Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 200 (2022) 418–427

Procedia Computer Science

www.elsevier.com/locate/procedia

3rd International Conference on Industry 4.0 and Smart Manufacturing

Generative adversarial networks with bio-inspired primary visual cortex for Industry 4.0

Vladyslav Branytskyi [a], Mariia Golovianko [a]*, Diana Malyk [a], Vagan Terziyan [b]

[a]*Depatment of Artificial Intelligence, Kharkiv National University of Radio Electronics, 61166, Kharkiv, Ukraine*
[b]*Faculty of Information Technology, University of Jyväskylä, 40014, Jyväskylä, Finland*

Abstract

Biologicalization (biological transformation) is an emerging trend in Industry 4.0 affecting digitization of manufacturing and related processes. It brings up the next generation of manufacturing technology and systems that extensively use biological and bio-inspired principles, materials, functions, structures and resources. This research is a contribution to the further convergence of computer and human vision for more robust and accurate automated object recognition and image generation. We present VOneGANs, a novel class of generative adversarial networks (GANs) with the qualitatively updated discriminative component. The new model incorporates a biologically constrained digital primary visual cortex V1. This earliest cortical visual area performs the first stage of human`s visual processing and is believed to be a reason of its robustness and accuracy. Experiments with the updated architectures confirm the improved stability of GANs training and the higher quality of the automatically generated visual content. The promising results allow considering VOneGANs as providers of high-quality training content and as enablers of future simulation-based decision-making and decision-support tools for condition-monitoring, supervisory control, diagnostics, predictive maintenance, and cybersecurity in Industry 4.0.

Keywords: Biologicalization; Industry 4.0; GAN; VOneGAN; primary visual cortex V1; hybrid CNN

1. Introduction

Biologicalization is an emerging trend in Industry 4.0 affecting digitization of manufacturing and connected processes. It brings up the next generation of manufacturing technology and systems that extensively use biological

* Corresponding author. Tel.: +38-057-702-13-37; fax: +38-057-702-13-37.
E-mail address: mariia.golovianko@nure.ua

1877-0509
10.1016/j.procs.2022.01.240

and bio-inspired principles, materials, functions, structures and resources [1]. One of the frontier approaches within this paradigm is the neuro-inspired Cognitive Manufacturing [2] aiming at copying the perception-reasoning-learning processes of the human brain into advanced industrial architectures. The recent progress done by the scientific community working on deep learning and convolutional neural networks (CNNs) is believed to bring innovations to the industrial environments and optimize manufacturing processes [3]. In particular, it can essentially facilitate industrial computer vision used in a wide range of Cognitive Manufacturing tasks: from remote sensing and industrial inspection to document processing.

However, artificial environments deployed in sandboxes of scientific labs are different from the real industrial environments, as well as the requirements to the neural systems trained for scientific purposes and the ones operating on the factory floor. The security issues of the latter are of the greatest concern, since all the vulnerabilities to adversarial attacks, acceptable in theory, pose a serious threat to the success of machine learning in practice [4, 5]. Thus, reliability of the underlying machine learning models in terms of their adversarial robustness becomes a cornerstone for computer vision systems in Industry 4.0 [6].

Two common approaches increasing machine learning robustness against adversarial attacks are (i) redesigning the learning process, such as, enhancing it with adversarial training on adversarial examples [7], or defensive distillation [8], and (ii) improving network architectures adding more robust features and components to them [9, 10].

In this research, we demonstrate both approaches following the biologicalization trend and contributing to the further convergence of computer and human visual perception aiming at more robust and accurate industrial machine learning.

First, we refer to a new recently appeared class of hybrid CNN image recognition architectures VOneNets containing a biologically constrained neural network that simulates more accurately primary visual cortex (V1) [11]. It was already reported that simulating a human primary visual cortex at the front of CNNs improves learning model robustness to adversarial image perturbations. The reason for this is the fact that the internal representation of images in CNNs and in the humans` visual system is still fundamentally different [12]. Humans' classification abilities for image recognition hugely outperform computer models from the perspective of robustness. Computer vision systems fail to recognize objects correctly within intentionally corrupted adversarial images while humans have no trouble with these. Including new digital components inspired by human vision system is a step towards narrowing the gap between the two systems.

Inspired by the success of VOneNets in discriminative tasks, we have developed the VOneGANs – a new class of generative adversarial networks (GANs) [13] containing digital version of primary visual cortex V1 simulating human schema of neural processing of generated samples. In this paper, we report the first promising results of VOneGANs training. They support our hypothesis that digital primary visual cortex V1 not only improves robustness of the discriminative models but also it can be used for essential improvement of generative neural models. This fact opens good industrial perspectives for already existing VOneNets and for the new VOneGANs architectures suggested in this paper. The first potential contribution of VOneGANs to Industry 4.0 is generation of good quality synthetic content. Finding or collecting an appropriate dataset is a highly demanding task, especially considering specific settings of each new industrial environment. Datasets augmentation with synthetic (artificially produced) samples has become an acknowledged solution to various problems related to intelligent models learning [14, 15]. Another advantage is that improved discriminator capability will also improve the quality of classification models used in Industry 4.0. Moreover, we foresee the VOneGANs to be stronger providers of not only clean highly realistic samples, but also new challenging adversarial training content or training infrastructure for other classifiers. In this case, VOneGAN plays a role of a generative adversarial trainer protecting other classifiers against adversarial attacks and improving their prediction accuracy [16].

