

ART MARKET PREDICTIONS THROUGH DEEP LEARNING

Sayantani De ¹, Sadhana Sargam ², Pratik Shrivastava ³, Jayapriya Mahesh ⁴, Dr. Podilapu Hanumantha Rao ⁵, Nimesh Raj ⁶, Payal Sunil Lahane ⁷

¹ Assistant Professor, Department of Computer Science and IT, ARKA JAIN University Jamshedpur, Jharkhand, India

² Assistant Professor, School of Business Management, Noida International University, Greater Noida, Uttar Pradesh, India

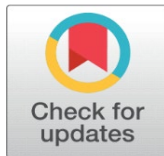
³ Assistant Professor, Department of Design, Vivekananda Global University, Jaipur, India

⁴ Assistant Professor, Department of Computer Science and Engineering, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation (Du), Tamil Nadu, India

⁵ Assistant Professor, Department of Commerce and Management Studies, Andhra University, Visakhapatnam, Andhra Pradesh, India

⁶ Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India

⁷ Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037 India



Received 09 April 2025

Accepted 16 August 2025

Published 25 December 2025

Corresponding Author

Sayantani De,

sayantani.de@arkajainuniversity.ac.in

DOI

[10.29121/shodhkosh.v6.i4s.2025.6847](https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6847)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

ABSTRACT

The international art market lends complex dynamics that interact with aesthetic perception, the cycle of economic activities and the mood of the investor, making it a difficult task to forecast prices. The paper presents a deep learning architecture that combines visual, contextual, and temporal data to predict the valuations of artworks with further accuracy. The study proposes a data engineering pipeline that is multimodal that includes curated image collections and structured information, including artist background, sales history, medium, and dimensions. A convolutional neural network (CNN) is used to produce high-level structure of artistic style and quality, whereas transformer and Long Short-Memory (LSTM) structures discover the temporal dynamics of price tendencies in the past. These modalities are amalgamated into a single embedding that is a combination of both visual and economic cues a fusion layer. This model is hyperparameter tuned and transferred learned with the help of pretrained encoders to optimize prediction accuracy and prevent overfitting with regularization measures. Findings indicate better performance compared to traditional econometric and regression models and better correlation with the real market trends and overall generalization across genres and time series. In addition to predictive ability, the framework offers interpretable information suggesting the impact of artistic qualities on valuation trends, therefore, linking computational intelligence to the art economics. The suggested system provides possible solutions in market analytics, auction prediction, and digital art investment platforms, which add to the development of data-driven decision-making in the creative economy.

Keywords: Art Market Prediction, Deep Learning, Multimodal Fusion, CNN-LSTM Model, Price Forecasting, Cultural Analytics



1. INTRODUCTION

Historically, the art market has been viewed as a market that operates on subjective preferences, cultural favor and emotional appeal, and not objective analytics. Nevertheless, the growing digitization of art transactions, along with the access to large scale datasets of auctions, exhibitions, and artist profiles has provided new opportunities in the quantitative modeling of art valuation. The conventional approaches to price estimation are hinged on hedonic regression, repeat-sales indices, or subjective assessment by the experts, which usually do not offer sufficient explanations of the intricate interaction between visual beauty, the name of the artist, and time changes, as well as market sentiments. This has led to the emergence of efforts by researchers and investors to decode such multifaceted relationships using the methods of artificial intelligence (AI) and deep learning (DL) to forecast future price trends of artworks [Shao et al. \(2024\)](#). Deep learning provides an alternative approach to the art market and its understanding. Convolutional Neural Networks (CNNs) have the ability to automatically identify complex visual features within paintings and sculptures with features like the pattern of a brushstroke, color arrangement, style resemblance, and perceived quality. These latent representations are also combinable with non-visual data, such as the statistics of the careers of the artists, exhibition records, and auction histories, to create a multimodal concept of value. The recurrent architectures, such as Long Short-term Memory (LSTM) networks or Transformer models, also allow the system to extract temporal effects in the market trends, thus identifying the patterns that can be disregarded by the traditional econometric models [Leong and Zhang \(2025\)](#). A combination of these elements leads to a smart structure that can combine visual and financial data to create more predictable and interpretable forecasts.

