Detecting Sensitive Data with GANs and Fully Convolutional Networks

Marcin Korytkowski[1,2], Jakub Nowak[1], and Rafał Scherer[1,2](✉)

[1] Czestochowa University of Technology, al. Armii Krajowej 36, Czestochowa, Poland
{marcin.korytkowski,jakub.nowak}@pcz.pl
[2] Intigo Ltd., Haryana, India
rafal.scherer@pcz.pl
http://intigo.ai/

Abstract. The article presents a method of document anonymization using generative adversarial neural networks. Unlike other anonymization methods, in the presented work, the anonymization concerns sensitive data in the form of images placed in text documents. Specifically, it is based on the CycleGAN idea and uses the U-Net model as a generator. To train the model we built a dataset with text documents with embedded real-life images, and medical images. The method is characterized by a very high efficiency, which enables the detection of 99.8% of areas where the sensitive image is located.

1 Introduction

Nowadays, many entities and private persons want to protect their data against leakage. It can be information about both health and company secrets, e.g. research works. The subject of the processing of sensitive data is also extremely important in the context of EU regulations, e.g. Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and the criminal and financial liability of persons creating and processing such collections of information. It is also worth noting that the theft of sensitive data may be used to assess the health condition of politicians or other decision-makers. In this paper, we present a system for detecting documents containing sensitive data, also in the cases where they are intentionally hidden there.

For obvious reasons, the classification of data into one of two classes: with and without sensitive data must be automatic. In a situation where nowadays even small entities process gigabytes of information daily, a human is not able

The work was supported by The National Centre for Research and Development (NCBR), the project no POIR.01.01.01-00-1431/19.
The project financed under the program of the Polish Minister of Science and Higher Education under the name “Regional Initiative of Excellence” in the years 2019–2023 project number 020/RID/2018/19 the amount of financing PLN 12,000,000.

N. T. Nguyen et al. (Eds.): ACIIDS 2023, LNAI 13995, pp. 273–283, 2023.
https://doi.org/10.1007/978-981-99-5834-4_22

to manually verify the content of processed files. In this article, we propose a solution that fully automates this process based on machine learning techniques. The task facing the system is to detect and remove sensitive data from the documents being processed. It is based on the CycleGAN idea and uses the U-Net fully convolutional neural model as a generator. To train the model we built a dataset with text documents with embedded real-life images, and medical images.

The rest of the paper is organized as follows. In Sect. 2 we describe shortly models used. Our method for is described in Sect. 3. Section 4 presents the results of experiments on the dataset of sensitive documents we created and Sect. 6 concludes the paper.

2 Related Work

Currently, anonymization with the use of neural networks is used primarily to detect specific phrases in the text [11,13]. Unlike the text, we will not analyze the context of the text, but the actual information contained in a visual form of images. The detection of sensitive data such as the human face has been very well developed, among others, thanks to the huge amount of data available on social network channels [1,2]. In the studies cited, very good results were achieved with the use of convolutional neural nets [16]. Simirarly, well structured data are well processed by various neural networks, even anomalies in data can be easily detected [5,6]. Our work is aimed at detecting the places of occurrence of sensitive data, such as: results of magnetic resonance imaging, X-rays, etc. However, in the case of this type of data, we usually face the problem of small amounts of data available for training. One of the ways to improve the operation of classification algorithms is by generating synthetic data on the basis of the available set [12]. In our research, we tackle the presented problem differently. We want a generative adversarial neural network (GAN) [7] to be able to distinguish between sensitive data itself. To put it simply, the GAN is supposed to generate images without sensitive data. The general GAN diagram is presented in Fig. 1 and of the proposed model in Fig. 2.

The presented solution is based on the CycleGAN network model [17]. In its original application, the network was designed to convert graphic images. Among other things, they trained the model to convert horse images to zebra images, and city landscapes at night to city landscapes by day. The great advantage of CycleGAN is that this model can be trained without paired examples, i.e. it does not require sample photos before and after conversion to train the model. For example, it is not necessary to provide the same image of a horse turned into a zebra. The model architecture consists of two generator models: one generator (GeneratorA) for generating images for the first class (ClassA) and a second generator (GeneratorB) for generating images for the second class (ClassB):

GeneratorA → ClassA (documents requiring anonymization),

GeneratorB → ClassB (documents that do not require anonymization).

