Patterns of the Past:

EXPLORING ART AND ARCHITECTURE THROUGH VISUAL INTELLIGENCE

Presenters: Dr. Atreyee Sinha and Dr. Sugata Banerji

Atreyee Sinha
Edgewood College
asinha@edgewood.edu

Sugata Banerji
Lake Forest College
banerji@lakeforest.edu

LAKE FOREST COLLEGE

IARIA

EDGEWOOD COLLEGE

Who Are We?

- I'm a Professor of Computer Science at Edgewood College, Madison, Wisconsin, and do research in the field of Computer Vision.

Atreyee Sinha
Edgewood College

- My co-presenter, Dr. Sugata Banerji, is an Associate Professor of Computer Science at Lake Forest College, Lake Forest, Illinois. He shares the same research interests.

Sugata Banerji
Lake Forest College

Introduction And Overview

- Interaction with Audience

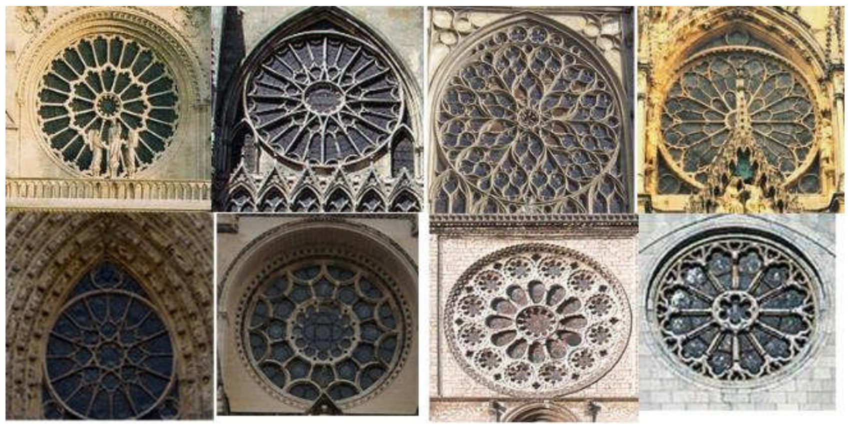
- Tutorial Goals:
  - Understand fundamental computer vision concepts.
  - Learn techniques for analyzing and classifying art and architectural styles.
  - Discuss some interesting results and future directions for research.

- Welcome Video from asynchronous online presenter, Dr. Sugata Banerji.

What are Patterns?

- A **pattern** is a *repetitive* design.
- When it is *distinctive* as well, it can help us recognize a class.
- Patterns can be found in literature, art, architecture, town planning, and other areas.
- Here, we talk only about visual patterns.
- We use them to recognize art and architecture from the past.

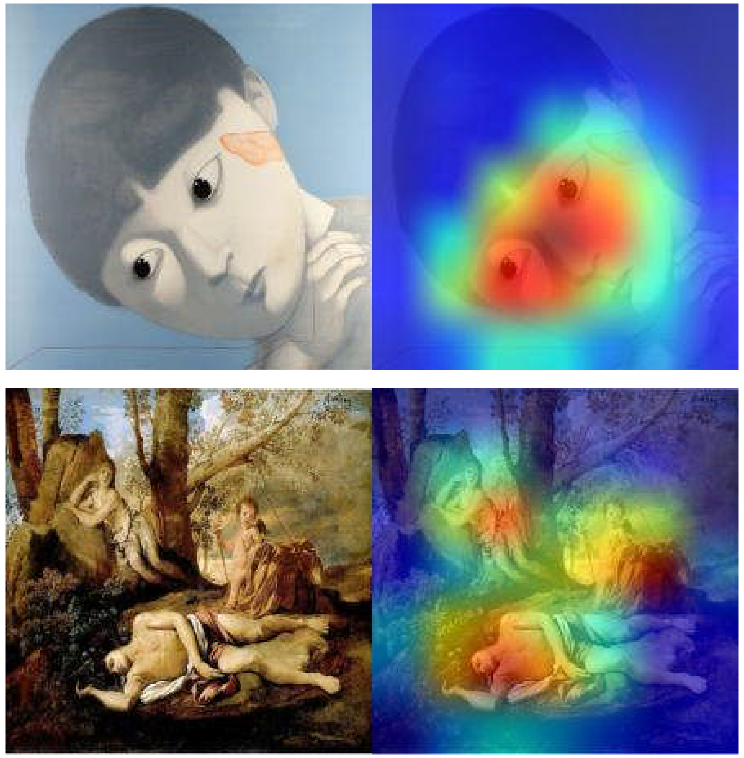
Architecture Patterns



Images taken from the Architectural Styles dataset.
Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, "Architectural style classification using multinomial latent logistic regression," in *ECCV*, 2014.

- Can tell us about the architectural style
- Can tell us about the time and place that a building was made

Art Patterns



Original images taken from the Paintings-100 dataset.
E. Knizhnik, B. Rivera, A. Sinha, and S. Banerji, "**Paintings-100: A Diverse Painting Dataset for Large Scale Classification**," in *Proceedings of the Ninth International Conference on Advances in Signal, Image and Video Processing (SIGNAL 2024)*, March 10-14,

- Can tell us the artistic style
- Can tell us the name of the artist

Other Patterns

- Repetitive patterns can be a distinguishing feature of a city
  - Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros. 2012. **What makes Paris look like Paris?** *ACM Trans. Graph.* 31, 4, Article 101 (July 2012)
- Traditionally, patterns were found by various computer vision methods
- Today, Neural Networks dominate the field

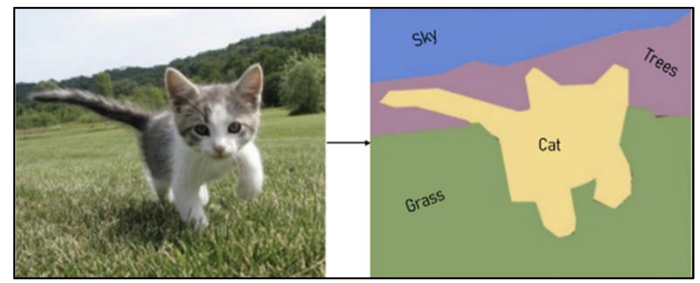
http://graphics.cs.cmu.edu/projects/whatMakesParis/paris_sigg_reduced.pdf

What is Computer Vision?

LAKE FOREST COLLEGE

- How computers "see" and interpret images.
- Basic image processing steps:
  - Image acquisition,
  - preprocessing,
  - feature extraction,
  - etc.
- Some Computer Vision Problems:
  - Classification
  - Retrieval
  - Semantic Segmentation
  - Object Detection

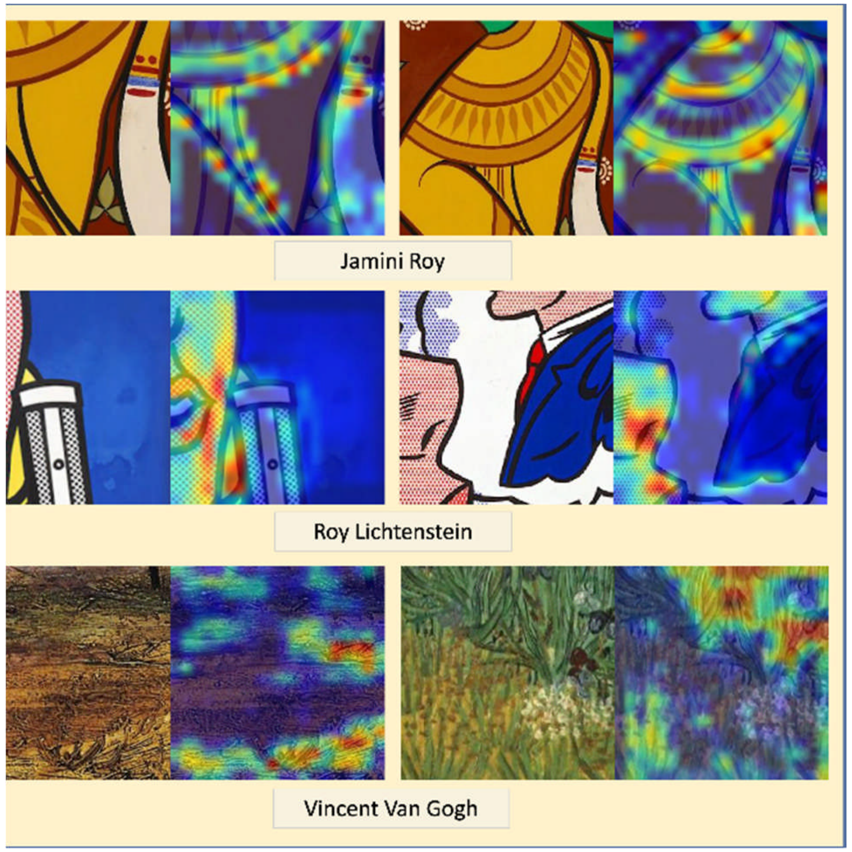
https://medium.com/analytics-vidhya/beginners-guide-to-object-detection-algorithms-6620fb31c375