The rest of the paper is organized as follows: in Section 2, we report on related work in the field of biologically-inspired architectures for computer vision, particularly, neural networks containing simulation of the human primary visual cortex; Section 3 describes the methodology of our research; in Section 4, we introduce a new class of GANs VOneGANs containing a primary visual cortex V1 neural network and demonstrate our experiments with different VOneGANs; and we, finally, conclude in Section 5.

The source code, datasets and additional information are available online at: https://github.com/Adversarial-Intelligence-Group/vone-gan.

2. Methodology

Our *research hypothesis* is that GANs, enhanced by simulation of primary visual cortex image processing, can be successfully used in Industry 4.0 and are more effective than existing pure GAN architectures in performing generation and classification tasks.

To confirm this hypothesis, we design and implement a new architecture of GANs with a discriminative component imitating image processing in human primary visual cortex more precisely than previously existing convolutional architectures. Empirical evidence in favor of our hypothesis is collected by running the newly developed architectures on three different datasets: MNIST [17], CIFAR-10 [18], and our own dataset Conveyor-V3 (see Fig.1.) containing 4400 images of the industrial inter-roll conveyor system used for transportation and sorting of plastic cassettes with heterogeneous loads at the experimental facility within the NATO SPS project "Cyber-Defense for Intelligent Systems" (http://recode.bg/natog5511).

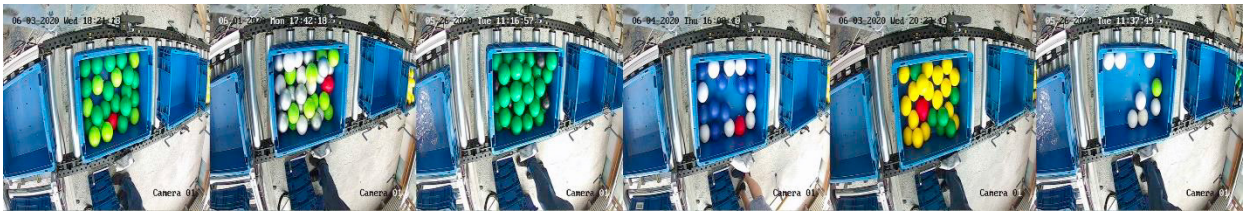

Fig. 1. Image samples from dataset Conveyor-V3 containing 4400 images of the industrial inter-roll conveyor system.

The performance of the newly suggested models is evaluated based on traditional machine learning metrics demonstrating accuracy of the model on both clean and adversarial inputs, including:

- *Classification accuracy* – describes the discriminator`s ability to predict the correct image class. Calculated as the number of correct predictions in the class divided by the total number of predictions.
- *Binary accuracy* – describes the discriminator`s ability to distinguish between original and generated examples. Calculated as the percentage of correct predictions (generated and original images) between all predictions.

The analysis of the models` effectiveness in industrial context is made based on the measured performance of the model in application to two tasks: realistic clean images generation and challenging adversarial images` generation.

3. Related work

Modelling cognitive processes by mimicking those of human brain or other natural systems is not new for the computer science community: artificial intelligence systems have often been developed guided by neural science discoveries. Current state-of-the-art object recognition is mainly based on CNNs, which are believed to be built from the knowledge of the functional organization of the natural ventral visual pathway, a hierarchical brain structure consisting of the areas V1-V2-V4-IT [19]. However, CNNs, accurate on clean input data and even in some cases superior to humans in performance [20], can be relatively easily fooled by imperceptibly small, crafted perturbations and struggle to recognize objects in corrupted images, while those are easily recognized by humans [21, 22, 23]. This shows that the organization of human visual cortex is not well understood by now and the digital models are still significantly different from the natural ones.

There have been numerous attempts to bring architectures of NN more in line with biology [24, 25, 26], including those focused on extending and modifying CNNs` traditional architecture and training techniques with elements believed to resemble those from human or animal brain [27, 28, 29, 30]. Much effort has been made to digitize more precisely the earliest cortical visual area which performs the first stage of visual processing – primary visual cortex (V1) [31, 32, 33]. In [11] it is argued that accuracy of the learnt model under adversarial attacks is strongly

correlated to the V1 explained variance, i.e. the ability of the model to explain the responses of single V1 neurons, which act as local features and edge detectors.

Recent studies also confirm that there is a direct positive impact of the adoption of the V1 mechanisms inherent to natural vision on the improved adversarial robustness of CNN [34].

Among new architectures, VOneNets [11] show the most promising results in terms of robustness – hybrid CNNs with a more neuro-biologically precise V1 Neural Network (VOneBlock) in front of a classic CNN. VOneBlock is a specific mathematically parametrized model based on biological convolutions simulated by Gabor filters (GBF) [35].