Generator models perform image translation, which means that the image conversion process depends on the input image, particularly an image from another

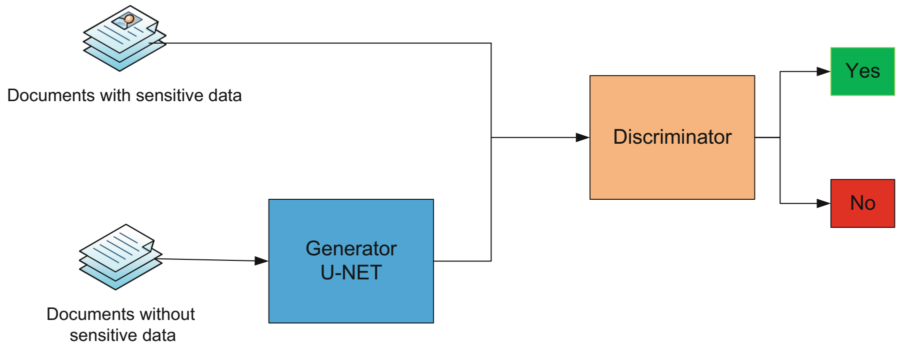

Documents with sensitive data
Documents without sensitive data
Generator U-NET
Discriminator
Yes
No

Fig. 1. General diagram of the GAN network.

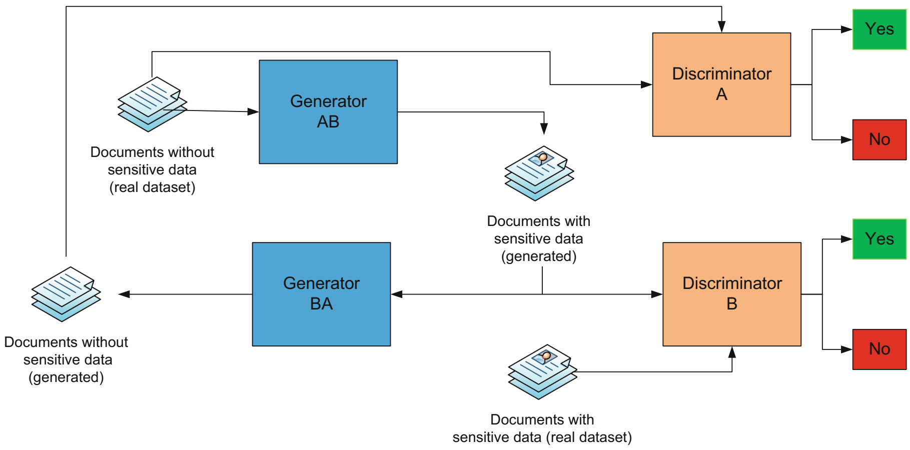

Documents without sensitive data (real dataset)
Generator AB
Discriminator A
Yes
No
Documents with sensitive data (generated)
Generator BA
Discriminator B
Yes
No
Documents without sensitive data (generated)
Documents with sensitive data (real dataset)

Fig. 2. General diagram of the proposed model: GAN network and the network used to build the training sets.

domain. Generator-A takes an image from ClassB as the input and ClassB takes an image from ClassA as the input:

ClassB → GeneratorA → ClassA,
ClassA → GeneratorB → ClassB.

Each generator has its own dedicated discriminator model. The first discriminator model (DiscriminatorA) takes the true images from ClassA and the generated images from GeneratorA and predicts whether they are true or false. The second discriminator model (DiscriminatorB) takes the true images from ClassB and the generated images from GeneratorB and predicts whether they are true or false. The discriminator and generator models are trained in an adversarial zero-sum process, much like normal GAN models. Our solution uses also the U-Net model [14] as a generator, which was initially used, inter alia, to detect neoplasms in medical images. The architecture stems from the so-called "fully convolutional network" [10] (Fig. 3).

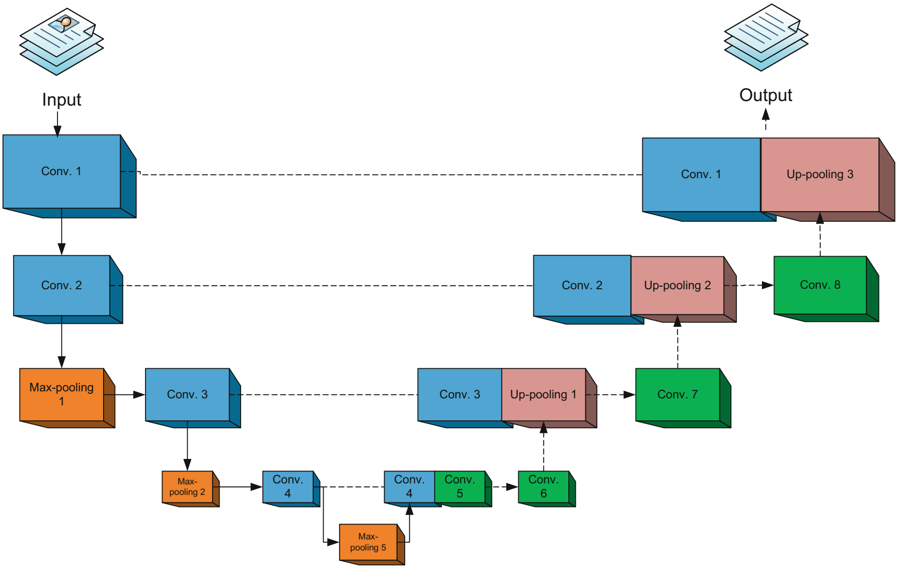


Fig. 3. General U-Net model.