https://datahacker.rs/020-overview-of-semantic-segmentation-methods/

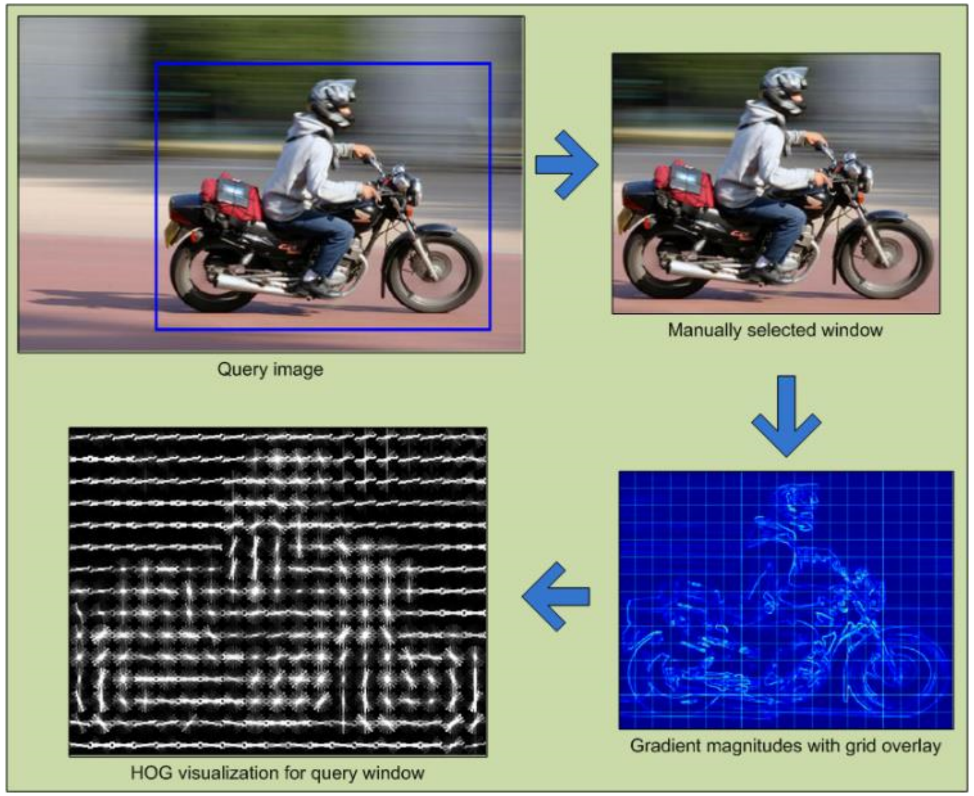
Relevance to Art and Architecture

- Can reveal patterns and insights beyond easy human perception.
- Examples: Style analysis, object detection, 3D reconstruction.
- Key Techniques:
  - Image segmentation
  - Feature extraction: Identifying distinctive features
  - Machine learning algorithms: Training models to classify styles.



How do we turn images into numbers?

- Until the last decade, we used feature vectors
  - HOG
  - LBP
  - SIFT
  - GIST
  - DPM
  - TinyImage
  - Color histograms
  - … and many, many more.



Current Approach - Deep Learning

- A supervised machine learning technique that uses multi-layered convolutional neural networks (CNN).
- Uses multiple layers to progressively extract higher-level features from the raw input.
- Currently the most successful method for classifying images with complex patterns
- There are countless different architectures, used for different purposes.

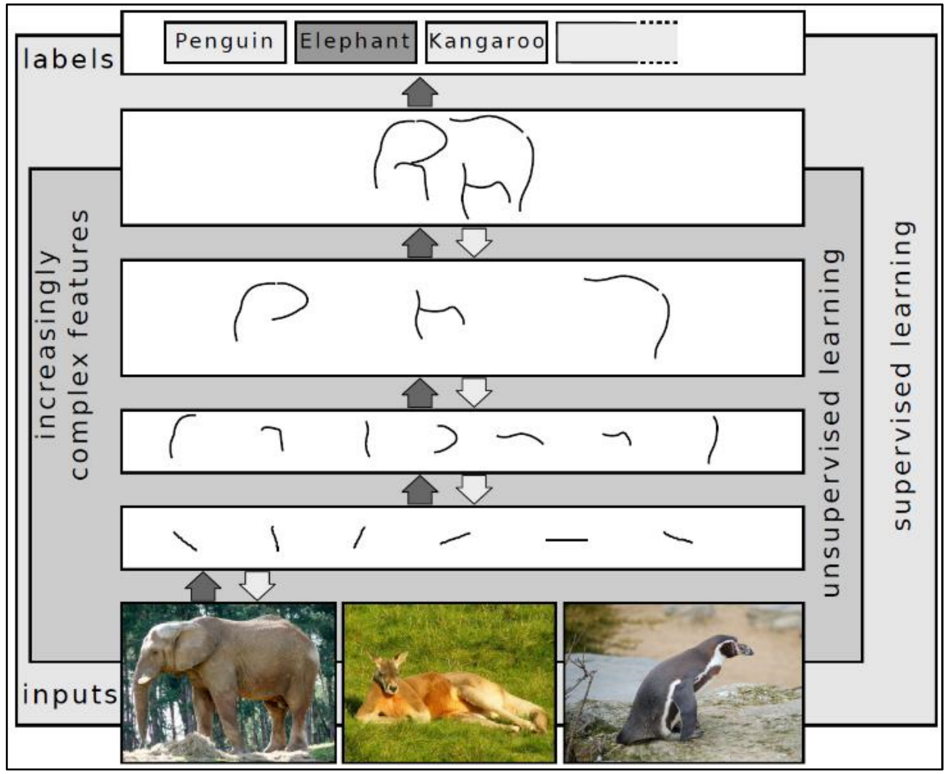


Image by Sven Behnke - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=82466022

LAKE FOREST COLLEGE

Architecture Classification – An Example

- R. D. Meltser, S. Banerji, and A. Sinha, "**What's that Style? A CNN-based Approach for Classification and Retrieval of Building Images**", in *Proceedings of the International Conference on Advances in Pattern Recognition*, December 27-30, 2017, Bangalore, India.

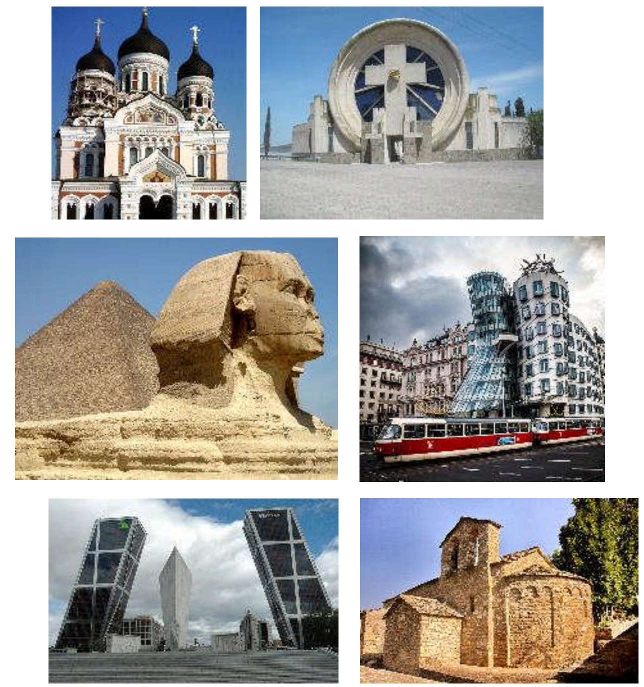


IARIA

EDGEWOOD COLLEGE

Introduction – The Problem

- Classification of buildings based on architectural style
- Retrieval of images having the same architectural style
- Potential applications:
  - tourism industry,
  - historical analysis of building styles,
  - cinema and theater set design,
  - architecture education,
  - geo-localization, etc.
- A comparative evaluation of different conventional classification techniques by [1] for architectural style classification clearly suggests the need for more powerful visual features for architectural style classification and retrieval tasks.



[1] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, "Architectural style classification using multinomial latent logistic regression," in ECCV, 2014.

Architecture Classification

Classifying Architectural Styles:

Historical periods: Gothic, Baroque, Modernist.

- Techniques: Building shape analysis, facade feature detection, structural pattern recognition.
- Example: Classifying Gothic cathedrals based on architectural elements.

Analyzing Architectural Structures:

Ancient monuments: Identifying building techniques and materials.

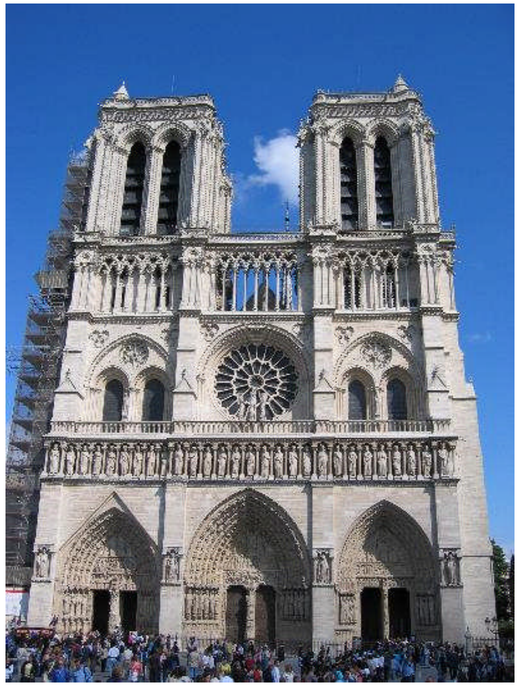
- Modern buildings: Analyzing design patterns and structural innovations.
- Techniques: 3D modeling, structural analysis.

Urban Planning and Design:

Analyzing city layouts and building densities.

- Techniques: Aerial image analysis, spatial data processing.

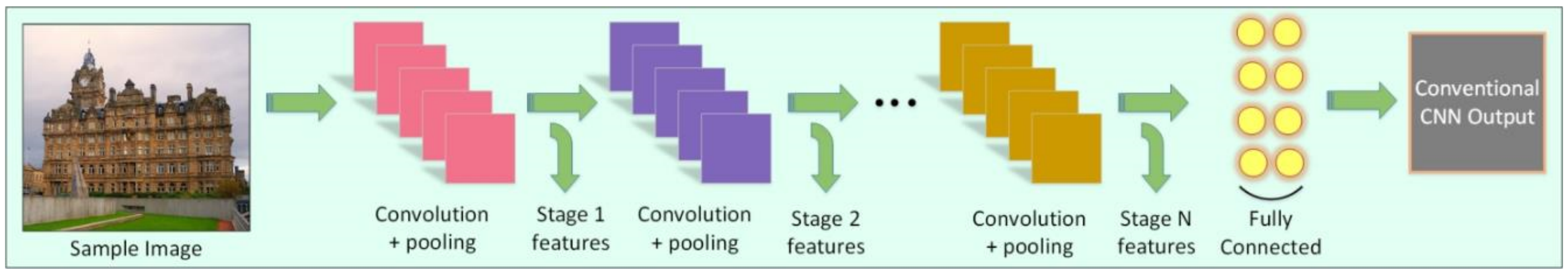
Which One is Gothic?



