DOI: 10.47743/ejpar.2023-2-2

# ENHANCING PAINTER IDENTIFICATION THROUGH STATE-OF-THE-ART ARTISTIC IMAGE RECOGNITION TECHNIQUES

**EMILIAN VIOREL MIHĂILĂ**
*Faculty of Computer Science, "Alexandru Ioan Cuza" University*
*Iasi, Romania*
*emilian.mihaila@info.uaic.ro*

**ADRIAN IFTENE**
*Faculty of Computer Science, "Alexandru Ioan Cuza" University*
*Iasi, Romania*
*adiftene@info.uaic.ro*

**Abstract**
*In this work, we will present a project through which we aim to recognize an artist whether we are in front of the painting in a museum, or whether we are on the street or in our own homes. For this, by simply uploading an image to the application, it will recognize the artist with considerable accuracy and provide us with certain details about the painting. itself, about the artist and about the current that influenced this painting.*
**Keywords:** *Painter identification, ResNet, MobileNet.*
**JEL Classification:** C38.

## 1. EXISTING WORK

In (Keren, 2002), the author uses the discrete cosine transforms, a mathematical method for analysing and representing compressed pictures, and it serves as the foundation for the classification process. By using this technique, the classifier can successfully differentiate between various types of paintings based on their distinctive visual characteristics. Using a diverse set of picture characteristics and image modifications to represent three distinct art domains, (Shamir et al., 2010) presents a method for automatically recognizing nine different artists. The method makes use of image features. The authors achieved a classification accuracy of 77% by using a dataset consisting of 360 training photos and 153 test images to categorize paintings according to the artists who created them.

ArtHistorian is a categorization and indexing system based on content that shows images using a set of criteria with tiny dimensions and simply understandable global data (Gunsel, Sariel and Icoglu, 2005). The authors employed 290 distinct paintings with styles ranging from two to twenty-ten to evaluate their algorithms' capacity to categorize data. The authors say that by utilizing this basic model, they were able to achieve noteworthy results, including accuracy gains of 86,51% with a plain Bayesian classifier.

In (Shen, 2009), the author presents a comprehensive analysis of diverse techniques employed for artist identification and furnishes a structured approach for categorizing classical Western paintings. The corpus comprises a total of 1,080 images, which have been sourced from 25 renowned classical painters. The framework proposed is founded on both global and local texture attributes, in addition to global colour and form attributes. The employment of a radial basis neural network in classification yields an identification accuracy of 69.7%.

According to (Viswanathan, 2017), this solution was able to achieve a prediction performance of 90% with a margin of 3 artists for each artwork. This information can be found in the research that was conducted. In (Lombardi, Cha, and Tappert, 2004), the authors obtained an accuracy rate of up to 85% for certain artists following the use of a combined total of 100 train data and 80 test data. In spite of this, the authors were eventually able to arrive at an accuracy rate of 71.9% after expanding the margin of error to cover all artists.

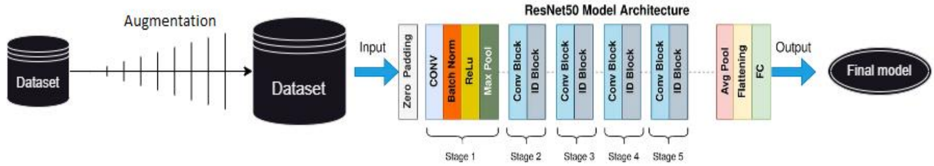
When we look at the papers that have been supplied up to this point, we are able to see very clearly the differences that exist, distinctions that belong not only to the dataset but also to the augmentations and the development of the model that is being discussed here. A further important distinction would be that all of the components are still in the stage of inquiry; as a result, there is no application that has been constructed to serve the end client of the project. This is one of the main differences between the two types of projects. In addition to the two factors that were just covered, there is also the question of accuracy, which is an area in which our project shines, both in terms of the number of individual artists whose predictions were accurate and in terms of the overall number of accurate estimates for each artist.

