# Painting vs Photography Classification

Bindu Ardhya Janagani
*Department of CS*
*New Jersey institute of technology NJ ,US*
bj266@njit.edu

Suchithra Udeda
*Department of CS*
*New Jersey institute of technology NJ ,US*
su252@njit.edu

Posa Venkata Swamy Akash
*Department of CS*
*New Jersey institute of technology NJ ,US*
vp724@njit.edu

Neeraj Bhargav Rondla
*Department of CS*
*New Jersey institute of technology NJ ,US*
nr559@njit.edu

Venkata Sai Kiran Raju
*Department of CS*
*New Jersey institute of technology NJ ,US*
vs768@njit.edu

Arashdeep Kaur
*Department of CS*
*New Jersey institute of technology NJ ,US*
ak3257@njit.edu

***Abstract*—The growing digitization of visual content has made it necessary to develop efficient techniques for distinguishing between various image formats, especially photographs and paintings. Conventional methods of image processing frequently fail to capture the subtleties included in artistic and photographic content.Because these various visual features are difficult for conventional picture classification methods to generalise across, more sophisticated machine learning techniques must be investigated.In this paper the Convolutional Neural Networks (CNNs) are utilised in the present research to perform the classification of paintings and photos. The objective is to create models that can reliably and accurately discriminate between these two types of visual media. The difficulties lie not only in the differences among the various categories but also in the possible feature overlaps that may cause problems with classification.The aim of this research is to help address the problems that presently exist.**

***Keywords*—Convolutional Neural Networks, Image Classification, Automated Classification, Machine Learning, Deep Learning, VGG16 Architecture, Cross-Validation, Neural Network Evaluation, Image Processing, Art Authentication, Digital Curation.**

## I. Introduction

A new era of potential and problems in image categorization has been brought about by the widespread digitization of visual content, especially when it comes to differentiating between paintings and photographs. In order to automate the classification of artistic and photographic visual media, this research explores the complex universe of these mediums and the possibilities presented by Convolutional Neural Networks (CNNs). Differentiating between paintings—which have a wide range of brushstrokes, palettes, and creative styles—and photographs—which realistically capture moments—presents a special challenge that goes beyond the scope of conventional image processing techniques. Our research focuses on assessing how well CNNs perform in effectively classifying a variety of visual content because machine learning techniques—in particular, their employment in CNNs—have demonstrated impressive effectiveness in image recognition. By contrasting a standard CNN model with a VGG16-based.The figure below gives an estimate to the number of images to class.

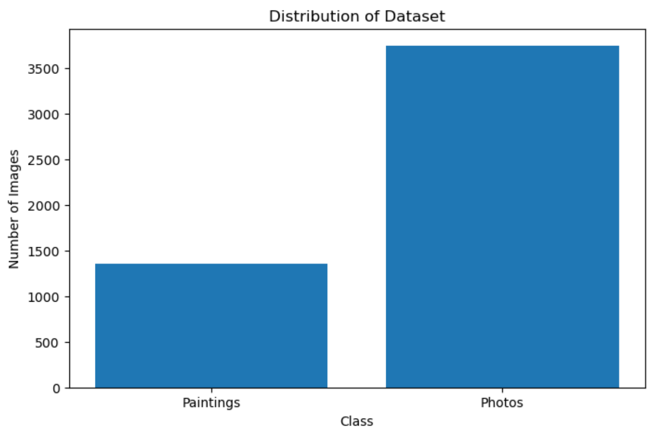


Fig1.Distribution of Dataset

## II. Incentive For Set Up Of Model

The ability to automate and improve a variety of processes is the motivation behind building up a model, especially a Convolutional Neural Network (CNN) in the context of picture categorization.CNNs are very good at identifying patterns and characteristics in images,
which makes them ideal for tasks like telling paintings from photos. The model is able to learn and recognise fine details that are difficult for conventional approaches to pick up on.The requirement for human categorization and analysis is decreased by automated picture classification. When it comes to sorting and organising visual content, this can save a lot of time and labour costs—especially when dealing with enormous datasets.CNNs are capable of achieving great accuracy in image classification tasks when properly trained. The consistency of the model guarantees that the categorization remains objective across many data sets and is not impacted by subjective influences.

