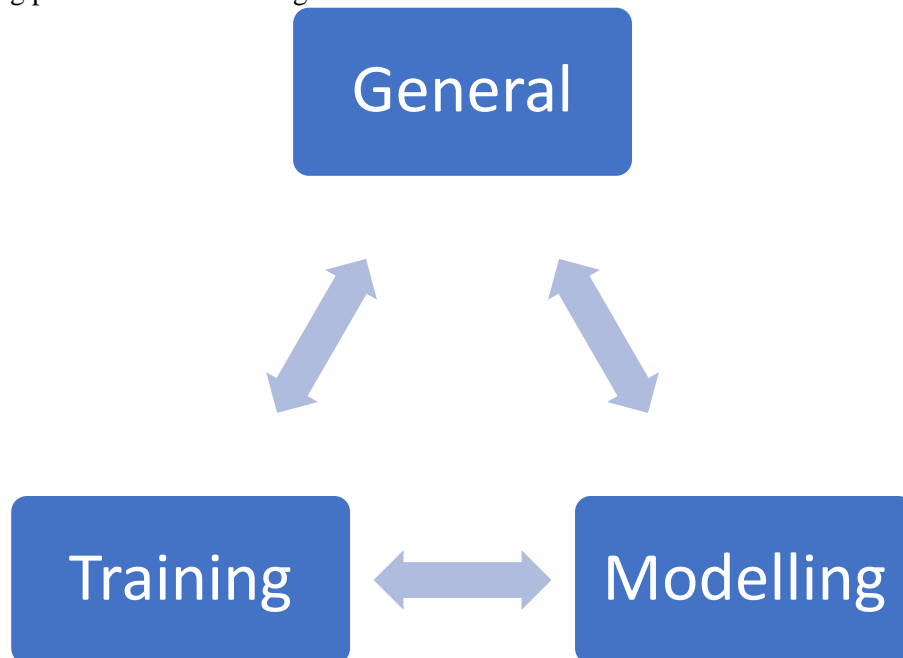


Figure 1: Structure of compressed images restoring resolutions software for video surveillance systems

3.2. Packages using for creating software

The following libraries were used in the developed software: Keras (version 2.4.3), NumPy (version 1.18.5), OpenCV-Python (version 4.4.0.42), Matplotlib (version 3.3.2), Library tqdm (version 4.49.0). We will briefly describe their main functional properties that were used to develop compressed images restoring resolutions software for video surveillance systems.

Keras [20] is an open-source library that we used to create an interface to a neural network. Keras is an interface to the TensorFlow library. The main functionality that has been used is the convenience of creating fast experiments with deep neural networks, it is characterized by modularity and advancement. We have used Keras tools to build neural networks, add new layers, activation functions, optimizers and many other tools to make working with images and text data easier. In addition to standard neural networks, Keras supports convolutional and recursive neural networks. Other common levels of utilities are supported, such as dropouts, batch normalization, and pulling. One of the important and necessary features of the library for working with video is the ability to use distributed deep learning models on clusters of graphics processing units (GPUs) and tensor blocks (TPUs).

NumPy [21] is a library for the Python programming language, used by us to extend the functionality of working with multidimensional arrays and matrices. NumPy is open-source software. Using NumPy in Python has created a high level of functionality of the developed program, when processing operations with arrays or matrices.

OpenCV (Open Source Computer Vision Library) [22] is a library of programming functions focused mainly on real-time computer vision. The library is cross-platform and free to use under the Apache 2 open source license. We have used OpenCV for GPU acceleration of real-time operations.

Matplotlib [23] is a library for the Python programming language, used by us to create and display graphics. This library provides an object-oriented API for embedding graphics into applications. Its functionality is quite extensive, and a big advantage of Matplotlib is the support of the NumPy library, which provides a wider range of uses. To avoid conflicts with installed libraries, we used the Anaconda software package, which allows us to create a compact development environment and avoid conflicts in installed packages.