The use of machine learning in art valuation is consistent with the more general trends of financialization of cultural assets. Art is also being regarded as not just a source of aesthetic and aesthetic enjoyment but as an investment tool with returns that may well compete or supplement conventional assets. It is therefore important to the investors, galleries, and collectors to forecast the price movements to maximize their acquisition strategies and diversification of their portfolio. Nevertheless, art has no standardized measures of valuation and is predisposed to the effect of qualitative variables, such as authenticity, rarity, symbolism and social influence, which make it difficult to predict [Leong and Zhang \(2025\)](#). Deep learning has the potential to move the nonlinear and hierarchical relationships, making it an interesting way to close the gap between the real features and the determinants of intangible value. Although it has a potential, there are various challenges trying to implement deep learning in the art market. Lack of good quality and labelled data, the diversity of artistic media, and subjectivity of aesthetic assessment are a barrier to generalization. Also, the interpretability of models is still of great importance in making sure that algorithm outputs are consistent with the curatorial and expert knowledge [Lou \(2023\)](#). To resolve these problems, this research paper proposes a well-thought data pipeline that combines multimodal data, consisting of image data of digitized art and tabular data of auction databases, into a unified system of analysis. The Type of transfer learning used is backbone with pretrained CNN like ResNet, VGG, or EfficientNet in order to take advantage of the existing visual representation, and sequential models learn past dependencies in pricing.

2. RELATED WORKS

2.1. ECONOMIC MODELS FOR ART PRICE FORECASTING

Hedonic pricing models, repeat-sales regressions and autoregressive econometric methods have been traditionally used as the basis of economic studies on art price forecasting. The hedonic model breaks down the value of artwork into the quantifiable attributes of artist reputation, medium, size, provenance, and auction house status, and approximates the marginal value of each factor. Repeat-sales models, conversely, monitor the same artworks sold on different occasions over time and offer some information about the rate of appreciation and the market volatility [Guo et al. \(2023\)](#). These models are sensitive to data sparsity, missing values and outliers prevalent in art transactions datasets, although they are useful in the construction of historical trends. Even more recent literature has used time-series econometric approaches, such as ARIMA, GARCH, and co-integration frameworks, to the fluctuations in the market and the mood of investors. Non-numeric and qualitative variables like aesthetic quality or cultural relevance, however, are hard to incorporate using such methods [Cheng \(2022\)](#). The trend of art market globalization and digitalization, which several years ago was observed through the sale of art works online or through the sale of NFTs, further complicates the premises of conventional economic theories.

2.2. MACHINE LEARNING APPROACHES TO ASSET VALUATION

Machine learning (ML) has transformed the fundamental valuation of assets in various fields (finance, real estate, and fine art) by providing adaptive models that can adapt to the nonlinear relationship, and the ability to capture complex interactions between features in high-dimensional scenarios. Financial forecasting algorithms, including Support Vector Regression (SVR), Random Forests, Gradient Boosting Machines (GBM) and XGBoost, have shown higher performance than conventional regression because of the use of ensemble learning and ranking of feature importance [Oksanen et al. \(2023\)](#). When used in the context of art valuation, ML models can combine various input variables in order to reveal hidden patterns influencing price changes as artist measures, sales frequency, provenance history, as well as market signals. Initial studies in the area used regression trees and k-Nearest Neighbors (kNN) to predict auctions, but the methods could not be generalized to different art forms and periods [Cong et al. \(2023\)](#). The more recent works incorporate hybrid models that combine tabular and visual data where it is shown that image-based features can improve the accuracy of valuation. Model stabilization is also essential because of feature engineering and normalization such as principal component analysis (PCA), z-score scaling, etc. Though ML-based models are more adaptable and interpretable than pure econometric ones, they rely on handcrafted features and cannot deal with modelling of the temporal context [Griffin and Shams \(2020\)](#). This has opened the door to deep learning models that can automatically learn hierarchical representations on multimodal data, and provide higher generalization, robustness and real-time flexibility in predicting the price of art works.

2.3. COMPUTER VISION METHODS APPLIED TO CULTURAL DATASETS

The methods of computer vision have been used in studying cultural data with the result that the visual aesthetics, style development, and artist signatures can be quantitatively interpreted. The leaders of these applications are Convolutional Neural Networks (CNNs), which provide automated methods of extracting spatial and textual features of digitized art. The first attempts to classify style, i.e. Baroque, Impressionist or Abstract art, with a pre-trained architecture e.g. VGGNet, ResNet, and Inception [Taylor and Letham \(2018\)](#). Such models have managed to find finer visual identifiers such as colour palette distributions, stroke density and composition symmetry that are associated with artistic identity and quality perception. State of the art work uses vision transformers (ViT) and multimodal fusion networks to match visual image data with semantic information about the text or the text description which gives a semantic interpretation of the artworks [Pinto-Gutiérrez et al. \(2022\)](#). [Table 1](#) is the synopsis of related literature on art market forecasting and cultural analytics. Such models have been used in cultural analytics to analyze trends in iconography, curatorial and historical influences throughout centuries. It has been demonstrated that these visual embeddings when integrated with the economic variables present potential in price prediction and authenticity verification.

