Mazurkiewicz Jacek, 0000-0002-7708-907X
Wrocław University of Science and Technology, Wrocław, Poland, Jacek.Mazurkiewicz{at}pwr.edu.pl

Softcomputing system for painting art style verification

Keywords

art style recognition, convolution neural networks, deep learning, image recognition

Abstract

The chapter discusses the foundations for the system to verify and recognize the art style. Such a system seems to be interesting for the first step of the painting fraud identification and following the possible path of the different style influence for the final shape of the masterpiece. The approach presents image recognition using convolutional neural networks. These networks, due to their structure resembling the sight apparatus, and due to their efficiency in the case of two-dimensional data, are very often used for image recognition. Style recognition in art is currently a hot topic in machine learning circles. Courtesy of museums and galleries, there are now many databases available on the web that can be used in scientific work. The classes on which the network was to be tested are styles in history that are associated by the layman with the subject of art. This group includes Renaissance, Baroque, Romanticism, Neo-classicism, Surrealism, Cubism, Art Nouveau, Abstract Expressionism, Pop Art, and Impressionism. These classes were described in the work in terms of parameters that could affect the learning of the neural network. The networks were tested to determine the best parameters for identifying artistic styles. Networks with changing filter values, stride, and pooling parameters, and by selecting various additional layers were tested. The most important parameter was overfitting, which had to be prevented. As a result, networks peaked at 40% in Top1 and 80% in Top3. For smaller data, this result was optimistic for further research in recognizing other parameters, as well as using networks that were previously taught specific characteristics of styles, such as frequently used motifs or colors.

1. Introduction

Humanity has always tried to express itself through creativity. Images have always been the closest due to the simplicity of copying the world. Each person sees it individually as it is understood more psychologically than geometrically, and this gives many ways of expression. Because of them, the styles change throughout history. Through the style of each painting, we can discover a lot about the period in history and the matters which concerned humanity at the time. The investigation of art styles helps catalog databases and understand the culture of the past generations. People are drawn to the past just as much as to the future. Art was always an inspiration both for artistic and non-artistic individuals. Even in Middle Ages, the sculptures from ancient Greece and Rome were noticed and appreciated in the old continent. Then, the Renaissance took even more inspiration for them to create the most magnificent sculptures and paintings of their own. No one would have wanted to copy the old masters. Every reputable artist has a call to be extraordinary and create the art style which later would be recognized as theirs. Nowadays it is the same. Museums are still open and the rich still want to have an original piece from a Baroque master in their home. The times changed, and the world has completely taken over electronic devices are not expressing art in the same way. Recently more and more artists use computers to help themselves in the process of creation and so the storing of art became easier by using modern databases and searching systems.

Safety and Reliability of Systems and Processes, Summer Safety and Reliability Seminar 2023.

DOI: 10.26408/srsp-2023-12.

This process is also applicable to the identification of art. Beforehand using only the knowledge of the past, now it can be simplified by machine learning. Although the networks were introduced in the 1980s (Culurciello, 2017), the process of discovering their methods and usage is still growing. Convolutional neural networks were based on vision processing in a living organism. In the 1960s, it was discovered that each receptive field (single neurons) responds to stimuli (Culurciello, 2017). Their size and location vary systematically across the cortex to form a complete map of visual space. The cortex in each hemisphere represents the contralateral visual field. The first convolutional network appeared in 1998 (Viswanathan, 2017), it had seven layers and was classifying digits from 32x32 pixel images. A few years later, it was modified and applied for medical purposes to detect breast cancer mammograms. In the early twenty-first century, the feed-forward structure was extended to convolutional neural networks. Also, machine learning started using GPU which made impressive progress in the matter of convolutional networks. (Culurciello, 2017) Considering the variety of methods available in machine learning, the number of usages can be countless. Each request containing enough data can be suitable to create a helpful tool. Considering the style of art and usage of such methods, the work of custodians may become easier and more precise, depending on the network they would use.

In this chapter, we want to examine how neural networks can be applied to the recognition of art styles. We believe it could be an interesting subject for custodians, art historians, or scientists. This may help them not only recognize the style with some certainty but also compare and mark the similarities and differences between styles or artists (Singh, 2017; Viswanathan, 2017). The chapter can be extended to help them in authenticating and determining the timeline of paintings. The human mind and eyes can be affected by the power of suggestion while a machine will into consideration only the visible elements. Because of that, the recognition tool can correct mistakes of the past and improve the value of art identification. It can become an irreplaceable assistant for every expert specialized in this field.

We assume it could be also applied in art schools while studying the differences in the history of art and also to challenge the students to create their painting in a certain style and then check if the neural network confirms their attempts. Furthermore, we believe it can be a great tool to introduce to not people not interested in art. Combining technology and culture can be a way to show more people the beauty of art and not treat it as boring.

2. Aims and assumptions

The question we would like to answer in this chapter is how well art styles in paintings could be recognized by convolutional neural networks. In other words, we try to provide the foundations for the system to verify and recognize the art style. Such a system seems to be interesting for the first step of the painting fraud identification and following the possible path of the different style influence for the final shape of the masterpiece. We would like to base the research on the geometric features of the paintings, including gradients and colors.

The goal to achieve is to have comparable results to research papers by others but using a less complicated model (Zaki, 2017). The efficiency of the neural network should get close to 50% for Top1 prediction (Lecountre et al., 2017). The network will be created using Python programming language supported by Keras and TensorFlow machine learning libraries (Pai, 2017; Eclipse, 2018). The input data will be an RGB set containing 100 images per 10 classes. We want to test it with:

- cropped piece of high-resolution painting,
- cropped piece of the normalized image,
- normalized image.

The classes used in the project will be the most popular and recognizable art styles. Through that assumption, the data we should collect is within easy reach and of good quality. For the best results, images were chosen by hand to make the range of different qualities wider and to include the most important pieces of art from given periods. The neural network will be tested with different models and parameters to reach the best identification results.

In this chapter, we would like to describe certain styles and their impact on society at the time of creation as much as in the twenty-first century. Then we will cover the subject of creating datasets and inputs to neural networks. The neural network will be the next part of the chapter, including the experiments and results gathered throughout them. Afterward, we should end it with conclusions and suggestions for future research.

3. Styles in painting

According to Austrian art historian, Ernst Gombrich art style is “any distinctive, and therefore recognizable, way in which an act is performed, or an artifact made or ought to be performed and made” (Gombrich, 1968). Every single piece is capable of style, which is a set of decisions made by the maker. It is unavoidable to have a style because of the impressions of a certain period. The art style is separated by art movements, periods, places, or groups of artists. It shows individual characteristics of art and it is often categorized later in history. A lot of styles are also divided into early, middle, and late stadiums which can be interpreted as slow shaping of the used methods or subjects. Recent history had a big increase in styles which made historians avoid the classification as far as it is possible (Hockney & Gayford, 2016). Nonetheless, the motives and trends are helping humans understand art and place it in a certain period. While discovering history, humanity had plenty of sources to understand our ancestors better. Furthermore, the symbolism and themes of paintings are now common knowledge and are widely known.