Which is American Foursquare?
Which is American Craftsman?

Proposed Method

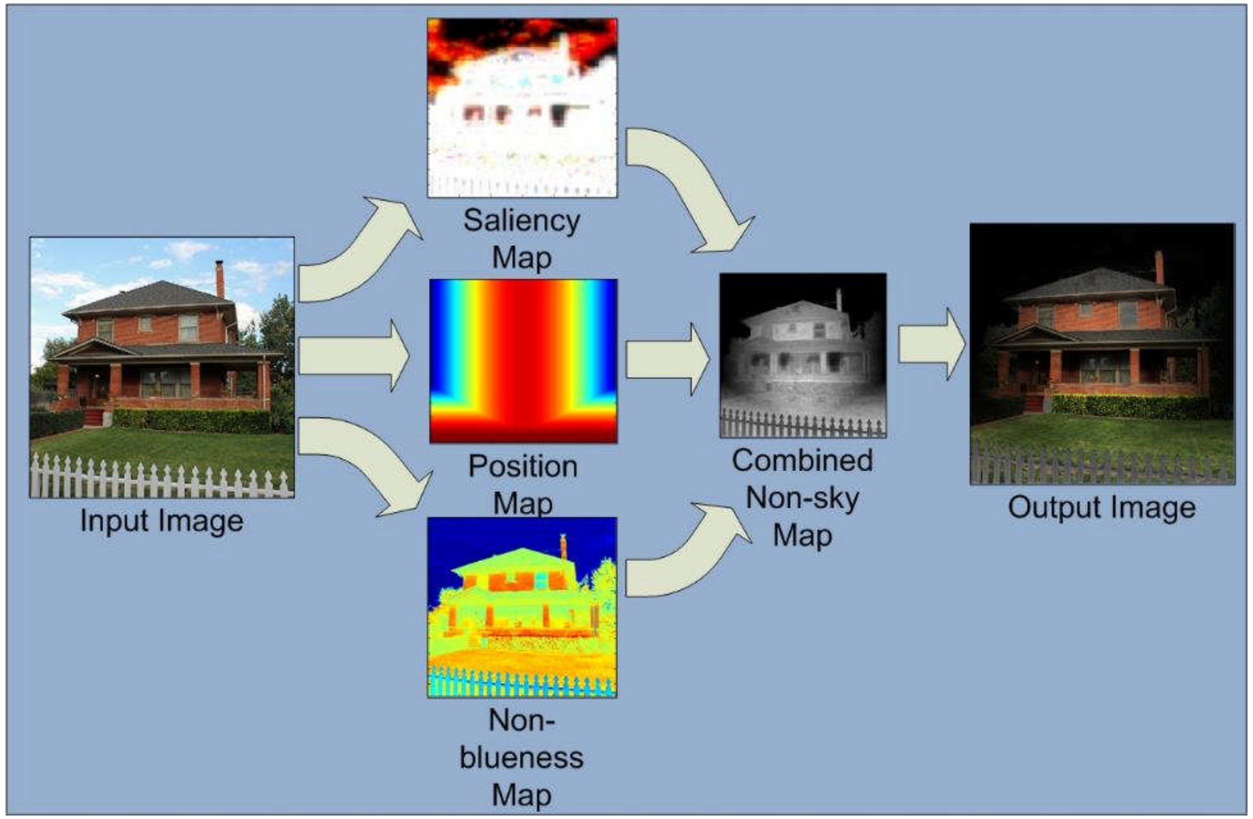
- We extract CNN features using the OverFeat image feature extractor [3]
- OverFeat is trained on ImageNet and has 21 layers
- The first 16 layers are convolution and pooling layers
- The last 5 are fully connected classifier layers
- We extract features from the outputs of layers 1-16 and work with them



[3] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks. ICLR (2014)

Sky-detection – preprocessing

- We use saliency, position and color cues to detect the sky region and remove it from each image.



The Dataset

- We evaluate our representation and classification techniques on the challenging Architectural Style Dataset [1]
- The dataset consists of 4794 building images from 25 different architectural styles
- These images have been downloaded from the Wikimedia collection and feature an extensive selection from different eras.
- The names of the categories and the number of images in each are shown in Table 1.

TABLE 1. CLASS NAMES AND NUMBER OF IMAGES IN EACH CLASS

Class Name	Image Count
Achaemenid architecture	69
American craftsman style	195
American Foursquare architecture	59
Ancient Egyptian architecture	256
Art Deco architecture	366
Art Nouveau architecture	450
Baroque architecture	239
Bauhaus architecture	92
Beaux-Arts architecture	191
Byzantine architecture	111
Chicago School architecture	153
Colonial architecture	177
Deconstructivism	213
Edwardian architecture	79
Georgian architecture	154
Gothic architecture	109
Greek Revival architecture	327
International style	207
Novelty architecture	212
Palladian architecture	113
Postmodern architecture	163
Queen Anne architecture	425
Romanesque architecture	107
Russian Revival architecture	165
Tudor Revival architecture	162

[1] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, “Architectural style classification using multinomial latent logistic regression,” in ECCV, 2014.

LAKE FOREST COLLEGE

DATASET IMAGES


Achaemenid


American Craftsman


American Foursquare


Ancient Egyptian


Art Deco


Art Nouveau

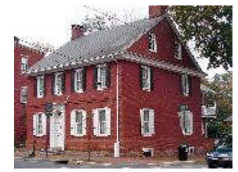

Baroque


Bauhaus


Beaux-Arts


Byzantine


Chicago School

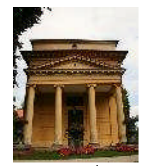

Colonial


Deconstructivism


Edwardian

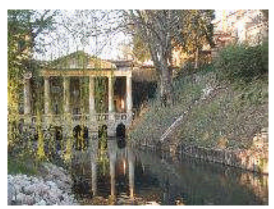

Georgian

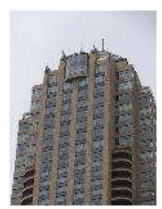

Gothic


Greek Revival


International

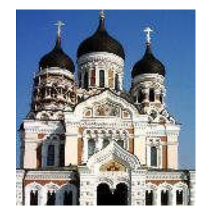

Novelty


Palladian


Postmodern


Queen Anne


Romanesque


Russian Revival


Tudor Revival

Classification and Retrieval

- We use three different classifiers
  - K-nearest neighbors (KNN) – no training phase
  - Enhanced Fisher Model (EFM) [4] – 30 training images per class
  - Support Vector Machine (SVM) - 30 training images per class
- Images not used for training are used for testing

- For the retrieval task, the test images are used as queries for retrieving nearest neighbors.
- All images other than the query are used as the retrieval set

[4] C. Liu and H. Wechsler, "Robust coding schemes for indexing and retrieval from large face databases," IEEE Transactions on Image Processing, vol. 9, no. 1, pp. 132–137, 2000.

Classification Results

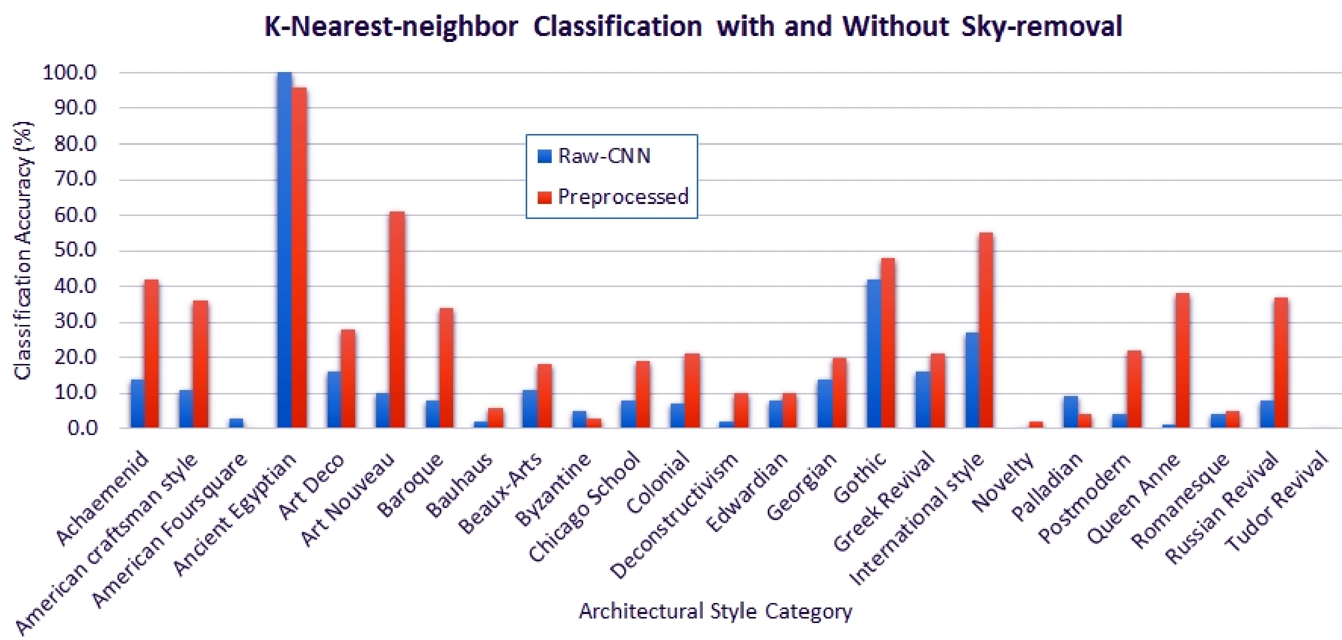
TABLE 2. Comparison of Classification Performance (%) between the best-performing CNN Layers using different classifiers on the Architectural Style Dataset

CNN Features Used	KNN	SVM	EFM-KNN
Layer 9 raw	13.2	43.0	59.7
Layer 12 raw	10.0	45.5	62.0
Layer 16 raw	9.1	47.9	**63.9**
Layer 21 raw	41.9	33.7	55.2*
Layer 21 skyless	35.9	28.2	50.3
Layer 9 skyless	25.4	39.1	56.3
Layer 12 skyless	23.4	41.5	58.3
Layer 16 skyless	22.7	42.4	59.2

*17% more than the MLLR+SP method proposed by [1].

