UDC 004.032.26

NEURAL NETWORK TRANSFER freestyle in portrait images using neural networks MECHANISM OF ATTENTION *

SA Berezin [1], VM Volkova [2]

[1] *630073, Russian Federation, Novosibirsk, pr. Karl Marx, 20, Novosibirsk State Technical University, a student of the Department of Theoretical and Applied Informatics. Mail-E: sergeyberezin123@gmail.com*

[2] *630073, Russian Federation, Novosibirsk, pr. Karl Marx, 20, Novosibirsk State Technical University, Assistant Professor, Department of Theoretical and Applied Informatics. Mail-E: volkova@ami.nstu.ru*

arbitrary style transfer problem is to create a new, not previously existing image, by combining the two image data: the original and stylistic. The original image forms the structure, basic geometrical lines and shapes of the resulting image, while image stylistic sets the color and texture of the result. The word "random" in this context refers to the absence of a single, pre-unlearned style. For example, a convolutional neural network capable of carrying the new style only after their retraining or additional training on new data volume are not considered to solve such problem, while the network based on attention mechanism capable of producing such a transformation without a retraining - yes. The original image may be another, e.g., a photograph, and stylistic - a picture of a famous artist. The resulting image in this case will be the scene depicted in the original photo, made in the style of this painting. Modern style arbitrary transfer algorithms can achieve good results in this problem, but the processing of portrait images of people the result of such algorithms is or is not acceptable, because of excessive distortion of facial features, or bland, not having the characteristics of style image. In this work the considered However, in the processing of portrait images of people the result of such algorithms is or is not acceptable, because of excessive distortion of facial features, or bland, not having the characteristics of style image. In this work the considered However, in the processing of portrait images of people the result of such algorithms is or is not acceptable, because of excessive distortion of facial features, or bland, not having the characteristics of style image. In this work the considered

* *Manuscript received December 20, 2019*

etsya approach to solving this problem by using a combination of architecture of deep neural networks with the mechanism of attention and convolution segmentation network, carry forward the style, taking into account the image content of a particular segment: with a bright predominance of the style of the form for the background of the image, and the prevalence of substance over form in portrait portion of the image containing the human face image is directly and / or shapes.

Keywords: machine learning, deep learning, neural networks, image processing, style transfer, segmentation, convolutional neural network mechanism of attention.

INTRODUCTION

machine learning methods in general and neural networks in particular, today find their application in a large variety of areas [1, 9-15]. One such area is the image processing. Let us consider the problem of the transfer of style, related to this field.

Moving images are artistic style themselves to create a new image containing characteristic patterns of global and local image of style, and wherein the retaining structure of the image source [1].

Architecture based on convolutional neural network successfully cope with the transfer of [2] or more [3] in advance learned styles, however going beyond the prefabricated templates they require full retraining.

Previous attempts to solve this problem with the adaptive normalization technique data (adaptive instance normalization) [4], but a subjective evaluation of the results was far from ideal.

With the release papers authored by L. Sheng, Z. Lin, J. Shao, X. Wang [6] and Dae Young Park, and Kwang Hee Lee [1] in 2018 year, which describes a new approach to solving this problem, there was significant breakthrough in the field. The proposed architecture Avatar-Net [6] and modified in SANet architecture [1] solutions using a focus mechanism (attention mechanism) allows appropriately reconstruct the characteristic pattern portable style for each portion of the processed image with the contents of this section by comparing the relations, such as proximity (identity loss), between the contextual and stylistic images.

However, in the processing of portrait images in any of above given methods inevitably arises a problem of distortion of facial features: those geometric transformation learned neural network making pattern similar in style to the stylistic image completely distort the shape of an oval face, eyes, mouth, resulting in a loss of recognition in the portrait individual.

The object of study of this work was to find ways to solve this problem, allowing to save as expressive geometric and color transformations produced when transferring the style and features of a human face recognition.

1. METHOD

The solution proposed in this paper consists in splitting of the problem into two parts: the transfer of the style on the background of the original image with a greater weight of form over content, thus preserving the bright colors and geometric transformations, and the transfer of style to the face and figure of a man with a predominance of weight content over form. Visualization solutions is presented in Figure 1.

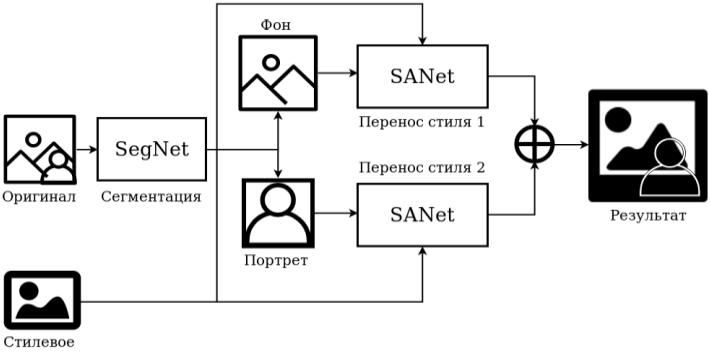

Fig. 1 - Scheme of the solution architecture

For dividing an image into two semantic areas: the background and the figure of a man, that is, to solve the segmentation problem, the neural network has been applied SegNet architecture [5], trained on a set of Microsoft COCO data [7].