Generators learn to cheat discriminators better, and discriminators learn better to detect false images. Together, the models find the equilibrium during the training process. In addition, generator models are regulated not only to create new images in the target class, but instead, to create translated versions of the input images from the source class. This is achieved by using the generated images as input to the appropriate generator model and comparing the output image with the original images. The transmission of the image by both generators is called a cycle. Together, each pair of generator models is trained to better recreate the original source image, which is referred to as cycle consistency.

3 Training the Anonymisation Model

In the presented model, the role of the generator is performed by a modified U-Net network inspired by the work [14]. In our structure, compared to the original concept, the activation functions have been changed from ReLU to SELU [9]. This change contributed to reducing the model's learning time from approximately 21.5 h to 19. However, in the conducted research, no significant impact on improving the learning outcome was observed. The U-Net model has also been downsized compared to the original by removing convolutional layers within the same dimension. In the original work, the creators used two convolutional layers after each max-pooling process, while in our solution, we limited it to one such operation. The reason for this approach was to create a network that processes input signals as quickly as possible, which directly translates to minimizing computational requirements.

The training sequence was built on the basis of text documents with the embedded medical data. Now we will describe the input data. The training dataset for the GAN requires two classes of objects — that require anonymisation and do not require anonymization. The diagram of the network is shown in Fig. 2. The non-anonymised part contained only a grayscale image obtained from a text document (e.g. WORD, PDF) along with random images from the ImageNET dataset [4] inserted in random places. The part requiring anonymisation was additionally provided with data in the form of X-ray images and computed tomography. The image was saved as a JPEG file with the size of 1024 × 1024 pixels. The chosen image size was dictated by the compromise between the model accuracy and processing speed.

Table 1. Experimental results for various input image sizes.

Input image size	Batch processing time For 100 images [seconds]	Image recognition result
256 × 256	8.3	79.8%
400 × 400	12.1	97.3%
1024 × 1024	21.6	99.8%
2048 × 2048	57.2	99.8%

Analyzing the results collected in Table 1, it can be concluded that satisfactory image recognition performance is already achieved at the image size of 400 × 400 px. Based on the obtained results, it is evident that to achieve the best outcomes, a network capable of analyzing images at a resolution of 1024 × 1024 px should be employed. Increasing the size twice to 2048 × 2048 px does not improve the already high classification accuracy significantly but increases the computation time of the system.

The method of generating training and testing images is presented in Fig. 4. In this case, the input signal is the text appropriately converted into bitmap maps derived from textual documents. During the generation of samples, consecutive textual documents are provided, which are converted into jpg images along with medical images from a medical database or the ImageNet dataset. When training the GAN network, three types of documents are distinguished: documents containing data in the form of images combined with medical photos (recognized as sensitive images), documents combined with images from the ImageNet database, or documents without any images, saved as files that do not contain sensitive data.

The portion of documents that did not required anonymisation did not contain images from the ImageNET database. For any document, the size of the image could not be greater than 80% of the width and height, and less than 20%. The inserted images were additionally scaled by a random factor.

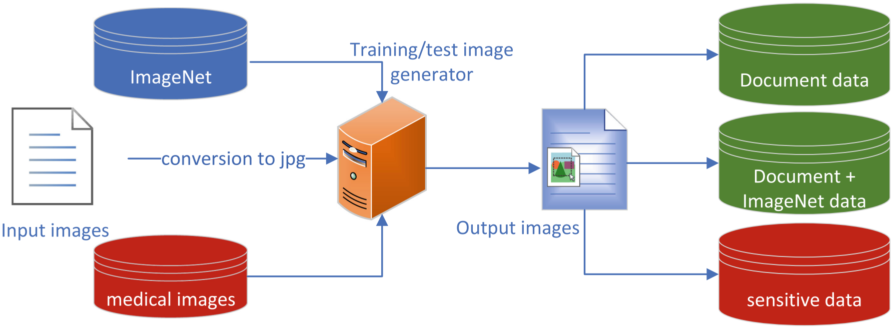


Fig. 4. Generating training and testing images.

The output of the GAN is also a grayscale image with the size of 1024 × 1024 pixels. In the GAN model, the discriminator as the evaluating part of the generator is also important.

