MLPR - AI 3011

Art Authentication Using ResNet-50 and GMM-Based Anomaly Detection

Deep Feature Statistical Modeling for One-Class Artistic Authentication

Ayushmaan, Ronak Tiwari, Shloka Srivastava

The Growing Problem of Art Forgery

CURRENT LANDSCAPE

Global art forgery costs billions annually.

Traditional authentication:

- Expert-based (subjective)
- Chemical analysis (slow, expensive)

Need: Fast, scalable, objective authentication system.

BBC

Home News Sport Business Technology Health Culture Arts Travel Earth Audio Vid

German police seize fake Picassos in multi-million euro forgery raid

24 October 2025 Share

Dearbail Jordan

Art crime // News

Auctioneer admits he helped create fake Basquiats seized by FBI in museum raid

Art crime // News

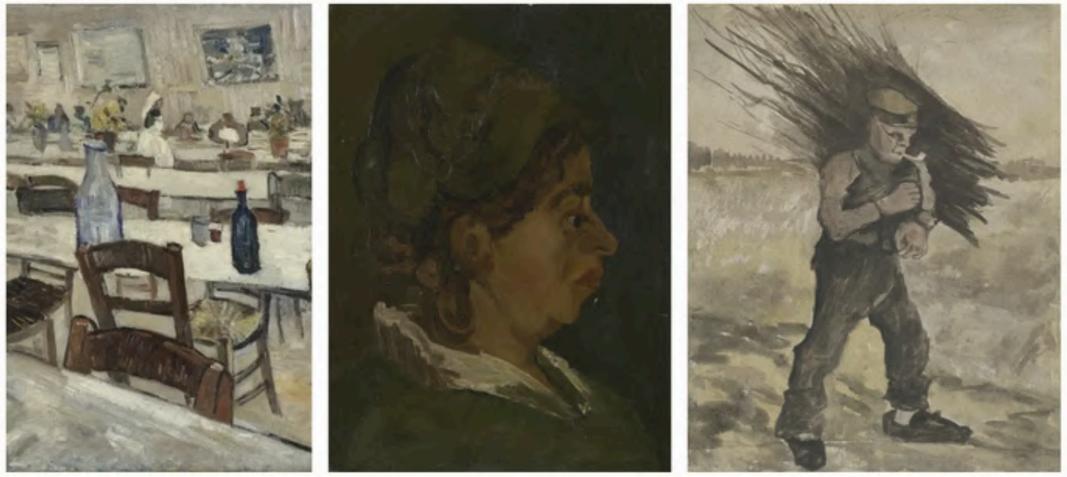
Pennsylvania man sentenced to prison for fraud scheme involving forged works by Picasso, Basquiat and Warhol

Art market // News

'Everything was fake but the money': forgers in Versailles chair scandal await sentencing

Van Gogh Museum exposes three early fakes

A vase of summer sunflowers in a late autumn scene proved a giveaway



Left: FAKE: *Interior of a Restaurant* (detail) (mid 20th century), private collection
Centre: FAKE: *Head of a Woman* (1902-09), private collection, France

Literature Review

Choudhury, 2021, Automated Identification of Painters Over WikiArt Image Data Using Machine Learning Algorithms

Applied **ResNet-50** on WikiArt for multi-class impressionist painter identification, achieving **75% accuracy** via style features; limited to 10 artists and supervised setup.

IMPLEMENTATION

- **Framework & Library:** The transfer learning pipeline was developed and executed using the PyTorch library.
- **Data Allocation:** Implementation was performed on a dataset split consisting of 3,800 training, 500 validation, and 700 test images.
- **Preprocessing & Augmentation:** Employed transformations and data augmentations (similar to ResNet-18) to enhance model generalization and robustness.

	precision	recall	f1-score	support
0	0.77	0.71	0.74	70
1	0.80	0.70	0.75	70
2	0.77	0.71	0.74	70
3	0.66	0.79	0.72	70
4	0.85	0.87	0.86	70
5	0.69	0.59	0.64	70
6	0.62	0.64	0.63	70
7	0.84	0.80	0.82	70
8	0.74	0.86	0.79	70
9	0.77	0.83	0.80	70
accuracy			0.75	700
macro avg	0.75	0.75	0.75	700
weighted avg	0.75	0.75	0.75	700

Figure 10(a) ResNet-50 Classification Report

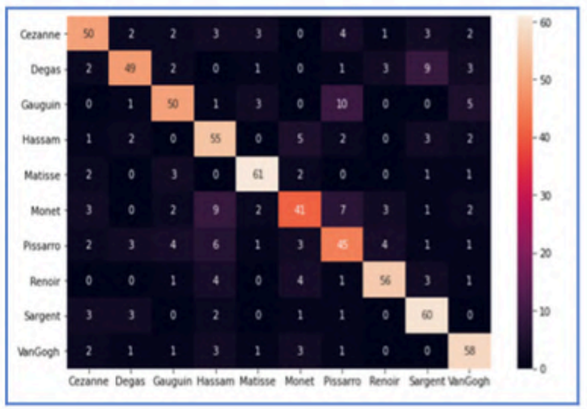
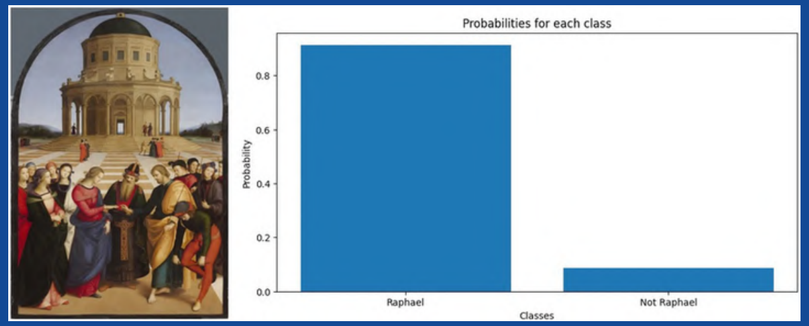


Figure 10(b) ResNet-50 Confusion Matrix

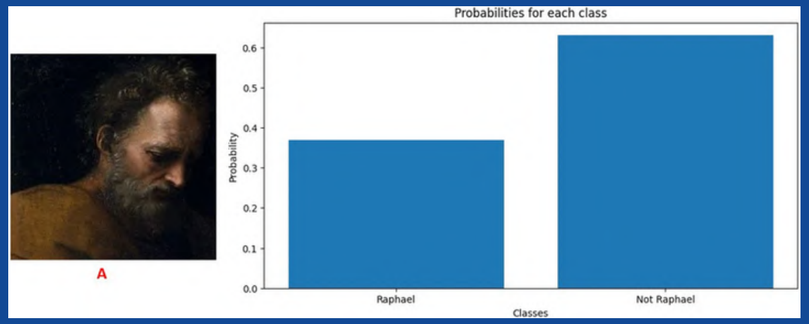
Gencarelli et al., 2023, Deep transfer learning for visual analysis and attribution of paintings by Raphael

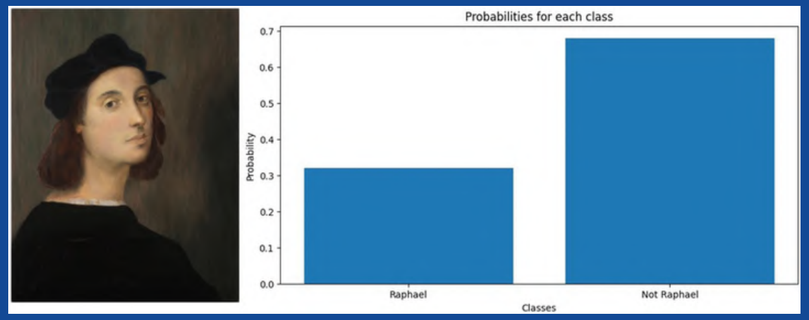
Used **ResNet-50 + SVM** for binary Raphael attribution, integrating edge detection for enhanced accuracy on authenticated vs. similar artists.