Based on the various factors that we have considered; we have reasons to believe that the research we conducted and the efforts we exerted have yielded a significant discovery. This discovery can be incorporated into a program that is capable of effectively fulfilling all of its intended functions. Our confidence in this conclusion is supported by the careful analysis we have conducted, taking into account various variables and factors that could impact the outcome of our research.

## 2. PROPOSED SOLUTION

During the process of building this project to determine the artist of a painting based on a photograph of the work, we utilized a number of different augmentation strategies to improve the quality of the training data before introducing it to the model, as shown in Figure 1.

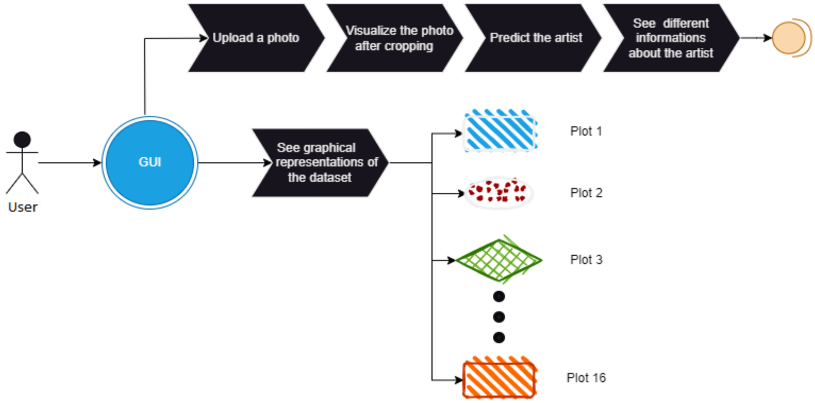
To improve the variety and unpredictability of the training set, augmentation entails applying alterations to the pictures in order to achieve this goal. These transformations can include *blurring, adding salt and pepper noise, rotation,* and *zooming*.




Source: computed by authors
**Figure 1. Dataset augmentation and the use of it**

The model is able to handle pictures collected from a variety of angles or orientations with the assistance of rotation augmentation, which enables it to learn elements that are consistent regardless of where the artwork is placed. The data is improved by zooming in, which presents the model with multiple scales of the painting. This allows the model to be more adaptive to the varying picture sizes that occur in real-world situations.

During the process of developing this application, we made use of two widely used convolutional neural network designs, namely ResNet (Targ, Almeida and Lyman, 2016) and MobileNet (Sinha and El-Sharkawy, 2019), in order to evaluate how well each one does its job of recognizing the artist. Both ResNet and MobileNet (Mohapatra et al., 2021) are examples of deep learning models that are well-known for their efficiency in the performance of image classification tasks. In general, the incorporation of various augmentation strategies as well as the use of ResNet and MobileNet models was an essential stage in the development of my project. This stage contributed to the model's capacity to identify artists based on paintings in a manner that was more accurate and resilient.




Source: computed by authors
**Figure 2. App architecture use case**

The user's experience using the application, presented in the diagram above (Figure 2), is intended to be both simple and enlightening throughout the entirety of the interaction process. When the user first opens the program, they are given two primary alternatives to choose from: (1) *upload a photo* to the application or (2) *examining dataset plots*.

If the user decides to *upload a photo*, they have the option of choosing an image from their own device to use the program's cropping feature to isolate the artwork and its key characteristics. This phase of cropping guarantees that the model's attention is entirely directed toward the pertinent artwork and removes any superfluous or irrelevant background material. After the image has been analysed, the program makes use of an advanced machine learning model to make a prediction regarding the artist based on the attributes that were extracted. After the prediction has been made, the user is shown further information about the artist who has been recognized. In addition to this, the application provides information on the art movement or style to which the artist belongs, providing a historical and cultural context for the artwork. Users are able to cultivate a more profound respect and comprehension of the artist's contribution to the world of art as a result of reading this material.

