Dear Students,

I forgot about individual tasks. According to curriculum you have individual tasks in this semester.

You can choose the topic from the list below or you can come up with a task yourself.

Topics:

1. The study of edge detectors. Implementation of the most famous algorithms for detection edges.

2. Geometrical transformation of the image (Data augmetation). Implementation 5 or more methods for geometrical transformation.

3. Tempate matching. Implementation algorithm for finding small picture (pattern) on a big image.

4. The sudy of adaptive methods of image thresholding (binary segmentation). Implementation 2 or more adaptive methods for binary segmentation (Isodata, Li, mean, minimum, Yen).

5. Finding objects with nessesery shape in the image (for instance, circle, rectangle, etc.).

6. Implementation of random walker segmentation.

7. Implementation of watershed segmentation.

8. Implementation of object tracking.

9. Implementation of histogram of oriented gradients (HOG).

You should write the report for your topic to the enad of this semester.

If you want take any another topic (for instance,  implementation of face detector or object tracking) you can take it.

If you have a topic for your diploma which is connected with image processing you can take it.

I give you freedom for choosing and preparing your work. I try to find example of report and send you soon.

If you have any questions we can discuss today.

*Best regards,*

*Zyuzin Vasily Viktorovich*

The Ministry of education and science of the Russian Federation

Ural Federal University named by the first President of Russia B.N.Yeltsin

Institute of Radioelectronics and Information Technologies

Faculty of Radioelectronics and Information Systems

**NEURAL NETWORK FOR IMAGE PROCESSING IN A SPECIFIC STYLE**

Project supervisors: V.V.Zuzin

D.P.Zarifullina

Student of group RI-440004: M.V.Gabdulkhanov

Yekaterinburg

2017

CONTENTS

[INTRODUCTION 2](#_Toc502271644)

[1 SYNOPSIS OF MACHINE LEARNING 4](#_Toc502271645)

[1.1 Historical reference to segregation from AI 4](#_Toc502271646)

[1.2 Forms of learning 6](#_Toc502271647)

[1.2.1 Unsupervised learning 7](#_Toc502271648)

[1.2.2 Supervised learning 9](#_Toc502271649)

[1.2.3 Reinforcement learning 9](#_Toc502271650)

[2 ARTIFICIAL NEURAL NETWORK AND DEEP LEARNING 11](#_Toc502271651)

[3 STYLE TRANSFER WITH DEEP LEARNING 14](#_Toc502271652)

[3.1 Introduction 14](#_Toc502271653)

[3.2 The principle of image generation 15](#_Toc502271654)

[3.3 Mathematical methodology 16](#_Toc502271655)

[CONCLUSION 19](#_Toc502271656)

[LITERATURE REVIEW 21](#_Toc502271657)

[REFERENCES 22](#_Toc502271658)

# INTRODUCTION

Artificial intelligence (AI) is the future. Artificial intelligence is science fiction. Artificial intelligence is already a part of our life. All these statements are true, it all depends on what you mean by artificial intelligence. When the professional go-game player was beaten by program, the developer of it was asked for explanation of such succeed. The answer was "Artificial intelligence, machine learning and deep learning". The easiest way to explain this statement is the idea of artificial intelligence, which appeared first - the largest area, then later inside it appeared machine learning, and, finally, the deep training that caused today's rapid development of AI - being a sub-domain of artificial intelligence and machine learning.

Machine learning - a set of methods that allows a computer to find initially unknown connections and patterns in the arrays of data. This large subsection of artificial intelligence is mathematical discipline that uses mathematical statistics sections, numerical optimization methods, probability theory, discrete analysis, and extracts knowledge from the data.

A neural network – a mathematical model, this is one of the possible outcomes of machine learning process, its software or hardware representation, based on the principle of the organization and functioning of biological neural networks. ANN is a system of connected and interacting with each other simple processors (artificial neurons). Such processors are usually fairly simple, each one only deals with signals that receive and send to other, sometimes. Nevertheless, connected to large enough network, with an interaction such simple processors together can solve fairly complicated tasks.