In contrast to classical computer vision approaches, parameters in VOneBlock first layers are not learned: they are fixed to approximate the empirical data of evolution-optimized actual primate V1 neural response [36]. That approach supports the idea that human cognitive possibilities rely on the combination of both learned and innate mechanisms emergent under the influence of natural evolution [37]. By deploying such an architectural change, VOneNets improve their robustness to various types of adversarial attacks and common image corruptions, as well as results interpretability allowing their decision processes to be better understood, refined, and overridden when necessary [38].

Deep convolutional architectures success leads naturally to the further active development of neural generative architectures, i.e., GANs [13], and further introduction of CNN and GAN-supported innovations in Industry 4.0 [39, 40, 41, 42]. In the first place, GANs enable development of novel simulation-based decision-making and decision-support industrial tools, such as, digital twins [43, 44], and maintain their robustness and accuracy by means of adversarial training [45, 46].

4. VOneGAN: a GAN with a V1 neural network front-end in the discriminator

4.1. VOneGANs architecture

GANs traditional architecture implies a contest between a generative deep neural network (generator) learning to create synthetic content with the same joint probability distribution as a given dataset of real samples; and a discriminative deep neural network (discriminator) predicting boundaries of the reality and classifying samples as either original or generated (synthetic). Putting two learning models in confrontation solves a problem of a strong teacher, since both networks act as constantly developing challengers for the opponent, thus, synchronously co-evolving during the training. Some architectures rely on the idea of reaching better generation quality by primarily enhancing the discriminative component [46, 47].

Designing VOneGANs we also pursue this vision and presume the higher quality of image generation due to a more robust discriminator. We increase the robustness by simulating human neural processing of images in the earliest cortical area of the brain – primary visual cortex V1. As it is suggested in [11], V1 is based on the classical linear-nonlinear-Poisson (LNP) model (see Fig.2.) consisting of:

- A convolutional layer – a Gabor Filter Bank (GFB) Gabor filters with fixed weights constrained by empirical data [36]. It convolves the RGB input images with Gabor filters considerably more heterogeneous than those found in standard CNNs.
- A nonlinear layer (applying rectified linear transformation for simple cells, and spectral power of a quadrature phase-pair for complex cells).
- A stochastic layer (a “stochasticity” generator with variance equal to mean).

A neural network simulating V1 with biologically constrained parameters is incorporated before an updated traditional deep CNN. A symbiosis of a deep CNN and the digital V1 creates a hybrid V1+CNN model. Our test runs in the industrial settings of the abovementioned experimental facility equipped with the inter-roll conveyor confirm the higher accuracy of the hybrid V1+CNN on clean input and its robustness under adversarial attacks. The performance comparison of V1+ResNet-50 network and pure ResNet-50 trained on Conveyor-V3 dataset (see Fig. 3.) is shown on Figure 3.