RESULTS

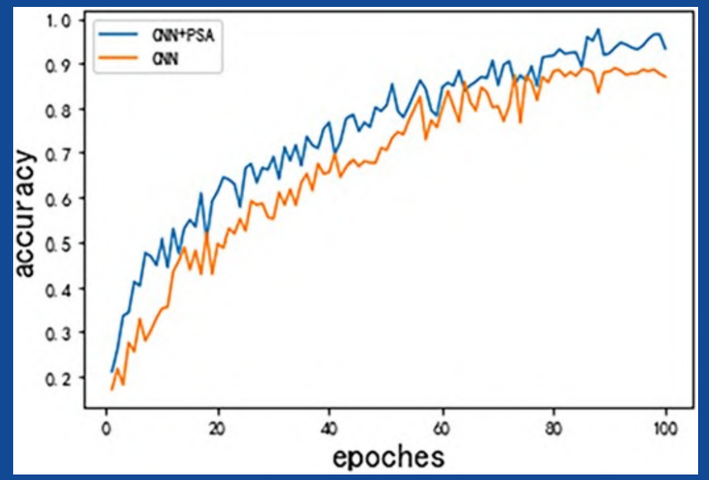
- **Hybrid Verification Pipeline:** Combines deep features from transfer learning with a secondary SVM classifier for robust attribution.
- **Edge Feature Integration:** Introduced a second layer of scrutiny using edge feature comparison to better distinguish authentic works from imitations.
- **Threshold-Based Scrutiny:** Established a precise edge feature threshold (T) based on the narrow statistical range found in Raphael's authenticated works





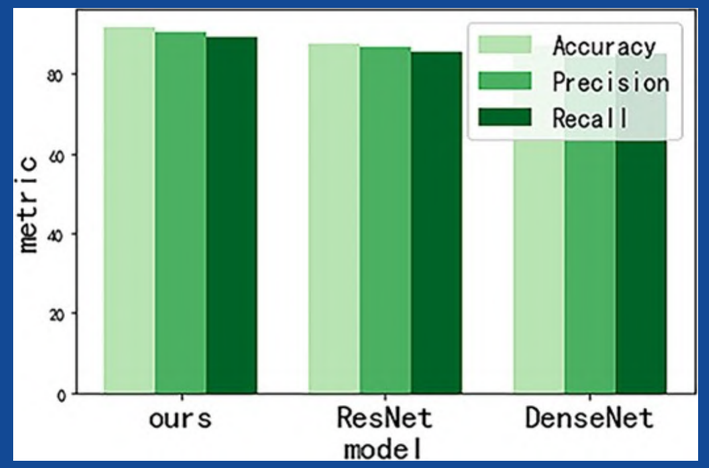
Afifi et al. , 2025 Research on automatic classification method of artistic styles based on attention mechanism convolutional neural network

CNN for Rubens workshop attribution, capturing micro-stylistic features on curated datasets; high accuracy but artist-specific.



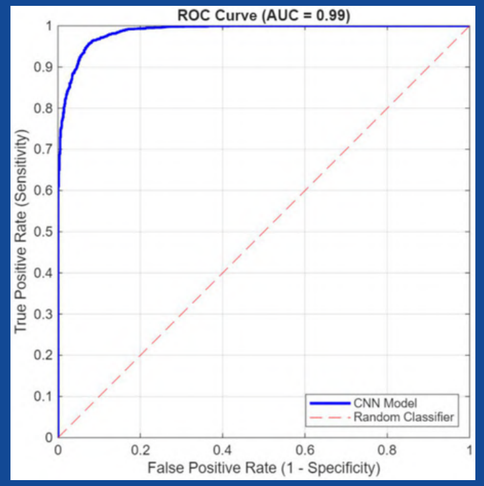
RESULTS and IMPLEMENTATION

- **High Classification Performance:** Achieved 91.52% accuracy, 90.49% precision, and 89.09% recall on the WikiArt dataset, outperforming ResNet and DenseNet.
- **Improved Feature Extraction:** PSA enhanced spatial awareness and multi-scale feature learning, leading to better artistic style classification.
- **Future Scalability:** Future improvements include refining the PSA module and integrating additional attention mechanisms for stronger generalization and accuracy.



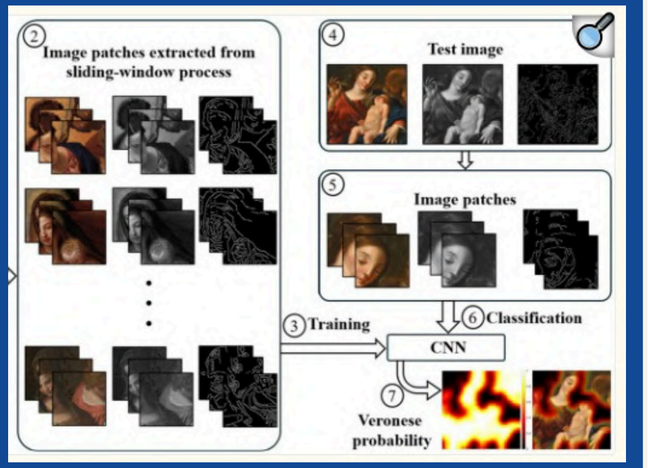
Sabino et al. , 2026, Painting authentication using CNNs and sliding window feature extraction

CNN with sliding windows for Veronese authentication, **94.5% accuracy** via stylistic coherence heatmaps.



RESULTS

- **Methodological Innovation:** Developed a patch-based CNN pipeline utilizing sliding window feature extraction and multichannel inputs (RGB, grayscale, and edge maps) to address data scarcity in stylistic analysis.
- **Authentication Accuracy:** Successfully distinguished authentic Veronese works from non-Veronese paintings, achieving painting-level probabilities between 80.3% and 99.9% for genuine works.
- **Visual Fingerprinting:** Authentic works exhibited broad, contiguous high-probability regions in probability heatmaps, whereas non-Veronese paintings showed fragmented, localized high-probability zones.
- **Robust Validation:** Implemented painting-level cross-validation to prevent data leakage and used localized patches with regularization to enhance the model's stylistic fidelity and generalization.



Literature Review Limitation & Project Extension

Problem 1: "Closed-Set" Classification

- **Previous Gap (Choudhury, 2021):** Models were forced to classify every image into a known artist category, making them unable to detect forgeries.
- **Our Extension:** We transitioned to Open-Set Anomaly Detection. The model rejects "UNKNOWN" paintings that fall outside the learned statistical boundaries.

Problem 2: Global Information Loss

- **Previous Gap (Sabino et al., 2026):** Resizing large canvases to small fixed dimensions discards the fine-grained brushstroke detail needed for authentication.
- **Our Extension:** We implemented Dual-Pooled Patching (Mean + Max pooling). This preserves both the overall style and the unique "textural peaks" that characterize a master's hand.

Problem 3: Threshold Subjectivity

- **Previous Gap (Gencarelli et al., 2023):** Reliance on manual, case-by-case threshold tuning for specific features like edge comparison.
- **Our Extension:** We established an Objective Statistical Boundary. Using GMM log-likelihood scores and a k=2.5 rejection threshold, the system provides automated, data-driven verdicts.

Problem 4: Specificity vs. Generalization

- **Previous Gap (Afifi et al., 2025):** Research often focuses on a single artist, making the pipeline difficult to scale to museum-wide archives.
- **Our Extension:** Our architecture generalizes across 22 diverse masters simultaneously by using Diagonal Covariance GMMs to manage high-dimensional data efficiently.