3.1. Style recognition

Art identification is a wide process that can take a lot of time. In the process, the art historian compares the given subject with other examples from the alleged artist. Because of the expert knowledge of the historian, it can be examined through its theme, material, and technique and be compared to the artists’ habits. Then the investigation is taken over by photographing the painting in normal, ultraviolet, and infrared light, with different illuminations to reveal irregularities on the surface.

Then the layers of the painting are examined. A painting is made up of four layers: support, ground, paint, and varnish. The materials used are often characteristic of the period of creation. A lot of materials can be re-created, which is the reason why the expert has to be careful in their research. The age of the wood or canvas on which the painting was made (support) can be carried out by dendrochronology or carbon-14 dating. Then the ground layer is measured in used pigments. A lot of them were used in a certain period and then to be replaced. The interesting part of identification is uncovering of underdrawings with infrared reflectography. It can be very helpful because it reveals early sketches or the parts of the concepts that later changed. Many artists have their manner of painting or drawing and that can be a confirmation of authenticity (Hockney & Gayford, 2016). The styles themselves are often compared to the other paintings from the period and other pieces by the same artist. Plenty of painters were involved in more than one style so the exact timing can be crucial to determine the style. Furthermore, the style is not only determined by the technical details such as colors and methods of painting. Often the topic of the art is as important. In many periods, certain society was struggling with problems shared and remembered by the paintings. Symbols and allegories are as important. A lot of styles have certain motives which were duplicated due to their popularity among people. For instance, Romantic paintings have a lot of nature, cliffs, and ships and Renaissance focused on Roman and Greek gods as much as the symbolic representation of Christian religion (Zaki, 2017).

In the experiment in 2012 approximately 1000 paintings by 34 artists were used and let the algorithm analyze similarities in the visual content of the paintings without any human guidance. The computer provided a network of similarities between painters that is largely in agreement with the perception of art historians (Shamir & Tarakhovsky, 2012).

The average person can make the broad identification between modern and classical styles, they have difficulty telling the difference between closely related schools of art such as Early and High Renaissance or Mannerism and Romanticism. The experiment showed that machines can outperform untrained humans.

3.2. Description of used styles

3.2.1. Renaissance

Paintings at this time were done in a highly Christian style. Following the Middle Ages’ subjects of sacred paintings, art still focused on the scenes from Bible, especially the child and mother motives. The rediscovery of Greek and Roman masters inspired Renaissance artists to take other gods as subjects as well. Other subjects used at that time are portraits of the nobility which were made explicitly to earn a salary from painting. Only the wealthy could afford a painting of themselves and the skilled painter was usually under their protection as an artist for such service or at least lavishly rewarded. Paintings often included so-called continuous narrative – it had more than one scene on a single canvas. The color palette is usually bright, pastel even, with very slow grading. The shadows were more naturalistic, and the perspective became linear (called Renaissance by historians). Paintings started looking more realistic, with a lot of copying from nature included. Artists started to use models and shape faces on the canvas as individuals and painting objects more realistically. The number of details increased, and surfaces were painted with precision on every fold of cloth or every pile on the floor (Zaki, 2017; Hockney & Gayford, 2016) (Fig. 1).

Figure 1. Raphael – *Madonna in the Meadow*.

3.2.2. Baroque

Subjects were covering very different subjects from landscapes through still life to dramatic and emotional moments. The most important color here was black. Baroque was focusing on showing chiaroscuro most diversely. Plenty of pieces are showing bright objects on black or very dark backgrounds. Especially saints who were still very popular subjects, are shown almost white to make them look vivid in comparison to the dark background. This was supposed to signal their purity and sanctity. People were also included in dramatic situations, often concerning death. Featuring the still life had similar colors. Most of these were supposed to bring thoughts about death and life to the audience. Skulls or dead people or animals were often pictured, right next to fruit or cut flowers (Zaki, 2017) (Fig. 2).

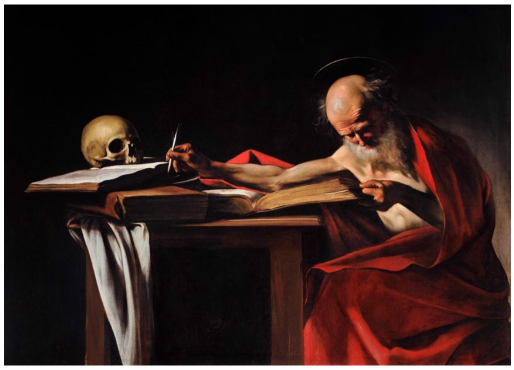

Figure 2. Caravaggio – *Saint Jerome Writing*.

3.2.3. Neoclassicism

The opposite of the baroque movement, the time of this period was similar. The subjects of them were focused mostly on classical culture, Greek and Roman, due to the rediscovery of ancient art. In that period, the ruins of Herculaneum and Pompei were excavated which resulted in bigger interest in ancient Rome. The art itself was trying to copy the old masters by the representation of perfected human bodies in mythology and scenes from the ancient era. The chiaroscuro was balanced throughout the whole painting and the colors were taken from the entire palette (Zaki, 2017) (Fig. 3).

Figure 3. Jacques-Louis David – *The Oath of Horatii*.

3.2.4. Romanticism

This period was a response to Industrial Revolution which not only produced major changes in industry but also in philosophy and culture. The movement was to show the opposite of science and reason. Romantic art was an emotional comeback to rural landscapes. The paintings often portrayed nature or emotional events. The XIX century was full of revolutionary fights and the artists painted their representation of these. Emotions were not only supposed to be included in the paintings but also induce them. For that reason, many paintings had scary images or elevated moments. Colors are also important. Although there was more than one idea, the usage should fit the emotional agenda as much as possible. This is why grotesque, scary figures were the dark and important person of a revolution was as bright as possible. The brushstrokes were to increase the emotions that accompanied the artist while creating his masterpiece. The paintings portrayed powerful events, they were visible and showed as much emotional behavior as possible (Zaki, 2017) (Fig. 4).

Figure 4. Caspar David Friedrich – *Wanderer Above the Sea of Fog*.

3.2.5. Impressionism

It was developed mainly by Claude Monet and moved by his colleagues in the French Painters' Society. Due to the industrial progress, they could paint outside which lead to a lot of landscapes captured. This period was focused on stains and splotches not only in visual arts but also in music and literature. If looked at from proximity, the paintings seemed like a complete mess, but the mess made sense when the eye connected the stains to the whole picture. The outlines were never clear, all the colors were mixed to get colors usually bright and vibrant, mostly marking the existence of an object rather than expressing it. Subjects were mainly about capturing a moment, something that will never repeat. There were a lot of sunsets or sunrises as well as some performances, especially ballet. If the people were captured, they were usually the middle class (Zaki, 2017) (Fig. 5).

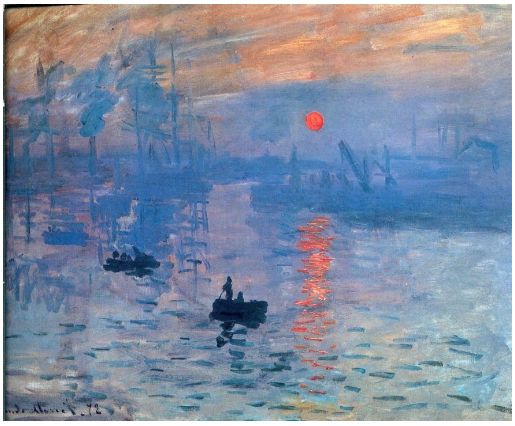

Figure 5. Claude Monet – *Impression, Sunrise*.