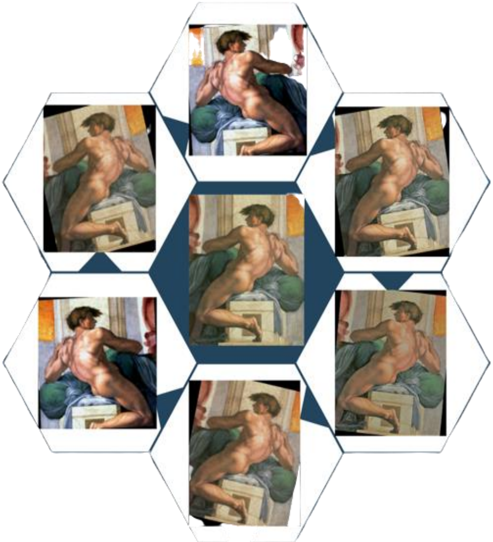
Users also have the option to *investigate the dataset plots* that are contained within the program itself. These plots provide visual representations of many features of the dataset, such as the distribution of artists, the frequency of different art movements, or any other relevant statistical analysis. For example, one plot may show the frequency of art movements, while another plot may show the frequency of distinct art movements. Users are able to get insights into the dataset as a whole by utilizing this function, as well as investigate patterns or trends among artists and art movements.

The user may engage with the program on a global scale by either uploading and evaluating their own photo or examining the plots of the dataset. In both instances, the objective is to give users an experience that is both entertaining and informative. This will allow users to explore and learn more about artists, the artwork they created, and the larger art movements in which they participated.

## 2.1 Image Augmentation

We start with 8,450 paintings from 50 different artists taken from the Best Artworks of All Time dataset. On average, each artist has contributed approximately 170 paintings to our collection. However, it is important to note that some artists have provided more paintings than others. In fact, there is one artist who has only contributed 10 paintings, which is significantly less than the average. We would need another artist to contribute an equivalent quantity of paintings, like 330, to balance the aforementioned one. The issue of uneven and poorly balanced numbers of paintings among different artists has led us to explore various augmentation techniques to significantly increase the size of our dataset. By doing so, we hope to address the inconsistencies and ensure that our dataset is more representative and comprehensive. The practice of attributing paintings to artists is not primarily aimed at increasing the number of works by well-known artists. Rather, it is intended to boost the number of paintings associated with artists who have produced only a limited number of works. In other words, the focus is on those artists who have a relatively small body of work to their name. By utilizing this approach, we are able to address not only the discrepancies associated with the distribution of paintings but also furnish the model with a greater amount of information for both training and testing. This increased amount of data will ultimately lead to a more precise and accurate model, resulting in a solution that produces favourable statistics and outcomes.

The end outcome, which could be taken into consideration in the final work, was to multiply the number of existing photographs by seven to arrive at a total of 59,150 pictures, with an average of 1,183 pictures produced by each artist. This increase is advantageous both in terms of the accuracy of the model and by reproducing some of the common conditions of photography in a museum. These factors include a loss of focus, different lighting, or even the function of the phone that many people still have active, which is the mirroring of the image. All of these variables have been taken into consideration in order to recreate the scenarios that have been described and educate the model using them as well. While the operations may seem simple, they are in fact fundamental and highly effective in the process of training the model. These operations form the building blocks upon which more complex operations can be built and are essential in ensuring that the model is able to learn and adapt to new data. Despite their simplicity, these operations have been rigorously tested and proven to be highly effective.