## III. Literature Review

The review of the literature indicates that deep learning techniques are replacing manually established measures in the process of differentiating between paintings and photos. Previous efforts that used pre-analyzed image attributes obtained up to 94.82% precision. This study's suggested VGG-based CNN achieves an AUC ROC of 0.99 or higher, demonstrating notable progress in automated image categorization.[1].

In order to solve issues with broad picture categorization, this literature review examines visual attributes for separating paintings from images. Colour edges, distinct colours, pixel saturation, and spatial-color distribution are among the suggested visual characteristics. Furthermore, Gabor filter-based texture-based features improve discrimination. By using neural networks in Gabor and RGBXY spaces, the study achieves a notable level of classification accuracy.[2]

This review of the literature looks on complexity-based metrics that use entropy estimates, edge detection, and compression to differentiate between paintings and photos. Using a dataset of 5235 photos, the study achieves a 94.82% classification success rate, surpassing standard feature sets. Understanding of picture classification is advanced by this study, which highlights the usefulness of complexity measures.[3]

## IV.Proposed Methodology

The suggested approach entails normalising pixel values, standardising image dimensions, and curating a varied dataset of artworks and photos. There will be two Convolutional Neural Network (CNN) models used: a VGG16-based CNN that has been pre-trained on ImageNet and a baseline CNN with a more straightforward design. For a rigorous evaluation, the training will use k-fold cross-validation. Metrics such as accuracy and loss will be used to evaluate the performance of the model, and visualisation techniques will be employed to improve interpretability. We will discuss ethical issues related to biases in the dataset. The baseline and VGG16-based CNNs' disparities in accuracy and processing efficiency will be emphasised through comparative study. The study intends to investigate useful applications in the fields of art and provide directions for further investigation, guaranteeing a thorough comprehension of automated classification's potential and constraints in relation to paintings and other artwork.

1. Dataset Collection:

Assemble a varied collection of paintings and photos from numerous sources, making sure that many artistic genres, styles, and photographic settings are represented.As you divide the dataset into training and testing sets, make sure that the proportion of paintings and photos is balanced.

2. Preprocessing Data:

In order to guarantee consistency in model input, resize every image to a consistent format (e.g., 64x64 pixels).To effectively train a model, normalise the values of the pixels to the interval [0, 1].Grayscale photos should highlight important details in both artworks and photos.

3. Model Selection:

To do a comparative analysis, use two CNN models:Initial CNN: A convolutional and pooling layer architecture that is streamlined and tailored to a particular job.VGG16-Based CNN: Apply extra layers to the VGG16 architecture, which has been pre-trained on ImageNet, to enable it to adapt to the task of classifying images and paintings.

4. Model Training:

Train both models on the training dataset using a categorical cross-entropy loss function and a suitable optimizer (such as Adam).Reduce the likelihood of overfitting by assessing the model's robustness using k-fold cross-validation.

5. Model Evaluation:

Utilising criteria like accuracy, loss, precision, recall, and F1 score, assess the models on the testing dataset.

Examine the outcomes to obtain a better understanding of the models' performance, including ROC curves and confusion matrices.

6. Interpretability and Visualisation:

To show the regions of images that affect classification and to explain the judgements made by the model, apply approaches such as gradient-weighted class activation mapping (Grad-CAM).Examine misclassifications to identify possible problems and areas for improvement.

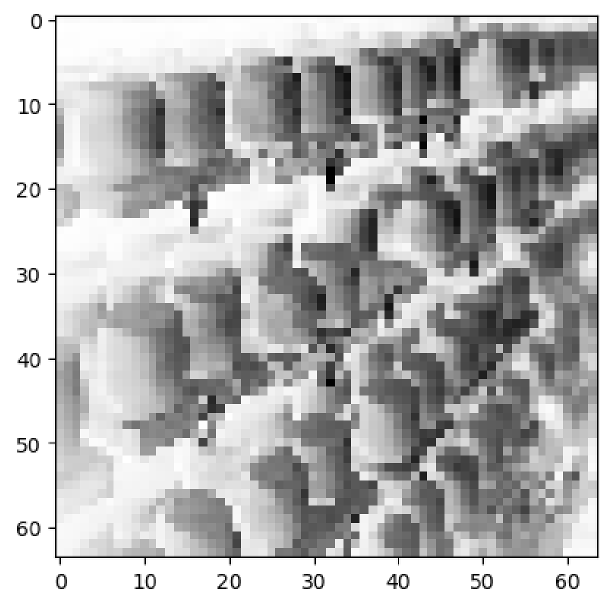


Fig2 Prediction sample

7. Assessment via Comparative Analysis:

Analyse the variations in classification accuracy between the CNNs based on VGG16

8. Ethical Considerations:

Take into account any gender, cultural, or stylistic inequalities when addressing potential biases in the dataset and models.Consider the moral ramifications of automated categorization in artistic situations, placing a focus on justice and openness.