[1] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, "Architectural style classification using multinomial latent logistic regression," in ECCV, 2014.

Classification Results



A Comparison of the class-wise classification performance between the layer 9 raw-CNN features and the pre-processed (sky-removed) CNN features. Both the features use a KNN classifier.

Classification Results

LAKE FOREST COLLEGE

TABLE 3. Comparison of Classification Performance (%) between our best-performing CNN Layer and other methods as reported by [8] on the Architectural Style Dataset

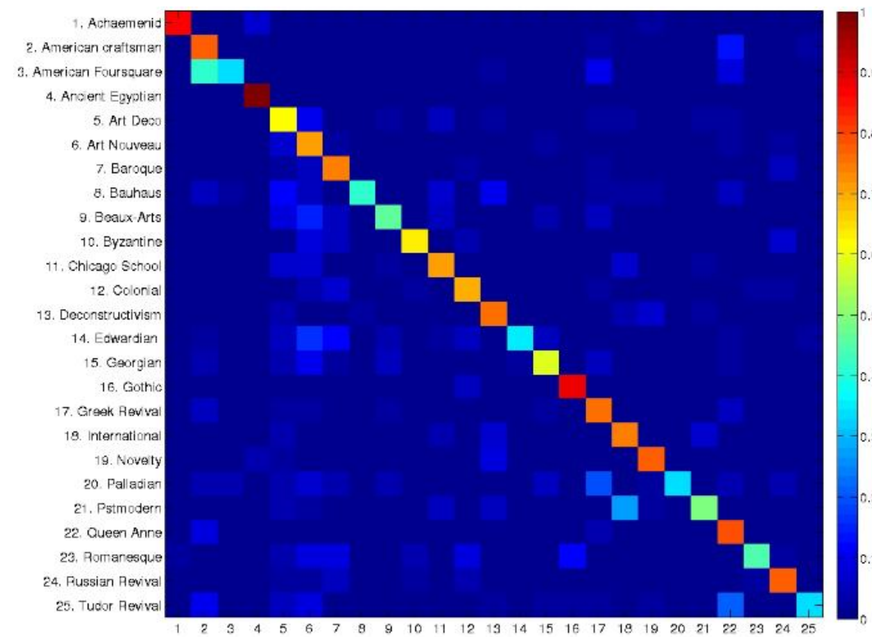
Method Used	25-Class Classification Rate(%)
GIST	17.39
SP	44.52
OB-Partless	42.50
OB-Part	45.41
DPM-LSVM	37.69
DPM-MLLR	42.55
MLLR+SP [8] *	46.21
Layer16 Raw+EFM	**63.90**

* [1] As per the numbering on this slide show.

[1] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, "Architectural style classification using multinomial latent logistic regression," in ECCV, 2014.

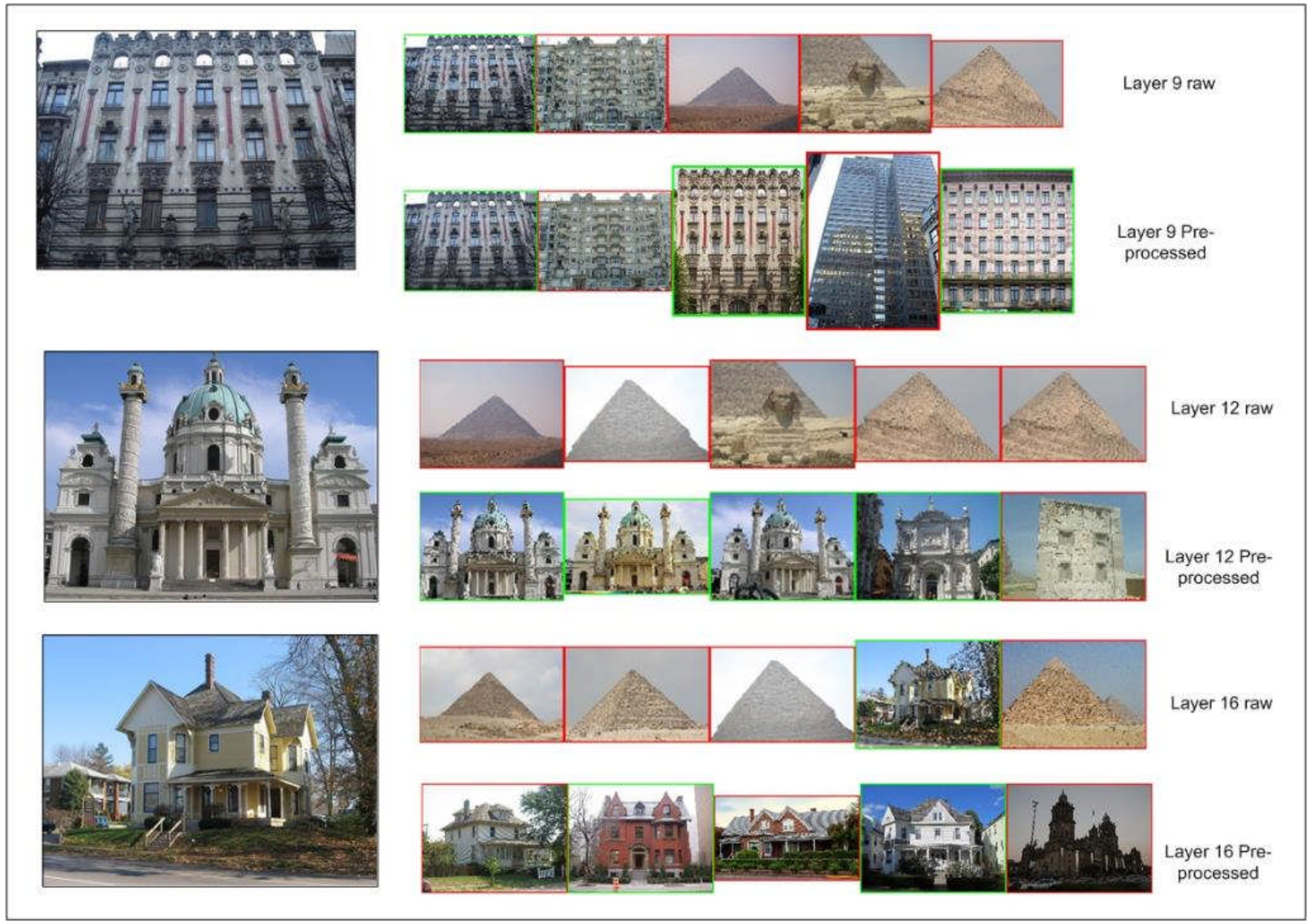
Classification Results

Confusing images



The confusion matrix for architectural style classification using Layer 16 CNN features and EFM-KNN classifier. The rows show the real style categories and the columns show the assigned style categories.

Retrieval Results



Some results from the retrieval task. For each query, the top row shows 5 nearest neighbors retrieved by the raw-CNN representation. The lower row shows the retrieval set obtained after pre-processing.

Key Takeaways

- We have proposed an image representation based on features extracted from intermediate layers of a pre-trained CNN.
- Our proposed representation performs better than traditional final-stage CNN features at both retrieval and classification tasks.
- CNN-based features vastly outperform traditional Computer Vision features such as GIST and DPM

Art Analysis and Classification

Analyzing Paintings:

Style classification

Artist Recognition

- Techniques: Color histograms, texture analysis, edge detection.

Analyzing Sculptures:

3D reconstruction (Creating digital models from images)

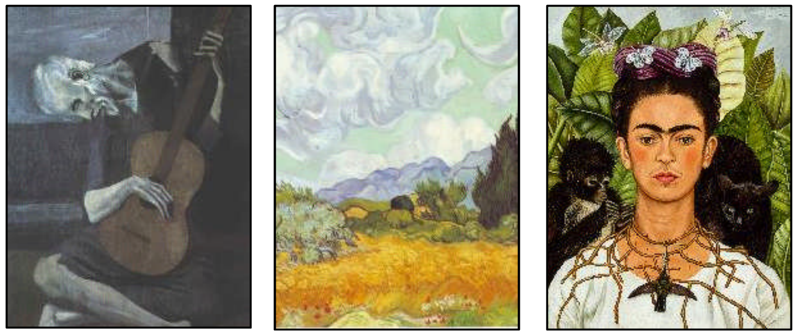
- Techniques: Shape analysis, surface texture mapping.
- Example: Analyzing ancient Greek sculptures for stylistic features.

Other Forms of Visual Art:

Textile analysis (Identifying patterns and origins)

- Ceramic analysis: Classifying pottery styles.
- Techniques: Pattern recognition, shape matching.

Painting Classification/ Artist Recognition



Do you know who painted these?

Can you tell if these were painted by the same artist?

Artist/ style recognition is a challenging problem

- Too few training images
- Variance in data
- Emphasis is not on the subject, but something else

Potential applications:
- museum work,
- painting theft investigation,
- forgery detection,
- art education,
- other applications

Algorithms and Methods Used

- Data augmentation (style transfer)
  - This concept combines (or multiply, in a more technical term) natural images from large datasets such as PASCAL VOC/ImageNet, and images with specific artistic styles, such as Impressionism or Realism.