Source**:** computed by authors

**Figure 3. The results of data augmentation**

As depicted in the image above (Figure 3), it is evident that the original image placed at the centre has undergone a series of transformations, resulting in six different augmented versions. Each of these versions is unique in its own way and showcases the diverse possibilities of image augmentation techniques. For two of them, the contrast level was changed to simulate distinct warm or cool lights that may place these paintings in the limelight. Additionally, a random rotation between -10 and 10 degrees was picked to simulate the erroneous orientation of the phone when the photo was shot. This was done to simulate the effects of shaky hands. For the other two, a Gaussian blur was used to simulate the lack of focus on the phone; the so-called salt and pepper [5] filters were added to simulate the noise that older generations of cameras produce; and the paintings were randomly rotated between -15 and 15 degrees. This is how the number 2 came about, as the only difference between these two is the degree to which the paintings have been rotated. Additionally, a Gaussian blur [6] was applied to the last two, and auto contrast was utilized. In addition to this, the picture was cropped by using a random value between 0 and the modulus of the value between height and width. This was done after the previous step had been completed. After that, the picture was flipped horizontally by a random number of degrees ranging from -5 to 5, and finally, a mirroring operation was carried out to reproduce the function described above.

To further enhance the training process, an additional augmentation technique is applied to the data. In addition to the augmentations mentioned earlier, such as rotation, scaling, and translation, a horizontal flip is also performed on each image during the parsing stage. This technique involves flipping the image horizontally, which effectively doubles the amount of training data available. By incorporating this technique into the training process, the model is able to learn from a more diverse range of images and improve its ability to accurately classify new data.

## 2.2 Image Cropping

As the development of the app progressed, another issue arose. It was discovered that there was a specific image in a museum that posed a challenge for the future user. Despite the app's capabilities, it would not be possible for the user to completely eliminate all the distracting elements surrounding the image. One such element was the presence of people in the frame, which proved to be a significant obstacle. It is quite common to find ourselves in a room where various exhibits are on display, surrounded by other people. However, it is important to remember that not everyone present may have the same intentions. Some individuals may wish to capture a specific moment in time through photography or other means, but this can be difficult when others are not considerate of their desire to do so. Therefore, it is important to be mindful of others and their intentions when visiting such spaces. Capturing a painting through photography can be quite challenging, especially in popular museums where the number of visitors can easily reach thousands. In such cases, it's practically impossible to have an ideal setting for photography, and monopolizing a painting for this purpose may not be feasible.

One of the best ideas was cropping the painting to emphasize its unique features that are relevant to prediction. This approach aims to avoid any interference with the artwork's main characteristics that are unrelated to the prediction task. By focusing solely on the relevant aspects of the painting, the predictive model can make accurate predictions without being influenced by irrelevant information. The process of extracting a well-defined rectangle can be quite challenging, and it's not uncommon to encounter difficulties along the way. However, with perseverance and dedication, it's possible to arrive at a solution that meets all the necessary criteria. In this case, despite initial setbacks, you were able to devise an optimal solution that effectively addressed the problem at hand. To successfully implement the plan, it is important to follow a series of steps. These steps will guide the user through the process and ensure that everything is done correctly. The following are the steps that should be taken during implementation:

- Convert the image to grayscale;
- Apply a bilateral filter to reduce noise while preserving edges;
- Detect faces in the photo using OpenCV's face detection algorithm (Khan et al., 2019);

- Create a mask to exclude any faces outside from the painting detection;
- Apply the mask to the photo to exclude any faces from the painting detection;
- Detect edges in the filtered image using the Canny algorithm (Song, Zhang, and Liu, 2017);
- Dilate the edges to connect any gaps and close contours;
- Find contours in the dilated image;
- Select the contour with the largest area that is not too large;
- Exclude any contours that overlap with the detected faces;
- Create a mask for the painting by filling in the selected contour;
- Apply the mask to the original image to extract the painting;
- Get the bounding box of the painting;
- Crop the painting from the photo.

Based solely on the visual evidence presented through the images we tested and experimented with, it appears that the algorithm is functioning as intended. Specifically, the algorithm seems to successfully identify and extract the primary image from within a given frame.