3.2.6. Art nouveau

Smaller art movements developed under the fame of impressionism. It had noticeable contours and little shading. The most common reason for the paintings was to be decorative rather than to be hung on walls. This is why it had to be easy to reproduce them and catch the attention of people. Letters were also often used which was a rather new move in the visual art. Portrait shape was more commonly used so the painting would fit its purpose as an image in a book or newspaper. The colors were light and softly shaded. The geometric ornaments were also included, often connected to nature. A lot of painters from art nouveau were also connected to other movements such as symbolism. It was a gap between fine art and applied art (Zaki, 2017) (Fig. 6).

3.2.7. Surrealism

Most famous in Catalonia (Spain) with its precursor Salvador Dali, surrealism was a description of imagination. The objects were never something seen in reality but rather a dream. Focusing on that, strange shapes and creatures could be seen in each painting. Surrealists believed that they should extend the human mind rather than copy what already was in nature. Their art was supposed to surprise them.

Figure 6. Alphonse Mucha – *Gismonda*.

The movement was also strongly experimental and tried to join unconsciousness with consciousness. The huge diversity noticeable between surrealists can imply that there were no rules applying to the movement despite two manifestos that were supposed to shape the rules. The most important however was the freedom in the art and abstract thinking outside the given reality.
Techniques used by them were rather to evoke a psychic response than to get a certain result. One of them, frottage, was described as rubbing with graphite over wood or other grained substances. The other – grattage – is scrapping the canvas (Zaki, 2017) (Fig. 7).

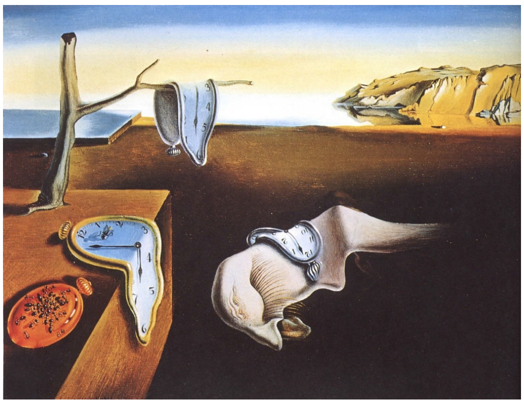

Figure 7. Salvador Dali – *The Persistence of Memory*.

3.2.8. Cubism

The style emphasized a flat, two-dimensional plane, rejecting the rules of perspective or chiaroscuro. Cubists did not want to copy nature but present new realities in paintings as fragmented objects. The color scale was nearly monochromatic not to distract the viewer from the artist's primary interest – the form. A lot of representations were combined with letters to make the painting more of a puzzle. In the later phase, color became more interesting as an expression of surface or texture. The subject could have been anything, from human figures to landscapes, but instead of painting them from one perspective, cubists composed together fragments from different vantage points, watching the object from every possible angle. Furthermore, they also took into consideration barriers of space and time so movement or change of state can be noticed in the painting. This style differs from any other from the past and cannot be mistaken. Understanding it takes time and patience to connect the fragments to the actual object (Zaki, 2017) (Fig. 8).

Figure 8. Pablo Picasso – *Seated Woman*.

3.2.9. Pop art

Art became opposed to fine art by taking inspiration from mass culture such as advertisements or comic books. Often using irony to emphasize the banal and kitschy parts of culture, pop art was trying to point out bad decisions, mistakes, and ugliness. A lot more artists decided to reproduce their pieces on a wider scale by using their images or designs in the industry as labels or logos. The important part was also the ability to copy their images or the ability to add different colors to the same design. Mostly characterized by uniform backgrounds with just as flat objects, pop art was supposed to be visible and clear in reception. Although the subjects varied as much as possible and the technique could be different in every country, it was always simple without any shading, using as few colors as possible, and often used in any media possible, including music albums or soup cans (Zaki, 2017) (Fig. 9).

Figure 9. Andy Warhol – *Big Campbell Soup Can*.

3.2.10. Abstract expressionism

After World War II, artists had to get creative to give the message to the world. Especially in the United States, where censorship was very rough, the idea of abstract painting was supposed to get through censors but be understandable to insiders. Expressionism was not indicated to be similar in style but rather to express the message and be interpreted through elements common to certain artists. The movement came from surrealism, futurism, and cubism but was modeled differently in each country affected by it. It was a very political movement through art. It had no certain characteristics to follow but one: it had to make the viewer wonder. No matter whether it had only one line on the canvas or a lot of different paint stains, it had to give powerful meaning rather than be aesthetically pretty (Zaki, 2017) (Fig. 10).

Figure 10. Jackson Pollock – *Number 3*.

* * *

The idea behind chosen styles was to use ones that are popular among less expert audiences. We wanted them to be commonly known so random people can distinguish the styles with small errors or at least tell which ones belong to the same group. Nonetheless, there are some strong similarities between a few of these. Baroque and High Renaissance might be completely mixed up by inexperienced eyes. Cubism, on the other hand, started from the Surrealism movement and there is plenty of identical features.

We would like to create a tool rather than be useful in the hands of laymen and establish the connection between technology and art. We wanted to make sure that by receiving the same information – only the canvas and a minor style description – a person can determine the difference in style as much as a machine. For an expert identification, a person would require more knowledge about the given art such as date of origin or artist name and his/her history. If this information was given to the machine, we believe it would have a much better performance than a human being. Furthermore, to create a more advanced art recognition tool using more classes, the machine would require more examples and it would increase the complexity of the system.