In the conducted research, synthetic data requiring anonymisation were inserted in random places. The network in the learning process retrieved document scans without inserted sensitive images as ClassA. As ClassB, the network received documents with sensitive (e.g. medical) images.

On the basis of the conducted research, the network, apart from detecting the document with the newly generated sensitive image, was able to remove the sensitive image from text documents without losing such elements as a stamp or a barcode. Examples of such objects are presented in Fig. 5.

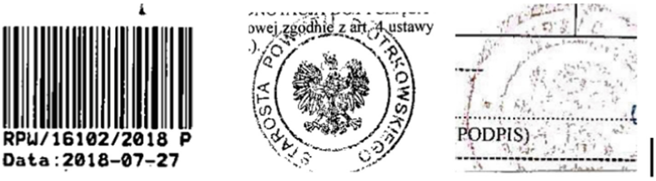


Fig. 5. Examples of objects placed in documents.

To prepare input data with sensitive medical objects we used our own X-ray images and images taken from the datasets described in [3,8] and [15]. Eventually, we created a dataset which composition is presented in Table 2.

Table 2. Dataset composition.

Type	Number of files	Size in MB (2048×2048)	Size in MB (1024×1024)	Size in MB (400×400)	Size in MB (256×256)
Number of files without images	20100	38592	4769	1923	588
Number of files with images from ImageNet	15000	29214	3632	1611	454
Number of files with sensitive medical images	15000	29112	3219	1022	410

4 Results

Through the operation of the GAN structure and the proprietary solution of inserting medical photos into the content of various text files, the first class of documents (containing sensitive data) was defined, which consisted of 30,000 graphic files. The second class of documents was created by artificially generating graphic files on the basis of the public ImageNET database (a total of 30,000 randomly selected files) of files.

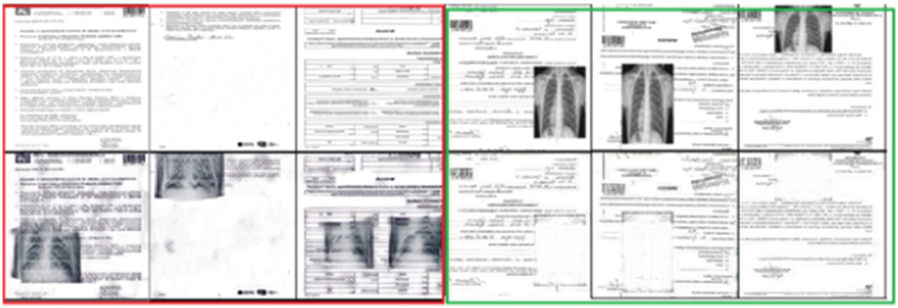
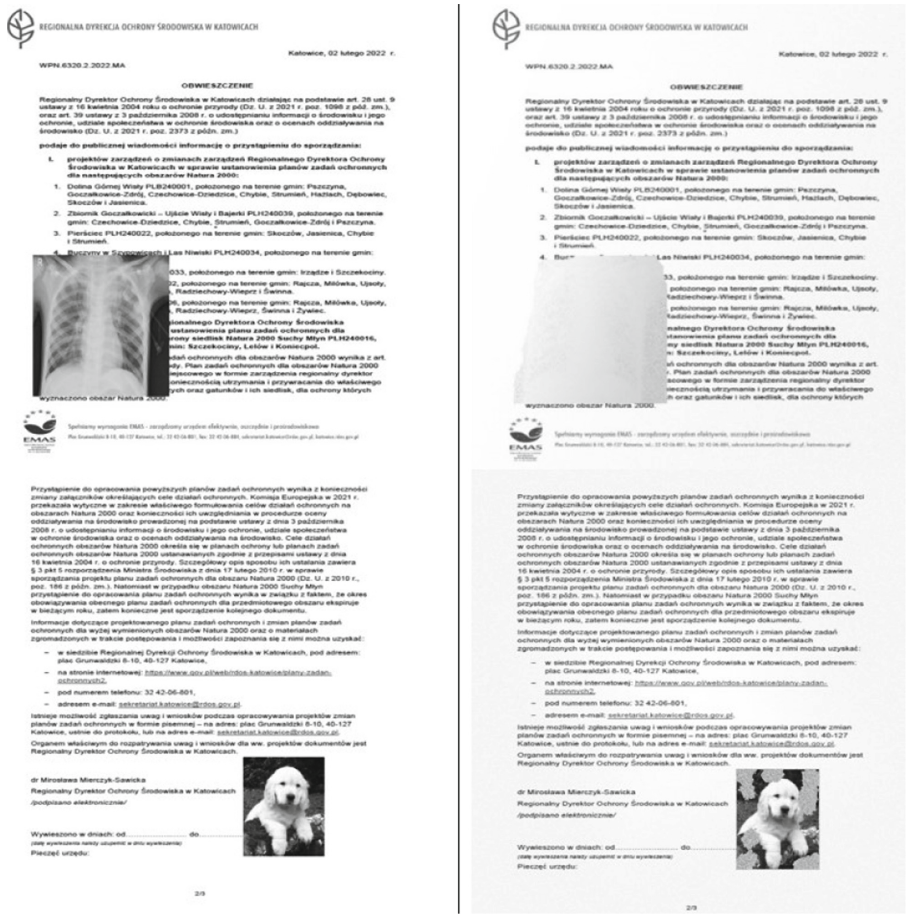


Fig. 6. Example images with sensitive objects inserted for training and removed by the model.