Contradiction: the use of machine learning in general and neural networks, and especially deep learning, in particular, consists in the fact that for the implementation and using such systems a well-qualified specialist is needed that can rationally design the network and train it, in conditions of poor data or in complete absence of this, on the one hand. Also, for successful learning process, you need a fairly powerful computer, which price can reach hundreds of thousands of rubles. On the other hand, if such conditions are satisfied, the result will be a system capable of speeding up and automating the work of a large number of people or able to do what people cannot do at all.

The objectives is artificial neural networks.

The subject of this research is deep learning, which is a part of broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms.

The purposes of the project is the exploration of neural networks and deep learning on the example of the problem of style transfer.

The hypothesis: is it possible to use deep neural networks to transfer a style from one image to another.

Allocated the following research questions:

* To learn about segregation from AI
* To study different learning forms
* To consider artificial neural networks
* To look through the solution for style transfer problem

# SYNOPSIS OF MACHINE LEARNING

## 1.1 Historical reference to segregation from AI

Already in the early days as a scientific discipline, there was a need for computer learning according to some data for that were made attempts to use different methods, including the similarity of the first neural networks - perceptrons and other models, evaluated much later in statistics (now machine learning depends on statistics, in the large range of different methods and tools).

Great dependence on data and observations, has become causal separation and machine learning into two separate branches of technology development. Inside, the rest of the work on learning by data, lead to the emergence of inductive logical programming, but the more static line went beyond this. By 1980, expert systems began to dominate, and the statistics constrained by the problems of obtaining data, their processing and presentation, began to lag. Neural networks, also rejected by the computer science and technology, broke away into another independent area, and to the 1980s the back propagation become the main success of it.

Machine learning, as an independent field, began a second life in 1990 with a rethinking of its goal, from the achievement artificial intelligence to solving practical problems. Such change affected the tools: statistical and probabilistic methods began to be used, which lead to the possibility of digitizing information and more easily distributing. Machine learning and data mining often use similar methods, but while machine learning is aimed at making predictions, based on the features and knowledge derived from the data, data mining looks for previously unknown properties of the same data. For intellectual data analysis, many methods of machine learning are used, but for different purposes; on the other hand, machine learning also uses data mining techniques as "uncontrolled learning" or as a pre-processing step to improve learning accuracy. The main difference is that machine learning attempts to infer the data and reproduce this data in its own way, while the date mining looks for patterns in the data.[2,4]

Artificial intelligence has become a part of our dreams and aspirations since scientists first pronounced this term at the Dartmouth conferences in 1956 and initiated the development of the field of AI. In the future, AI was perceived as a pledge of the bright future of our civilization, then as a strange invention of computer geniuses. Over the past two years, there has been a strong leap in the development of AI, especially after 2015. And the main reason is the wide distribution of graphics processors, which accelerate and lower the cost of parallel computing. This was also facilitated by the appearance of almost unlimited data storage capabilities and avalanche-like growth of data of various types (what is called "big data") - images, text, cartographic data, and so on.

As the result, machine learning took about twenty years to go from the subfield of the artificial intelligence to the super-perspective independent science, which had fallen into a hibernation for another decade, to rise on the wings of the multiply increased computing power of all types of computers.

## 1.2 Forms of learning

There are three types of feedback that determine the three main types of learning:

In unsupervised learning the agent learns patterns in the input even though no explicit feedback is supplied. The most common unsupervised learning task is clustering: detecting potentially useful clusters of input examples. For example, a taxi agent might gradually develop a concept of “good traffic days” and “bad traffic days” without ever being given labeled examples of each by a teacher.

In reinforcement learning the agent learns from a series of reinforcements—rewards or punishments. For example, the lack of a tip at the end of the journey gives the taxi agent an indication that it did something wrong. The two points for a win at the end of a chess game tells the agent it did something right. It is up to the agent to decide which of the actions prior to the reinforcement were most responsible for it.