The choice of this architecture is due to small time spent on processing frame (40 ms GPU Nvidia Tesla v100) and satisfactory for this task results.

To transfer style architecture used previously mentioned SANet (Style-Attentional Network), which is the best solution currently providing the highest quality results [1].

One of the key features of this model is the use of the mechanism of attention (attention mechanism) to detect patterns specific to each individual image.

For varying the ratio of the original image features to the features of style it was asked to enter two weighting coefficient. In fact, this has been achieved by merging the results of the two layers of the neural network with different coefficients:

$$\bullet\bullet\bullet = \bullet\bullet\bullet\bullet\bullet 4_1 * \bullet_1 + \bullet\bullet\bullet\bullet\bullet 4_2 * \bullet_2$$

where the coefficients $\bullet_1$ and $\bullet_2$ determine the degree of importance of the low-level and high-level features, respectively.

After isolation of the field with the image of a person is made a binary mask that covers this area. Each image obtained by the frame mask trimmed fragment. Then, the produced first (background) style transferred to the original image and the second (portrait) style transport. Thereafter, the resulting images are superimposed on each other.

2. RESULTS OF EXPERIMENTS

Background SANet neural network was trained on a set of data MSCOCO [7] as the original images and set WikiArt [8] as style images. Both sets contain about 80,000 images.

The objective function was chosen as a function of identity loss, designed to prioritize conservation structure of the image, rather than a change of style characteristics:

$$\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet = \lambda\bullet\bullet\bullet\bullet\bullet\bullet\bullet 1 (\| \bullet\bullet - \bullet 2\bullet \|_2 + \| \bullet\bullet - \bullet\bullet \|_2)$$

$$+ \lambda\bullet\bullet\bullet\bullet\bullet\bullet\bullet 2 \sum_{\bullet=1}^{\bullet} (\| \varphi\bullet(\bullet\bullet - \bullet\bullet) \|_2 + \varphi\bullet(\| \bullet\bullet - \bullet\bullet \|_2))$$

Where $\bullet\bullet$ and $\bullet\bullet$ designate two output generated from the same initial image and style, $\bullet\bullet$ and $\bullet\bullet$ there is the original image and style, each $\varphi\bullet$ denotes a neural network layer $\lambda\bullet\bullet\bullet\bullet\bullet\bullet\bullet 1$ and

λ ·····2 - gipperparametry with experimentally selected values 1 and 50, respectively. As a result, such a loss function allows to preserve the original structure of the image and style features transferred image in one and the same time [1].

Optimizer Adam (adaptive moment estimation) [9] with a step learning 0.0001 and packet size applied to 5. augmentation cutoff portion images.

Figure 2 and 3 show the result of the proposed phase-network.

Fig. 2 - Transfer to the portrait style. From left to right: original image stylistic image, the result of direct transfer of style, the result of the segmentation, the result of the transfer to the segmented style Bathing image combined result.

Fig. 3 - The result of the algorithm. Observance of the form seen chelo- lines vecheskogo face when styling.

CONCLUSION

The combination of all above techniques and methods make it possible to improve the quality of style transfer with an arbitrary image on human portraits in comparison with the original sets. The solution found was introduced in the final software package.

LIST OF REFERENCES

1. *Dae Young Park, Kwang Hee Lee.* Arbitrary Style Transfer with Style-Attentional Networks [Electronic resource] // arXiv: 1812.02342 [cs.CV] 6 Dec 2018 - URL: https://arxiv.org/abs/1812.02342

2. *Leon A. Gatys, Alexander S. Ecker, Matthias Bethge.* A Neural Algorithm of Artistic Style [Electronic resource] // arXiv: 1508.06576v2 [cs.CV] 2 Sep 2015 - URL: https://arxiv.org/abs/1508.06576

3. *Brandon Cui, Calvin Qi, Aileen Wang.* Multi-style Transfer: Generalizing Fast Style Transfer to Several Genres [Electronic resource] // Stanford University 2017 - URL: http://cs231n.stanford.edu/reports/2017/pdfs/401.pdf

4. *Xun Huang, Serge Belongie.* Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization [Electronic resource] // arXiv: 1703.06868 [Cs.CV] 20 Mar 2017 - URL: https://arxiv.org/abs/1703.06868

5. *Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla.* SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation

[Electronic resource] // arXiv: 1511.00561 [cs.CV] 2 Nov 2015 - URL: https://arxiv.org/abs/1511.00561

6. *L. Sheng, Z. Lin, J. Shao, X. Wang.* Avatar-Net: Multi-scale zero-shot style transfer by feature decoration. In Proc. CVPR, pages 8242-8250, 2018

7. *T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll'ar, and CL Zitnick.* Microsoft COCO: Common objects in context. In Proc. ECCV, pages 740-755. Springer, 2014.