4. State of art

In recent years, specialists who identify art styles received a lot of help from machines which could have helped them manipulate the painting in ways such as enlarging certain pieces or determining the age of the canvas (Bar et al., 2015). Although machines were often used to ease the work of historians, there is no record of recognizing art style using technologies other than neural networks. Due to their biological accuracy, they are crucial to the future of artificial intelligence and the recognition of data from a human perspective like images or sounds. Neural network art recognition is a procedure that surfaced a few years ago (Blessing & Wen, 2018). It happened because of the decision of a few art galleries which made their databases available to the public in digital form. Museums became more transparent, sharing their collections in virtual galleries and agreeing on using them for researchers. Recently, it can be found that people started using the data to recognize artists, the age of paintings, elements on the canvas, and so on. It is a great opportunity for any kind of art-related institution to have a machine recognize elements of a masterpiece and help them determine its details. Furthermore, researching art by machines is a very broad subject due to its reasons and manners. Google Arts and Culture shared a tool that searches for a face similar to a photograph given as input. A lot of people used it to find their twins saved in some less-known paintings (Chen, 2018). One of the most popular usages of convolutional neural networks is turning a photograph into a painting in a certain style. This way, people can create filters that turn their photographs into Van Gogh's lost masterpiece. It is done by processing an image through a neural network learned to affect the input in a specific way. The network is trained with paintings in the given style which formats the expected result. Nowadays, there are applications, especially mobile, that create filters transferring all kinds of artistic works. There are a lot of research papers about machine learning with convolution (Kurenkov, 2015). The research includes subjects similar to mine. We have found plenty of papers acknowledging the identification of art styles. Each of them had a different approach, different data sets, and results, depending on the details of the assumptions. Lecoutre, Negrevergne, and Yger described how such networks (Lecountre et al., 2017). They write about the accuracy they tried to improve by using popular neural networks used for art recognition. They used the whole WikiArt database with class distribution being not equal. Then they used AlexNet and ResNet which are convolutional neural networks to get the results. The ResNet has 53 convolutional layers which happen to work well with the given data but could overfit due to its size. Other network is smaller but does not receive as good results (Lecountre et al., 2017). Another research done to create a new way to categorize paintings used the SVM algorithm to pre-train the network and to create binary classifiers. Algorithms such as SMV, Naïve Bayes, and kNN were in use to choose five neural networks to classify the images. The Matlab framework VLFEAT was used. The precision they reached was set around 40% in different methods and tests (Bar et al., 2015). Sets were rather small – 200 paintings per artist, but compared to sources it is not easy to find more paintings by the same artist. The problem of the number of data and the overfitting was noticed. For different approaches, data in different batch sizes were tested, noticing that the more was given, the more problems the network has. In the two-class problem tested, over 90% accuracy was reached, and then, in the seven classes – over 80% which is impressive considering the diversity between each input (Blessing & Wen, 2018). Deep learning is commonly used for art recognition purposes. It is still being improved, but the engines already available are enough to get satisfactory results (Blessing & Wen, 2018). There is more than one approach to determining styles though we believe the most promising one is still a convolutional neural network. This approach receives the highest results and can be manipulated for a variety of purposes. Python seems to be a common tool for such practices. Convolutions seem to be more efficient in our assumptions and the dataset taken under consideration. These networks receive much higher percentage values in such experiments It will be highly effective in our research (Bar et al., 2015).

5. Dataset

5.1. Inputs preparation

People recognize paintings by their age, the objects or the entire subject of the painting, and the noticeable differences in the artist's style. Machine learning, however, does not focus on the subjects nor their connection to a certain period. Its main source of information is completely technical. It can either focus on colors and gradients or search for certain objects. We wanted our network to learn the wide idea of the style using as little a quantity of paintings as possible. For later research, the database should be expanded to receive the highest possible percentage of recognition (Pai, 2017; Kurenkov, 2015).

For this purpose, we decided to choose a dataset of 1000 paintings from 10 classes. Each class got an equal value of 100 paintings, chosen by hand to increase the diversity in the set. Through that, our dataset did not have more than 10% paintings by the same artist, the motives used in certain periods are widely discovered and not focused on only one. We also value the quality of the images. To test one of our assumptions, we needed paintings with as many details as possible. The dataset used for that purpose was wikiart.org which has a major database divided into different art styles. The classes were named from 0 to 9 in such manner: 0 – Abstract Expressionism, 1 – High Renaissance, 2 – Baroque, 3 – Romanticism, 4 – Impressionism, 5 – Art Nouveau, 6 – Surrealism, 7 – Cubism, 8 – Neoclassicism, 9 – Pop Art.

Apart from the style identification, we also used the name of the artist in the image to help in gathering the data. Through that, we were able to quickly check if the painting was already in the set and if one artist is not used too many times. The data was then split to train and test sets. The default value was set to 75 images for training and 25 images for testing images per class however it is expected to test other splits in the experiment section (90-training / 10-testing and 70-training / 30-testing). This way it was possible to check the practical and functional quality of the system. We used the train_test_split function from the Sklearn library, which divided the data into training and testing batches. We set the percentage of tests to 25%. The random state was set to 42, this parameter was used to make sure each run of the network will have equal outputs. It is a great addition especially when the network is trained on more than one machine. This way the division between the training and testing inputs will always be the same. Another parameter was a shuffle. Shuffling was needed to make sure that the network will not learn one style at a time. The results could be affected by it and it is a normal process to shuffle the data before learning (Pedregosa et al., 2016).

5.2. Preprocessing and data augmentation

Our experiments were focused on two different approaches. The first idea was to cut fragments of a painting and process them through the network without normalization. Because of that, we tried to include only high-resolution images which had the most data about the movement of a brush and the artist's painting methods. This was supposed to separate the styles where the brushstrokes were chaotic and vast from the small and detailed ones. For instance, Impressionism’s main rule was to create paint stains and use plenty of paint. On the contrary, artists in the Renaissance believed in balance and detail so their movements were barely noticeable. Unfortunately, it did not give us satisfying results. Due to not consistent sizes and a lot of noticeable damage on painting (cracks seen on scanned canvas, the result of aging), We had to give up that idea and turn to the more common way of creating inputs. To gather less damaged data, we decided to normalize them by resizing them. The small helped us pull out the information and decrease the impact of noise. One of the most common practices is creating convolutional network input of size 100x100 pixels – Figure 11.

For the first tests, we resized the images to 200 pixels in smaller dimensions and then cropped four fragments 100x100. This way dataset of 100 per class was enlarged four times. This might or might not improve the behavior of the network.

The second process was to simply cut the square out of the painting. For that reason, We used the same algorithm but without cropping the four parts. This set was supposed to tell whether parts of paintings and bigger inputs set is working better or worse to the whole image but with four times smaller set. While training the network, each square was inputted with a label of a class it belonged to. After setting the images to the input matrix, we used Keras’s method ImageDataGenerator to generate more data by rotating the input images. ImageDataGenerator class is used to generate batches of tensor image data with real-time data augmentation. The class uses augmentation methods to help with pre-processing the data right before learning the network. (Chollet, 2015) The class has a lot of different kinds of methods which makes Keras a very helpful tool in designing convolutional neural networks. It is common knowledge that the more data is available, the better-quality results we can expect. This is why augmentation is useful in the process. It is possible, however, that processing the same data too many times may receive different results, so it has to be balanced.