The first discriminator model (Fig. 2) (DiscriminatorA) takes the true images from ClassA and the generated images from GeneratorA and predicts whether they are true or false. The second discriminator model (DiscriminatorB) takes the true images from ClassB and the generated images from GeneratorB and predicts whether they are true or false. The discriminator and generator models are trained in an avdersarial zero-sum process, much like the standard GAN models. Figure 6 in the red box shows the effect of generating data by the network during the learning process. An extremely interesting feature of the trained GAN network, as described above, is that it can be used to remove sensitive medical data (a kind of graphical anonymisation). The effect of this structure used for this purpose is shown in Fig. 6 in a green frame. The input data is on the first line, and the net result (output) is on the bottom line. The percentage of correctly selected sensitive data was 99.8%. The percentage of incorrectly selected insensitive data was 29.3%. That was calculated based on the number of similar pixels.



REGIONALNA DYREKCJA OCHRONY ŚRODOWISKA W KATOWICACH

Katowice, 02 lutego 2022 r.

WPN.6320.2.2022.MA

OBWIESZCZENIE

Regionalny Dyrektor Ochrony Środowiska w Katowicach działając na podstawie art. 28 ust. 9 ustawy z 16 kwietnia 2004 roku o ochronie przyrody (Dz. U. z 2021 r. poz. 1098 z póź. zm.), oraz art. 39 ustawy z 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko (Dz. U. z 2021 r. poz. 2373 z późn. zm.)

podaje do publicznej wiadomości informację o przystąpieniu do sporządzania:

I. **projektów zarządzeń o zmianach zarządzeń Regionalnego Dyrektora Ochrony Środowiska w Katowicach w sprawie ustanowienia planów zadań ochronnych dla następujących obszarów Natura 2000:**

1. Dolina Górnej Wisły PLB240001, położonego na terenie gmin: Pszczyna, Goczałkowice-Zdrój, Czechowice-Dziedzice, Chybie, Strumień, Hażlach, Dębowiec, Skoczów i Jasienica.
2. Zbiornik Goczałkowicki – Ujście Wisły i Bajerki PLH240039, położonego na terenie gmin: Czechowice-Dziedzice, Chybie, Strumień, Goczałkowice-Zdrój i Pszczyna.
3. Pierściec PLH240022, położonego na terenie gmin: Skoczów, Jasienica, Chybie i Strumień.
4. Buczyny w Szypowicach i Las Niwiski PLH240034, położonego na terenie gmin:

033, położonego na terenie gmin: Irządze i Szczekociny.

2, położonego na terenie gmin: Rajcza, Milówka, Ujsoły, , Radziechowy-Wieprz i Świnna.

6, położonego na terenie gmin: Rajcza, Milówka, Ujsoły, , Radziechowy-Wieprz, Świnna i Żywiec.

ionalnego Dyrektora Ochrony Środowiska ustanowienia planu zadań ochronnych dla rony siedlisk Natura 2000 Suchy Młyn PLH240016, min: Szczekociny, Lelów i Koniecpol.

dań ochronnych dla obszarów Natura 2000 wynika z art. dy. Plan zadań ochronnych dla obszarów Natura 2000 iejscowego w formie zarządzenia regionalny dyrektor oniecznością utrzymania i przywracania do właściwego ych oraz gatunków i ich siedlisk, dla ochrony których wyznaczono obszar Natura 2000.