DataSet

WikiArt (Kaggle)

It is a dataset containing **verified digitized paintings from multiple artists** and artistic movements. Includes metadata such as artist name, genre, and style, making it suitable for computational analysis of artistic patterns.

Images were compiled from publicly available museum archives and online art collections, and are used strictly for academic research and non-commercial analysis.

Statistics

- Total images: 81,444 paintings
- Total unique artists: 1,178
- Total genres: 27

Filtering Process

Only artists with **>= 400 images** are retained. Every qualified artist is then capped at exactly 4**00 images, randomly sampled.**

Final dataset:

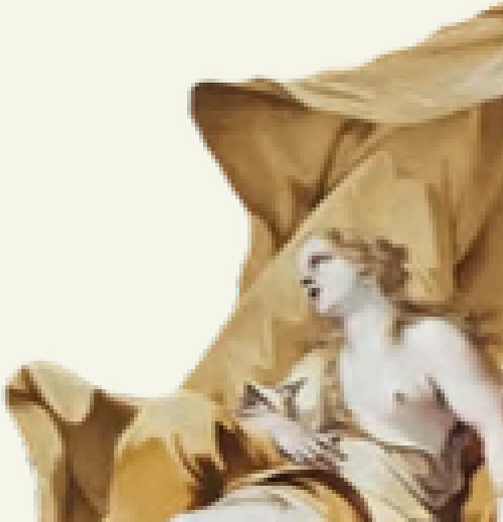
- 8,800 images
- 22 artists

Train–Test Split

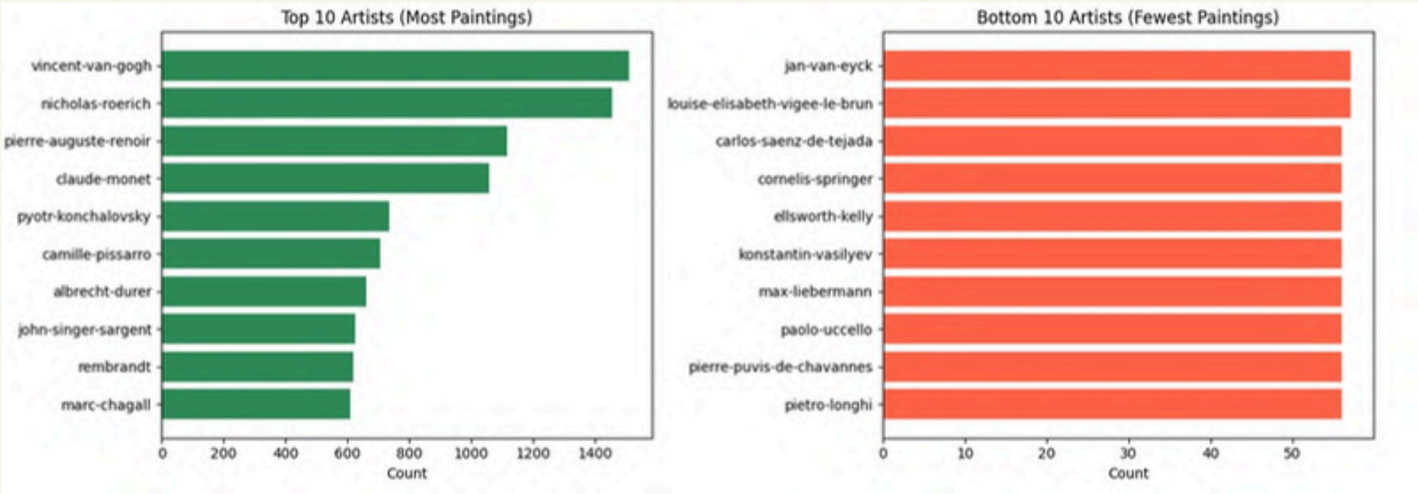
- Training set: 6,160 images (70%)
- Test set: 1,320 images (15%)
- Validation set: 1,320 images (15%)

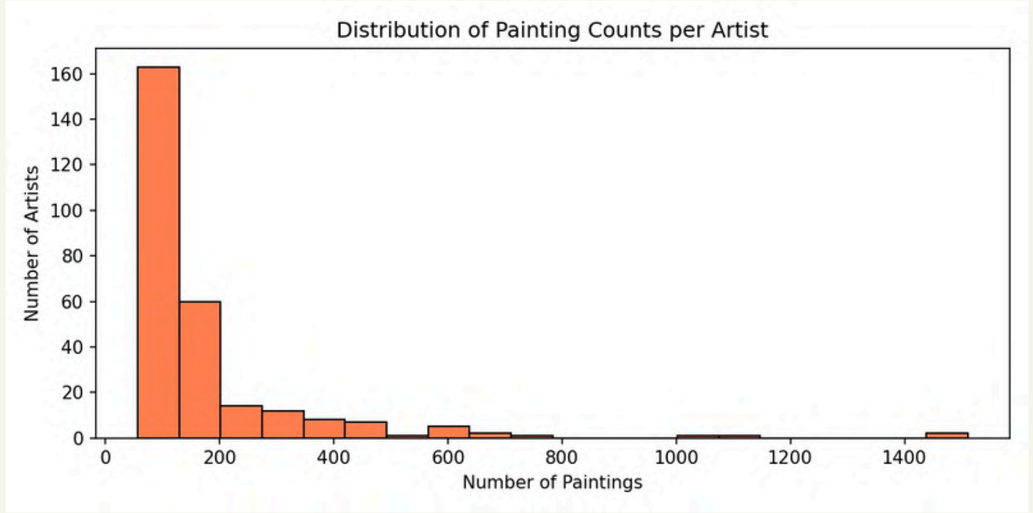
We chose this dataset because:

1. It is **widely used in existing literature** and research papers, including the ones we referred to, making it a reliable benchmark.
2. It offers **diversity in artistic styles and artists**, which helps in building robust models.
3. It is well-suited for deep learning approaches (e.g., CNNs like ResNet) due to its **structured labeling and sufficient size.**

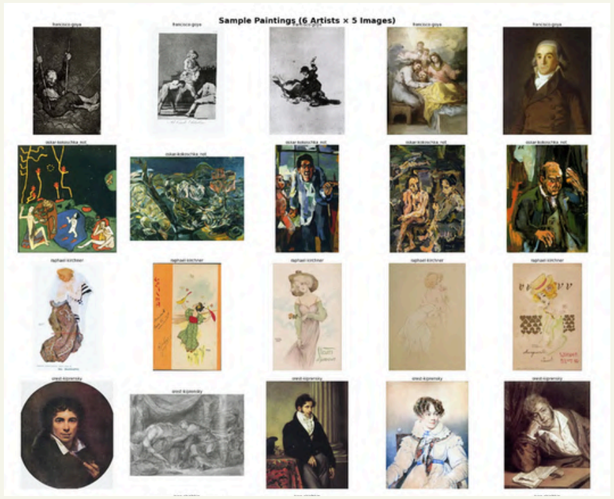


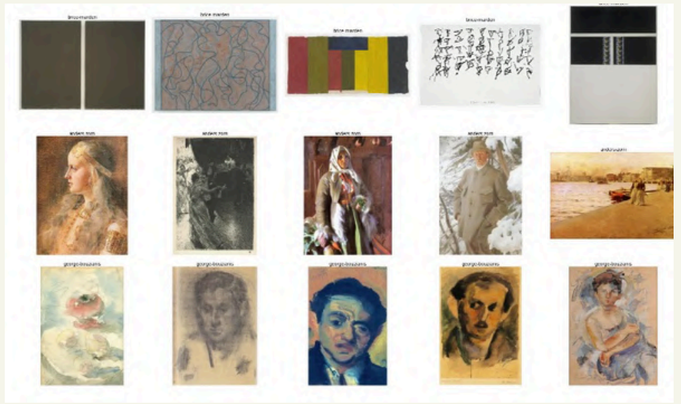
Data Diversity & Imbalance

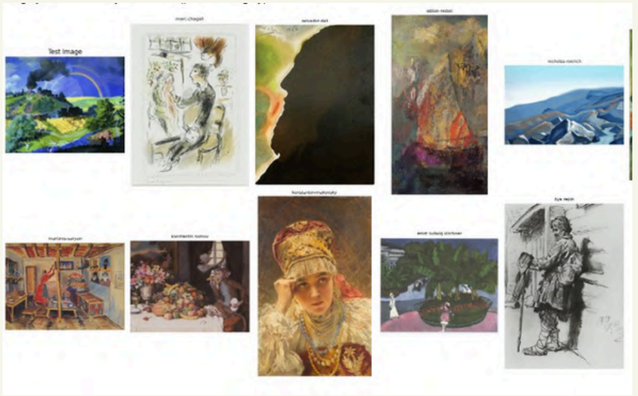




Visualizing the WikiArt imbalance: Our 'Filter-and-Cap' strategy was essential to prevent the model from being biased toward prolific artists like Van Gogh.