Source**:** computed by authors

**Figure 4. The results of image cropping at the level of mobile application**

This primary image is defined as the largest image present, with smaller elements being excluded from consideration. In addition to its ability to process images with standard frames, the algorithm is also capable of handling images with paintings that feature irregular frames, as demonstrated in Figure 4. It is not very common to come across paintings that have frames of various shapes. However, having a frame that complements the painting's shape and this implementation can prevent the artwork from being cropped poorly and the overall aesthetic from being compromised. This is especially important when considering the accuracy of the painting's presentation and how it will be perceived by viewers.

## 2.3 Flickr API

The Flickr API is an incredibly useful resource for developers who want to tap into the vast collection of images and metadata available on the Flickr platform (Van Zwol, 2007). With this powerful tool, developers can easily access a wealth of information about photos, including their *titles*, *descriptions*, *tags*, and more. This makes it possible to create all sorts of innovative applications and services that leverage the rich content available on Flickr. Whether you're building a photo-sharing app, a search engine, or a recommendation system, the Flickr API provides a flexible and powerful way to access the data you need.


Source**:** computed by authors

**Figure 5. Images retrieved by Flickr API**

The viewer is provided with the option to examine photographs linked not just to the artist but also to the particular movement that the artist was involved in at the time. In the image that was just provided to us (Figure 5), it is quite apparent that by beginning with the name of the artist, we are able to quickly extract five additional photographs that are related to the term that was defined. We get the name of the artist as well as five photos relating to the art movement, so that we may have a more comprehensive understanding of the artist and the origins of the movement.

By seamlessly incorporating the image links that we had retrieved into our application's user interface, we were able to craft a display that not only highlighted the artist and their masterpiece but also offered valuable insights into the art movement they were associated with. The end result was not meant to catch the user's attention but to provide them with a wealth of information about the artwork and its historical context. With this approach, we were able to create a truly immersive experience for the user, one that not only showcased the beauty of the artwork but also provided them with a deeper understanding of its significance.

## 3. MODELS

The process of developing a model that can accurately recognize a painting is a complex one that involves multiple stages of modifications and fine-tuning. Each step is crucial in ensuring that the final result is valid and can be considered a good variant in the development of the application. Through careful analysis and testing, the model is refined to achieve the highest level of accuracy possible. This level of precision is essential in ensuring that the application can be relied upon to recognize paintings with a high degree of certainty. Overall, the development of such a model is a challenging but rewarding process that requires a great deal of expertise and dedication. In the following section, we will delve into the detailed steps that were taken to enhance the accuracy of the model. Each step was carefully considered and implemented to address any potential issues that could have impacted the model's accuracy. By following these steps, we were able to significantly improve the model's accuracy and ensure that it was able to provide reliable and accurate results. So, let's dive into the steps taken to enhance the accuracy of the model:

- Load the pre-trained model.
- Finish off with more layers.
- When training, include the ReduceLROnPlateau parameter.
- Train all of the model's layers.
- Freeze some core ResNet layers.
- Add ReduceLROnPlateau when training.
- Add also EarlyStopping (Bai et al., 2021) when training.
- Train the previously determined layers once more.
- Combine the records of the two types of training.
- Plot the training graph.
- Print the correctness of the data on both the train and the cross validation.
- Save the model.
- You may print off some helpful graphs to understand how the accuracy relates to each individual artist.

Although ResNet (He et al., 2016) and MobileNet (Sinha and El-Sharkawy, 2019) were designed using different architectural ideas, they have a few things in common. Both models extract features using convolutional layers and are based on deep convolutional neural networks (CNNs). They make use of batch normalization in order to increase the stability of training, and they make use of ReLU activation functions in order to add non-linearity. In addition, both architectures are trained with different variations of stochastic gradient descent (SGD), and they use anti-overfitting methods like weight decay and dropout regularization. Both of these methods are used to prevent the models from becoming too accurate. The best results were obtained with ResNet and were around 99.96% accuracy on training data and around 93.28% accuracy on cross validation data.