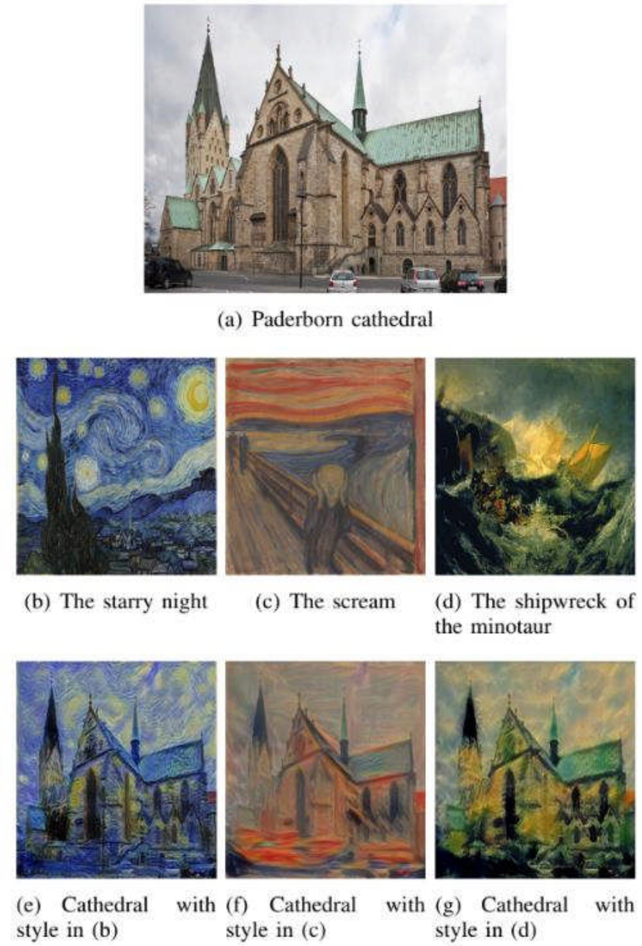
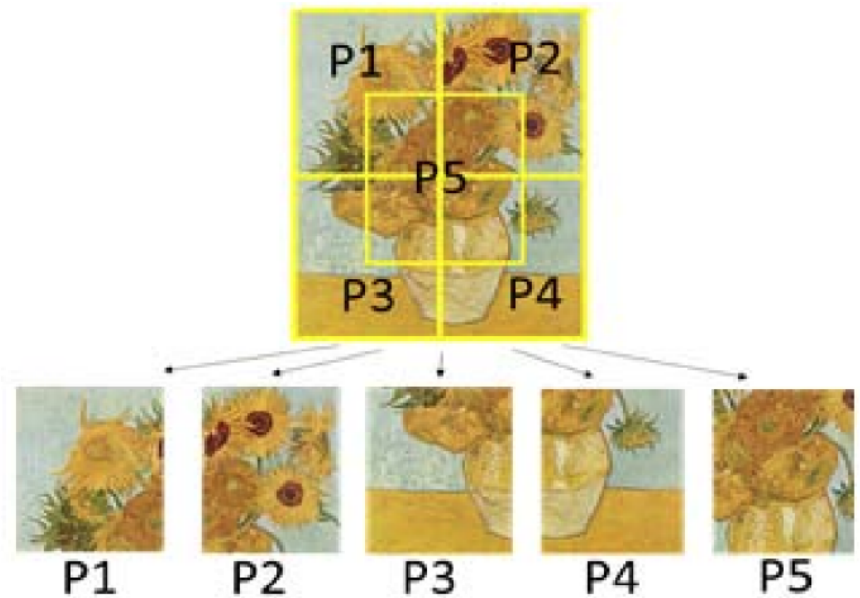


Image from Smirnov & Eguizabal (2018)

Algorithms and Methods Used

- **Transfer learning with convolutional neural networks (CNNs)**
  - To benefit from the valuable features obtained from some pre-trained CNNs and save computational cost, all three papers adopt transfer learning methods on different CNN models to classify fine-art paintings.

- **Image patches or segmentation**
  - Rodriguez et al. (2018) suggest that each image from the dataset should be equally segmented into five pieces as the CNN model's input.



An example of segmenting the original image into five equal size pieces. Image from [Rodriguez et al. (2018)]

Challenges

A significant challenge is the shortage of large datasets of labeled digitized artworks.

Without substantial data, even a recognized classification model will not yield a good classification result.

The designs of the classification models and the implementations of model training are equivalently critical as well.

- Some Example Datasets:
  - Painting 91 dataset
  - TICC Printmaking Dataset
  - Web Gallery of Art (WGA)
  - WikiArt

Our Work – Curating the Paintings-100 Dataset

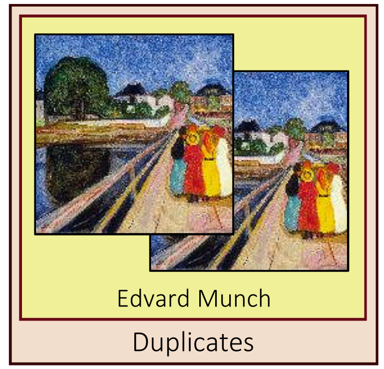
- We started with a dataset of 91 painters in the Painting-91 dataset [1]
  - The dataset consisted of 4266 paintings from 91 different artists
  - The number of images per artist varied ranging from 31 (Frida Kahlo) to 56 (Sandro Botticelli)
  - The images were all collected from the Internet

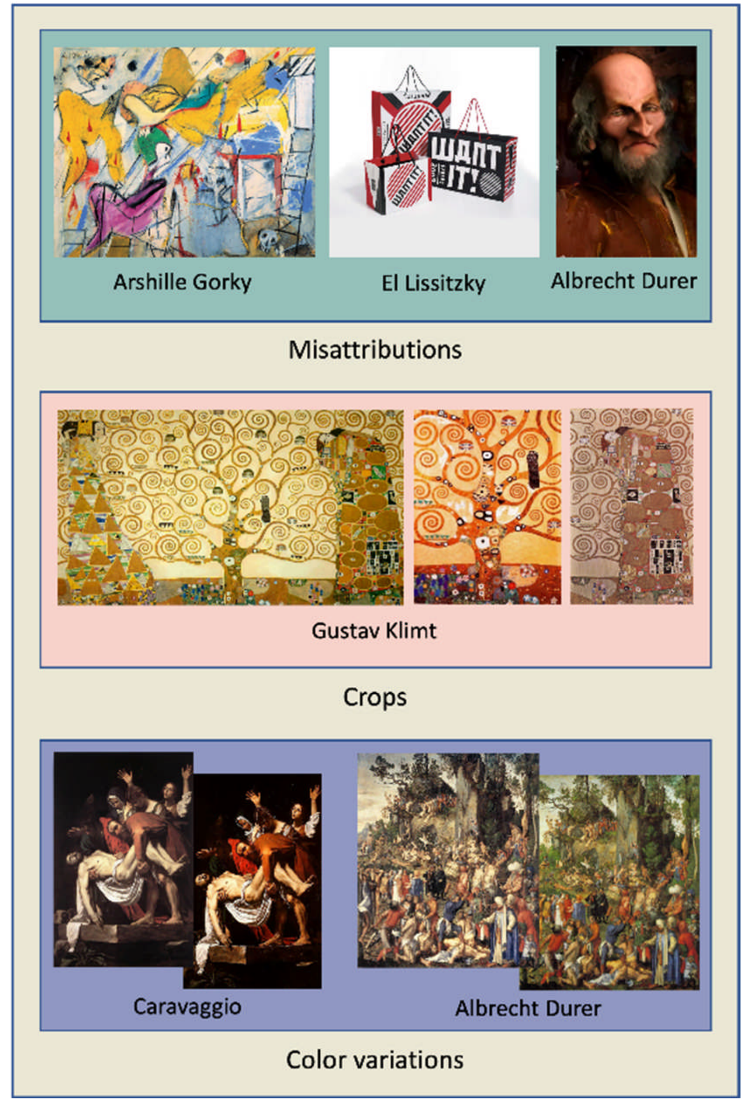
[1] Khan, F.S., Beigpour, S., de Weijer, J.V., Felsberg, M.: Painting-91: A large scale database for computational painting categorization. MVAP(2014)

The Dataset Problems

- The images were very low resolution
- It had
  - Tiny images (~260px)
  - Misattributions
  - Partial paintings
  - Duplicates (with variations)

✓ We removed these errors
✓ Increased resolution (~1500px)





Enhancing the Dataset

- We used Google reverse image search to find high resolution version of each image, manually
  - Wasn't always available easily, especially for newer artists
- Made sure duplicates are recorded, and slightly cropped and color-changed versions are replaced by exact copies
- Read the description to make sure the images are correct
- Read up about the painter to discover different versions etc.
- Increased the number of classes/images per class (where needed)

Time consuming work!