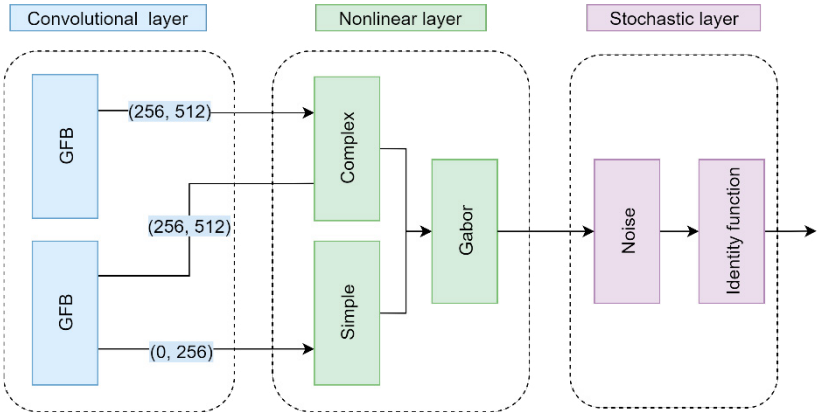

Fig. 2. Model of primary visual cortex V1

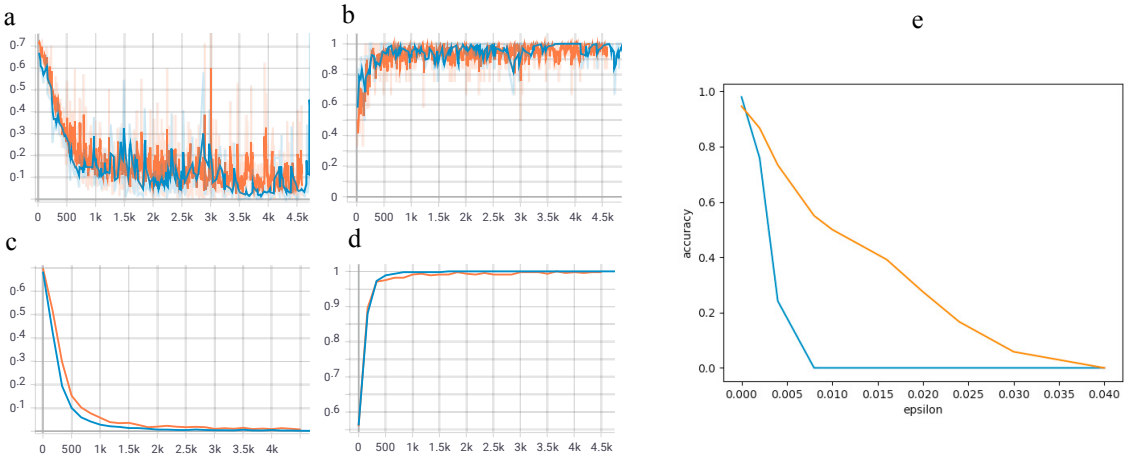

Fig. 3. Performance evaluation of the hybrid CNN over a number of iterations based on ResNet-50 model combined with a V1 network recognizing objects in the industrial inter-roll conveyor system. Blue line – pure ResNet-50 model, orange line – V1+ResNet-50 model. (a) loss on training dataset; (b) top-1 accuracy on training dataset; (c) loss on validation set; (d) top-1 accuracy on validation dataset; (e) adversarial accuracy of the classifier under PGD-20 attack with different values of perturbation l∞ norm, denoted as epsilon.

In this research, we transfer the advantages of more human-like image processing onto generative models. The most obvious solution is to introduce the hybrid V1+CNN model into the GAN discriminator.

We coin VOneGANs as a novel class of GANs applying the hybrid V1+CNN model for performing discrimination of the generated samples, thus, simulating neural processing in natural primary visual cortex. To verify our claims, we collect empirical evidence for the new architecture performance based on several popular GAN architectures believed to have a high potential in Industry 4.0: AC-GAN [48] (see Fig. 4), and RobGAN [49] (see Fig. 5). Both belong to class-conditional image synthesis models with an auxiliary classifier in the discriminator that, besides deciding whether data is generated or original, output the class label for the training data. During the adversarial game, the model improves not only its generative and discriminative but also classification ability.

To adapt the basic CNN discriminator to V1 we add a bottleneck – a transition layer that follows V1 and compresses 512 channels of V1 to the number of channels in the basic GAN discriminator.

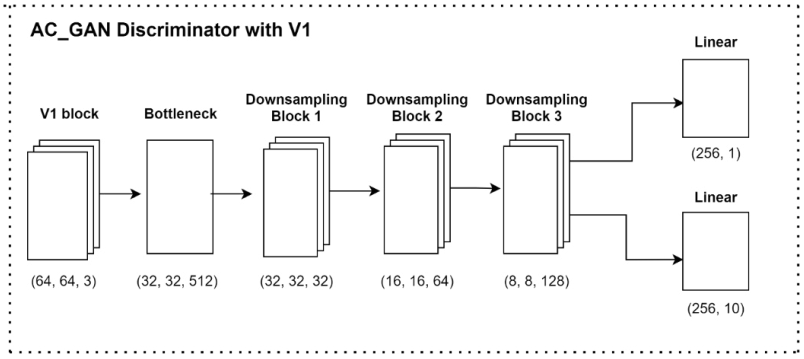

Fig. 4. An example of the hybrid discriminator architecture based on AC-GAN.

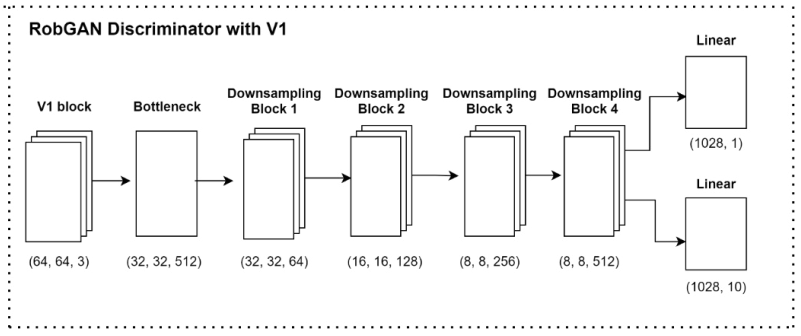

Fig. 5. An example of the hybrid discriminator architecture based on RobGAN.

4.2. Training GANs with a V1 neural network front-end in the discriminator

The discriminator being more robust to various noises, distortions, and adversarial attacks due to its specific architectural features is a challenging and an inconvenient rival for the generator from the very first epochs. The generator does not have the opportunity to fool the discriminator easily, and it is forced to learn faster anticipating the GAN convergence.

Generative abilities of VOneGANs commonly used for realistic synthetic visual content creation are tested on CIFAR-10 (see Fig. 6) and MNIST (see Fig. 7.).