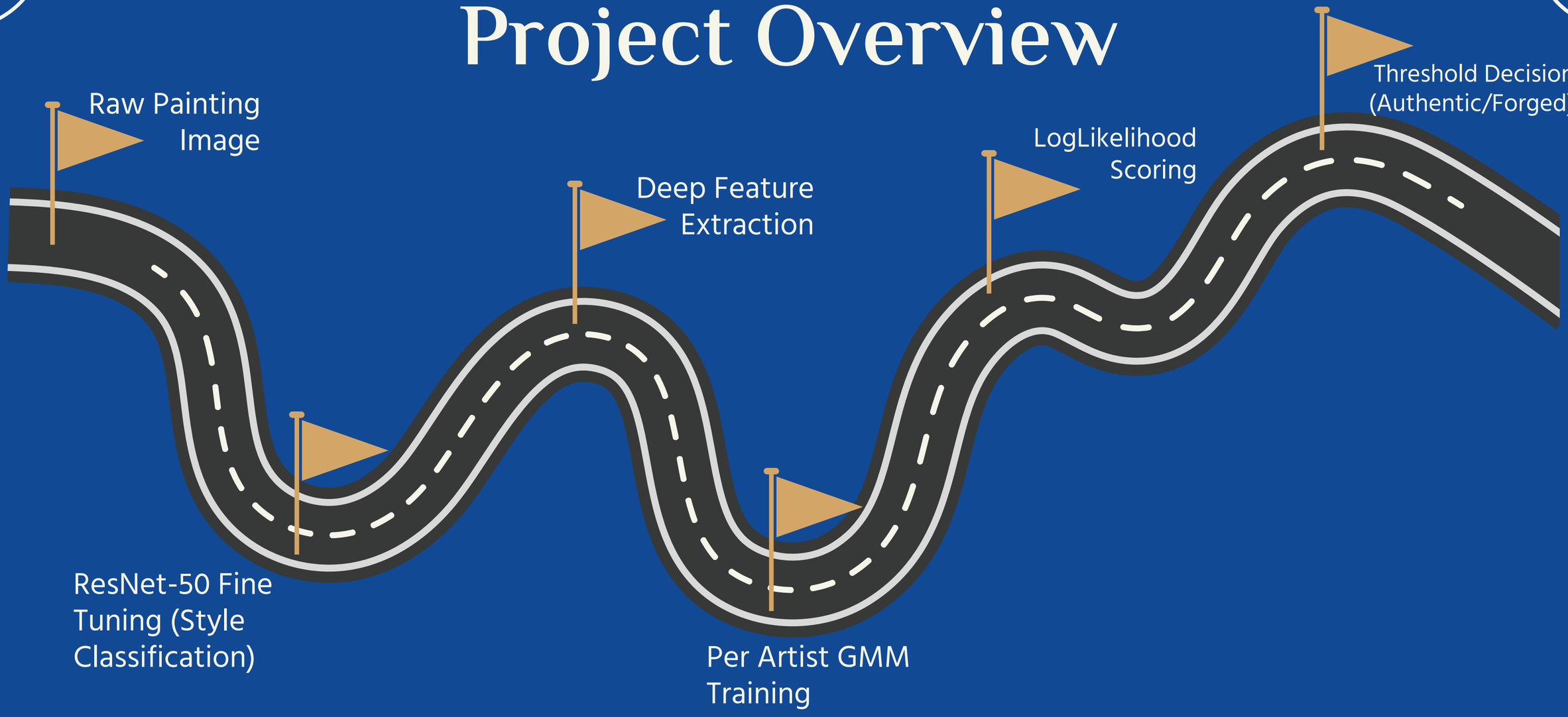






A sample of the training manifold: The model must internalize a wide spectrum of textures, from dark Mannerist portraits to vibrant Impressionist landscapes.

Project Overview



The CNN learns **what makes each artist distinctive**, and the GMM learns **how tightly clustered** those features are, enabling rejection of out-of-distribution samples.

Feature Extraction

Model Architecture: ResNet-50

A **50-layer deep convolutional neural network** featuring residual connections. It consists of an

1. Initial 7×7 stem 2. Four residual stages. 3. Global average pooling

initialized with **ImageNet pre-trained weights** learned from 1.2 million images.

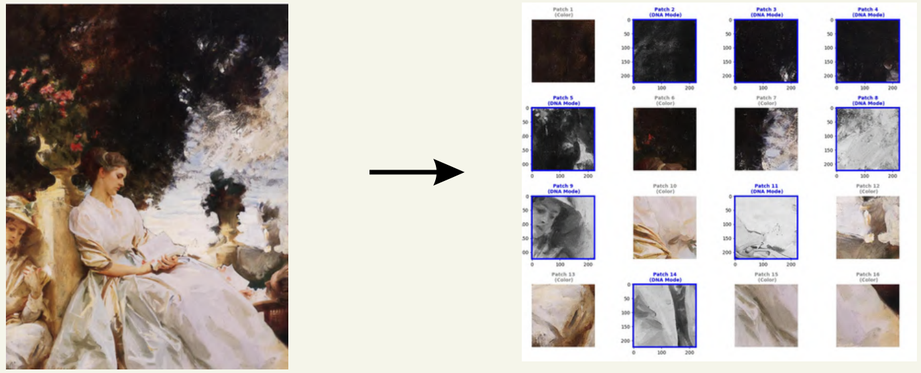
Why we chose ResNet- 50?

- **Robust Feature Hierarchy:** ResNet-50 provides a clear hierarchy of features from basic edges in early layers to complex textures in later ones.

- **Transfer Learning Efficiency:** By using ImageNet pre-trained weights, the model already understands fundamental visual concepts (textures, shapes, and gradients).

Feature Preprocessing

- **Geometric Uniformity:** All images were resized and center-cropped to **224×224 pixels** to maintain a standard input manifold for the ResNet backbone.
- **Normalization:** Pixel values were scaled using **ImageNet mean and standard deviation** to align with the pre-trained weights of the feature extractor.
- **Augmentation (Training Only):** Used *RandomResizedCrop* and *ColorJitter* to make the features robust against photographic variations and reproduction quality.



Dimensionality Strategy

- **4096-d Style Vector:** We engineered a high-dimensional embedding by concatenating **Mean Pooling** (average style) and **Max Pooling** (textural peaks).
- **Diagonal Covariance Strategy:** To handle the "Curse of Dimensionality," we used *covariance_type='diag'* in our GMMs. This reduced the estimated parameters from ~16M to ~4096 per component.
- **Implicit Feature Importance:** The pipeline relies on the **ResNet-50 hierarchy**:
  - **Early Layers:** Capture low-level textures and brushstrokes.
  - **Deep Layers:** Capture high-level semantic style fingerprints.
- **Interpolation:** No traditional data interpolation (like regression) was required, as the patch-extraction stride (112 pixels) ensured 50% overlap for continuous feature coverage.