We Added More Painters

We Added More Painters

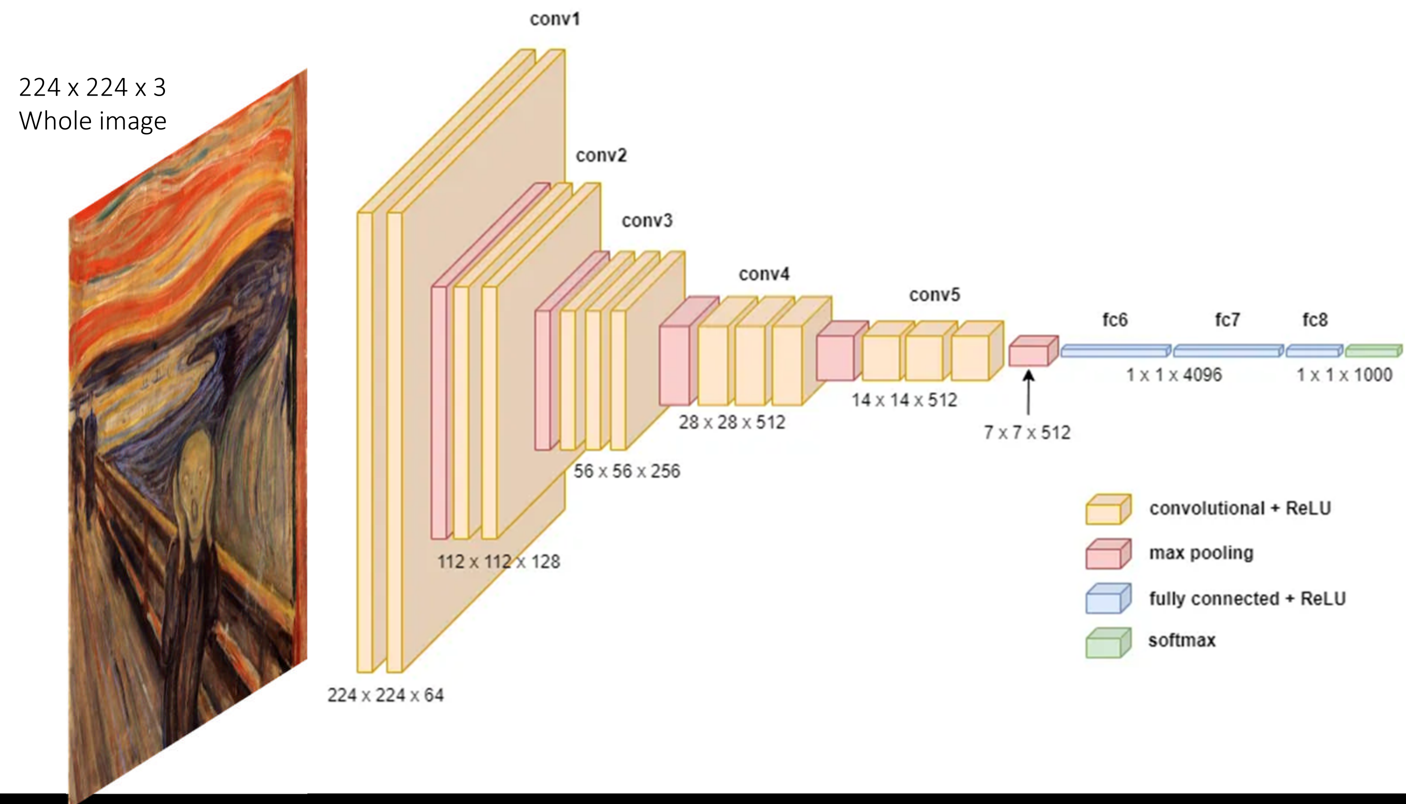
The Ukiyo-e style class was added to the existing 13

Artist	Nationality	Style
Amrita Sher-Gil	Hungarian-Indian	Several
Jamini Roy	Indian	Indian folk art
Julie Mehretu	Ethiopian American	Several
Katsushika Hokusai	Japanese	Ukiyo-e
Kitagawa Utamaro	Japanese	Ukiyo-e
Rafiy Okefolahan	Cape Verdean	Contemporary multimedia
Raja Ravi Varma	Indian	Indian realism
Utagawa Hiroshige	Japanese	Ukiyo-e
Zhang Xiaogang	Chinese	Surrealism

Methodology

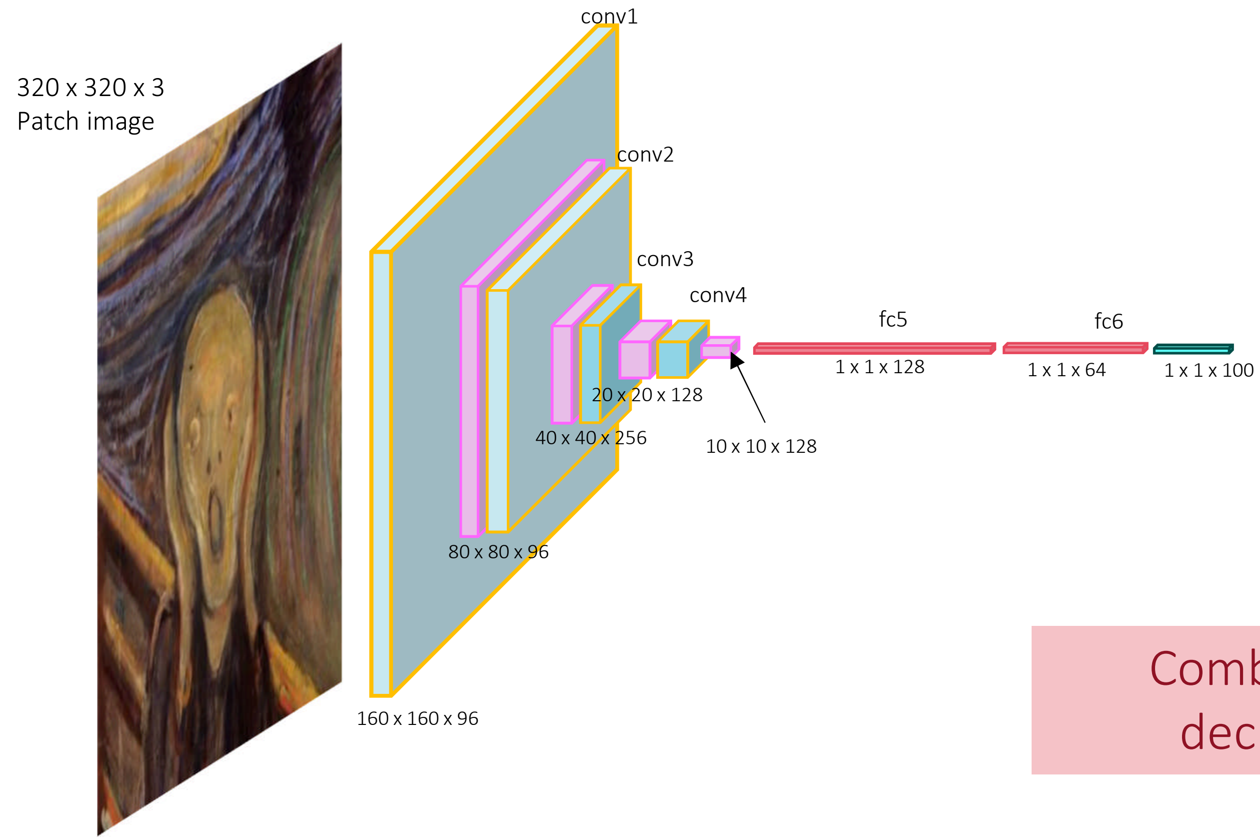
Train a CNN on the whole images

- **Challenge**: too few images for training CNN
- **Tool used**: CNN (VGG-16) pretrained on ImageNet, fine-tuned on whole images