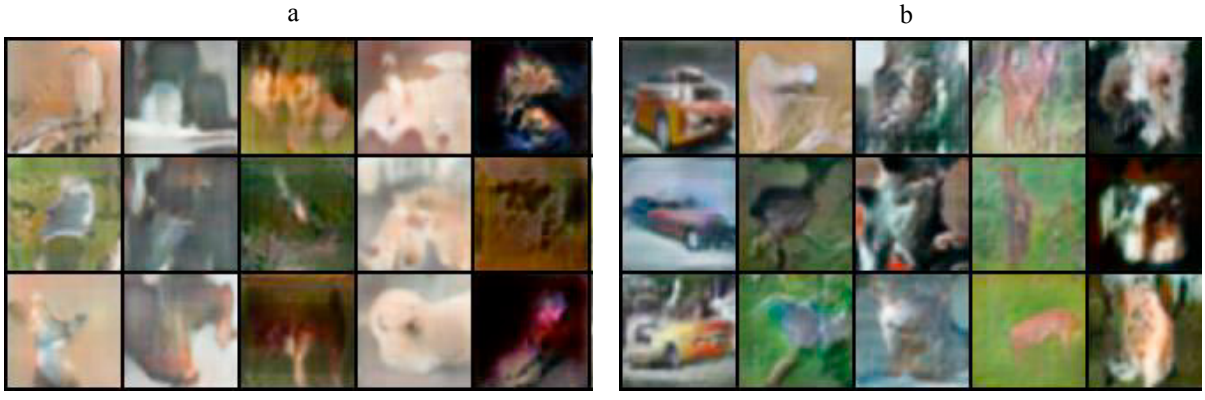

Fig. 6. Newly generated images by GAN trained with a 64 channels CNN in the discriminator during 220 epochs on CIFAR-10 dataset. (a) Pure RobGAN; (b) V1+RobGAN.

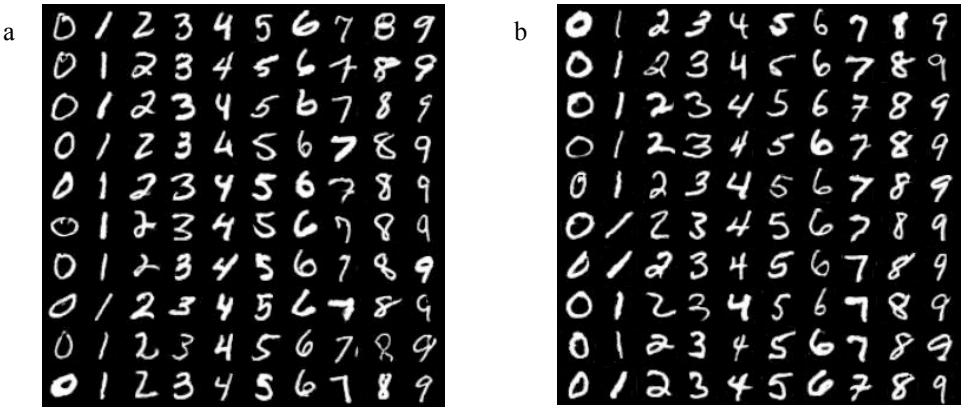

Fig. 7. Newly generated images by GAN trained with a 32 channels CNN in the discriminator during 220 epochs on MNIST dataset. (a) Pure RobGAN; (b) V1+RobGAN.

The key metrics confirm the improved stability of the training process, along with the increased perceptual realism of the generated images (see Fig. 8, Fig. 9). For the human eye, images generated without V1 are more distorted and harder to distinguish.