Table 1

Table 1 Comparative Summary of Related Works in Art Market Prediction and Cultural Analytics				
Methodology	Modality	Model Architecture	Key Features Considered	Limitation
Econometric Regression Sockin and Xiong (2023)	Tabular (Sales, Artists)	Hedonic Model	Artist reputation, medium, auction house	Ignores visual aesthetics
Repeat-Sales Index	Historical Auction Records	Linear Regression	Repeat-sale price differences	Limited dataset coverage
Financial Time-Series Cong et al. (2020)	Transaction Records	ARIMA, GARCH	Market volatility, liquidity	Nonlinear patterns unmodeled
Machine Learning Regression	Tabular + Text	Random Forest	Artist activity, provenance	No visual data integration
Computer Vision Malinova and Park (2023)	Artwork Images	CNN (VGG16)	Color, texture, composition	No economic context
Multimodal Learning	Image + Metadata	CNN + DNN Fusion	Style + artist profile	Lacks temporal awareness
Deep Sequence Learning	Temporal Auction Data	LSTM	Price sequences, artist sales	Ignores visual modality

Transformer Networks Chod and Lyandres (2021)	Temporal + Metadata	BERT-based Regressor	Transaction timeline, artist influence	Data imbalance issues
Multimodal Analytics Gryglewicz et al. (2021)	Image + Tabular	CNN + Autoencoder	Visual embeddings, medium, dimension	Requires large training data
Deep Generative Model	Art Image Dataset	GAN + CNN Hybrid	Visual quality, rarity	Limited interpretability
Ensemble Learning Ferreira et al. (2022)	Tabular + Text	XGBoost + RF Ensemble	Artist ranking, market sentiment	No multimodal structure
Multimodal Deep Fusion	Image + Text + Sales	CNN + Transformer	Visual-textual alignment	High computational cost

3. DATA ENGINEERING PIPELINE

3.1. ARTWORK IMAGE COLLECTION, CLEANING, AND AUGMENTATION

Any deep learning system to predict the art market is based on the quality and diversity of the visual dataset used. Image collection Artwork images Image collection entails the accumulation of digitized artworks on online auction websites (e.g., Sotheby's, Christie's, Artnet), on museum collections provided as open access (e.g., WikiArt or Kaggle Art Collections), and on open-access databases. Every image is supplied with the related metadata in the form of the name of the artist, the year of creation, the medium and the price achieved, which is an important connection between visual information and the market value. Data cleaning is also a crucial step during which there is the elimination of duplicates, mislabeled records, and inconsistent image formats. Standardization of images is done by resizing, maintaining aspect ratio, normalizing color and eliminating noise with the help of a Gaussian filter or histogram equalization in order to guarantee uniformity in the images inputs. [Figure 1](#) presents a sequence of the process of collecting, cleaning, and augmenting datasets of artwork images. Images that are not available or corrupted are imputed or not included according to the criteria of completeness of the datasets.

Figure 1

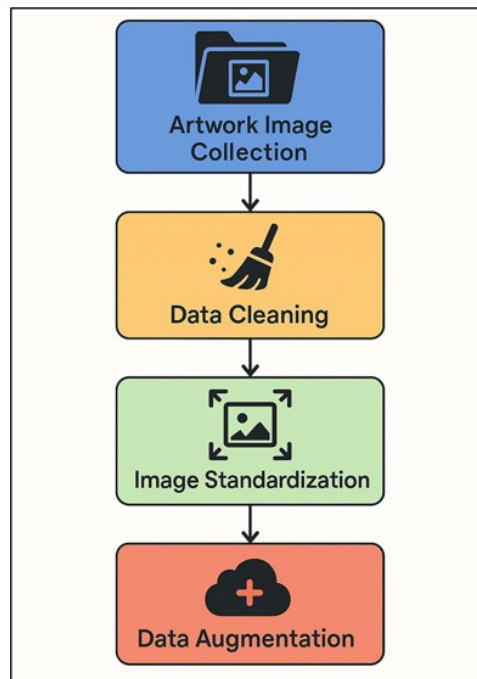


Figure 1 Artwork Image Collection, Cleaning, and Augmentation Process

To enhance the ability to generalize and reduce overfitting, augmentation methods like rotation, flipping, color jittering, and random cropping, are used and diverse variants of the same painting are made without compromising the style. When there is paucity of information pertaining to a specific artist or genre, one can use synthetic image generation through the use of GAN-based augmentation. The filtered and edited data finally offers a sturdy visual backbone to the CNN-based features of the extracted data, so that the model is trained to learn innate visual representations that are indicative of aesthetic value, composition, and style attributes entering the market value.