## 4. CONCLUSIONS

The purpose of this project was to create an artistic image recognition system that would be able to determine the name of the painter based on a photograph of a work of art displayed in a museum. In order for the system to accomplish this objective, it made use of two well-known convolutional neural network (CNN) models. These models are ResNet and MobileNet. Throughout the whole of the research, substantial progress was made in identifying artists based on pictures of their paintings. The employment of ResNet and MobileNet models produced outstanding results, with an accuracy rate that was above 90% for the bulk of the artists included in the dataset. These findings can be found in the dataset. This accuracy is superior to the results that have been published in other relevant publications, which demonstrates that the dataset that was utilized in this study is better. In addition, the creation of an integrated application that allows users to submit images, crop them, and acquire forecasts of the artist proved to be a vital contribution to the project. The program not only correctly identifies the artist, but it also offers users other information, such as an overview of the art movement that was significant in the development of the artist's work, relevant photos obtained using the Flickr API.

In conclusion, the research was successful in developing an artistic image identification system that is able to identify the creator of a painting based on a photograph of the work shown in a museum. The employment of ResNet and MobileNet models, in conjunction with a dataset of superior quality, led to an accuracy rate that was more than 90% for the majority of artists. The addition of an application, which included detailed information on the artist and the work they had created, made the user's whole experience even more enjoyable.

**References**

1) Bai, Y., Yang, E., Han, B., Yang, Y., Li, J., Mao, Y., ... and Liu, T. (2021). Understanding and improving early stopping for learning with noisy labels. *Advances in Neural Information Processing Systems*, 34, pp. 24392-24403.
2) Gunsel, B., Sariel, S., Icoglu, O. (2005). Content-based access to art paintings. In *IEEE International Conference on Image Processing 2005*, 2, pp. II-558.

3) He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778.
4) Keren, D. (2002). Painter identification using local features and Naive Bayes. In *2002 International Conference on Pattern Recognition*, *IEEE*, 2, pp. 474-477.
5) Khan, M., Chakraborty, S., Astya, R., Khepra, S. (2019). Face detection and recognition using OpenCV. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE,* pp. 116-119.
6) Lombardi, T., Cha, S. H., Tappert, C. (2004). A graphical user interface for a fine-art painting image retrieval system. In *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*, pp. 107-112.
7) Mocanu, A.A., Iftene, A. (2021). How the Events in the Life of Painters Influence the Colors of their Paintings. In *Proceedings of the 23rd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*, pp. 105-112.
8) Mohapatra, S., Abhishek, N. V. S., Bardhan, D., Ghosh, A. A., Mohanty, S. (2021). Comparison of MobileNet and ResNet CNN Architectures in the CNN-Based Skin Cancer Classifier Model. *Machine Learning for Healthcare Applications*, pp. 169-186.
9) Shamir, L., Macura, T., Orlov, N., Eckley, D. M., Goldberg, I. G. (2010). Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Transactions on Applied Perception (TAP)*, 7(2), pp. 1-17.
10) Shen, J. (2009). Stochastic modeling western paintings for effective classification. *Pattern Recognition*, 42(2), pp. 293-301.
11) Sinha, D., El-Sharkawy, M. (2019). Thin MobileNet: An enhanced MobileNet architecture. In *2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON)*, *IEEE*, pp. 280-285.
12) Song, R., Zhang, Z., Liu, H. (2017). Edge connection-based Canny edge detection algorithm. *Pattern Recognition and Image Analysis*, 27, pp. 740-747.
13) Targ, S., Almeida, D., Lyman, K. (2016). ResNet in ResNet: Generalizing residual architectures. *arXiv preprint arXiv:1603.08029*.
14) Van Zwol, R. (2007). Flickr: Who is looking? In *IEEE/WIC/ACM International Conference on Web Intelligence (WI'07),* pp. 184-190.
15) Viswanathan, N. (2017). Artist identification with convolutional neural networks. *Standford193CS231N Report*.