Training Pipeline

PHASE 1:

Clasification Training

Data Pipeline

- **Source:** WikiArt large-scale dataset.
- **Filter:** Artists with **≥200 paintings (~55 artists).**
- **Balance:** Capped at **200 images/artist** (11,000 total) to ensure unbiased learning.
- **Split:** 70% Train | 15% Val | 15% Test.

Training Protocol

- **Architecture:** ResNet-50 with ImageNet pre-trained weights.
- **Fine-tuning:** All 50 layers optimized for 20 epochs.
- **Optimizer:** SGD (lr=0.001, momentum=0.9) using CrossEntropyLoss.
- **Augmentation:** Training (Random) | Validation (CenterCrop).

PHASE 2:

Embedding Extraction

Extraction Pipeline

- **Input Layer:** Verified training images from the balanced WikiArt subset.
- **Patch Extraction:** Canvas divided into 224×224 patches using a stride of 112, capped at 48 patches per painting.
- **Feature Backbone:** ResNet-50 architecture with the final Fully Connected (FC) classification head removed.
- **Local Representation:** Extraction of a 2048-dimensional embedding for every individual patch.

Encoding and Normalization

- **Initial Normalization:** L2 normalization applied to each 2048-d patch embedding for forensic consistency.
- **Global Pooling:** Sequential application of Mean Pooling and Max Pooling across all patches.
- **Feature Fusion:** Concatenation of pooled results into a robust 4096-dimensional style vector.
- **Final Output:** L2 normalization of the 4096-d vector; results stored as embeddings.npy and labels.npy.

Training Pipeline

PHASE 3:

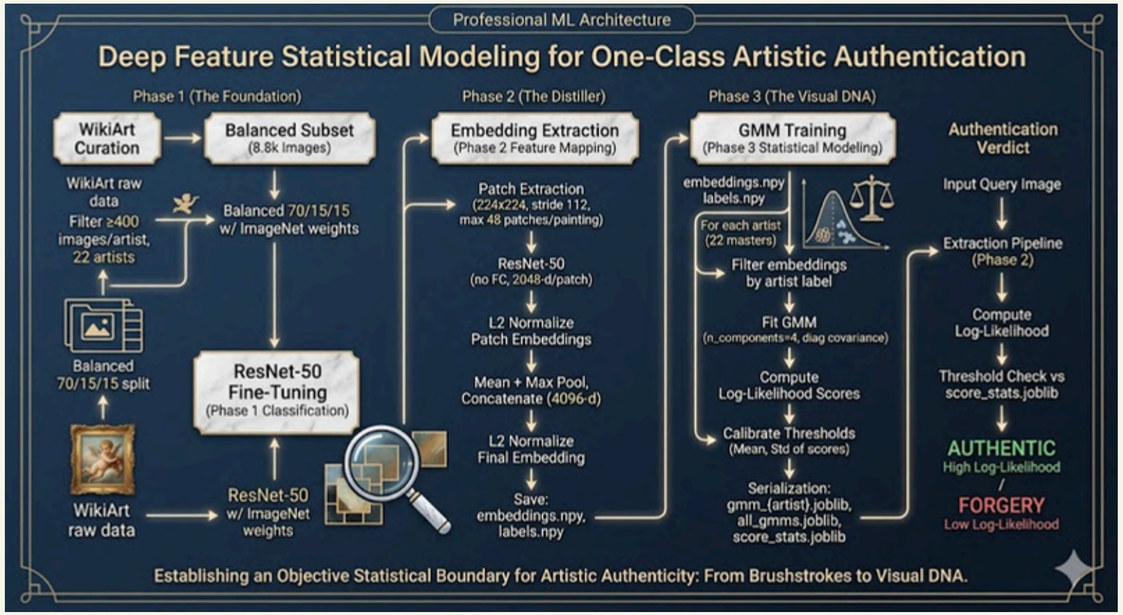
GMM Training

Embedding Extraction

- **Input Data:** Utilizes the pre-computed *embeddings.npy* and *labels.npy* from Phase 2.
- **Artist Isolation:** Systematically filters the 4096-d embeddings for each of the **55 unique artists.**
- **GMM Implementation:** Fits a **Gaussian Mixture Model** for every artist with:
  1) **Components:** *n_components* = 4 to model stylistic variance.
  2) **Covariance:** *covariance='diag'* for computational efficiency and stability
- **Artifact Generation:** Individual models are serialized as *gmm_{artist}.joblib*.

Threshold Calibration

- **Scoring Engine:** Computes **Log-Likelihood scores** for all authentic embeddings in the training set.
- **Statistical Baseline:** Calculates the **Mean** and **Standard Deviation** of these scores per artist.
- **Rejection Thresholds:** These statistics are saved as *score_stats.joblib* to define the mathematical boundary between "Authentic" and "Anomaly."
- **System Integration:** Final consolidation of all artist models into a master *all_gmms.joblib*.



AI Generated image of our Model Architecture

Performance Metrics

```
======================================================================
CLASSIFICATION REPORT (Figure 10a)
======================================================================
                     precision    recall  f1-score   support

      albrecht-durer       0.80      0.92      0.85        60
      boris-kustodiev      0.76      0.68      0.72        60
    camille-pissarro       0.80      0.75      0.78        60
       childe-hassam       0.77      0.92      0.84        60
        claude-monet       0.72      0.82      0.77        60
         edgar-degas       0.83      0.72      0.77        60
       eugene-boudin       0.90      0.93      0.92        60
      gustave-dore         0.95      0.93      0.94        60
         ilya-repin        0.75      0.82      0.78        60
     ivan-aivazovsky       0.95      0.97      0.96        60
       ivan-shishkin       0.88      0.93      0.90        60
  john-singer-sargent      0.90      0.77      0.83        60
        marc-chagall       0.91      0.80      0.85        60
     martiros-saryan       0.87      0.78      0.82        60
    nicholas-roerich       0.81      0.70      0.75        60
       pablo-picasso       0.75      0.77      0.76        60
        paul-cezanne       0.86      0.85      0.86        60
pierre-auguste-renoir      0.90      0.87      0.88        60
   pyotr-konchalovsky      0.93      0.93      0.93        60
    raphael-kirchner       0.86      0.93      0.90        60
         rembrandt         0.83      0.92      0.87        60
    vincent-van-gogh       0.78      0.78      0.78        60

           accuracy                           0.84      1320
          macro avg       0.84      0.84      0.84      1320
       weighted avg       0.84      0.84      0.84      1320
```

Final Quantitative Summary: 84.02% accuracy in high-dimensional style classification.

0.8399
Precision

How often predicted artists are correct.

0.8376
Recall

How many real artist paintings are correctly found.