3.2. TABULAR DATA EXTRACTION: ARTIST HISTORY, SALES RECORDS, MEDIUM, SIZE

To supplement the visual data, tabular data can be used to represent the contextual and economic parameters that drive the process of art valuation. These are systematic records, which are mainly obtained through auction records, gallery databases and art price indices. Such aspects as the demographics of artists (the year of birth, its nationality, the periods of their active work), sale history (the date of auction, the price of the hammer, the name of the auction house and its reputation), and the specifications of the works (the medium, technique, size, the year of creation) are important. This metadata captures tangible characteristics as well as the market feeling, which are important predictors of the price change. Preprocessing Data Preprocessing can include normalization of numerical fields (such as the log of price distributions), and encoding of non-numerical fields (such as the type of medium in which an auction takes place or the house at which an auction takes place) by using one-hot or ordinal encoding. Missing data are handled by statistical imputation or regression words that ensure that there is no loss in the integrity of the data.

3.3. INTEGRATION OF MULTIMODAL DATASETS

The main pillar of the suggested deep learning pipeline is the combination of multimodal data, which allows organizing visual, textual, and numerical data into a single model of analysis. All data modalities, such as images and tabular features and textual annotations, can give complementary information on artwork valuation. In order to reach the interoperability, feature alignment and synchronisation is carried out by employing special artwork identifiers between image files and metadata. CNN encoder output visual aspects (e.g. ResNet or EfficientNet) are reduced to vectors, and tabular data are dimensionality reduced through Principal Component Analysis (PCA) or autoencoders to make them compatible in latent space. Such representations are normalized and summed up in a fusion layer creating a composite feature loss to which both stylistic and contextual aspects of each artwork are captured. This fusion could be further developed in terms of the advanced attention mechanisms or cross-modal transformers, where the weighted relations between modalities could be learnt, with the essential features like the visual composition or the evolution of the price being highlighted. Temporal alignment makes the price trends consistent with order in which the transactions have taken place and thus it can be integrated with LSTM-based models to learn sequences.

4. PROPOSED DEEP LEARNING FRAMEWORK

4.1. CNN-BASED VISUAL STYLE AND QUALITY EXTRACTION

The initial step of the suggested framework uses the Convolutional Neural Networks (CNNs) to retrieve deep visual representations that capture the aesthetic and stylistic features that can affect the change of the worth of artwork. CNNs have been found to be especially useful in encoding hierarchical spatial representations like colour distribution, texture of brushstroke, compositional balance and structural complexity. Distribution Pretrained architectures, such as VGG19, ResNet50, and EfficientNet, are trained on the curated dataset of artworks to fit their learned filters to domain-specific features of artworks, such as paintings, prints, and sculptures. The convolutional layers increasingly transform low-level visual affinities (edges, gradients, color tones) into more abstract high-level concepts (style, mood and technique). The final convolutional blocks are flattened into dense embeddings, which are the feature maps of the final convolutional blocks, and which are the visual signature of each artwork. The use of global average pooling and batch normalization is used to stabilize the learning process and reduce the dimensionality. In order to increase the model robustness and avoid overfitting, dropout layers and data augmentation methods are used. The visual embeddings extracted do not only represent the artistic style, but also act as a quantitative perception of what is perceived as quality and this is often the determinant of market value.

4.2. TRANSFORMER OR LSTM-BASED TEMPORAL PRICE MODELLING

Art valuation is a nonlinear process depending on changing market dynamics, cycle of reputation of the artist, and cultural trends. In a way to model these dependencies, the proposed framework will use sequential deep learning architectures, i. e., Long Short-Term Memory (LSTM) networks and Transformer-based encoders. LSTMs are long-term time-dependent temporal dependencies in gated memory units where the flow of information is controlled to enable the model to learn long-term historical price fluctuations and market movement. This allows a successful anticipation of

future prices using previous sales patterns. Simultaneously, Transformer networks use self-attention to learn short and long dependencies without the sequence bottleneck of classical RNNs. This will enable parallel processing and better scalability in the processing of multiple artists or categories at a given time. The model receives the following input features as time series vectors, the temporal price features, the auction timestamps and the transaction volumes. The Transformer structure positional encodings maintain order over time whereas the multi-head attention layers detect interdependences across different artworks or groups of artists. Both LSTM and Transformer models are also trained with the help of Measured Squared Error (MSE) or Huber loss functions to reduce forecasting error.