Spełniamy wymagania EMAS - zarządzamy urzędem efektywnie, oszczędnie i prośrodowiskowo

Plac Grunwaldzki 8-10, 40-127 Katowice, tel.: 32 42-06-801, fax: 32 42-06-884, sekretariat.katowice@rdos.gov.pl, katowice.rdos.gov.pl

Przystąpienie do opracowania powyższych planów zadań ochronnych wynika z konieczności zmiany załączników określających cele działań ochronnych. Komisja Europejska w 2021 r. przekazała wytyczne w zakresie właściwego formułowania celów działań ochronnych na obszarach Natura 2000 oraz konieczności ich uwzględniania w procedurze oceny oddziaływania na środowisko prowadzonej na podstawie ustawy z dnia 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko. Cele działań ochronnych obszarów Natura 2000 określa się w planach ochrony lub planach zadań ochronnych obszarów Natura 2000 ustanawianych zgodnie z przepisami ustawy z dnia 16 kwietnia 2004 r. o ochronie przyrody. Szczegółowy opis sposobu ich ustalania zawiera § 3 pkt 5 rozporządzenia Ministra Środowiska z dnia 17 lutego 2010 r. w sprawie sporządzania projektu planu zadań ochronnych dla obszaru Natura 2000 (Dz. U. z 2010 r., poz. 186 z późn. zm.). Natomiast w przypadku obszaru Natura 2000 Suchy Młyn przystąpienie do opracowania planu zadań ochronnych wynika w związku z faktem, że okres obowiązywania obecnego planu zadań ochronnych dla przedmiotowego obszaru ekspiruje w bieżącym roku, zatem konieczne jest sporządzenie kolejnego dokumentu.

Informacje dotyczące projektowanego planu zadań ochronnych i zmian planów zadań ochronnych dla wyżej wymienionych obszarów Natura 2000 oraz o materiałach zgromadzonych w trakcie postępowania i możliwości zapoznania się z nimi można uzyskać:

- w siedzibie Regionalnej Dyrekcji Ochrony Środowiska w Katowicach, pod adresem: plac Grunwaldzki 8-10, 40-127 Katowice,
- na stronie internetowej: https://www.gov.pl/web/rdos-katowice/plany-zadan-ochronnych2,
- pod numerem telefonu: 32 42-06-801,
- adresem e-mail: sekretariat.katowice@rdos.gov.pl.

Istnieje możliwość zgłaszania uwag i wniosków podczas opracowywania projektów zmian planów zadań ochronnych w formie pisemnej – na adres: plac Grunwaldzki 8-10, 40-127 Katowice, ustnie do protokołu, lub na adres e-mail: sekretariat.katowice@rdos.gov.pl.

Organem właściwym do rozpatrywania uwag i wniosków dla ww. projektów dokumentów jest Regionalny Dyrektor Ochrony Środowiska w Katowicach.

dr Mirosława Mierczyk-Sawicka
Regionalny Dyrektor Ochrony Środowiska w Katowicach
/podpisano elektronicznie/

Wywieszono w dniach: od........................ do........................
(daty wywieszenia należy uzupełnić w dniu wywieszenia)
Pieczęć urzędu:

2/3

REGIONALNA DYREKCJA OCHRONY ŚRODOWISKA W KATOWICACH

Katowice, 02 lutego 2022 r.

WPN.6320.2.2022.MA

OBWIESZCZENIE

Regionalny Dyrektor Ochrony Środowiska w Katowicach działając na podstawie art. 28 ust. 9 ustawy z 16 kwietnia 2004 roku o ochronie przyrody (Dz. U. z 2021 r. poz. 1098 z póź. zm.), oraz art. 39 ustawy z 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko (Dz. U. z 2021 r. poz. 2373 z późn. zm.)

podaje do publicznej wiadomości informację o przystąpieniu do sporządzania:

I. **projektów zarządzeń o zmianach zarządzeń Regionalnego Dyrektora Ochrony Środowiska w Katowicach w sprawie ustanowienia planów zadań ochronnych dla następujących obszarów Natura 2000:**

1. Dolina Górnej Wisły PLB240001, położonego na terenie gmin: Pszczyna, Goczałkowice-Zdrój, Czechowice-Dziedzice, Chybie, Strumień, Hażlach, Dębowiec, Skoczów i Jasienica.
2. Zbiornik Goczałkowicki – Ujście Wisły i Bajerki PLH240039, położonego na terenie gmin: Czechowice-Dziedzice, Chybie, Strumień, Goczałkowice-Zdrój i Pszczyna.
3. Pierściec PLH240022, położonego na terenie gmin: Skoczów, Jasienica, Chybie i Strumień.
4. Bu Las Niwiski PLH240034, położonego na terenie gmin:

33, położonego na terenie gmin: Irządze i Szczekociny.

położonego na terenie gmin: Rajcza, Milówka, Ujsoły, Radziechowy-Wieprz i Świnna.

położonego na terenie gmin: Rajcza, Milówka, Ujsoły, Radziechowy-Wieprz, Świnna i Żywiec.

nalnego Dyrektora Ochrony Środowiska stanowienia planu zadań ochronnych dla ny siedlisk Natura 2000 Suchy Młyn PLH240016, n: Szczekociny, Lelów i Koniecpol.

ań ochronnych dla obszarów Natura 2000 wynika z art. . Plan zadań ochronnych dla obszarów Natura 2000 acowego w formie zarządzenia regionalny dyrektor ecznością utrzymania i przywracania do właściwego h oraz gatunków i ich siedlisk, dla ochrony których wyznaczono obszar Natura 2000.