In supervised learning the agent observes some example input–output pairs and learns a function that maps from input to output. In component 1 above, the inputs are percepts and the output are provided by a teacher who says “Brake!” or “Turn left.” In component 2, the inputs are camera images and the outputs again come from a teacher who says “that’s a bus.” In 3, the theory of braking is a function from states and braking actions to stopping distance in feet. In this case the output value is available directly from the agent’s percepts (after the fact); the environment is the teacher.[2]

In practice, these distinctions are not always so crisp. In semi-supervised learning we are given a few labeled examples and must make what we can of a large collection of unlabeled examples. Even the labels themselves may not be the oracular truths that we hope for. Imagine that you are trying to build a system to guess a person’s age from a photo. You gather some labeled examples by snapping pictures of people and asking their age. That’s supervised learning. But in reality some of the people lied about their age. It’s not just that there is random noise in the data; rather the inaccuracies are systematic, and to uncover them is an unsupervised learning problem involving images, self-reported ages, and true (unknown) ages. Thus, both noise and lack of labels create a continuum between supervised and unsupervised learning.

### 1.2.1 Unsupervised learning

Unsupervised learning (self-learning, spontaneous learning) is one of the ways of machine learning, in which the test system spontaneously learns to perform the task without the intervention of the experimenter. From the point of view of cybernetics, this is one of the types of cybernetic experiment. As a rule, this is only suitable for tasks in which descriptions of a set of objects (learning sample) are known, and it is required to detect internal relationships, dependencies, patterns existing between objects.[3]

Unsupervised learning is often contrasted with supervised learning, when for each learning object the "correct answer" is forcibly given, and it is required to find the relationship between the stimuli and the reactions of the system.

Despite numerous applied achievements, supervised learning was criticized for its biological implausibility. It is difficult to imagine a learning mechanism in the brain that would compare the desired and actual output values, performing feedback correction. Teaching without a teacher is a much more plausible model of learning in the biological system. Developed by Kohonen and many others, it does not need a target vector for exits and, therefore, does not require comparison with predefined ideal answers.

In this case, a set of N observations of the random vector X is fed into the input. The goal is to restore the regularities of this distribution, without resorting to the expert's help to establish the correct answers and the degree of correspondence of the prediction. The measurement of X is usually much larger than in teaching with the teacher, and the patterns in the data are more complex than simple location estimates. But when learning without a teacher all available data can be used directly for training, without the need for fragmentation into training, test and validation samples.

Generally, these descriptive statistics attempt to characterize X-values, or collections of such values, where Pr(X) is relatively large. Principal components, multidimensional scaling, self-organizing maps, and principal curves, for example, attempt to identify low-dimensional manifolds within the X-space that represent high data density. This provides information about the associations among the variables and whether or not they can be considered as functions of a smaller set of “latent” variables. Cluster analysis attempts to find multiple convex regions of the X-space that contain modes of Pr(X). This can tell whether or not Pr(X) can be represented by a mixture of simpler densities representing distinct types or classes of observations. Mixture modeling has a similar goal. Association rules attempt to construct simple descriptions (conjunctive rules) that describe regions of high density in the special case of very high dimensional binary-valued data. [3]

With supervised learning there is a clear measure of success, or lack thereof, that can be used to judge adequacy in particular situations and to compare the effectiveness of different methods over various situations.

Lack of success is directly measured by expected loss over the joint distribution Pr(X, Y ). This can be estimated in a variety of ways including cross-validation. In the context of unsupervised learning, there is no such direct measure of success. It is difficult to ascertain the validity of inferences drawn from the output of most unsupervised learning algorithms. One must resort to heuristic arguments not only for motivating the algorithms, as is often the case in supervised learning as well, but also for judgments as to the quality of the results. This uncomfortable situation has led to heavy proliferation of proposed methods, since effectiveness is a matter of opinion and cannot be verified directly.

### 1.2.2 Supervised learning

Given a training set of N example input–output pairs (x1, y1),(x2, y2),...(xN , yN ) , where each yj was generated by an unknown function y = f(x), discover a function h that approximates the true function f, this is the standard task of supervised learning:

Here x and y can be any value; they need not be numbers. The function h is a hypothesis. Learning is a search through the space of possible hypotheses for one that will perform well, even on new examples beyond the training set. To measure the accuracy of a hypothesis we give it a test set of examples that are distinct from the training set. We say a hypothesis generalizes well if it correctly predicts the value of y for novel examples. Sometimes the function f is stochastic—it is not strictly a function of x, and what we have to learn is a conditional probability distribution, P(Y | x).