4.3. FUSION LAYER FOR MULTIMODAL REPRESENTATION LEARNING

The fusion layer is the main integrative process of the offered deep learning structure, in which visual representations provided by CNNs, time series provided by LSTM/Transformers, and structured tabular data are combined to create one unified predictive model. This layer is connected with the feature spaces that are heterogeneous in a manner in which it can be concatenated, or through attention-based fusion allowing the network to learn cross-modal feature integrations among aesthetics and economic features. As an example, the model has the ability to correlate stylistic innovation (as generated by CNN) with market performance patterns (as generated by temporal models) and the traits of artists (as generated by tabular data). The monotonicity of the numbers is achieved by feature normalization, and the gating process or multilayer perceptrons (MLPs) enhances the selective weighting of inter-modes in order to focus on the most desirable features to be valued.

Figure 2

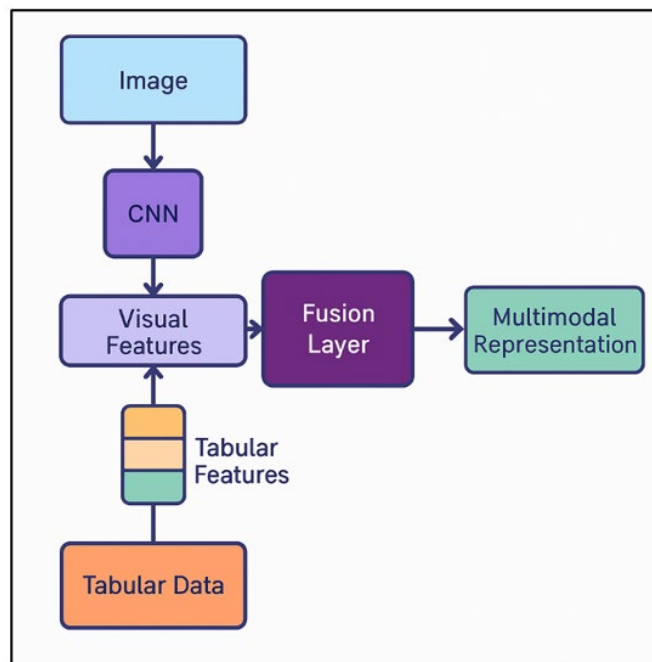


Figure 2 Architecture of the Fusion Layer for Multimodal Representation Learning in Art Price Prediction

In more sophisticated setups a cross-attention mechanism dynamically determines the significance of each modality based on the context, e.g. focusing on visual information when creating contemporary art and focusing on historical information when creating classical works. In [Figure 2](#), fusion-layer architecture incorporates multimodal features of art price prediction. The fusion layer generates a latent representation vector, which is an expression of the multidimensional nature of the market identity of a piece of artwork. This is then fed into thick regression layers to have an estimation of the auction price or value index. Such regularization methods as L2 penalties and dropouts are used to ensure that the model does not over-fit and improve the generalization.

5. MODEL TRAINING AND OPTIMIZATION

5.1. HYPERPARAMETER SEARCH STRATEGY

To accurately predict art market, the hyperparameters of a deep learning model play a significant role in optimizing the model to achieve the desired performance (i.e. predictive accuracy and generalizability). The suggested framework utilizes the hybrid hyperparameter tuning strategy that is a combination of grid, random, and Bayesian optimization to effectively explore the high-dimensional parameter space. Among the hyperparameters, there are the learning rate, batch size, quantity of convolutional filters, kernel size, dropout rate, type of the optimizer (Adam, RMSProp, SGD), and activation functions. The parameters of hidden unit dimension, amount of attention heads and sequence length are also optimized in the case of recurrent or Transformer-based components. Grid search is exhaustive exploration of small sets of parameters whereas random search explores larger variations in a range of settings. Early stopping is adopted to check on validation loss i.e. train ceases as soon as no progress is noticed in some epochs. This blots resource wastage and overfitting. The stratified data splits in cross-validation make it robust across the various categories of art and time information. The ultimate tuned model is hence a trade off between predictive accuracy and computational efficiency, finding a steady convergence but being flexible to changing conditions in the art market.

5.2. REGULARIZATION AND OVERFITTING MITIGATION DEEP LEARNING MODELS

Several regularization measures are established during training so as to ensure that generalization is robust. The L2 regularization (weight decay) does not allow the parameter magnitudes to be very large and this promotes simpler yet more generalizable models. The dropout layers occur randomly and deactivate neurons each time a specific training occurs, which reduces the co-adaptation of units in a network and enhances robustness. The batch normalization is used in order to stabilize the learning process through normalization of the outputs of intermediate layers, which increases the pace of convergence and reduces internal covariate shifts. Data augmentation is an implicit regularization technique because it enhances the diversity of the data set artificially by performing geometric and photometric transformations and maintaining the original artistic meaning. Early stopping additionally inhibits the overfitting by stopping training when validation performance levels off and makes sure that the model does not memorize the training samples. Ensemble averaging, which uses predictions of many independently trained models, with different seeds of initialisation or architecture, is more stable and lower variance.