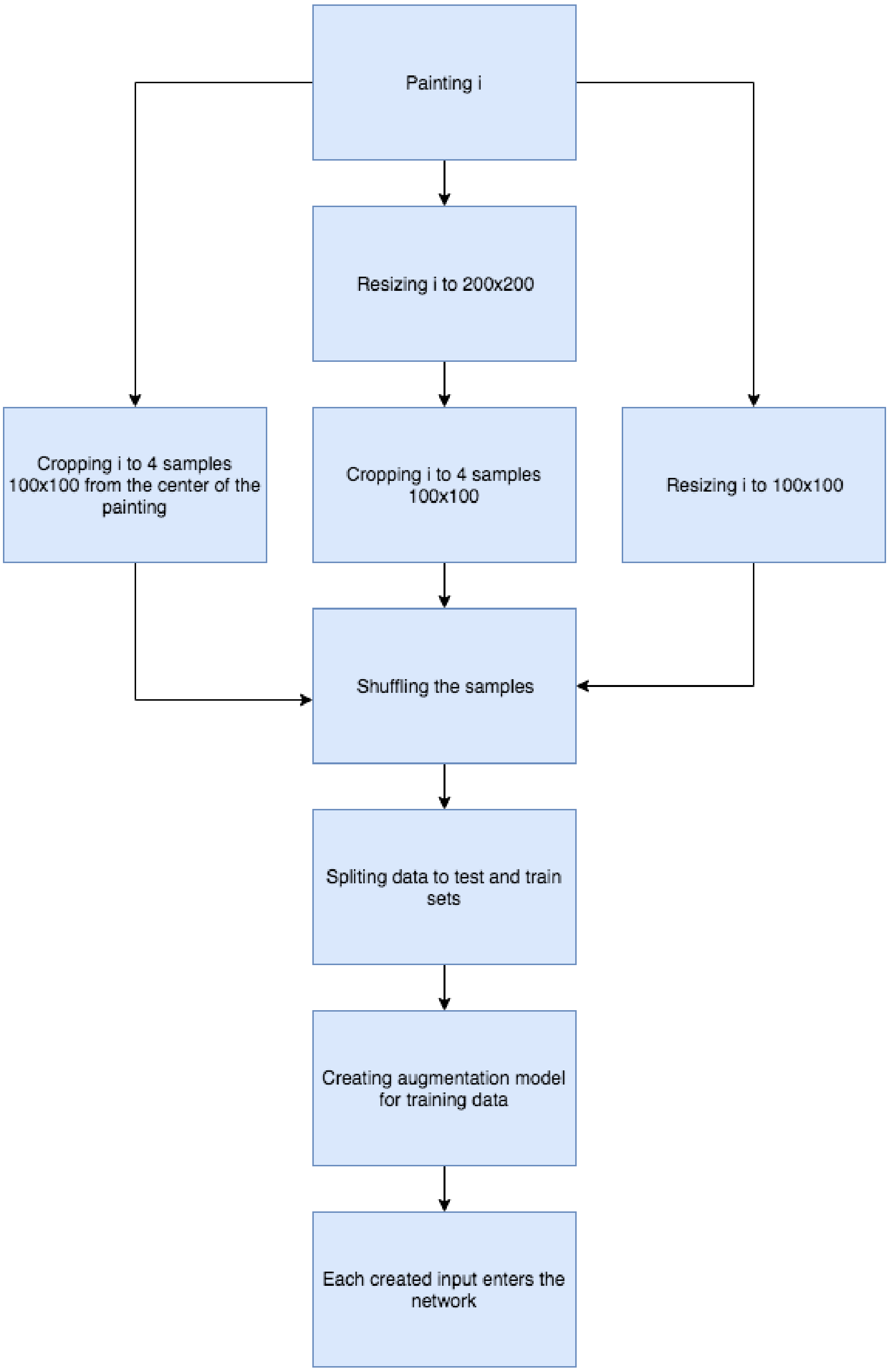

Figure 11. Preprocessing steps block diagram.

6. Convolution neural network

6.1. Overview

Convolution itself is a mathematical operator, which is often used within signal processing to simplify equations. It processes two functions and modifies them. In the matter of neural networks, it processes the network's inputs. The image usually is a three-dimensional matrix composed of three sets of width and height of the image, each of them of different color in RGB model. The input is processed by the kernel. It is also known as a mask in image processing and it is used to alter the image. It can blur, sharpen the image, or be used to detect edges, etc. In CNN, it is used as a filter that creates the map of features (Hearty, 2016). The structure of CNN is common for many popular neural networks. Like any other neural network, it consists of neurons that have learnable weights and biases. Its shape, however, is an acyclic graph with fewer and fewer nodes, where the output of one layer is connected to the next one. Although, it is common for some networks such as multi-layered perceptron, the rest of the convolutional structure is focused on images (now also on sounds) (Karpathy & Stanford, 2018). Convolutional Neural Network is made up of identical neurons which is not common in regular neural networks. All neurons have equal parameters and weights. It reduces the number of parameters controlled by the network which makes the network more efficient. The connection between nodes is limited to local connection patterns. In other words, the inputs of each node connect only with neighboring receptors. In images, it means spatial neighboring. Layers can be assembled as a representation of data, the lower layer, the more abstract the features. If enough layers are connected they can cover the whole image. Weights are shared by all the nodes of each layer. It should prevent learning through nodes of the unpredictable set of local parameters. Filters in the convolutional layer work in one set and process input data together. The standard training is done with the backpropagation algorithm, which is a common procedure in neural networks. The goal of backpropagation is to minimalize the cost function, also known as the error function. To do that, it goes backward and changes the weights of the neurons to fit the model more precisely (Nielsen, 2015). See – Figure 12. Repeating the training is crucial to learn each kernel. One epoch equals one iteration through all the inputs. Meaning, the more epochs, the more times the inputs are passed through the network. The convolutional layer's parameters consist of a set of learnable filters. Every filter has four parameters: size, depth, stride, and pooling. During the forward pass, the network convolves each filter across the width and height of the input volume and computes dot products between the entries of the filter and the input at any position. As it convolves the filter over the width and height of the input volume we will produce a 2-dimensional activation map that gives the responses of that filter at every spatial position. The network will learn filters that activate when they see some type of visual feature such as an edge of some orientation or a blotch of some color on the first layer. Apart from perceptron (fully connected), every layer is partially connected, and each layer uses many feature maps. This helps create an expanded set of filters.

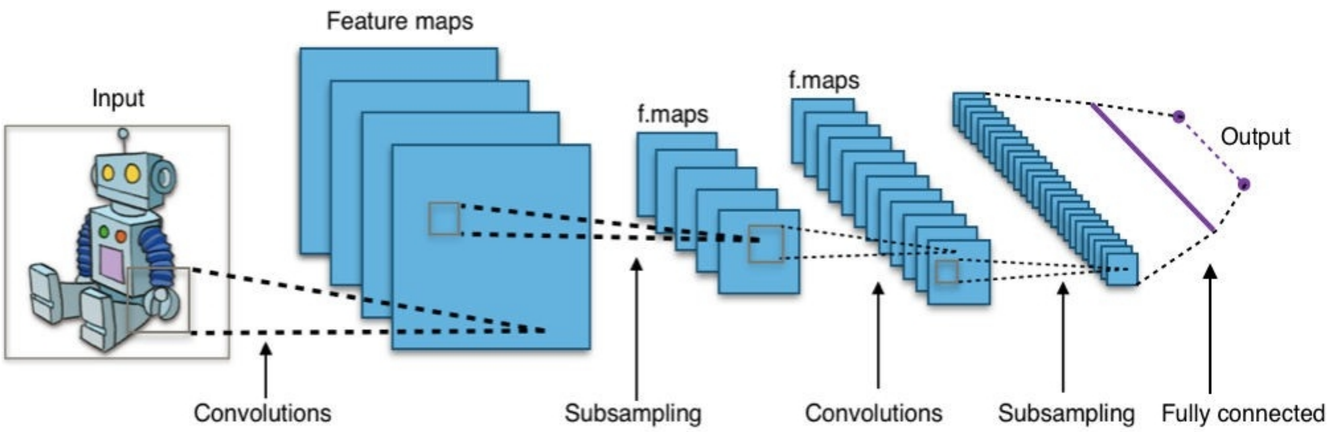

Figure 12. Typical CNN structure.