9. Practical Applications:

Talk about how the models can be used in digital curation, art authentication, and other pertinent fields.Examine the possibility of engaging in interdisciplinary partnerships with engineers, art historians, and experts in cultural heritage.

10. Future Directions:

Make suggestions for future study directions, such as improving the model architecture, adding new features, and investigating a variety of datasets.Take into account the suggested methodology's adaptability and scalability for larger applications which makes them ideal for tasks like telling paintings from photos.

The model is able to learn and recognise fine details that are difficult for conventional approaches to pick up on.The requirement for human categorization and analysis is decreased by automated picture classification. When it comes to sorting and organising visual content, this can save a lot of time and labour costs—especially when dealing with enormous datasets.CNNs are capable of achieving great accuracy in image classification tasks when properly trained. The consistency of the model guarantees that the categorization remains objective across many data sets and is not impacted by subjective influences.

## V. Implementation

The challenge is to create an automated system that can discriminate between photos and paintings. The objective is to train Convolutional Neural Networks, models that can correctly categorise pictures as either paintings or photographs given a diversified dataset comprising both types of visual information. Capturing the complex elements and stylistic subtleties present in creative works is a problem that calls for sophisticated computer vision techniques.

Let D stand for the carefully selected image dataset, where each picture,$X_i$ is labeled as $y_i$ (0 for paintings, 1 for photographs). The task is to learn a mapping F:,$X_i$ to $y_i$ in the training process.

1.Data Processing

Resizing image:$X_i$= resize($X_i$,64x64)
Normalizing pixel values:$X_i$=$X_i$/255
Grayscaling
Flattening

2.Architecture of the Model

A convolution operation is applied to the input by a convolution layer. A filter, also known as a kernel, computes dot products by swiping it over the input image, also known as a feature map.The process of convolution at a point (i, j) is defined as follows if F is the k x k filter and I is the input image:

$$(F * I)(i,j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} F(m,n) \cdot I(i+m, j+n)$$

3.Loss function
Categorical cross-entropy is frequently used for problems with classification such as the one in your script

$$\text{Cross-Entropy}(y, \hat{y}) = -\sum_i y_i \log(\hat{y}_i)$$

where y is the expected probability distribution and y is the actual tag in one-hot encoded form.
4.Activation Function

These functions give the network non-linear characteristics.
• Softmax and ReLU (Rectified Linear Unit) are typical examples.

ReLU: $\text{ReLU}(x) = \max(0, x)$

For classification jobs, Softmax: The softmax function for component i of an output layer with n units is as follows:

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

5.Optimization

The backpropagation method of computing gradients for updating the weights.If is the rate of learning, Wis the amount of weight, $\frac{\partial L}{\partial W}$ is the loss

gradient The weight update for L in relation to W is provided by:
$W := W - \eta \frac{\partial L}{\partial W}$

These equations serve as the basis for understanding how CNNs function, absorb information, and generate predictions.

6.Model Evaluation

For comparative analysis we have used two different classification algorithms i.e. Random Forest and Decision Trees. The same preprocessing methods were followed i.e. converting taking target size 64*64, Flattening the image, converting images to grayscale, Normalizing the data and One Hot Encoding.

The objective is to achieve precise and reliable painting and photo classification by optimising the model variables \thetaθ to minimise the category cross-entropy loss function. This formulation offers a framework for using and assessing the suggested approach.

## VI. Results

These outcomes show how well the suggested methodology works to automate the classification of images and paintings. The results open the door for more developments and applications in the fields of artificial intelligence and art analysis by providing insightful information.

Fig3 shows two graphs, with the left one representing the training accuracy of a model over time, increasing to above 95%. The right graph shows the training loss decreasing sharply, suggesting good model performance.