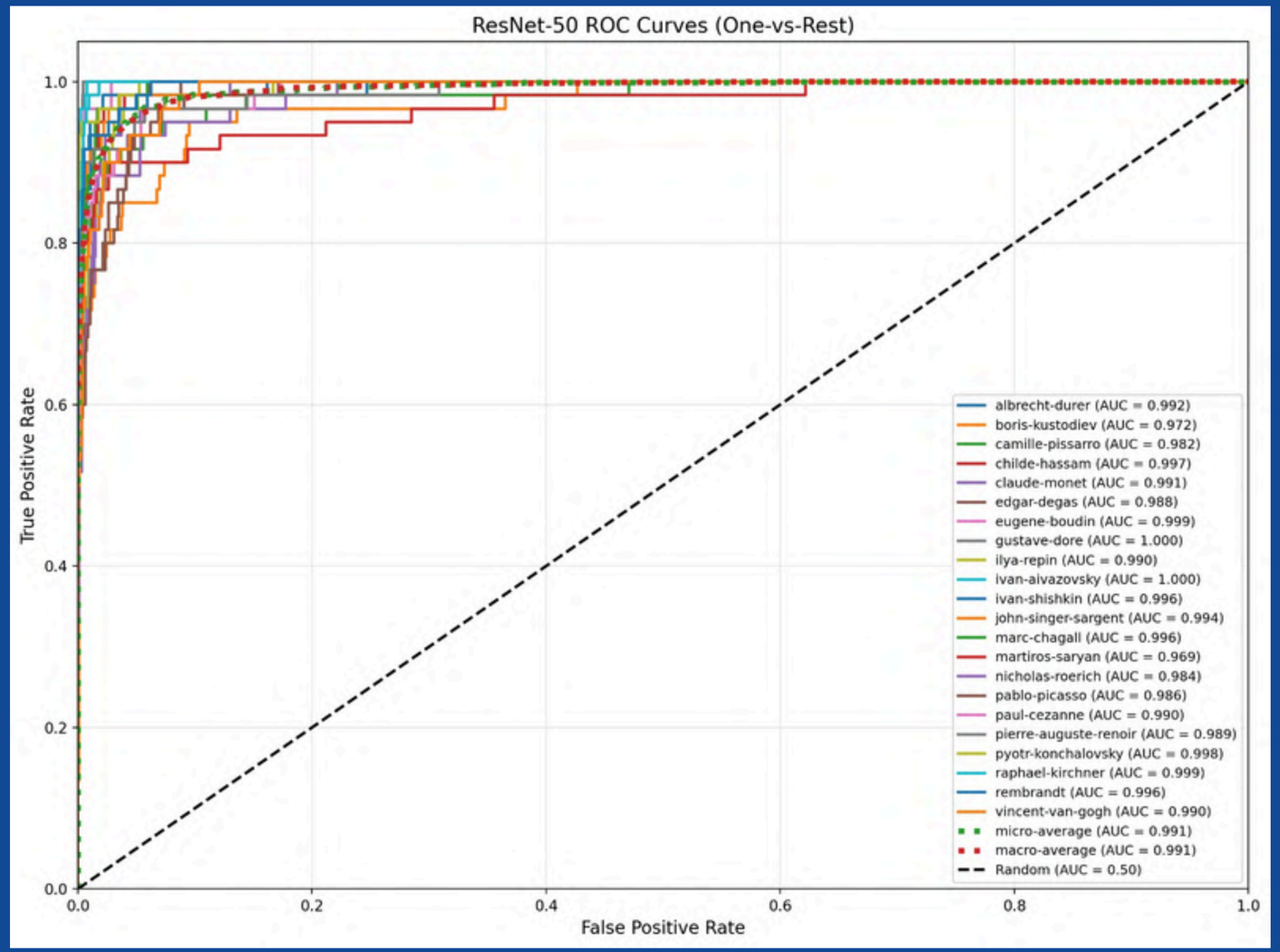
0.84
F1-Score

Balance between precision and recall.

0.5379
Log Loss (Cross Entropy)

Measures confidence quality of predictions (lower is better).

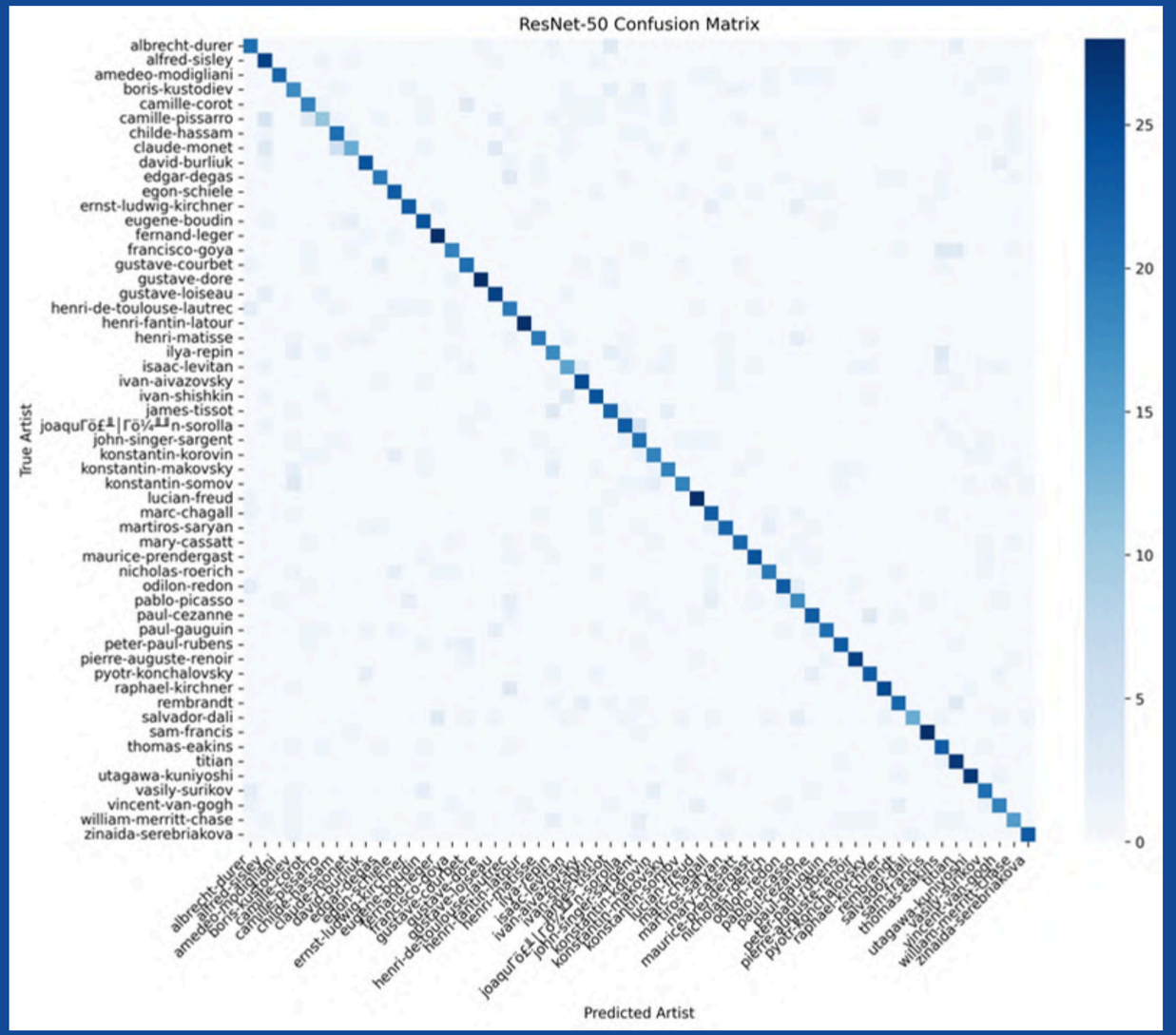
ROC Analysis



One-vs-Rest Performance: The near-perfect AUC scores (0.97–1.00) demonstrate that the ResNet backbone provides an elite feature space for artist differentiation.

ResNet-50 Confusion Matrix

22 * 22



Error Analysis: Most classification overlaps occur between artists of similar movements (e.g., Monet and Pissarro), proving the model is learning stylistic nuances.