The depth is the number of nodes in the layer that are connected with the same output. If the depth parameter is too low, it will achieve poor results, because the network will not have enough channels to describe input data. On the other hand, too high a depth parameter will hurt the network's efficiency. Stride describes the distance between the centers of neighboring filters. The lower the number, the more possibly the filters will overlap. If it equals 1, it will mean that each pixel will be taken under consideration separately as a center of the filter. For bigger stride values, the overlapping is smaller and output data is reduced. It should be balanced and personalized for each network since it has a different impact on each dataset. Pooling might seem avoidable, but it has an important purpose. It is the parameter responsible for supplementing pixels surrounding the input image with zeros. It usually occurs when filters reach the edge of the image. For getting better results, it is usually better to focus on the values of the parameters rather than trying to fit the image. If it was the case, all parameters would have to be calculated so they do not reach outside the input data. The other purpose of pooling is to reduce the size of the data. If more than one line of pixels is turned to zeros, then the data is taken as a smaller percentage of the whole. Pooling layers have three types of operations: max-pooling, mean-pooling, and sum-pooling. Each of these is required in different circumstances and should be chosen wisely. They are used on feature maps and they aggregate values into a single one. Without this manipulation, the number of features would be enormous.

Choosing pooling and stride parameters has to be done thoughtfully and not overlooked, so the network is focusing on the information we care about. If these parameters were set randomly, it might happen that the network would not use the data at all. There are some different ways of determining the parameters. Instead of standard neuron output described by the function called tanh (sigmoid), the network used Rectified Linear Units (ReLU). They are proven to train faster than the standard output function and also enable to train of the network on bigger datasets. It also makes the pre-learning unnecessary for CNN. But ReLU is not a flawless system. It can make the neurons not activate throughout the whole learning process (Karpathy, Stanford, 2018).

6.2. Practical usage

Due to minor complexity and a lot of libraries, Python became the most popular language used in machine learning. This is also the reason why it was chosen in this research. It is also way faster than other environments that could have been used for this task. For instance, MATLAB is a powerful tool, but its processing takes very long, and it cannot be used for free in development. In the neural networks agenda, there is a handful of frameworks for Python. For computing, we decided to use TensorFlow because of its simplicity and frequent updates. The framework is a powerful engine created by Google. It is used by many famous companies such as Twitter, eBay, Nvidia, Uber, and even Coca-Cola. TensorFlow enables computations on any processor, including graphics, and has a lot of support in machine learning of all kinds. And most importantly – it is open-source (Abadi et al., 2015). Above the engine, we decided to use Keras as a front-end framework. Keras is a simple library that uses the power of TensorFlow and creates human-understanding code. With this tool, we implemented models which later were used in the neural network. The above libraries were the main source of computing, but the program consisted also of Numpy and Matplotlib packages which are used for computing and drawing diagrams in Python. Numpy is often used for linear algebra and *N*-dimensional arrays, which is highly necessary for machine learning. Matplotlib is a tool for the results. With it, we could plot the results on a diagram to watch the behavior of built networks and make decisions about why should be changed (Chollet, 2015).

6.3. Approach

We decided to use the available Keras methods to change the parameters of the network. It refers to:

- filter size,
- pooling size,
- stride size,
- number of convolutional layers,
- number of hidden layers,
- number of epochs,
- activation type.

To create a correct model, the network requires information on what size of input it should expect:

- height = 100,
- width = 100,
- depth = 3.

This was passed in parameters to the Kears input_shape variable. The Sequential-Keras function is an initiation of the model and means that it has a linear stack of layers. After that function, we added the layers and the parameters to the model by using the Add-Keras method. We had to determine what type of layers and parameters we need for our network beforehand to test them (Chollet, 2015). We used two different types of layers: convolutional and regular called Dense (Krizvsky et al., 2018). Flatten method was also used to quite literally flatten the output. The convolutional layer's output was three dimensional and to use them as inputs in the regular layer they had to be set in one dimension. To do that they were simply turned into a vector that had a length equal to the multiplication of width, height, and depth of the output. Keras offers a lot of activation functions.

The ones used in our network were tested to reach the best effects and it appears the ReLU function worked the most efficiently. However, for the last layer, softmax activation was used. It is often used to work with multiple classes and gives the probability of belonging to each class which sums to one. The MaxPooling2D method was used to fill the input with zeros if filters were placed further than the actual data. This also determined the number of stride parameters (distance between filter centers) (Chollet, 2015).

7. Experiments

The next step was designing the right network to process data. Paintings, even ones from the same class, do not have much in common, so there was a lot of knowledge involved to improve the parameters of the network to fit our data. In Table 1 all datasets used in learning are presented. Dataset 3 was checked to determine how the network will behave with bigger images. The rest were described earlier in this chapter.

Table 1. Datasets used in experiments

Dataset	Classes to identify	Number of inputs	Image size	Additional information
1	10	4000	100	Image normalized to 200 pixels and cropped to 4 pieces
2		1000	100	Image normalized to 100 pixels
3		1000	250	Image normalized to 250 pixels

Through this approach, we had convolutional layers modeled and we could test their parameters by simply changing numbers in the code. Since ReLU activation proved to be the best, we decided not to change this parameter (as explained in the previous section). The experiments were mainly focused on adjusting the parameters of the network, especially the number of filters, their size, size of stride, and pooling. To improve a promising network, we was testing different numbers of dense and convolutional layers, then checked how it will work with more epochs. Epochs, however, were the least to change since the trend could be predicted from smaller numbers, too.
The first experiments were done to establish how complex the network should be to receive the best results. First tests (shown in Table 2) were done as classic network examples and then the network was extended to more convolutional and dense layers. Hidden layers were very highly useful to balance the network's results. The outputs were based on less amplitude and were closer to one another. The number of epochs tested was mainly 100 and 200. Small impact and easy to predict result in smaller numbers made me decide on using only numbers in intervals bigger than 50. The examples in one epoch were determined as in the dataset in Table 1, however, the proportion between the test and train data would vary. It was 25% of test data and 75% of train data as starting value but then we tested more combinations to determine which approach is better.

Table 2. First three examples of tested networks' convolutional layers

Network no	Conv layer no	Number of filters	Filter size	Pooling size	Stride size
1	1	32	2	2	2
	2	64	2	2	2
	3	128	2	2	2
2	1	32	2	2	2
	2	32	2	2	2
	3	64	2	2	2
3	1	32	2	2	2
	2	32	2	2	2
	3	64	2	2	2
	4	128	2	2	2

The overfitting problem was the most important to overcome in our research. A lot of first tests had incredibly large overfitting (Fig. 13) and we wanted to receive results as little affected by it as possible (Table 3).