When the output y is one of a finite set of values (such as sunny, cloudy or rainy), the learning problem is called classification, and is called Boolean or binary classification if there are only two values. When y is a number (such as tomorrow’s temperature), he learning problem is called regression. (Technically, solving a regression problem is finding a conditional expectation or average value of y, because the probability that we have found exactly the right real-valued number for y is 0.)[11]

### 1.2.3 Reinforcement learning

Consider, for example, the problem of learning to play chess. A supervised learning agent needs to be told the correct move for each position it encounters, but such feedback is seldom available. In the absence of feedback from a teacher, an agent can learn a transition model for its own moves and can perhaps learn to predict the opponent’s moves, but without some feedback about what is good and what is bad, the agent will have no grounds for deciding which move to make. The agent needs to know that something good has happened when it (accidentally) checkmates the opponent, and that something bad has happened when it is checkmated—or vice versa, if the game is suicide chess. This kind of feedback is called a reward, or reinforcement. In games like chess, the reinforcement is received only at the end of the game. In other environments, the rewards come more frequently. In ping-pong, each point scored can be considered a reward; when learning to crawl, any forward motion is an achievement. Our framework for agents regards the reward as part of the input percept, but the agent must be “hardwired” to recognize that part as a reward rather than as just another sensory input. Thus, animals seem to be hardwired to recognize pain and hunger as negative rewards and pleasure and food intake as positive rewards. Reinforcement has been carefully studied by animal psychologists for over 60 years.

Rewards were introduced before, where they served to define optimal policies in Markov decision processes (MDPs). An optimal policy is a policy that maximizes the expected total reward. The task of reinforcement learning is to use observed rewards to learn an optimal (or nearly optimal) policy for the environment. Above the agent has a complete model of the environment and knows the reward function, here we assume no prior knowledge of either. Imagine playing a new game whose rules you don’t know; after a hundred or so moves, your opponent announces, “You lose.” This is reinforcement learning in a nutshell. [3]

In many complex domains, reinforcement learning is the only feasible way to train a program to perform at high levels. For example, in game playing, it is very hard for a human to provide accurate and consistent evaluations of large numbers of positions, which would be needed to train an evaluation function directly from examples. Instead, the program can be told when it has won or lost, and it can use this information to learn an evaluation function that gives reasonably accurate estimates of the probability of winning from any given position. Similarly, it is extremely difficult to program an agent to fly a helicopter; yet given appropriate negative rewards for crashing, wobbling, or deviating from a set course, an agent can learn to fly by itself.

Reinforcement learning might be considered to encompass all of AI: an agent is placed in an environment and must learn to behave successfully therein. To keep the chapter manageable, we will concentrate on simple environments and simple agent designs. For the most part, we will assume a fully observable environment, so that the current state is supplied by each percept. On the other hand, we will assume that the agent does not know how the environment works or what its actions do, and we will allow for probabilistic action outcomes. Thus, the agent faces an unknown Markov decision process.[8]

In the end we have three general available branches of constructing systems of machine learning, each one of them is suitable for specified class of problems and consists of own statistical methods and foundations.

# 2 ARTIFICIAL NEURAL NETWORK AND DEEP LEARNING

The most widespread machine learning was applied in the field of computer vision, although now there is still a lot of manual coding. For example, you have to write classifiers such as gradient detection filters so that the program can determine where the object started and where it ended; form definition filters to, say, determine whether the object has eight sides; classifier for recognizing the letters "S-T-O-P." Of all these manually created classifiers, algorithms were created that helped to understand the image and understand that this is a STOP sign ("Movement without stopping is prohibited").

This already gives the best results. However, on a foggy day when the sign is poorly visible, or when it is partially blocked by a tree, the probability of error is high. Therefore, until recently, computer vision and image recognition still could not compete with humans.[10]

Time and correct learning algorithms changed the state of things, resulting in another algorithmic approach, separated from machine learning at an early stage of development, - artificial neural networks. ANN from the point of view of machine learning, neural network is a special case of methods of pattern recognition, discriminant analysis, clustering methods, etc. Neural networks are based on our knowledge of brain biology, namely, on the connections between neurons. But unlike the biological brain, where one neuron can communicate with any other within a certain distance, artificial neural networks have discrete levels, connections and directions of data distribution.