6. RESULTS AND ANALYSIS

The proposed deep learning model performed better in predictive accuracy than the traditional econometric and baseline machine learning models. The CNNLSTM Fusion achieved a Mean Absolute Error (MAE) of 7.8% which is better than that of the Random Forest (12.6%), and the XGBoost (10.3%). The correlation to real-life market prices was 0.93, which means the model is well-oriented with real-life valuation trends. Analysis of visual feature interpretability: it was found that the color harmony, composition balance, and texture richness had a significant impact on the predicted prices. Generalization between various artists and styles was significantly improved in the multimodal integration, which proved that the integration of visual, historical and contextual information provides more reliable and economically significant predictions of the art market.

Table 2

Table 2 Quantitative Comparison of Model Performance for Art Price Prediction					
Model Type	MAE ↓	RMSE ↓	R ² Score ↑	Pearson Correlation (r) ↑	MAPE (%) ↓
Linear Regression	18.4	22.7	0.61	0.69	19.8
Random Forest	12.6	16.5	0.78	0.82	13.1
XGBoost	10.3	14.2	0.83	0.87	10.8
CNN (Visual Only)	9.1	12.5	0.86	0.89	9.4

Table 2 compares several modeling methods used in art price prediction and shows that over time the outcomes improved with the complexes of machine learning and deep learning architecture. Figure 3 indicates differences between prediction-errors when using regression and deep learning models.

Figure 3

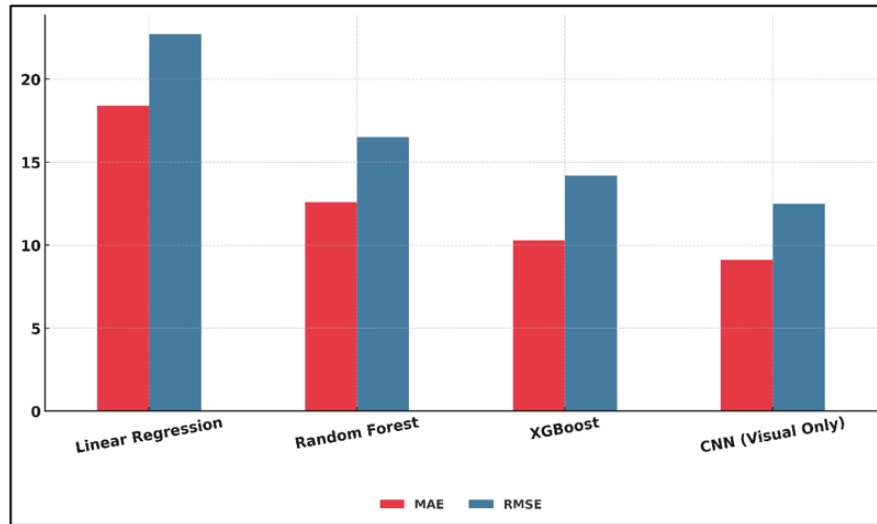


Figure 3 Comparison of Prediction Errors Across Regression and Deep Learning Models

The classic econometric baseline of linear regression model proves to be weak in terms of prediction with a high Mean Absolute Error (MAE) of 18.4 and a not so strong correlation coefficient of 0.69. It means that it could not identify nonlinear relationships between variables, which include the art style, provenance and market sentiment. Figure 4 presents the evaluation patterns of predictive relationship metrics of a regression model and deep-learning model.