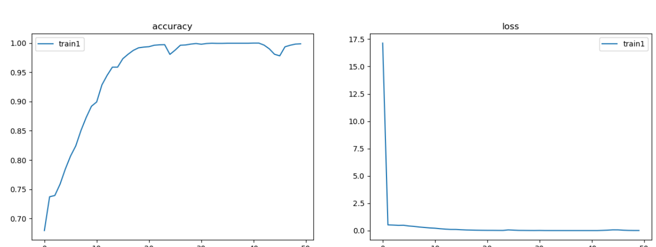


Fig3.Plot for accuracy and loss

The following is a Receiver Operating Characteristic (ROC) curve for a classification model that displays the true positive rate against the false positive rate. With an area under the curve (AUC) of 0.75, the accuracy is moderate.

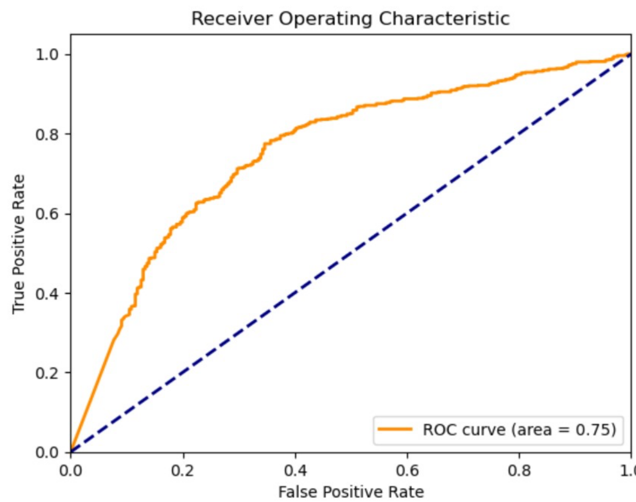


Fig4.ROC curve and ROC area for each class

## VII. Conclusion

In conclusion, this study examined the effectiveness and constraints of using convolutional neural networks (CNNs) towards the computerised categorization of images and paintings. The initial CNN and a VGG16-based system were compared, and this comparison provided insightful information about how well each model performed. A detailed grasp of the features impacting classifications was made possible by the interpretability of the model's conclusions. The significance of mitigating biases in datasets and guaranteeing fairness in computerised art analysis is emphasised by ethical issues. The study's applicability to real-world circumstances is indicated by its real-world uses in art verification and curation. Although there have been significant advancements, difficulties still exist, which calls for caution when implementing automatic classification systems in various settings.

## VIII. Future Work

More work should be done on improving model architectures, adding new features, and examining larger and more varied datasets in future studies. The ethical development of computerised art analysis will be aided by addressing cultural prejudices and improving interpretability. Working together with art historians and other cultural specialists can provide domain-specific knowledge to the models. Bridging the difference between research and real-world implementation will require investigating scalability and real-time applications. CNNs will remain relevant and flexible in the ever-changing field of visual analysis of content if they continue to explore the relationship between automated classification and developing photography trends and artistic styles.

## References


[1] J*ose Miguel Lopez-Rubio, Miguel A. Molina-Cabello,*
*Gonzalo Ramos-Jiménez & Ezequiel Lopez-Rubio*,' *Classification of Images as Photographs or Paintings by Using Convolutional Neural Networks'*, LNTCS,volume 12861,Aug. 2021.

[2] Florin Cutzu, Riad Hammoud, Alex Leykin "Estimating the photorealism of images: Distinguishing paintings from photographs" Department of Computer Science Indiana University Bloomington, IN 47408,IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition · July 2003 .

[3] Adrian Carballal1,Antonino Santos,Juan Romero,Penousal Machado,Joa o Correia,Luz Castro,"Distinguishing paintings from photographs by complexity estimates",The Natural Computing Applications Forum , Dec. 27, 2016.

[4] Carballal, A., Santos, A., Romero, J., Machado, P., João, C.: Distinguishing paintings from photographs by complexity estimates. Neural Comput. Appl. 30, 1957–1969 (2016)

[5] Athitsos, V., Swain, M.J., Frankel, C.: Distinguishing photographs and graphics on the World Wide Web. In: 1997 Proceedings IEEE Workshop on Content-Based Access of Image and Video Libraries, pp. 10–17 (June 1997). https://doi.org/10.1109/IVL.1997.629715

[6] B.S. Manjunath and W.Y. Ma. Texture features for brows- ing and retrieval of image data. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.

[7] V. Athitsos, M.J. Swain, and C. Frankel. Distinguishing photographs and graphics on the World Wide Web. Work- shop on Content-Based Access of Image and Video Libraries (CBAIVL '97) Puerto Rico, 1997.