GMM Confusion Matrix

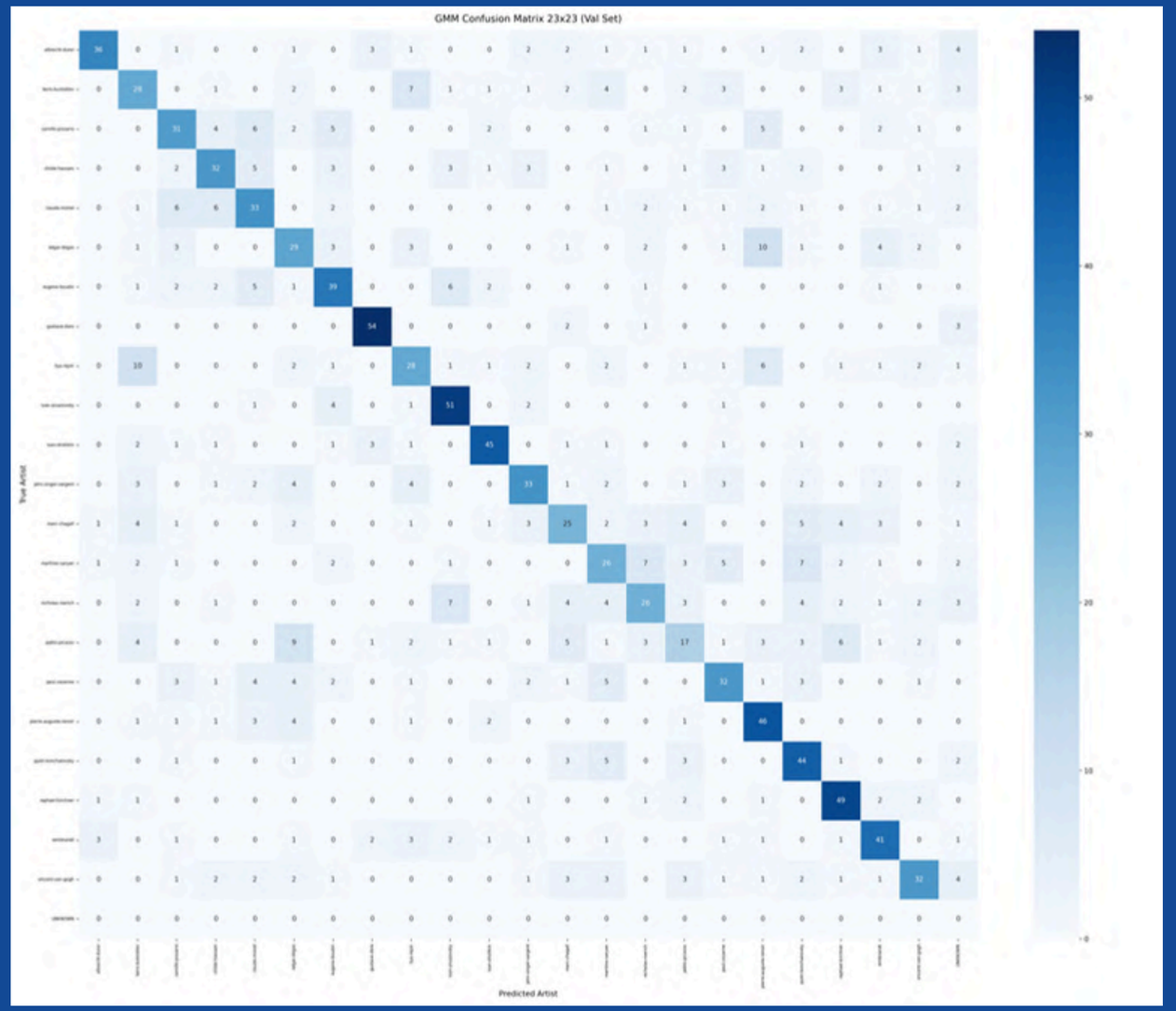
23 * 23

Specificity: 0.9813

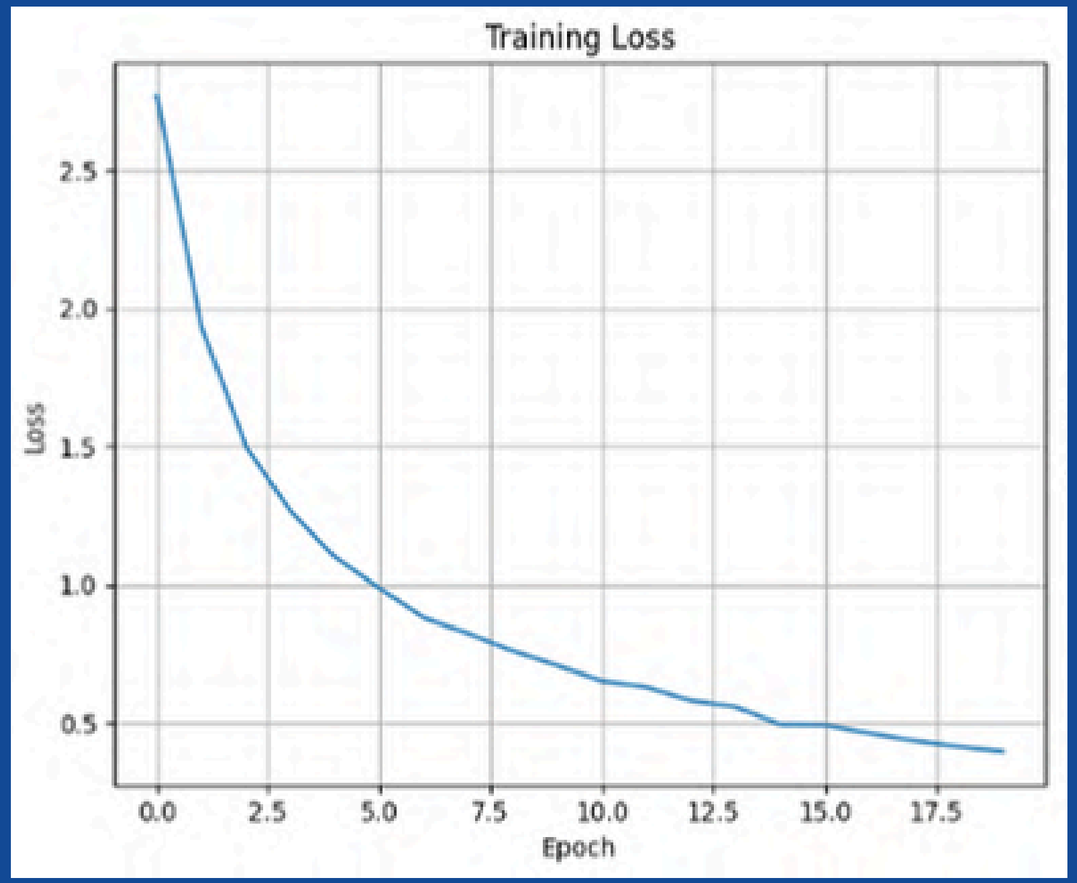
Ability to correctly reject paintings that do NOT belong to a given artist