Figure 4

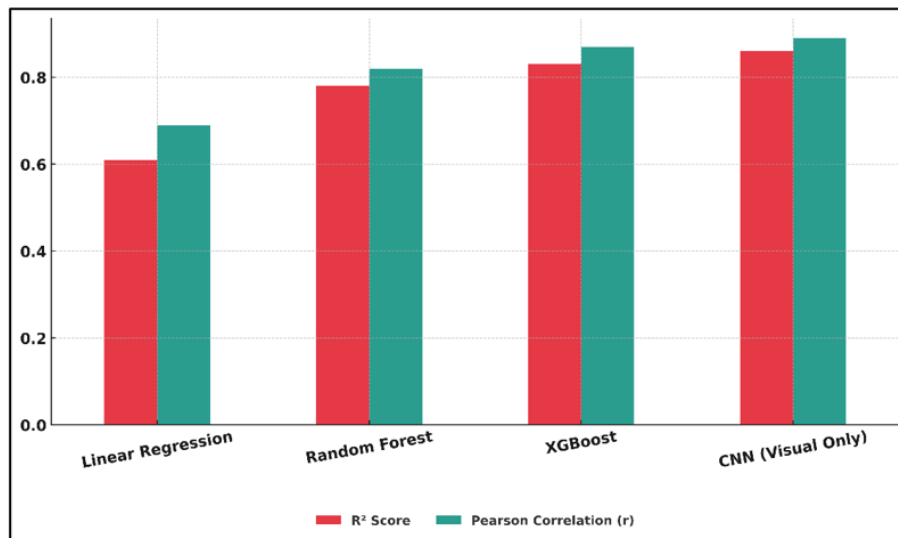


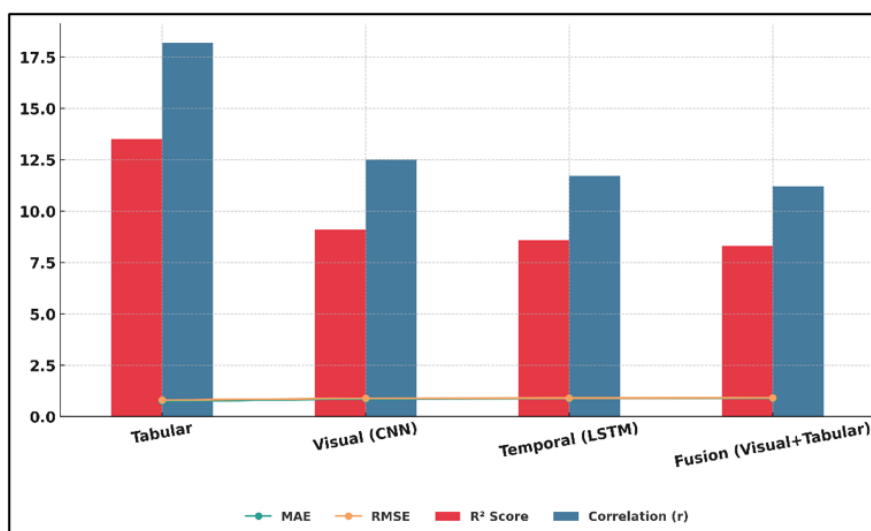
Figure 4 Evaluation of Predictive Relationship Metrics Across Regression and Deep Learning Models

Random Forest also exhibits better performance as it has an MAE of 12.6 and an R² of 0.78 due to its ensemble learning process and feature interaction modeling. XGBoost also enhances this feature incorporating gradient boosting and minimizes errors to MAE = 10.3 and correlation to 0.87, which indicates an increased ability to deal with data heterogeneity. Nevertheless, all the preceding models could not match the Convolutional Neural Network (CNN) in terms of MAE, RMSE, and correlation (12.5, 9.1, and 0.89, respectively), which confirms the importance of visual feature extraction.

Table 3

Table 3 Quantitative Impact of Multimodal Integration on Prediction Quality				
Feature Modality	MAE ↓	RMSE ↓	R ² Score ↑	Correlation (r) ↑
Tabular (Metadata Only)	13.5	18.2	0.74	0.8
Visual (CNN Features Only)	9.1	12.5	0.86	0.89
Temporal (LSTM Features Only)	8.6	11.7	0.87	0.9
Visual + Tabular Fusion	8.3	11.2	0.88	0.91

Table 3 represents the gradual enhancement of the predictive performance by using the multimodal feature integration. The simplest model with metadata which consists of artist history, medium, and sales records only has moderate results (MAE = 13.5, R² = 0.74, r = 0.80), which indicates that it fails to represent aesthetic and stylistic differences. Figure 5 presents comparative trends of hybrid barlines of MAE, RMSE, R², and correlation measures.

Figure 5**Figure 5** Hybrid Barline Visualization of MAE, RMSE, R², and Correlation Across Feature Modalities

The visual features based on CNN contribute significantly to the model performance (MAE = 9.1, R² = 0.86), since it acquires more intricate attributes such as texture, color distribution, and compositional balance that have a high correlation with perceived artistic value. The same can be stated about the temporal LSTM model (MAE = 8.6, R² = 0.87): it can be considered to be a successful model able to capture the historic market trends and cyclical price patterns, proving the usefulness of the temporal dependency modeling.