Spełniamy wymagania EMAS - zarządzamy urzędem efektywnie, oszczędnie i prośrodowiskowo

Plac Grunwaldzki 8-10, 40-127 Katowice, tel.: 32 42-06-801, fax: 32 42-06-884, sekretariat.katowice@rdos.gov.pl, katowice.rdos.gov.pl

Przystąpienie do opracowania powyższych planów zadań ochronnych wynika z konieczności zmiany załączników określających cele działań ochronnych. Komisja Europejska w 2021 r. przekazała wytyczne w zakresie właściwego formułowania celów działań ochronnych na obszarach Natura 2000 oraz konieczności ich uwzględniania w procedurze oceny oddziaływania na środowisko prowadzonej na podstawie ustawy z dnia 3 października 2008 r. o udostępnianiu informacji o środowisku i jego ochronie, udziale społeczeństwa w ochronie środowiska oraz o ocenach oddziaływania na środowisko. Cele działań ochronnych obszarów Natura 2000 określa się w planach ochrony lub planach zadań ochronnych obszarów Natura 2000 ustanawianych zgodnie z przepisami ustawy z dnia 16 kwietnia 2004 r. o ochronie przyrody. Szczegółowy opis sposobu ich ustalania zawiera § 3 pkt 5 rozporządzenia Ministra Środowiska z dnia 17 lutego 2010 r. w sprawie sporządzania projektu planu zadań ochronnych dla obszaru Natura 2000 (Dz. U. z 2010 r., poz. 186 z późn. zm.). Natomiast w przypadku obszaru Natura 2000 Suchy Młyn przystąpienie do opracowania planu zadań ochronnych wynika w związku z faktem, że okres obowiązywania obecnego planu zadań ochronnych dla przedmiotowego obszaru ekspiruje w bieżącym roku, zatem konieczne jest sporządzenie kolejnego dokumentu.

Informacje dotyczące projektowanego planu zadań ochronnych i zmian planów zadań ochronnych dla wyżej wymienionych obszarów Natura 2000 oraz o materiałach zgromadzonych w trakcie postępowania i możliwości zapoznania się z nimi można uzyskać:

- w siedzibie Regionalnej Dyrekcji Ochrony Środowiska w Katowicach, pod adresem: plac Grunwaldzki 8-10, 40-127 Katowice,
- na stronie internetowej: https://www.gov.pl/web/rdos-katowice/plany-zadan-ochronnych2,
- pod numerem telefonu: 32 42-06-801,
- adresem e-mail: sekretariat.katowice@rdos.gov.pl.

Istnieje możliwość zgłaszania uwag i wniosków podczas opracowywania projektów zmian planów zadań ochronnych w formie pisemnej – na adres: plac Grunwaldzki 8-10, 40-127 Katowice, ustnie do protokołu, lub na adres e-mail: sekretariat.katowice@rdos.gov.pl.

Organem właściwym do rozpatrywania uwag i wniosków dla ww. projektów dokumentów jest Regionalny Dyrektor Ochrony Środowiska w Katowicach.

dr Mirosława Mierczyk-Sawicka
Regionalny Dyrektor Ochrony Środowiska w Katowicach
/podpisano elektronicznie/

Wywieszono w dniach: od........................ do........................
(daty wywieszenia należy uzupełnić w dniu wywieszenia)
Pieczęć urzędu:

2/3

Fig. 7. Example input and output documents with sensitive and non-sensitive objects. We can observe the effect of anonimisation by the proposed system.

5 Automatic Sensitive Data Detection System

The previously described neural network model has practical applications in a document anonymization system. The primary task here is to protect client resources by verifying the presence of sensitive information within file resources in the context of GDPR or sensitive information for the company in terms of protecting confidential information.

It should be noted that the system only utilizes a portion of the trained CycleGAN model responsible for generating anonymized images. The remaining part of the network was implemented solely for the purpose of training the system. One practical challenge in the system's operation is the transmission of documents containing sensitive data. Therefore, several practical solutions have been devised to mitigate this risk. The concepts of such systems are presented in Fig. 8.