8. *F. Phillips and B. Mackintosh.* Wiki Art Gallery, Inc .: A case for critical thinking. Issues in Accounting Education, 26 (3): 593-608, 2011.

9. *Diederik P. Kingma and Jimmy Lei Ba.* Adam: A method for stochastic optimization. [Electronic resource] // arXiv: 1412.6980v9 2014. - URL: https://arxiv.org/abs/1412.6980v9

10. Chomutov SO, Serebryakov, NA Improving the quality of short-
Derivatives prediction
group electricity power supply points with agricultural pomo-

schyu machine learning tools // Scientific Bulletin of the NSTU.
- 2019. - № 3 (76). -
S. 149-168. - DOI: 10.17212 / 1814-1196-2019-3-149-168.

11. Gergeti OM, Mescheryakov RV Application of information Me-
todov, neural networks
and a genetic algorithm to solve the problem of the treatment regimen of choice // Scientific Bulletin
NSTU. - 2018. - № 3 (72). - S. 7-20. - doi: 10.17212 / 1814-1196-20183-7-20.

12. Voivod AA, Romannikov D. O. Synthesis of neural network for
decision logic-arithmetic problems // Tr. SPIIRAS. 2017. Vol. 54.
C. 205-223.

13. Romannikov D. O. On the transformation of Petri nets in neuronal
Network // Coll. scientific. tr. NSTU. 2016. No 4 (86). S. 98-103.

14. Dubenko YV, Dyshkant EE Neural network selection algorithm
methods for time series prediction // Bulletin of Astrakhan State Technical University. Series: Control, computer engineering and computer science. 2019. No 1. pp 51-60. DOI:

10.24143 / 2072-9502-2019-1-51-60.

15. Timofeev, AV, & Derin, OA (1). Design Concept
hierarchical neural networks for analysis of multi. *SPIIRAS Proceedings (* 10), 160-166.
https://doi.org/10.15622/sp.10.11

Berezin Sergey A., student of the Department of Theoretical and Applied Informatics of the Novosibirsk State Technical University, the lab technician business solutions based on artificial intelligence of the Moscow Physical-Technical Institute. The main direction of research - technology of machine learning. Mail-E: sergeyberezin123@gmail.com

Volkova Victoria M., Associate Professor, Department of Theoretical and Applied Physics, Novosibirsk State Technical University, Candidate of Technical Sciences. The main direction of scientific researches machine learning, data mining. He has more than 30 publications. Mail-E: volkova@ami.nstu.ru

Neural arbitrary style transfer for portrait images using the attention mechanism *

SA Berezin [1], VM Volkova [2]

[1] *Novosibirsk State Technical University, 20 Karl Marks Avenue, Novosibirsk, 630073, Russian Federation, undergraduate of the department of theoretical and applied informatics. Mail-E: sergeyberezin123@gmail.com*

[2] *Novosibirsk State Technical University, 20 Karl Marks Avenue, Novosibirsk, 630073, Russian Federation, PhD in engineering, associate professor of the department of theoretical and applied informatics. Mail-E: volkova@ami.nstu.ru*

Arbitrary style transfer is the task of synthesis of an image that has never been seen before, using two given images: content image and style image. The content image forms the structure, the basic geometric lines and shapes of the resulting image, while the style image sets the color and texture of the result. The word "arbitrary" in this context means the absence of any one pre-learned style. So, for example, convolutional neural networks capable of transferring a new style only after training or retraining on a new amount of data are not considered to solve such a problem, while networks based on the attention mechanism that are capable of performing such a transformation without retraining - yes. An original image can be, for example, a photograph, and a style image can be a painting of a famous artist. The resulting image in this case will be the scene depicted in the original photograph, made in the stylie of this picture. Recent arbitrary style transfer algorithms make it possible to achieve good results in this task, however, in processing portrait images of people, the result of such algorithms is either unacceptable due to excessive distortion of facial features,

* *Manuscript received on December 20. 2019*

or weakly expressed, not bearing the characteristic features of a style image. In this paper, we consider an approach to solving this problem using the combined architecture of deep neural networks with a attention mechanism that transfers style based on the contents of a particular image segment: with a clear predominance of style over the form for the background part of the image, and with the prevalence of content over the form in the image part containing directly the image of a person.