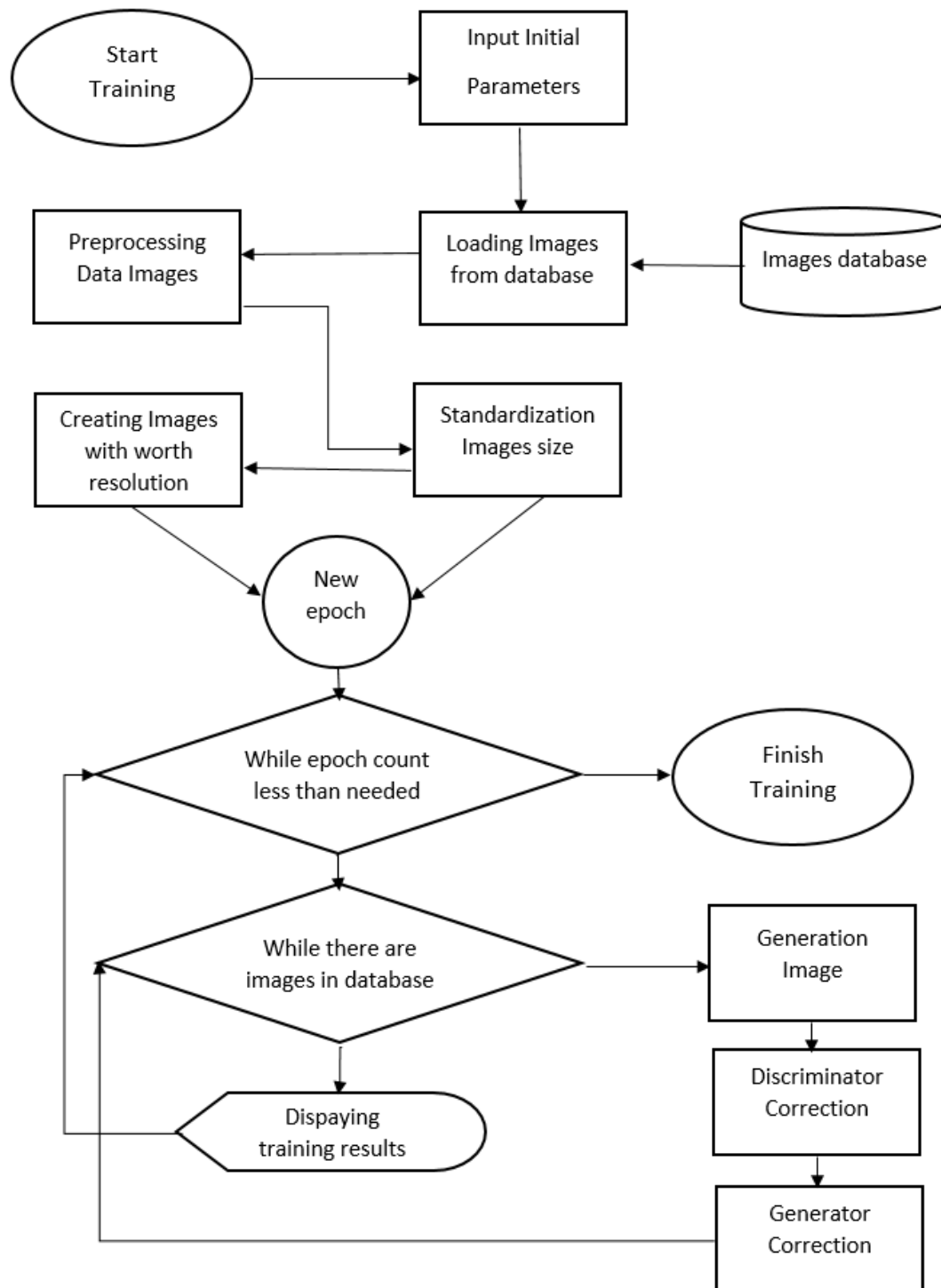


Figure 2: Detail training block-scheme of compressed images restoring resolutions software for video surveillance systems

4. Result and discussion

The network was trained on the dataset [24] on 7000 images. At the preprocessing stage, all images were reduced to the same size and an additional dataset was created, which was based on the original - it consisted of images of the dataset, the resolution of which was artificially reduced.

The generator, in the process of training, aimed to expand the resolution of images from the created dataset, and the discriminator compared the generated image with the original and adjusted the operation of the generator. The training was conducted for 500 epochs.

Four networks were trained:

- in the first network, pixel-by-pixel loss is selected as a loss function for the generator
- in the second network as a loss function for the generator selected perceptual loss, which is based on the network VGG16
- in the third network as a loss function for the generator selected perceptual loss, which is based on the network VGG19
- in the last network as a loss function for the generator selected perceptual loss, which is based on the network ResNet

GAN training is very complex and requires a high level of computing power. Therefore, to speed up the training process, it was conducted on a video card.

The first problem was the installation of the necessary versions of libraries that would allow for network training on GPUs. This is due to the fact that for different graphics processors there are different software and hardware architectures of parallel computing.

Another problem was the training itself. Due to the complex structure of the resulting networks and the number of filters, it became necessary to impose restrictions on the size of the input image. This is because the GPU has limited memory of its own.

5. Conclusion

The network topology with variations of loss functions with the help of which GAN networks were trained was proposed. An important part of network training is to assess the similarity between the restored image and the original. The paper proposes to use two variations of the loss function:

1. loss function, which estimates the pixel-by-pixel similarity of the restored image
2. loss function, which estimates the perceptual similarity of objects in the image. Perceptual similarity is assessed using previously trained networks, namely: VGG16, VGG19 and ResNet.

The results of the training showed that the network with perceptual loss function, which used the VGG16 network, performed the best. Images restored with this grid showed detail restoration without significant blurring of object boundaries and significant granulation effect. The network with the perceptual loss function, which used the ResNet network, showed the worst results. This is due to the fact that the ResNet network is very sensitive to the input parameters, which changed greatly after passing all the hidden layers of the proposed model. This technology is very promising and requires further research to increase the level of information security of financial transactions through video surveillance. Due to the complexity of the training and the limitations of the equipment available, the training was limited in size and in a relatively small number of epochs (500). Therefore, one of the main ways to improve the results is to increase the number of training epochs and increase the standard image size while reducing the image compression parameter.