Neural networks are not programmed in the usual sense of the word, they are learned. The possibility of learning is one of the main advantages of neural networks over traditional algorithms. Technically, learning is to find the coefficients of connections between neurons. In the process of learning, the neural network is able to detect complex dependencies between input data and output, and also perform generalization. This means that in case of successful training, the network will be able to return the correct result based on data that was not available in the training sample, as well as incomplete and / or "noisy", partially distorted data.

You can take an image, cut it into pieces and send it to the first level of the neural network. There they are processed by individual neurons and go to the second level of the network, and so on, to the final level, until the result is obtained.[5]

Each neuron assigns a weighting factor to the input data - to what extent they are correct with respect to the task being performed. The final result is determined by the sum of these weighting factors. Remember the example with the sign STOP. Signs of the image of this sign are studied by neurons - its octagonal form, red color, its clear letters, the size of a road sign, the presence or absence of movement. The task of the neural network is to determine whether this is a sign of STOP or not. Here we use the probability vector, a scientific guess based on weighting coefficients. In our example, the system can assume that the image is a STOP sign, a speed limit sign, a flying snake stuck in trees, and so on - and the neural architecture then tells the neural network whether it's right or not, based on the degree of neural network confidence in each Supposition.[10]

But until recently, scientists in the field of AI rather avoided neural networks. They appeared at the dawn of AI, but from the point of view of reason they gave almost nothing. The problem is that even the most basic neural networks required very powerful computations, and this was impractical. However, a small group of researchers led by Jeffrey Hinton of the University of Toronto was engaged in neural networks and eventually parallelized the algorithms for supercomputers, proving the correctness of the idea, but this happened after installing the graphics processors.

Let's return to our sign STOP: chances are great that the system will give out a lot of wrong answers in the learning process. It must be taught. She needs to feed hundreds of thousands, millions of images, until the weights of the input data are tuned so precisely that the correct answer will be given out almost every time - fog or no fog, sun or rain. At this stage, the network also learns to recognize the sign of STOP; or a cat - which Andrew Angie did in 2012 in Google.

The breakthrough of Angie was that he took these neural networks and made them huge, increasing the number of layers and neurons, and then passed through them a huge amount of data to train the system. In his case, these were images of 10 million YouTube videos. Angie added that very depth to the training.[10]

Today, machines that have undergone extensive training, in some cases, give results better than people: from recognizing cats and to signs of cancer in the blood and tumors in MRI images.

Deep learning, in fact, has widely introduced machine learning into practice. It breaks up tasks so that all kinds of machine assistance seem possible. Cars without drivers, the best preventive health care, film recommendations - all this already exists or almost is today.

# 3 STYLE TRANSFER WITH DEEP LEARNING

## 3.1 Introduction

In fine art, especially painting, humans have mastered the skill to create unique visual experiences through composing a complex interplay between the content and style of an image. Thus far the algorithmic basis of this process is unknown and there exists no artificial system with similar capabilities. However, in other key areas of visual perception such as object and face recognition near-human performance was recently demonstrated by a class of biologically inspired vision models called Deep Neural Networks. Here we introduce an artificial system based on a Deep Neural Network that creates artistic images of high perceptual quality. The system uses neural representations to separate and recombine content and style of arbitrary images, providing a neural algorithm for the creation of artistic images. Moreover, in light of the striking similarities between performance-optimised artificial neural networks and biological vision, our work offers a path forward to an algorithmic understanding of how humans create and perceive artistic imagery.