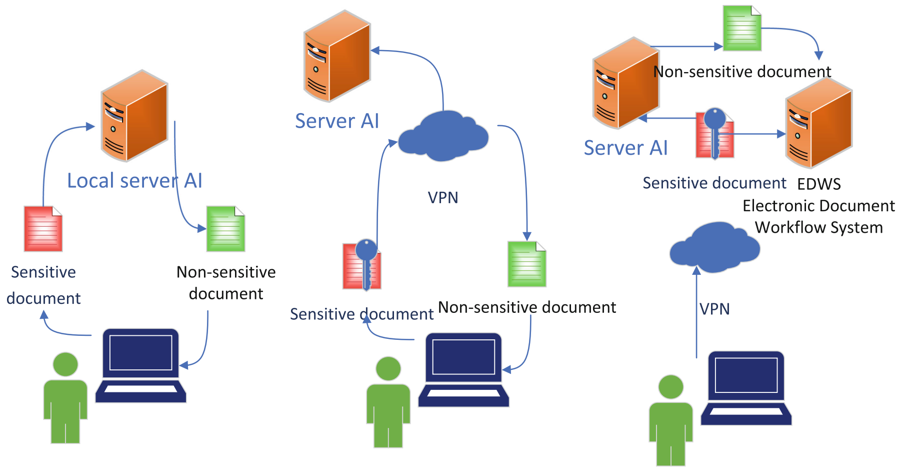


Fig. 8. Application of generative adversarial networks in anonymization systems.

The first solution pertains to a scenario where both the user's computer and the AI server with the GAN network operate within the same local network. In such cases, it is possible to directly transmit the documents to the AI server. In the second scenario, when the client cannot provide the necessary infrastructure for document processing, a remote AI server will act as the client's service. Communication with the user will take place through a dedicated VPN tunnel, and every sensitive document will be automatically encrypted. In this solution, there is no need for re-encrypting the anonymized documents. The last possible solution is when the client utilizes a document circulation system. In such cases, communication will occur between the system and the AI server. Depending on the infrastructure involved, an additional VPN connection between the servers providing the mentioned services or a local connection may be required.

6 Conclusions

We proposed a method to anonymise documents using generative adversarial neural networks and fully convolutional networks. The anonymisation concerns sensitive data in the form of images placed in text documents. It is based on the CycleGAN idea and uses the U-Net model as a generator. To train the model we built our own dataset with real-life MS Word and PDF text documents with embedded real-life images, and medical images. The method is characterized by a very high efficiency, which enables the detection of 99.8% of areas where the sensitive image is located. The method is able to remove the detected sensitive objects what is shown in Fig. 7.

References

1. Alhabash, S., Ma, M.: A tale of four platforms: motivations and uses of Facebook, Twitter, Instagram, and Snapchat among college students? Soc. Med.+ Soc. **3**(1), 2056305117691544 (2017)
2. Beaver, D., Kumar, S., Li, H.C., Sobel, J., Vajgel, P.: Finding a needle in haystack: Facebook's photo storage. In: 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10) (2010)
3. Cohen, J.P., Morrison, P., Dao, L.: Covid-19 image data collection. arXiv 2003.11597 (2020). https://github.com/ieee8023/covid-chestxray-dataset
4. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
5. Gabryel, M., Lada, D., Filutowicz, Z., Patora-Wysocka, Z., Kisiel-Dorohinicki, M., Chen, G.Y.: Detecting anomalies in advertising web traffic with the use of the variational autoencoder. J. Artif. Intell. Soft Comput. Res. **12**(4), 255–256 (2022). https://doi.org/10.2478/jaiscr-2022-0017
6. Gabryel, M., Scherer, M.M., Sulkowski, L., Damaševičius, R.: Decision making support system for managing advertisers by ad fraud detection. J. Artif. Intell. Soft Comput. Res. **11**(4), 331–339 (2021). https://doi.org/10.2478/jaiscr-2021-0020
7. Goodfellow, I., et al.: Generative adversarial nets. In: Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N., Weinberger, K. (eds.) Advances in Neural Information Processing Systems, vol. 27. Curran Associates, Inc. (2014) https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf
8. Kermany, D.S., et al.: Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell **172**(5), 1122–1131 (2018)
9. Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S.: Self-normalizing neural networks. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
10. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
11. Mosallanezhad, A., Beigi, G., Liu, H.: Deep reinforcement learning-based text anonymization against private-attribute inference. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 2360–2369. Association for Computational Linguistics, Hong Kong, China (2019). https://doi.org/10.18653/v1/D19-1240https://aclanthology.org/D19-1240
12. Röglin, J., Ziegeler, K., Kube, J., König, F., Hermann, K.G., Ortmann, S.: Improving classification results on a small medical dataset using a GAN; an outlook for dealing with rare disease datasets. Front. Comput. Sci., 102 (2022)
13. Romanov, A., Kurtukova, A., Shelupanov, A., Fedotova, A., Goncharov, V.: Authorship identification of a Russian-language text using support vector machine and deep neural networks. Future Internet **13**(1), 3 (2020)
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data **5**(1), 1–9 (2018)

16. Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K.: Convolutional neural networks: an overview and application in radiology. Insights Imaging **9**(4), 611–629 (2018)
17. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)