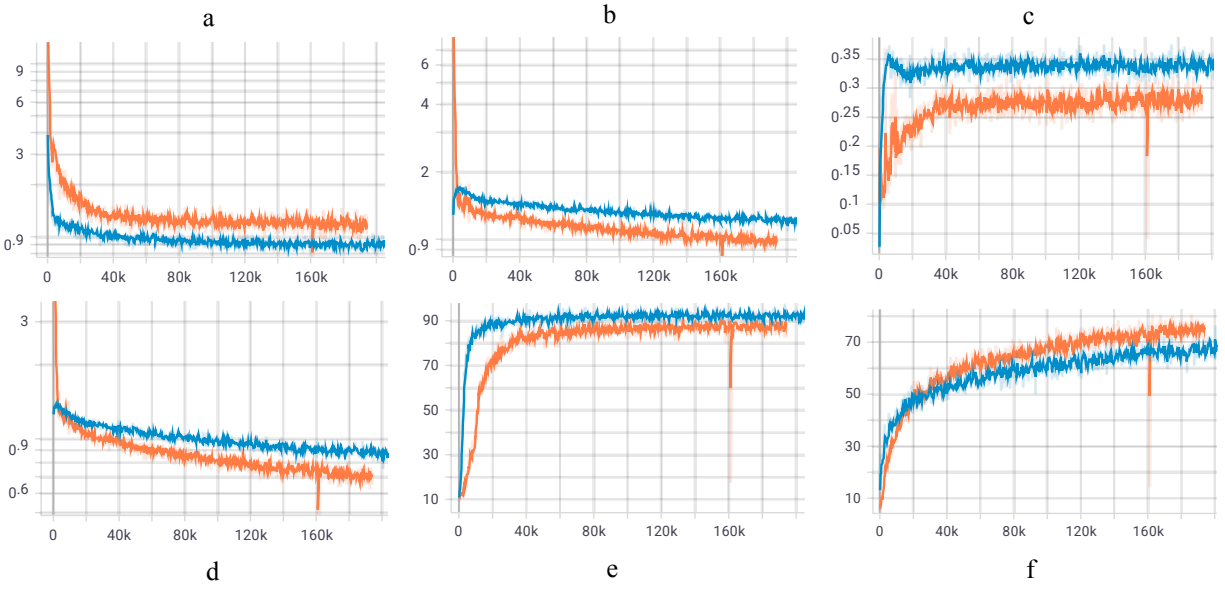

Fig. 8. The key training metrics of pure RobGAN (blue line) and V1+RobGAN (orange line) on CIFAR10 dataset. V1+RobGAN demonstrates improved classification accuracy on the original data with PGD attack, that indicates an increased robustness to the perturbation of discriminator over a number of training iterations. There is a small gap in the classification accuracy on the generated data due to the stronger generator. (a) Generator loss. (b) Discriminator loss. (c) Discriminator loss on generated images. (d) Discriminator loss on original images with PGD. (e) Classification accuracy on generated examples. (f) Classification accuracy on original examples with PGD.

Adding the third component, i.e., an adversarial attacker, to the single adversarial framework, we organize training for both generator and discriminator in the presence of adversarial attacks: the generator feeds generated images to the discriminator; meanwhile real images sampled from training set are pre-processed by an attacking algorithm before sending to the discriminator. Adversarial training is the technique commonly used to improve the robustness of the discriminator, as well as increase the convergence speed and lead to better generators (see Fig. 10). However, its task for GANs with an auxiliary classifier (such as, AC-GAN and RobGAN) is to develop increased robustness also in the

auxiliary classifier which can then be extracted from the discriminator to classify images independently from the GAN framework.

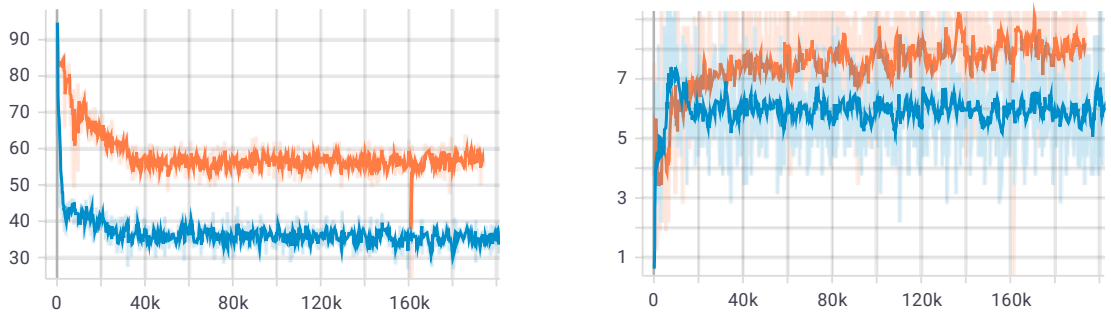

Fig. 9. The comparison of binary accuracy values confirm the improved ability of V1+RobGAN (orange line) on CIFAR10 dataset to distinguish between original with PGD and generated samples over the pure RobGAN (blue line). Left: Binary accuracy on original images with PGD over a number of iterations. Right: Binary accuracy on generated images over a number of iterations.

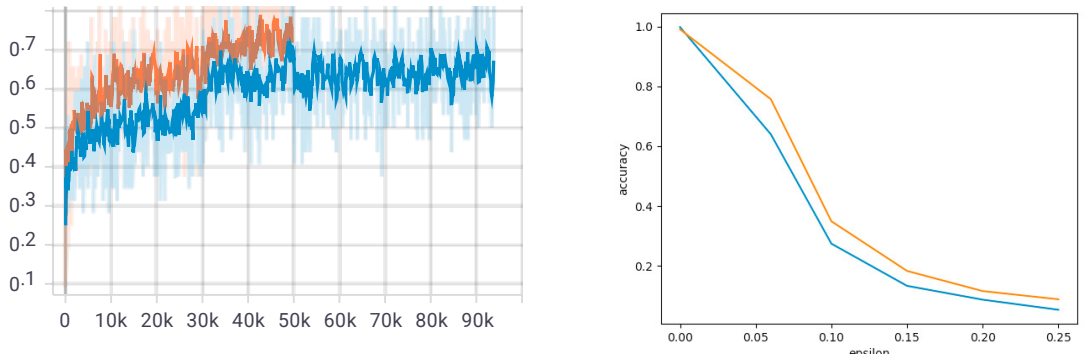

Fig. 10. Achieving better adversarial robustness of the V1+CNN discriminator after fine-tuning on CIFAR10 dataset. Left: Evaluation of Classification accuracy during the fine-tuning of V1+RobGAN (orange line) over a number of iterations and retraining pure RobGAN (blue line). Right: Comparison of V1+RobGAN (orange line) and pure RobGAN (blue line) under adversarial attack PGD-5 with different values of perturbation l∞ norm, denoted as epsilon.