The class of Deep Neural Networks that are most powerful in image processing tasks are called Convolutional Neural Networks. Convolutional Neural Networks consist of layers of small computational units that process visual information hierarchically in a feed-forward manner (Fig 1). Each layer of units can be understood as a collection of image filters, each of which extracts a certain feature from the input image. Thus, the output of a given layer consists of so-called feature maps: differently filtered versions of the input image. [1]

## 3.2 The principle of image generation

When convolutional neural networks learn to recognize objects, they develop an image representation that makes the object information more explicit in the processing hierarchy. Therefore, in the network processing hierarchy, the input image is converted into views that increasingly understand the actual contents of the image, rather than the detailed pixel values. We can directly visualize the information that each layer contains about the input image, by restoring the image from the function maps in this layer. Higher levels in the network capture high-level content in terms of objects and their location in the input image, but do not limit the exact pixel values ​​during recovery.

To get a representation of the style of the input image, a space of stylized features is used to obtain information about the texture. This function space is built on top of the filter responses at each level of the network. It consists of correlations between different filter responses over the spatial extent of the feature cards. Including correlation of features of several layers, we get a stationary scale representation of the input image, which captures its texture information, but not the global device.

A key feature is that the content and style representations in the convolutional neural network are shared. That is, we can manipulate both views independently of each other to create new images.[9]

Images are synthesized by searching for an image that simultaneously corresponds to the representation of the contents of the photograph and the presentation of the style of the corresponding work of art. At the same time, the global look of the original picture is preserved, as are the colors and local structures that make up global landscapes, appear as a cover.

The content and style of the image can not be completely separated. When synthesizing an image that combines the contents of one image with the style of another, there usually does not exist an image that ideally fits both conditions simultaneously. However, the loss function, which we minimize during the synthesis of images, allows us to influence the weights of style and content, so that we can smoothly adjust the emphasis either on restoring content or style. A strong emphasis on style will lead to the fact that the images will correspond to the appearance of the work, in more detail providing a textured version, but it is unlikely to save details of the contents of the image. On the contrary, the emphasis on content, can in detail transfer the contents of the image, but the style of the picture will be presented very weakly and not intrusively. For each pair of source images, you can adjust your balance between content and style to create a visually beautiful image.[6]

## 3.3 Mathematical methodology

To implement the ideas above, the program was based on the pre-engineered VGG-network (Visual Geometry Group), which is a convolutional neural network of 16 layers, which solves the problem of classifying images into a huge variety of classes.