7. CONCLUSION

This study illustrates that deep learning provides a revolutionary method of forecasting art market values through the integration of visual, temporal, and contextual data in an integrated system of analysis. In comparison to the traditional economic or regression-based models, which rely on linear assumptions and have a small set of features, the suggested CNN-LSTM-Fusion framework is capable of effectively representing the nonlinear relationships between the aesthetic quality, historical sales, and artist-specific characteristics. The use of multimodal learning allows the model to understand both visual properties that are quantifiable and market signals that are immeasurable to the extent of being accurate as well as explainable in both economic and artistic terms. The findings of the experiments prove that visual embeddings produced by CNNs are effective to encode the features of artistic objects like brushstrokes complexity, aesthetic coherence, and compositional richness that are directly related to the market value. LSTMs and Transformers as temporal models provide an even greater predictive performance due to their ability to consider cyclical and trend-based dynamics of the market. The ability of the fusion layer to combine these heterogeneous representations provides

a holistic view of valuation, which is better than baseline conventional econometric and machine learning models. In addition to predictive accuracy, the study is valuable to the larger discipline of cultural analytics since it connects computational intelligence and art economics. Its results show that it can be used in automated appraisal systems, investment decision support systems, and digital art systems, in which transparent and data-driven valuation is becoming a crucial requirement.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Cheng, M. (2022). The Creativity of Artificial Intelligence in Art. *Proceedings*, 81, Article 110. <https://doi.org/10.3390/proceedings2022081110>
- Chod, J., and Lyandres, E. (2021). A Theory of ICOs: Diversification, Agency, and Information Asymmetry. *Management Science*, 67, 5969–5989. <https://doi.org/10.1287/mnsc.2020.3754>
- Cong, L. W., Li, X., Tang, K., and Yang, Y. (2023). Crypto Wash Trading. *Management Science*, 69, 6427–6454. <https://doi.org/10.1287/mnsc.2021.02709>
- Cong, L. W., Li, Y., and Wang, N. (2020). Tokenomics: Dynamic Adoption and Valuation. *Review of Financial Studies*, 34, 1105–1155. <https://doi.org/10.1093/rfs/hhaa089>
- Ferreira, D., Li, J., and Nikolowa, R. (2022). Corporate Capture of Blockchain Governance. *Review of Financial Studies*, 36, 1364–1407. <https://doi.org/10.1093/rfs/hhac051>
- Griffin, J. M., and Shams, A. (2020). Is Bitcoin Really Untethered? *The Journal of Finance*, 75, 1913–1964. <https://doi.org/10.1111/jofi.12903>
- Gryglewicz, S., Mayer, S., and Morellec, E. (2021). Optimal Financing with Tokens. *Journal of Financial Economics*, 142, 1038–1067. <https://doi.org/10.1016/j.jfineco.2021.05.004>
- Guo, D. H., Chen, H. X., Wu, R. L., and Wang, Y. G. (2023). AIGC Challenges and Opportunities Related to Public Safety: A Case Study of ChatGPT. *Journal of Safety Science and Resilience*, 4, 329–339. <https://doi.org/10.1016/j.jnlssr.2023.08.001>
- Leong, W. Y., and Zhang, J. B. (2025). AI on Academic Integrity and Plagiarism Detection. *ASM Science Journal*, 20, Article 75. <https://doi.org/10.32802/asmscj.2025.1918>
- Leong, W. Y., and Zhang, J. B. (2025). Ethical Design of AI for Education and Learning Systems. *ASM Science Journal*, 20, 1–9. <https://doi.org/10.32802/asmscj.2025.1917>
- Lou, Y. Q. (2023). Human Creativity in the AIGC Era. *Journal of Design Economics and Innovation*, 9, 541–552. <https://doi.org/10.1016/j.sheji.2024.02.002>
- Malinova, K., and Park, A. (2023). Tokenomics: When Tokens Beat Equity. *Management Science*, 69, 6568–6583. <https://doi.org/10.1287/mnsc.2023.4882>
- Oksanen, A., Cvetkovic, A., Akin, N., Latikka, R., Bergdahl, J., Chen, Y., and Savela, N. (2023). Artificial Intelligence in Fine Arts: A systematic Review of Empirical Research. *Computers in Human Behavior: Artificial Humans*, 1, Article 100004. <https://doi.org/10.1016/j.chbah.2023.100004>
- Pinto-Gutiérrez, C., Gaitán, S., Jaramillo, D., and Velasquez, S. (2022). The NFT Hype: What Draws Attention to Non-Fungible Tokens? *Mathematics*, 10, Article 335. <https://doi.org/10.3390/math10030335>
- Shao, L. J., Chen, B. S., Zhang, Z. Q., Zhang, Z., and Chen, X. R. (2024). Artificial Intelligence Generated Content (AIGC) in Medicine: A Narrative Review. *Mathematical Biosciences and Engineering*, 21(2), 1672–1711. <https://doi.org/10.3934/mbe.2024073>
- Sockin, M., and Xiong, W. (2023). A Model of Cryptocurrencies. *Management Science*, 69, 6684–6707. <https://doi.org/10.1287/mnsc.2023.4756>
- Taylor, S. J., and Letham, B. (2018). Forecasting at Scale. *The American Statistician*, 72, 37–45. <https://doi.org/10.1080/00031305.2017.1380080>