5. Conclusions

In this research, we suggest a novel GAN architecture (VOneGAN) with the qualitatively updated discriminator. The new model imitates human visual processing more precisely than previously existing GANs due to the specific neural layers analogical to those in human primary visual cortex V1 built according to the classical linear-nonlinear-Poisson model but instantiated with the parameters approximating the empirical data of evolution-optimized actual primate V1 neural response [36]. We take our inspiration in VOneNets [11] which show the improved adversarial robustness in object recognition tasks. However, we, in this work, are in search of a solution for industrial tasks that require the generative function, such as, sophisticated data augmentation with synthetic samples, image to image translation, and adversarial robustness training.

Our hypothesis is that VOneGANs can be more successfully used in Industry 4.0 in comparison to existing pure GANs architectures. To prove the concept, we conduct a series of experiments testing the new architecture in respect to various tasks and datasets. We also create various settings for our experiments, engaging both experimental facilities equipped with industrial systems and machine learning sandboxes of research labs.

In this paper, we report the first promising results of VOneGANs training. They support our hypothesis that digital primary visual cortex V1 not only improves discriminative ability but can also be successfully used for the essential

positive change of generative neural models. We foresee good industrial perspectives for VOneGANs and suggest them as providers of high-quality training content and as enablers of future simulation-based decision-making and decision-support tools for condition-monitoring, supervisory control, diagnostics, predictive maintenance, and cybersecurity in Industry 4.0.

References

[1] Byrne, G., Dimitrov, D., Monostori, L., Teti, R., van Houten, F., & Wertheim, R. (2018). Biologicalisation: Biological transformation in manufacturing. CIRP Journal of Manufacturing Science and Technology 21: 1-32.

[2] Dumitrache, I., Caramihai, S. I., Moisescu, M. A., & Sacala, I. S. (2019). Neuro-inspired Framework for cognitive manufacturing control. IFAC-PapersOnLine, 52(13), 910-915.

[3] Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, 48, 144-156.

[4] Akhtar, N., & Mian, A. (2018). Threat of adversarial attacks on deep learning in computer vision: A survey. Ieee Access, 6, 14410-14430.

[5] Ibitoye, O., Shafiq, O., & Matrawy, A. (2019, December). Analyzing adversarial attacks against deep learning for intrusion detection in IoT networks. In 2019 IEEE Global Communications Conference (GLOBECOM) (pp. 1-6). IEEE.

[6] Kumar, R. S. S., Nyström, M., Lambert, J., Marshall, A., Goertzel, M., Comissoneru, A., ... & Xia, S. (2020, May). Adversarial machine learning-industry perspectives. In 2020 IEEE Security and Privacy Workshops (SPW) (pp. 69-75). IEEE.

[7] Shafahi, A., Najibi, M., Ghiasi, A., Xu, Z., Dickerson, J., Studer, C., ... & Goldstein, T. (2019). Adversarial training for free!. arXiv preprint arXiv:1904.12843.

[8] Papernot, N., McDaniel, P., Wu, X., Jha, S., & Swami, A. (2016, May). Distillation as a defense to adversarial perturbations against deep neural networks. In 2016 IEEE symposium on security and privacy (SP) (pp. 582-597). IEEE.

[9] Gu, S., & Rigazio, L. (2014). Towards deep neural network architectures robust to adversarial examples. arXiv preprint arXiv:1412.5068.

[10] Garg, S., Sharan, V., Zhang, B. H., & Valiant, G. (2018). A spectral view of adversarially robust features. arXiv preprint arXiv:1811.06609.

[11] Dapello, J., Marques, T., Schrimpf, M., Geiger, F., Cox, D. D., & DiCarlo, J. J. (2020). Simulating a primary visual cortex at the front of CNNs improves robustness to image perturbations. BioRxiv.

[12] Dodge, S., & Karam, L. (2017, July). A study and comparison of human and deep learning recognition performance under visual distortions. In 2017 26th international conference on computer communication and networks (ICCCN) (pp. 1-7). IEEE.

[13] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. arXiv preprint arXiv:1406.2661.

[14] Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018, April). Synthetic data augmentation using GAN for improved liver lesion classification. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 289-293). IEEE.

[15] Carlini, N., Athalye, A., Papernot, N., Brendel, W., Rauber, J., Tsipras, D., ... & Kurakin, A. (2019). On evaluating adversarial robustness. arXiv preprint arXiv:1902.06705.

[16] Lee, H., Han, S., & Lee, J. (2017). Generative adversarial trainer: Defense to adversarial perturbations with gan. arXiv preprint arXiv:1705.03387.

[17] LeCun, Y. (1998). The MNIST database of handwritten digits. http://yann. lecun. com/exdb/mnist/.

[18] Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images.

[19] Lindsay, G. W. (2020). Convolutional neural networks as a model of the visual system: past, present, and future. Journal of cognitive neuroscience, 1-15.

[20] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision (pp. 1026-1034).

[21] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199.

[22] Chakraborty, A., Alam, M., Dey, V., Chattopadhyay, A., & Mukhopadhyay, D. (2021). A survey on adversarial attacks and defences. CAAI Transactions on Intelligence Technology, 6(1), 25-45.

[23] Terziyan, V., Golovianko, M., & Gryshko, S. (2018). Industry 4.0 Intelligence under Attack: From Cognitive Hack to Data Poisoning. Cyber Defence in Industry, 4, 110-125.