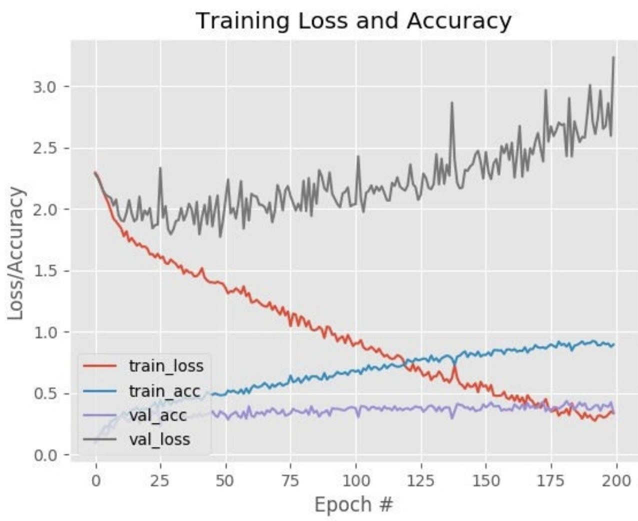

Figure 13. Results – network 1 (Table 2).

Table 3. First three experiments – details

Exp no	Dataset no	Network no	Number of epochs	Final percentage	Additional information
1	2	1	200	35%	Overfit
2	2	2	200	35%	Overfit
3	2	3	100	36%	Overfit

After a couple more adjustments, we were able to achieve this goal (Fig. 14). We took only the most promising networks and improved parameters to receive high results.

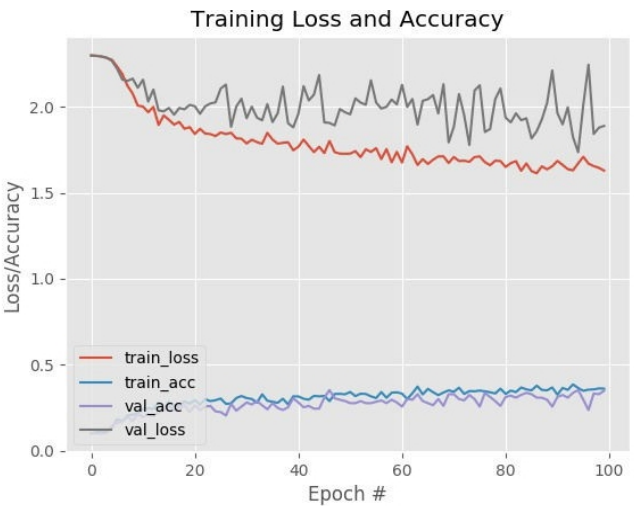

Figure 14. The best results – network 1 (Table 2).

8. Results

8.1. Full paintings input

Full painting inputs gave results with only 10% of overfitting. The same network settings trained by pieces of the painting received results with 70% of correct answers but with 20% of overfitting. Even network was set explicitly for that dataset, the overfit did not disappear. Since it was our main goal to reduce overfitting, after a couple of tests it completely disqualified the cropped database.
In Figures 15–17 (where: grey – accuracy of the network, red – loss function, blue – training set accuracy, purple – testing set accuracy, yellow – training set Top3 accuracy, green – testing set Top3 accuracy), the fluctuations are much bigger in the full painting network but the angle of the overfitting (red) function is much smaller.

8.2. Top1 and Top3 best results

Our networks oscillated in 30–40% interval Top1 and 70–80% intervals in the Top3 results. The best network reached 40% Top1 accuracy without highly overfitting. The Top3 tests were done on a smaller number of examples however their percentage was accordingly similar to Top1 results.

Figure 15. Network taught with full paintings.

Figure 16. Network taught with pieces of paintings.

Figure 17. Best Top1 network with 40% accuracy.

In the chart presented in Figure 17, this trend can be noticed, especially in the first five experiments.

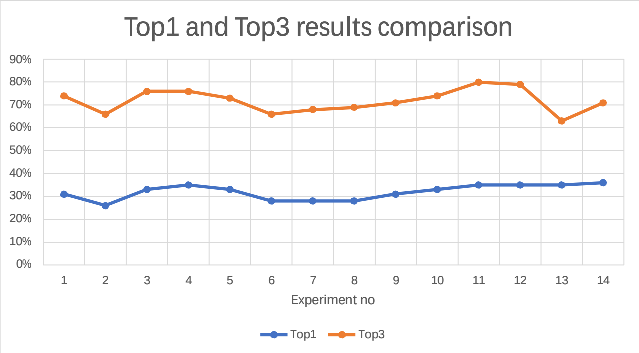

Figure 18. Top1 and Top3 results comparison.

8.3. Max pooling

Max pooling is used to down-sample a network representation. Through that, the nodes which were not activated or with values that do not give any meaning to the network are omitted. This is one of the ways to avoid overfitting and not deal with large networks.
Most of our tests had a max pooling function right after each convolutional layer. To test their importance in image recognition, we decided to learn a network without one or more of these functions. we believe the tests without one or two max-pooling functions can be promising since the results did not vary much from the network with max-pooling functions after each convolutional layer. However, they are important in image recognition since without them the network is completely damaged (Fig. 19–20).

Figure 19. Network without one MaxPooling function.

Figure 20. Network without three MaxPooling functions.

8.4. Dropout function usage

Dropout is yet another function in Keras which supposedly should decrease overfit. This is also often used in other architectures of neural networks. We discovered it while experimenting and it was very promising, especially when combined with MaxPooling functions. With the right dropout values, they may have a big impact on the success of the network without a big overfit value.
In Figures 21–22, we examined the same network with 0.1 and 0.2 valued dropouts. The values mean that accordingly 10% and 20% of the previous layer's results were deleted to make the network more flexible for different values.
The 20% dropout seemed to have better results and a smaller overfit angle. Top1 (purple line) values are growing but the Top3 (green line) is possible to get further from its trained value (yellow line).

Figure 21. Network with dropout equal to 0.1.

Figure 22. Network with dropout equal to 0.2.

8.5. Bigger input

We have also tested our networks for datasets with bigger pictures. The networks used for 100x100 samples did not have good results for them, but after the application of changes for larger sizes, the results were similar (Fig. 23). This can imply that networks will work with similar results no matter whether the input size would grow.

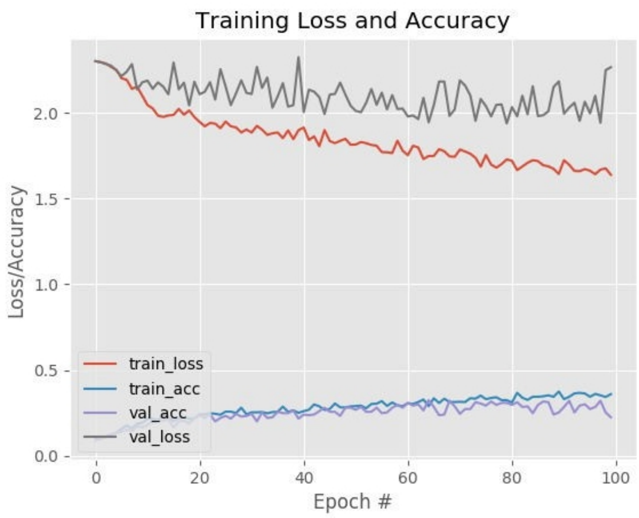

Figure 23. Diagram for dataset 3 (Table 1) with best-tested network.