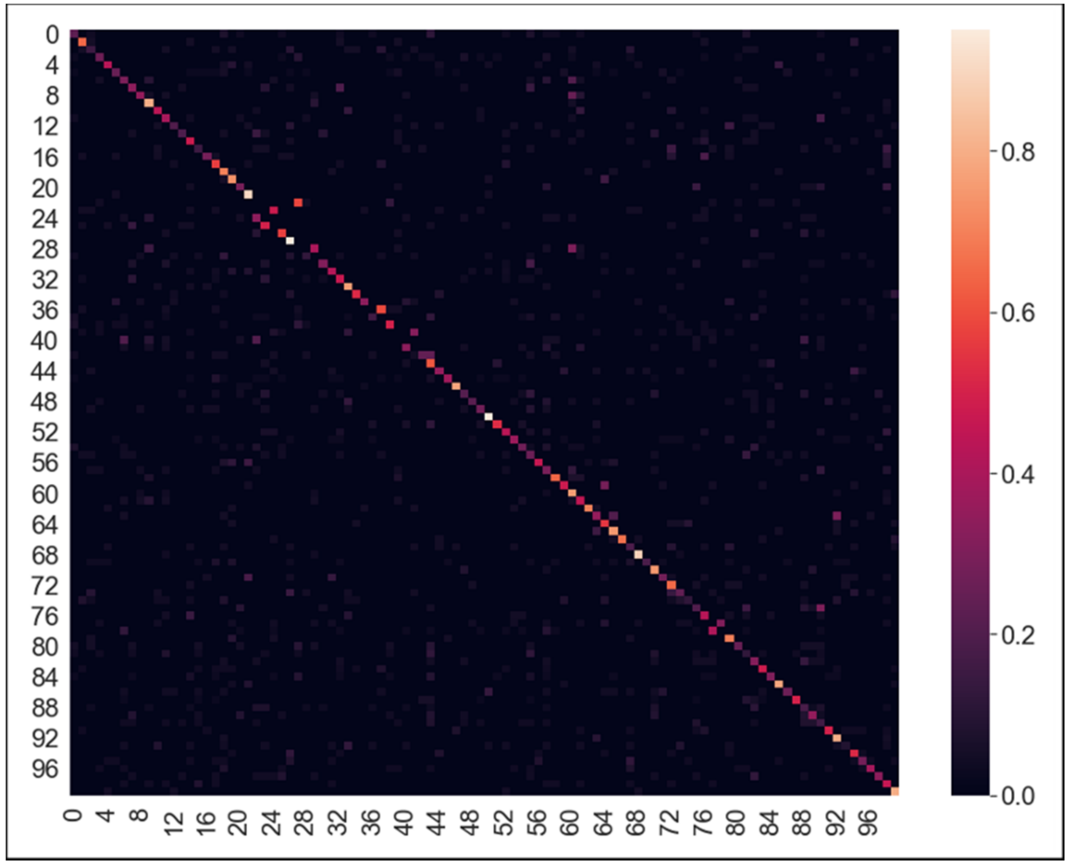
Methodology

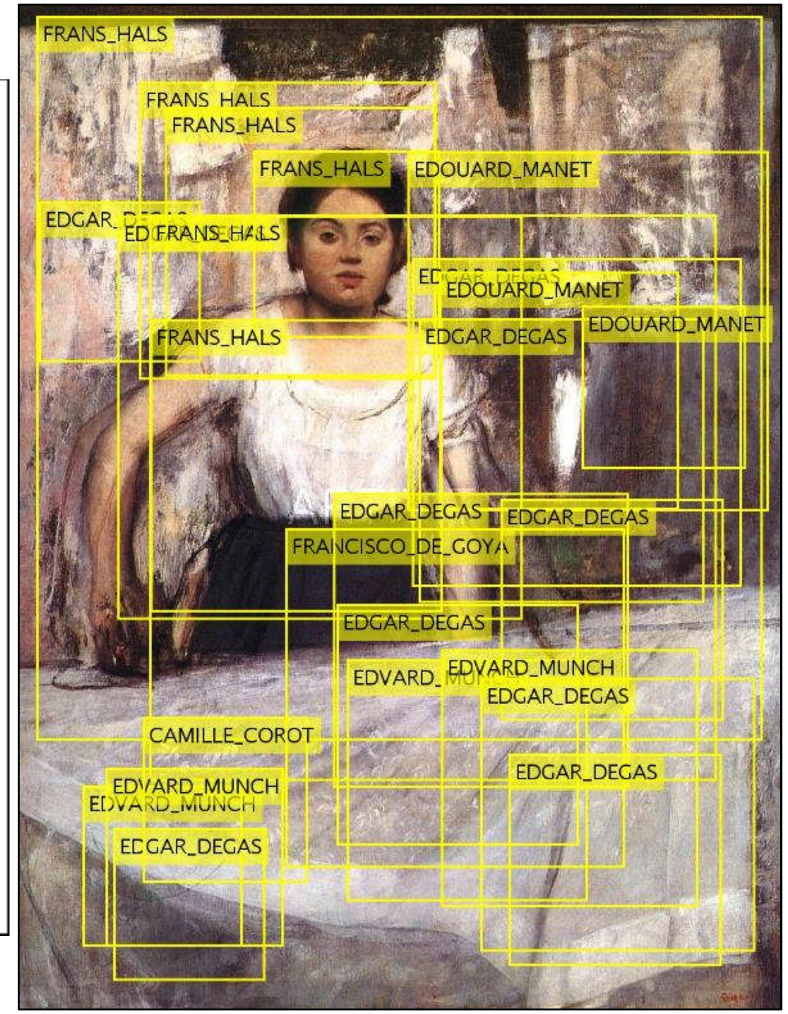
- Cut each image into multiple small patches randomly
- Train another CNN (designed from scratch) on the patches



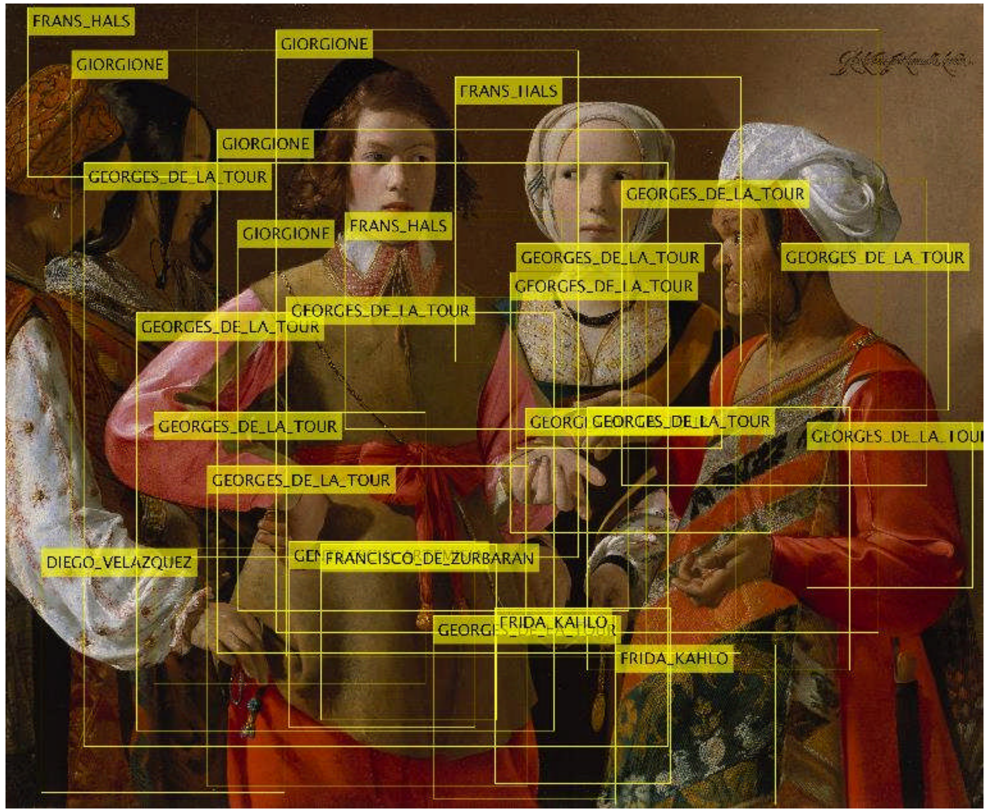
Combine the decisions!

Results


0
4
8
12
16
20
24
28
32
36
40
44
48
52
56
60
64
68
72
76
80
84
88
92
96
0.0
0.2
0.4
0.6
0.8


FRANS_HALS
FRANS_HALS
FRANS_HALS
FRANS_HALS
EDOUARD_MANET
EDGAR_DEGAS
FRANS_HALS
EDOUARD_MANET
FRANS_HALS
EDGAR_DEGAS
EDOUARD_MANET
EDGAR_DEGAS
EDGAR_DEGAS
FRANCISCO_DE_GOYA
EDGAR_DEGAS
EDVARD_MUNCH
EDGAR_DEGAS
CAMILLE_COROT
EDGAR_DEGAS
EDVARD_MUNCH
EDVARD_MUNCH
EDGAR_DEGAS

Results

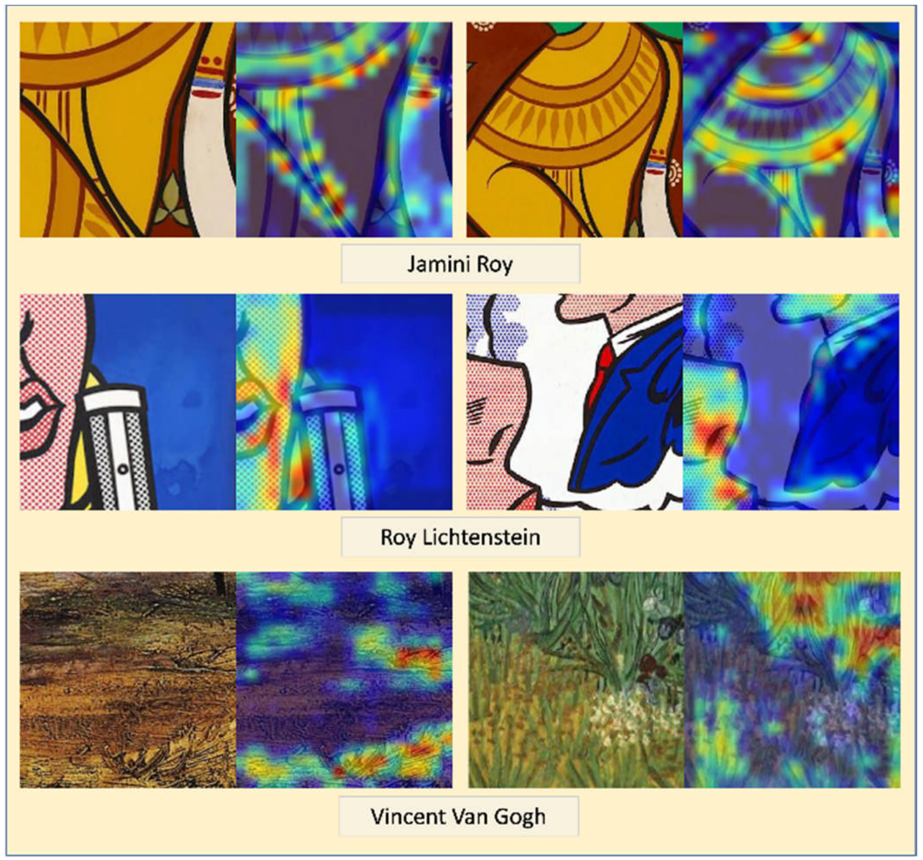


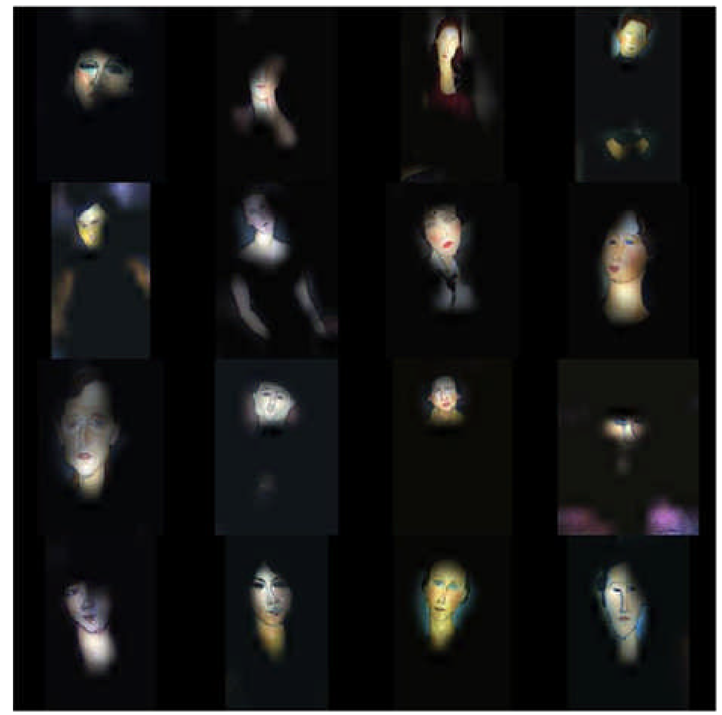
Real: Georges de la Tour. Whole label assigned: Diego Velazquez

Paintings by Diego Velazquez

Heatmap Analysis

What makes a Picasso look like a Picasso?