6. References

- [1] A. Kam, T. Plummer, G. Falco, D. Whyte, Survey of cyber security framework across industries, 2018 Winter Simulation Innovation Workshop, SIW, 2018

- [2] I. P. Moiseenko, N. O. Ryvak, Monitoring parameters of the level of economic security of Ukraine, *Actual Problems of Economics* 137(11) (2012) 95-102
- [3] I. P. Moiseenko, O. A. Martyniuk, Application of intelligent technologies in economic security system of a bank. *Actual Problems of Economics* 138(12) (2012) 234-238
- [4] W. Ebihara, Y. Mae, F. Abe, S. Hirooka, Development and future deployment of cloud-based video surveillance systems, *Hitachi Review* 63(8) (2014) 514-518
- [5] A. Skadins, M. Ivanovs, R. Rava, K. Nesenbergs, Edge pre-processing of traffic surveillance video for bandwidth and privacy optimization in smart cities, in: *Proceedings of the Biennial Baltic Electronics Conference, BEC, 2020*. doi:10.1109/BEC49624.2020.9276799
- [6] I. Izonin, R. Tkachenko, N. Kryvinska, K. Zub, O., Mishchuk, T. Lisovych, Recovery of incomplete IoT sensed data using high-performance extended-input neural-like structure. *Procedia Computer Science* 160 (2019) 521-526. doi:10.1016/j.procs.2019.11.054
- [7] I. Dronyuk, O. Fedevych, Traffic flows ateb-prediction method with fluctuation modeling using dirac functions, 718 (2017) 3-13 doi:10.1007/978-3-319-59767-6_1
- [8] M. Pasyeka, V. Sheketa, N. Pasiaka, S. Chupakhina, I. Dronyuk, System analysis of caching requests on network computing nodes, in: *Proceedings of the 2019 3rd International Conference on Advanced Information and Communications Technologies, AICT 2019*, pp. 216-222. doi:10.1109/AIACT.2019.8847909
- [9] P. Wang, P. Wang, E. Fan, Violence detection and face recognition based on deep learning. *Pattern Recognition Letters* (2021) 142 20-24. doi:10.1016/j.patrec.2020.11.018
- [10] S. Limkar, S. Hunashimarad, P. Chinchmalatpure, A., Baj, R. Patil, Potential of robust face recognition from real-time cctv video stream for biometric attendance using convolutional neural network, *Intelligent Data Engineering and Analytics* (2021) 11-20 doi:10.1007/978-981-15-5679-1_2
- [11] A. B. Deshmukh, N. Usha Rani, Optimization-driven kernel and deep convolutional neural network for multi-view face video super resolution, *International Journal of Digital Crime and Forensics* 12(3) (2020) 77-95. doi:10.4018/IJDCF.2020070106
- [12] O. Pomorova, T. Hovorushchenko, Artificial neural network for software quality evaluation based on the metric analysis, in: *Proceedings of IEEE East-West Design and Test Symposium, EWDTs 2013*, doi:10.1109/EWDTs.2013.6673193
- [13] C. Cai, Y. Wu, S. Li, The application of the dilated convolution based on small object detection, in: *Proceedings of the Chinese Control Conference, CCC, 2020*, pp. 7079-7083. doi:10.23919/CCC50068.2020.9189374
- [14] S. Hiriyannaiah, B. S. Akanksh, A. S. Koushik, G. M. Siddesh, K. G. Srinivasa, Deep learning for multimedia data in IoT *Multimedia Big Data Computing for IoT Applications* (2020) 101-129 doi:10.1007/978-981-13-8759-3_4
- [15] P. Yu, Image resolution enhancement technology based on deep neural network, in: *Proceedings of the Cyber Security Intelligence and Analytics*, pp. 687-693 doi:10.1007/978-3-030-43306-2_97
- [16] H. Wei, M. Laszewski, N. Kehtarnavaz, Deep learning-based person detection and classification for far field video surveillance, in: *Proceedings of the 2018 IEEE Dallas Circuits and Systems Conference, DCAS 2018*, doi:10.1109/DCAS.2018.8620111
- [17] J. I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al, Generative adversarial networks. *Communications of the ACM*, 63(11) (2014) 139-144. doi:10.1145/3422622
- [18] H. Chen, X. He, C. Ren, et al., CISRDCNN: Super-resolution of compressed images using deep convolutional neural networks, *Neurocomputing* 285 (2018) 204-2019.
- [19] C. Ledig, L. Theis, F. Huszar, et al., Photo-realistic single image super-resolution using a generative adversarial network, in: *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 105-114. doi:10.1109/CVPR.2017.19
- [20] Keras. URL: <https://keras.io/>
- [21] NumPy. URL: <https://numpy.org/>
- [22] OpenCV. URL: <https://opencv.org/>
- [23] Matplotlib: Visualization with Python. URL: <https://matplotlib.org/>
- [24] The USC-SIPI Image Database. URL: <http://sipi.usc.edu/database/>