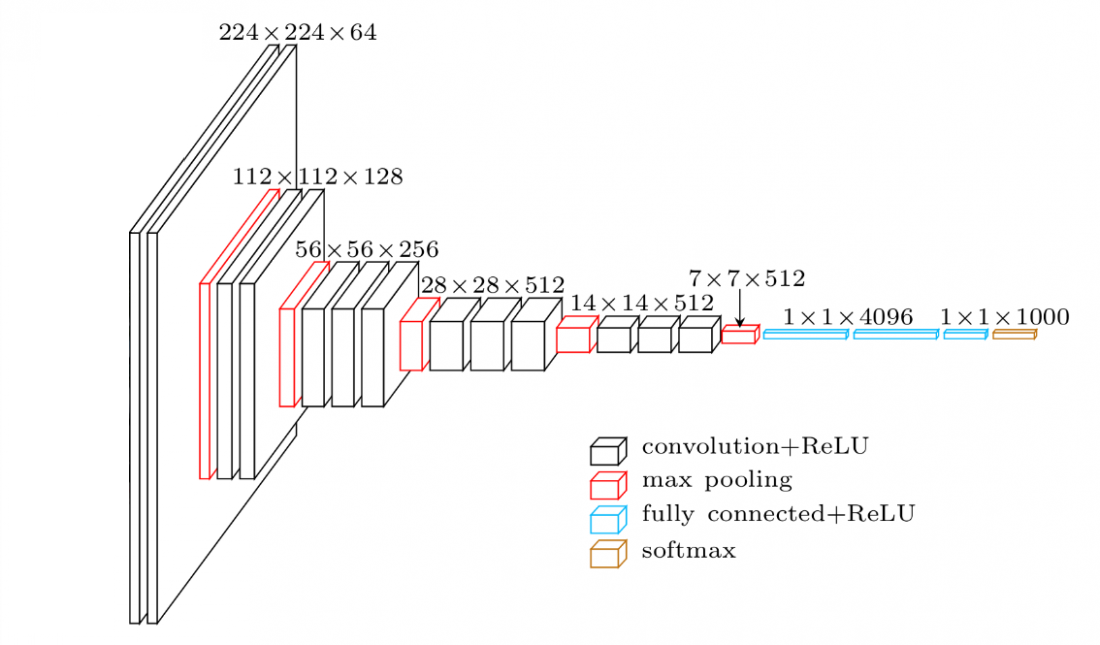


Figure 1 - Structure of the convolutional neural network VGG16

Normally, each layer in the network defines a non-linear filter bank, the complexity of which increases with the layer's position in the network. Therefore, a given input image is encoded in each CNN layer by filter responses to this image. A layer with filters has maps of functions of size , where is equal to the width of the map of objects taken equal to the height. Thus, the answers to the layer can be stored in the matrix , where l is the activation of the -th filter in the position in the layer . To visualize image information that was encoded at different levels of the hierarchy, a gradient descent projected on the white noise image is performed to find another image that corresponds to the characteristic responses of the original image.

Thus, let be the original image, and be the generated image, and and their corresponding representation of the features in layer . The square of the error between the two representations of the signs

(1)

The derivative of this loss by activations in the layer

(2)

from which the gradient with respect to the image can be calculated using standard back propagation of errors. Thus, you can change the original random image until it gives the same response in a specific CNN layer as the original image.

In addition to the CNN responses in each layer of the network, we created a style view that calculates the correlations between the various filter responses. These attribute correlations are given by the Gram matrix , where is the product between the vector map of signs and of layer :

(3)

To create a texture corresponding to the style of this image, gradient descent from the white noise image is used to find another image corresponding to the presentation of the original image style. This is done by minimizing the root-mean-square deviation between the Gram matrix records from the original image and the Gram matrix of the generated image. Let and the original image, and the generated image, and and - their corresponding representation of the style in the layer . The contribution of this layer to the total error

(4)

and general error

(5)

where are the weight coefficients of the contribution of each layer to the total error. The derivative of activations in layer can be calculated analytically:

(6)

Gradients with respect to activations in the lower layers of the network can be easily calculated using standard back propagation of errors.

In order to generate images that blend the contents of a photo with the painting style, the distance of the white noise image from the image content in one layer of the network is minimized and the presentation of the style in a number of layers is minimized. Let image and style. We minimize the loss function

(7)

where and are weight factors for content and style reconstruction, respectively.

Result of this actions is prepared systems the take to input content and style image and produce new one based on them. Unfortunately, shown output is not as good as some of the systems that are on sale, for example DeepArt or Prisma.

# CONCLUSION

This course work was aimed at giving short introduction into science of machine learning, at explaining the main principles of learning computer models from different data, at showing capabilities of artificial neural networks using and at demonstrating real working network based on deep learning architecture.

We can see that the general difference between artificial intelligence and machine learning with neural networks, is that the main power of machine learning hide in statistical methods and searching of initially unknown patterns in given data, and later building a predictions that can be used for analysis and benefit-sharing.

There are three master forms of learning simple models or complicated neural networks, They are supervised, unsupervised and reinforcement learning. They are applied for wide range of analytical, computer vision, recognition and much more various tasks, which we can not even imagine.

Neural networks could be applied in every human activity area, and deep learning can further improve it. It bases on using more convolutional layers in network, than usual, but with extremely increased demand of computing power.

Practical side of this research work, proves that deep neural network is used for style transfer. Resulting program was developed, based on the VGG16 network, transforming the image in accordance with the specified style. The validation on the test images showed that each pair of content-style requires adjustment as change in the weights, since some styles show more influence on the image, with the same scales. Also, tests have shown that the most effective set of style layers is:

['conv1\_1 / conv1\_1', 'conv2\_1 / conv2\_1', 'conv3\_1 / conv3\_1', 'conv4\_1 / conv4\_1', 'conv5\_1 / conv5\_1']

In the research work, the basic ideas of ​​using neural networks for combining content and the style of two images was shown. Result is not very well because of the following reason. Perhaps you need more processing power to perform more optimization iterations with smaller step sizes and higher resolution images. While doing the work, the TensorFlow computational library was used, and calculations were performed on the central processor, because of the inability to use the video card, due to the high demands of the algorithm for RAM, so a small number of iterations (<500) was used, which could not allow the program to be correctly configured and get better results.