9. Conclusion

The chapter presents the foundations for the system to verify and recognize the art style. The system seems to be interesting for the first step of the painting fraud identification and following the possible path of the different style influence for the final shape of the masterpiece. The image recognition was solved using convolutional neural networks. These networks, due to their structure resembling the sight apparatus, and due to their efficiency in the case of two-dimensional data, are very often used for image recognition. In many cases, they turn out to be the best possible solution. One of the largest image databases is wikiart.org, which was used as a database for work. The base was selected manually to obtain the highest quality and variety of images. Each of the ten classes was assigned 100 examples. The classes on which the network was to be tested are styles in history that are associated by the layman with the subject of art. This group includes Renaissance, Baroque, Romanticism, Neoclassicism, Surrealism, Cubism, Art Nouveau, Abstract Expressionism, Pop Art, and Impressionism. These classes were described in the work in terms of parameters that could affect the learning of the neural network. Then, data preprocessing and augmentation were presented. The data has been normalized to the values that the network accepts and the size of the inputs that the network will accept has been determined. By default, it was set to 100x100. In this way, the network was able to train on two types of training data: whole normalized images and images normalized to double value and cut into four parts. The networks that were used consisted of up to six convolutional layers with parameters selected for the best results, as well as up to five hidden layers. The networks were tested to determine the best parameters for identifying artistic styles. Networks with changing filter values, stride, and pooling parameters, and by selecting various additional layers, e.g. dropout, were tested. The most important parameter was overfitting, which had to be prevented. Overfitting in a network with a small amount of training data and a large range of variance between samples can have a very negative impact on the network, so it was important to keep the network from this impact.

Throughout the experimentation period, we noticed that hidden layers were for the reduction of spikes on the chart, and convolutional layers were for increasing the recognition. If these were balanced, there was a possibility of a network learned with a good percentage and small overfit. Nonetheless, in the case of too many layers, this may turn the other way. Compared to other research materials, we believe we reached stable results. In Rasta's research (Lecountre et al., 2017), the highest results were below 50% in Top1 and 80% in

Top3. Our network reached similar results using a smaller number of layers and smaller input sizes (Table 4).

Table 4. Results (Lecountre et al., 2017)

Architecture	Top1	Top3	Top5
Pre-trained AlexNet (6^{th} layer features)	0.378	0.627	0.733
Pre-trained ResNet50 accuracy	0.494	0.771	0.874

Machine learning is not yet ready to completely replace human beings in artistic style recognition. The study in this direction still requires more precision and tests to be able to fully recognize paintings. In our opinion, it is possible, but it might need much more precision. Comparing our first experiments with a smaller number of classes, the results were similar and the change in the number of inputs or size of them did not change much in the networks with a similar number of layers. Our guess would be, the network is either too simplified for the problem of style identification or the problem is too complicated to be handled without more complex data. In this, we mean especially the schemes used in different styles and the age of the painting or the artist's information. We strongly believe further samples would get much better results even without the determination of age or artist since the network will have such connections by itself. As a result, networks peaked at 40% in Top1 and 80% in Top3. These results from the perspective of other studies were similar (Viswanathan, 2017). For smaller data, this result was optimistic for further research in recognizing other parameters, as well as using networks that were previously taught specific characteristics of styles, such as frequently used motifs or colors.

The research should be broadened not only to teach the network the pixels and their values but also pre-train them with the whole idea of the style. For example, check how much is Romanticism by the existence of landscapes and nature in the painting or check the Baroque in how dark the painting is. we suppose even without the determination of the year of the painting, the network would be able to set the weights much better than simply focusing on given colors and gradients.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X. 2015. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, software available from http://tensorflow.org (accessed 1 June 2023).

Bar, Y., Levy, N., Wolf, L. 2015. *Classification of Artistic Styles Using Binarized Features Derived from a Deep Neural Network*. http://pdfs.semanticscholar.org/2b20/33af5ae4e705b90e970a586e0431678374b2.pdf (accessed 4 June 2023).

Blessing, A., Wen, K. 2018. *Using Machine Learning for Identification of Art Paintings*. https://pdfs.semanticscholar.org/1d73/0a452a5c03cc23f90d4fde71c08864f31c35.pdf (accessed 30 May 2023).

Chen, J. 2018. *Comparison of Machine Learning Techniques for Artist Identification*, Stanford University, USA.

Chollet, F. 2015. *Keras 2015*, www.keras.io (accessed 30 April 2023).

Culurciello, E. 2017. *The History of Neural Networks.* http://dataconomy.com/2017/04/history-neural-networks/ (accessed 20 May 2023).

Eclipse Deeplearning4j Development Team. 2018. *Deeplearning4j: Open-Source Distributed Deep Learning for the JVM*, Apache Software Foundation License 2.0. http://deeplearning4j.org (accessed 15 June 2023).

Gombrich, E. 1968. *Style, International Encyclopedia of the Social Sciences*, ed. D.L. Sills, xv, New York, 1968.

Hearty, J. 2016. *Advanced Machine Learning with Python*, Packt Publishing, USA.

Hockney, D., Gayford, M. 2016. *History of Pictures: From the Cave to the Computer Screen*, Thames & Hudson Ltd.

Karpathy, A., Stanford, C.S. 2018. *Convolutional Neural Networks for Visual Recognition*, cs231n.github.io/convolutional-networks/ (accessed 30 May 2023).

Krizvsky, A., Skutskever, I., Hinton, G. 2018. *ImageNet Classification with Deep Convolutional Neural Networks*, https://www.nvidia.cn/content/tesla/pdf/machine-learning/imagenet-classification-with-deep-convolutional-nn.pdf (accessed 10 June 2023).

Kurenkov, A.A. 2015. *'Brief' History of Neural Nets and Deep Learning*, http://www.andreykurenkov.com/writing/ai/a-brief-history-of-neural-nets-and-deep-learning/ (accessed 14 June 2023).

Lecountre, A., Negrevergne, B., Yger, A. 2017. *Recognizing Art Style Automatically in Painting with Deep Learning*. France. Proceedings of the Ninth Asian Conference on Machine Learning, PMLR 77: 327–342.

Nielsen, M. 2015. *Neural Network and Deep Learning.* Determination Press, http://neuralnetworksanddeeplearning.com/ (accessed 18 April 2023).

Pai, P. 2017. *Data Augmentation Techniques in CNN using TensorFlow*, https://medium.com/ymedialabs-innovation/data-augmentation-techniques-in-cnn-using-TensorFlow-371ae43d5be9 (accessed 10 June 2023).

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, 2016. E. *Scikit-learn: Machine Learning in Python*, http://scikit-learn.org/ (accessed 10 June 2023).

Shamir, L., Tarakhovsky. J.A. 2012. Computer analysis of art. *Journal on Computing and Cultural Heritage* 5(2), 1–11.

Singh, V. 2017. *Convolutional Neural Network for Image Classification*, www.completegate.com/2017022864/blog/deep-machine-learning-images-lenet-alexnet-cnn (accessed 28 May 2023).

Viswanathan, N. 2017. *Artist Identification with Convolutional Neural Networks.* Stanford University, USA.

Zaki, F., 2017. *Identify This Art*, http://www.identifythisart.com/ (accessed 20 April 2023).