[24] Puckett, A., Bollmann, S., & Ribeiro, F. (2020). Predicting the functional organization of human visual cortex from anatomy using geometric deep learning. Journal of Vision, 20(11), 928-928.

[25] Berco, D., & Shenp Ang, D. (2019). Recent progress in synaptic devices paving the way toward an artificial cogni-retina for bionic and machine vision. Advanced Intelligent Systems, 1(1), 1900003.

[26] Zhuang, C., Yan, S., Nayebi, A., Schrimpf, M., Frank, M. C., DiCarlo, J. J., & Yamins, D. L. (2021). Unsupervised neural network models of the ventral visual stream. Proceedings of the National Academy of Sciences, 118(3).

[27]Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. IEEE transactions on pattern analysis and machine intelligence, 29(3), 411-426.

[28]Huang, Y., Dai, S., Nguyen, T., Bao, P., Tsao, D. Y., Baraniuk, R. G., & Anandkumar, A. (2019). Brain-inspired Robust Vision using Convolutional Neural Networks with Feedback.

[29]Bertoni, F., Citti, G., & Sarti, A. (2019). LGN-CNN: a biologically inspired CNN architecture. arXiv preprint arXiv:1911.06276.

[30]Shakeri, M., Tsogkas, S., Ferrante, E., Lippe, S., Kadoury, S., Paragios, N., & Kokkinos, I. (2016, April). Sub-cortical brain structure segmentation using F-CNN's. In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) (pp. 269-272). IEEE.

[31]Lv, C., Xu, Y., Zhang, X., Ma, S., Li, S., Xin, P., ... & Ma, H. (2018, April). Feature extraction inspired by V1 in visual cortex. In Ninth International Conference on Graphic and Image Processing (ICGIP 2017) (Vol. 10615, p. 106155C). International Society for Optics and Photonics.

[32]Hu, Y., Qiao, K., Tong, L., Zhang, C., Gao, H., & Yan, B. (2018, March). A CNN-based computational encoding model for human V1 cortex. In 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI) (pp. 408-413). IEEE.

[33]Machiraju, H., Choung, O. H., Frossard, P., & Herzog, M. (2021). Bio-inspired Robustness: A Review. arXiv preprint arXiv:2103.09265.

[34]Vuyyuru Reddy, M., Banburski, A., Plant, N., & Poggio, T. (2020). Biologically Inspired Mechanisms for Adversarial Robustness. Center for Brains, Minds and Machines (CBMM).

[35]Jain, A. K., & Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. Pattern recognition, 24(12), 1167-1186.

[36]Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. Journal of neurophysiology.

[37]Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. Nature communications, 10(1), 1-7.

[38]Evans, B. D., Malhotra, G., & Bowers, J. S. (2021). Biological convolutions improve DNN robustness to noise and generalisation. bioRxiv.

[39]Kusiak, A. (2020). Convolutional and generative adversarial neural networks in manufacturing. International Journal of Production Research, 58(5), 1594-1604.

[40]Lv, W., Xiong, J., Shi, J., Huang, Y., & Qin, S. (2021). A deep convolution generative adversarial networks based fuzzing framework for industry control protocols. Journal of Intelligent Manufacturing, 32, 441-457.

[41]Zotov, E., Tiwari, A., & Kadirkamanathan, V. (2020, June). Towards a Digital Twin with Generative Adversarial Network Modelling of Machining Vibration. In International Conference on Engineering Applications of Neural Networks (pp. 190-201). Springer, Cham.

[42]Lee, Y. O., Jo, J., & Hwang, J. (2017, December). Application of deep neural network and generative adversarial network to industrial maintenance: A case study of induction motor fault detection. In 2017 IEEE international conference on big data (big data) (pp. 3248-3253). IEEE.

[43]Becue, A., Maia, E., Feeken, L., Borchers, P., & Praca, I. (2020). A New Concept of Digital Twin Supporting Optimization and Resilience of Factories of the Future. Applied Sciences, 10(13), 4482.

[44]Golovianko, M., Gryshko, S., Terziyan, V., & Tuunanen, T. (2021). Towards digital cognitive clones for the decision-makers: adversarial training experiments. Procedia Computer Science, 180, 180-189.

[45]Liu, X., & Hsieh, C. J. (2019). Rob-gan: Generator, discriminator, and adversarial attacker. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11234-11243).

[46]Terziyan, V., Gryshko, S., & Golovianko, M. (2021). Taxonomy of generative adversarial networks for digital immunity of Industry 4.0 systems. Procedia Computer Science, 180, 676-685.

[47]Schonfeld, E., Schiele, B., & Khoreva, A. (2020). A u-net based discriminator for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8207-8216).

[48]Odena, A., Olah, C., & Shlens, J. (2017, July). Conditional image synthesis with auxiliary classifier gans. In International conference on machine learning (pp. 2642-2651). PMLR.

[49]Liu, X., & Hsieh, C. J. (2019). Rob-gan: Generator, discriminator, and adversarial attacker. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11234-11243).