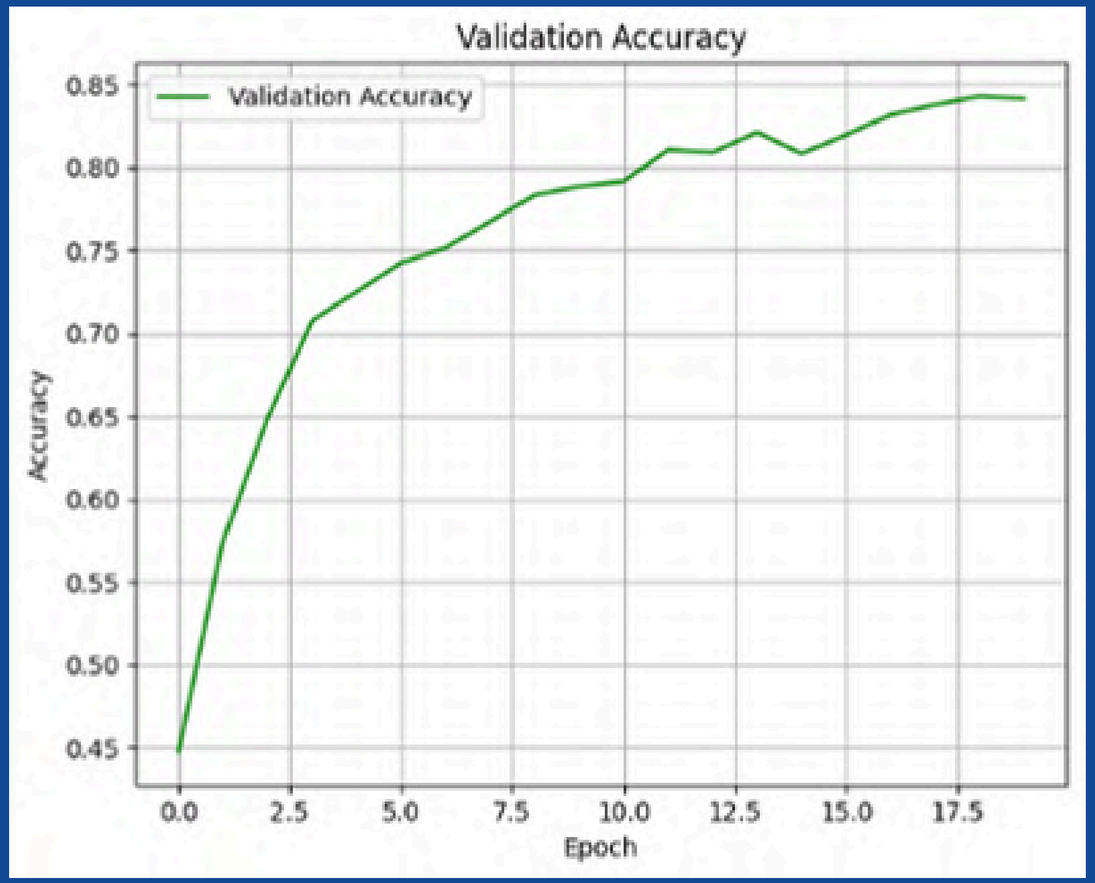
Rejection Rate : 32/1320

2.4% flagged as UNKNOWN



Training Loss & Vaidation Accuracy





Training loss steadily decreases while validation accuracy consistently improves, indicating effective learning and good generalization.

Challenges & Strategic Solutions

Scarcity of Forgery Data

The fundamental difficulty in art forgery detection as a machine learning problem is the **absence of a labeled forgery dataset**. Real documented forgeries are rare, legally sensitive, and not publicly available at scale.

The project addresses this by:

- Treating **low-likelihood images** as forgeries (one-class classification under GMM)
- Using **cross-artist samples** as implicit negative examples during the threshold calibration

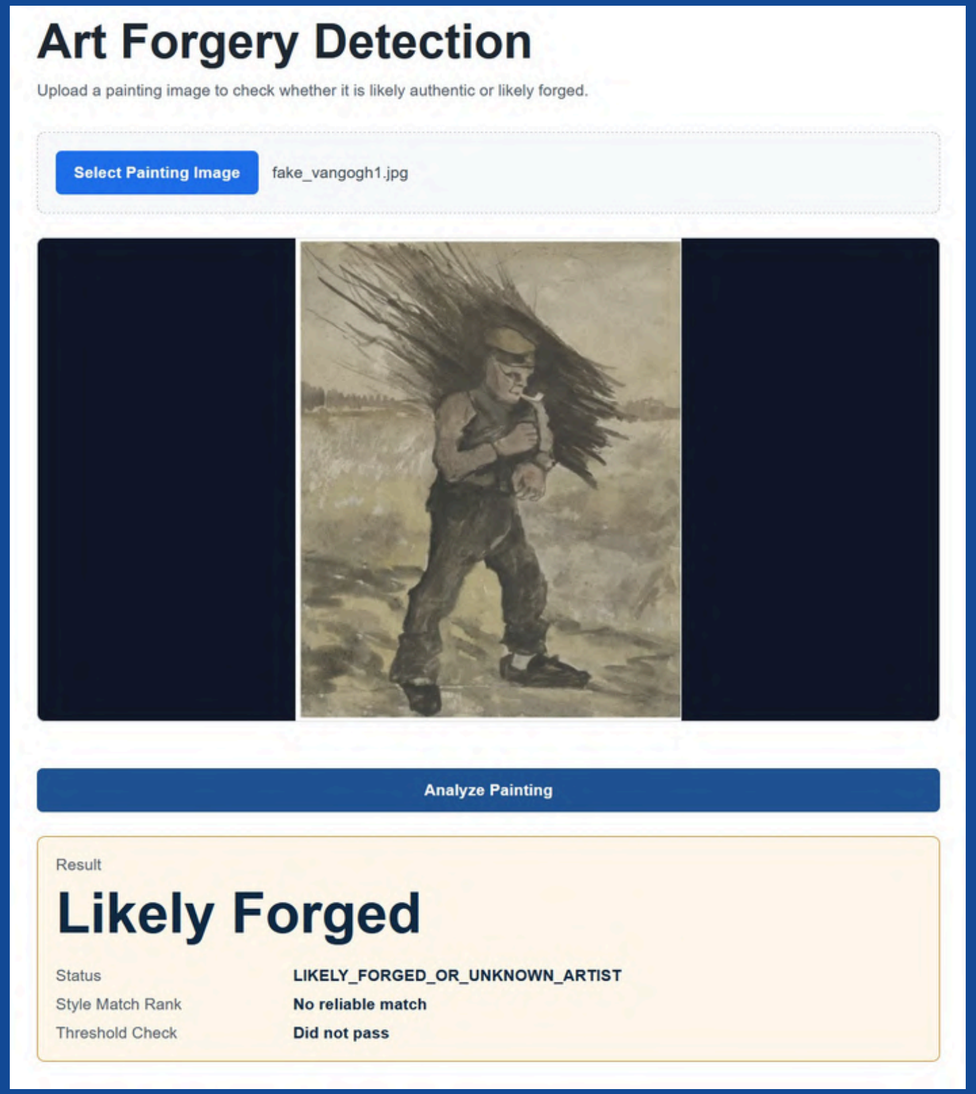
Inter-Class Style Similarity

The fundamental difficulty in artist classification is the **overlap in feature distributions between artists with very similar styles**. When stylistic characteristics closely resemble each other, their GMM likelihood distributions intersect, leading to ambiguity and confusion in classification.

The project addresses this by:

- **Tuning the rejection threshold** to better separate overlapping likelihood regions
- **Analyzing confusion patterns** to calibrate decision boundaries more effectively
- Considering **Vision Transformers (ViTs)** as a future improvement for stronger style representation and separation

Deployment



Deployment Setup

- **Frontend:** Next.js web interface for image upload and result display
- **Backend:** FastAPI inference server running the trained forgery detection pipeline
- **Model Pipeline:** ResNet-50 feature extraction + GMM artist-style matching
- **Input:** Painting image in JPG, PNG, JPEG, or WEBP format
- **Output:** Final screening verdict with closest artist matches
- **Verdicts:** Likely Authentic, Likely Authentic - Expert Review Recommended, Likely Forged

References & Sources

Research Papers:

Paper 1: Choudhury, 2021, Automated Identification of Painters Over WikiArt Image Data Using Machine Learning Algorithms

Paper 2: Gencarelli et al., 2023, Deep transfer learning for visual analysis and attribution of paintings by Raphael

Paper 3: Afifi et al. , 2025 Research on automatic classification method of artistic styles based on attention mechanism convolutional neural network

Paper 4: Sabino et al. , 2026, Painting authentication using CNNs and sliding window feature extraction

Dataset:

https://www.kaggle.com/datasets/steubk/wikiart

Github Repository:

https://github.com/AyushmaanWorks/Art-Forgery-Screening-tool