# LITERATURE REVIEW

The study of the topic actuality is very high, therefore the range of scientific articles and books are very wide. Books from the Springer Series in Statistics: An Introduction to Statistical Learning: with Applications in R” and " The Elements of Statistical Learning", which describe, in detail, the statistical methods and the learning methods based on them, became the foundation of this research. In this literature, the approach of deep learning covered not very detailed, because of this to the study this field of machine learning more closely were used next sources: " Learning Deep Architectures for AI ", covering the theoretical side, and " Deep Learning: Methods and Applications ", focused on application side. Practically every book about machine learning also tells us about artificial neural networks superficially, but for a deeper mastering, the following sources were observed "An Introduction to Neural Networks" and "Information Theory, Inference, and Learning Algorithms" providing information on the basic arrangement of neural networks and on more complicated varieties. Last three references immerse into deep learning field of neural networks, working with image representation and using of pretrained models to take required outputs from them.

# REFERENCES

1. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge «A Neural Algorithm of Artistic Style», 2015 (Access: <https://pdfs.semanticscholar.org/77c5/12cbb832436e1a35ad434e6bb3d763799763>)
2. Gurney K. “An Introduction to Neural Networks” London, 1997 (Access: <https://www.inf.ed.ac.uk/teaching/courses/nlu/reading/Gurney_et_al.pdf>)
3. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani “An Introduction to Statistical Learning: with Applications in R” (Access: <http://www-bcf.usc.edu/~gareth/ISL/ISLR%20First%20Printing.pdf> )
4. Stuart J. Russell, Peter Norvig “Artificial Intelligence: A Modern Approach”, 3rd edition, Prentice-Hall, 2003 (Access: <https://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&sqi=2&ved=0ahUKEwin0eep9sTTAhVEXSwKHeBVAqoQFgg7MAQ&url=https%3A%2F%2Ffaculty.psau.edu.sa%2Ffiledownload%2Fdoc-7-pdf-a154ffbcec538a4161a406abf62f5b76-original.pdf&usg=AFQjCNE-JSNtyAa-RIi9e4MzACPQwYgfYA&sig2=hzOgpMkM7q1zUbyz0rpjtg>)
5. Deng Li, Yu Dong "Deep Learning: Methods and Applications”, NOW Publishers, 2014 (Access: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/DeepLearning-NowPublishing-Vol7-SIG-039.pdf>)
6. Wei-Han Chang, Ming-Cheng Cheng, Chung-Ming Kuo, Guan-Da Huang «Feature-Oriented Artistic Styles Transfer Based on Effective Texture Synthesis», Ubiquitous International Volume 6, Number 1, 2015 (Access: <http://bit.kuas.edu.tw/~jihmsp/2015/vol6/JIH-MSP-2015-01-002> )
7. Ron Kohavi, Foster Provost "Glossary of terms. Machine Learning” 30, 1998 (Access: <http://ai.stanford.edu/~ronnyk/glossary.html> )
8. David J. C. MacKay. Information Theory, “Inference, and Learning Algorithms” Cambridge: Cambridge University Press, 2003 (Access: <http://www.inference.phy.cam.ac.uk/itprnn/book.pdf> )
9. Julia Patterson, B.A., «Investigation of Style Transfer on Digital Images», University of Dublin, Trinity College, 2012 (Access: <https://www.scss.tcd.ie/postgraduate/msciet/current/Dissertations/1112/Patterson> )
10. Yoshua Bengio “Learning Deep Architectures for AI”, NOW Publishers, 2009 (Access: <http://sanghv.com/download/soft/machine%20learning,%20artificial%20intelligence,%20mathematics%20ebooks/ML/learning%20deep%20architectures%20for%20AI%20%282009%29.pdf>)
11. Trevor Hastie, Robert Tibshirani and Jerome H. Friedman “The Elements of Statistical Learning”, Springer Jan 2013. (Access: <http://statweb.stanford.edu/~tibs/ElemStatLearn/printings/ESLII_print10.pdf> )