Keywords: machine learning, deep learning, neural networks, image processing, style transfer, segmentation, convolutional neural networks, attention mechanism.

REFERENCES

1. *Leon A. Gatys, Alexander S. Ecker, Matthias Bethge* - A Neural Algorithm of Artistic Style [Electronic resource] // arXiv: 1508.06576v2 [cs.CV] 2 Sep 2015 - URL: https://arxiv.org/abs/1508.06576

2. *Brandon Cui, Calvin Qi, Aileen Wang* - Multi-style Transfer: Generalizing Fast Style Transfer to Several Genres [Electronic resource] // Stanford University 2017 - URL: http://cs231n.stanford.edu/reports/2017/pdfs/401.pdf

3. *Xun Huang, Serge Belongie* - Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization [Electronic resource] // arXiv: 1703.06868 [Cs.CV] 20 Mar 2017 - URL: https://arxiv.org/abs/1703.06868

4. *Dae Young Park, Kwang Hee Lee* - Arbitrary Style Transfer with Style-Attentional Networks [Electronic resource] // arXiv: 1812.02342 [cs.CV] 6 Dec 2018 - URL: https://arxiv.org/abs/1812.02342

5. *Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla* - SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation [Electronic resource] // arXiv: 1511.00561 [cs.CV] 2 Nov 2015 - URL: https://arxiv.org/abs/1511.00561

6. *L. Sheng, Z. Lin, J. Shao, X. Wang* - Avatar-Net: Multi-scale zero-shot style transfer by feature decoration. In Proc. CVPR, pages 8242-8250, 2018

7. *T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll'ar, and CL Zitnick* - Microsoft COCO: Common objects in context. In Proc. ECCV, pages 740-755. Springer, 2014.

8. *F. Phillips and B. Mackintosh* - Wiki Art Gallery, Inc .: A case for critical thinking. Issues in Accounting Education, 26 (3): 593-608, 2011.

9. *Diederik P. Kingma and Jimmy Lei Ba.* Adam: A method for stochastic optimization. [Electronic resource] // arXiv: 1412.6980v9 2014. - URL: https://arxiv.org/abs/1412.6980v9

10. Khomutov SO, Serebryakov NA Povyshenie kachestva krat-
kosrochnogo prognozirovaniya elektropotrebleniya gruppy tochek postavki elektroenergii sel'khozproizvoditelei s pomoshch'yu instrumentov mashinnogo obucheniya [Increasing the quality of short-term load forecasting of the delivery point cluster of agricultural producers with a learning machine].
Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta - Science bulletin of the Novosibirsk state technical university, 2019, no. 3 (76), pp. 149-168. DOI: 10.17212 / 1814-1196-2019-3-149-168.

11. Gerget OM, Meshcheryakov RV Primenenie informatsionnyk hmetodov, neironnykh setei I geneticheskogo algoritma dlya resheniya zadachi vybora skhemy lecheniya [Applying information methods, neural networks and genetic algorithms for solving the problem of selecting a scheme of treatment]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta - Science bulletin of the Novosibirsk state technical university,* 2018, no. 3 (72), pp. 7-20. doi: 10.17212 / 1814-1196-20183-7-20.

12. Voevoda AA, Romannikov DO Sintez neironnoi seti dlia resh-
eniia logiko-arifmeticheskikh zadach. Trudy SPIIRAN, 2017, iss. 54, pp. 205-223.

13. Romannikov DO O preobrazovanii seti Petri v neironnuiu set '.
Sbornik nauchnykh trudov NGTU, 2016, no. 4 (86), pp. 98-103.

14. Timofeev, AV, & Derin, OA (1). The principles of building hier-
archical neural networks for the analysis of multi-images. Proceedings of SPIIRAS, (10) 160-166. https://doi.org/10.15622/sp.10.11

15. Dubenko Yu. V., Dyshkant EE Neural network algorithm for
choosing methods of time series forecasting. Vestnik of Astrakhan State Technical University. Series: Management, Com-puter Science and Informatics. 2019; 1: 51-60. (. In Russ) DOI: 10.24143 / 2072-9502-2019-1-51-60.

Berezin Sergey Andreevich, undergraduate of the Department of Theoretical and Applied Informatics, Novosibirsk State Technical University, technician of the laboratory of business solutions based on artificial intelligence of the Moscow Institute of Physics and Technology. The main direction of scientific research is machine learning technology.

Email: sergeybere-
zin123@gmail.com

Viktoriya M. Volkova. State Technical University Novosibisrk, department of theoretical and applied informatics, volkova@ami.nstu.ru , PhD in engineering, associate professor. Main areas of scientific interest: machine learn-

ing, data science, investigations of statistical criteria under conditions of basic assumptions violation. She is the author coauthor of more of over 30 research papers and teaching manuals.