

# Deep Learning-Based Approach for Distinguishing Between AI-Generated and Human-Drawn Paintings

A.I. Warnakulasooriya<sup>1\*</sup>, R.A.H.M. Rupasingha<sup>2</sup>, B.T.G.S. Kumara<sup>3</sup>

<sup>1</sup>*Dept. of Computing and Information Systems, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka, aiwarnakulasooriya@std.appsc.sab.lk*

<sup>2</sup>*Dept. of Economics and Statistics, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka, hmrupasingha@gmail.com*

<sup>3</sup>*Dept. of Data Science, Sabaragamuwa University of Sri Lanka, Belihuloya, Sri Lanka, kumara@appsc.sab.ac.lk*

With the increasing number of robust Artificial Intelligence (AI) art generation applications, more realistic AI-generated paintings (AIGP) are emerging, causing a critical impact on artists. Due to the wide acceptance of AIGP, the cultural, historical and monetary value of the real masterpieces of art is becoming doubtful along with the arising question of the significance of human painters and their mastered techniques. To protect the artist's rights, it is crucial to differentiate AIGPs from human-drawn paintings (HDP). Accordingly, the main objective of this research is to develop a Convolutional Neural Network (CNN) model that can automatically distinguish between AI-generated and human-drawn paintings without requiring human intervention. Without limiting to the pixel analysis like the previous literature, more features such as edge patterns, object arrangements, pattern distributions, etc. with also the gradient features were considered by the proposed model in the painting classification. A diverse set of 3000 paintings from the AI-ArtBench Dataset comprising 1500 AIGPs and 1500 HDPs across 10 different art themes was collected and pre-processed for this study. AIGPs were generated in equal amounts using Latent Diffusion and Standard Diffusion Models. The implemented CNN model achieved an optimum classification accuracy of 90% at the training data size of 10% while the ANN model exhibited 77% accuracy at the same training data size. Furthermore, models were compared using performance metrics such as precision, recall, F1-score, RMSE and MAE. Through a Gradient-weighted Class Activation Map (Grad-CAM), the key features that the CNN model uses to differentiate between AIGPs and HDPs were identified. These findings emphasize the potential capability of automated systems to detect AIGPs vs HDPs for art authentication purposes. In future works, the model's performance on different art styles and the key features specified for them will be analyzed.

**Keywords:** *AI-generated paintings, ANN, Classification, CNN, Human-drawn paintings*