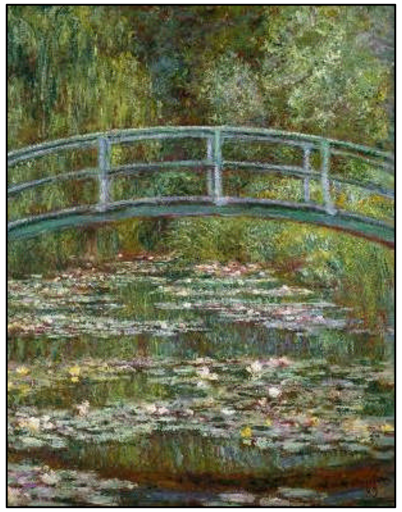
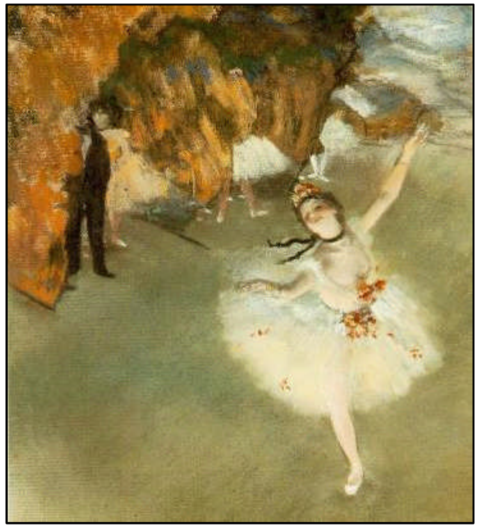
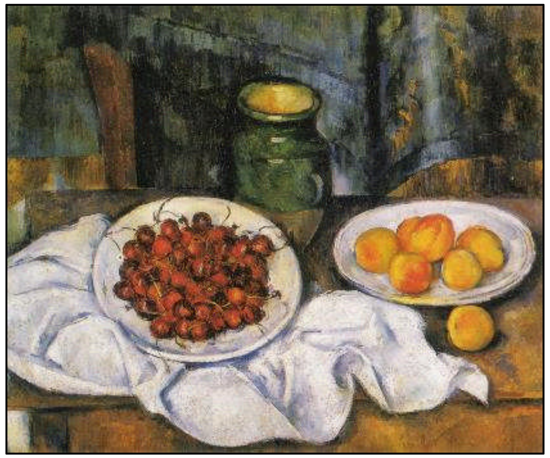




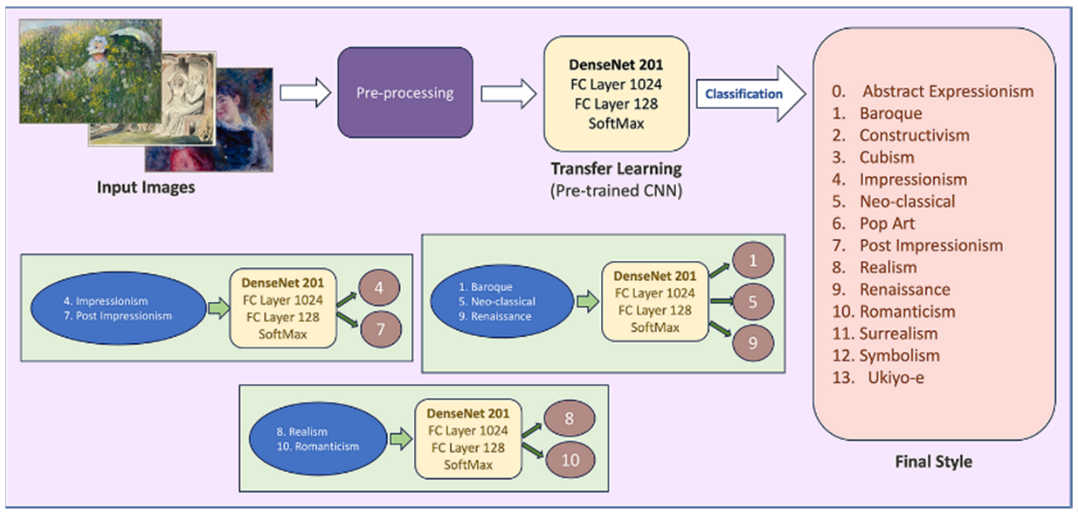
Amedeo Modigliani

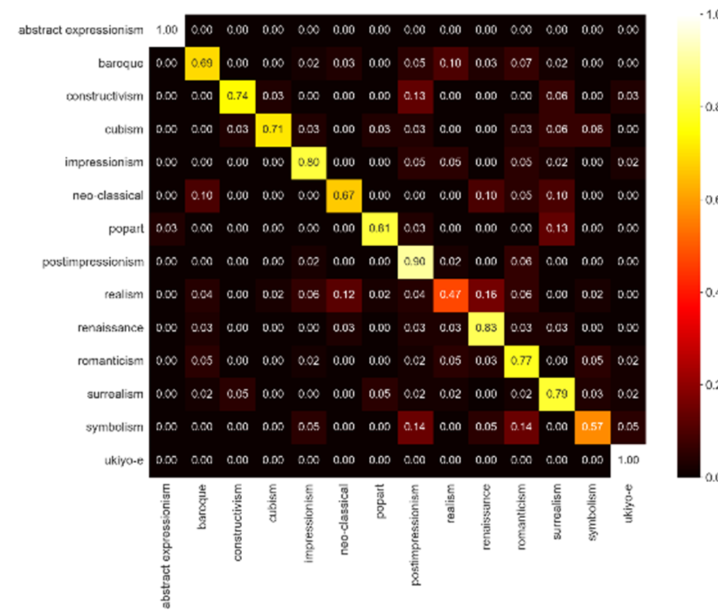
Our Current and Future Research

- Style Classification
- More heatmap analysis
- Interpretable results
- Try other techniques for generating boxes

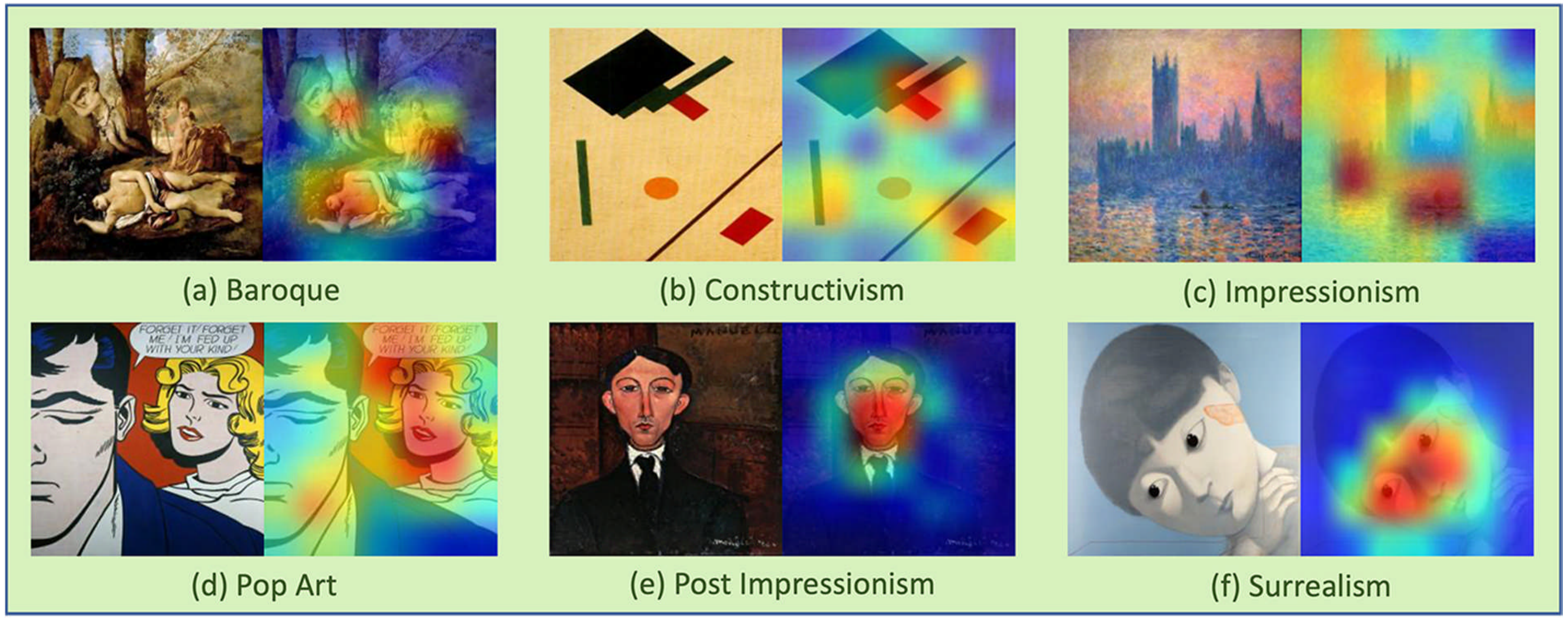
Style Classification (Work in Progress)





Heatmap Analysis and Interpretable Results

CNNs can detect which features of the image were most indicative of the style class that they were placed in.



(a) Baroque (b) Constructivism (c) Impressionism

(d) Pop Art (e) Post Impressionism (f) Surrealism

Window Selection Strategies

- We currently use randomly generated windows for the Artist recognition problem
- In future, we plan to use other strategies, like
  - Dense overlapping windows selected from a grid
  - Using a face detector to focus more on faces present in a painting
  - Using object detectors to draw mre windows containing full objects

Summary and Key Takeaways

- We curated a dataset of hi-resolution paintings from 100 diverse painters
- We used CNN-based techniques to explore the artist and style recognition problems
- We demonstrated that CNNs can recognize patterns present in art (and architecture) images that can lead to interpretable results.

Thank You!

Questions?

For further questions:

Atreyee Sinha
asinha@edgewood.edu

Sugata Banerji
banerji@lakeforest.edu