Research and Application of Vietnamese antiquities reconstructing and preserving Based on Deep Learning

1st Pham Thi Thao Nguyen
FPT University, Ho Chi Minh Campus
Vietnam
nguyenpttse170162@fpt.edu.vn

2nd Vo Dinh Nhat Truong
FPT University, Ho Chi Minh Campus
Vietnam
truongvdnse161235@fpt.edu.vn

3rd Tran Phuc Tho
FPT University, Ho Chi Minh Campus
Vietnam
thotpse170056@fpt.edu.vn

4th Nguyen Xuan Quang
FPT University, Ho Chi Minh Campus
Vietnam
quangnxse170087@fpt.edu.vn

5th Pham Phu Ngoc Trai
FPT University, Ho Chi Minh Campus
Vietnam
traippnse161809@fpt.edu.vn

6th Nguyen Quoc Trung
FPT University, Ho Chi Minh Campus
Vietnam
trungnq46@fe.edu.vn

***Abstract*—In the strong development of technology, the development and preservation of history should be focused. Besides, artificial intelligence, specifically deep learning has a great influence in the field of computer vision. Currently, deep learning techniques have developed and are capable of recovering images based on large amounts of data. This has great significance in the application to the field of historical preservation. This study focuses on the application of deep convolutional neural networks to recover ancient Vietnamese patterns and mainly focuses on completing the image data set of patterns on cobalt blue underglazed porcelains under the Nguyen Dynasty. This study includes the following tasks: (1) Collecting data on pattern images on Vietnamese Nguyen Dynasty Cobalt blue underglazed porcelains. (2) Present some image recovery models based on deep convolutional networks. (3) Experiment with the presented models on the collected pattern image data set. From there, evaluate and compare the experimental results and propose further research directions. The research will contribute to improving the ability to preserve and promote the value of Vietnamese cultural heritage, especially among young people. The results can also be applied in many different fields such as education, research, and tourism, contributing to improving understanding of Vietnamese history and culture.**

***Index Terms*—Convolutional neural networks, pattern recognition, data collection, inpainting, historical preservation**

I. INTRODUCTION

Although the practice of picture inpainting, or the art of restoring outdated and damaged photos, has been for a while, it has lately grown in favor as a result of recent advancements in image processing technology [1]. The development of deep learning techniques has revolutionized computer vision, enabling remarkable advancements in image recognition, restoration, and analysis. Researchers have successfully applied deep convolutional neural networks to various domains, including image inpainting, where missing or damaged parts of images are intelligently reconstructed. A study "Generative Image Inpainting with Contextual Attention" by Jiahui Yu et al. presents a deep generative model-based approach for image inpainting [2]. They introduce a fully convolutional neural network that can handle images with multiple holes of varying sizes and locations during testing. The key innovation is the integration of contextual attention mechanisms, allowing the network to explicitly utilize surrounding image features as references during training. By incorporating contextual attention, the model can effectively borrow information from distant spatial locations, resulting in improved structures and textures consistent with the surrounding areas. The authors evaluate their method on various datasets, including faces, textures, and natural images, and demonstrate that it generates higher-quality inpainting results compared to existing approaches. The paper "EdgeConnect: Generative Image Inpainting with Adversarial Edge Learning," was authored by Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Z. Qureshi, and Mehran Ebrahimi proposes EdgeConnect as an effective solution to this problem, leveraging generative adversarial networks (GANs) and edge information [3]. Ecem Sogancioglu, Shi Hu, Davide Belli, and Bram van Ginneken wrote "Chest X-ray Inpainting with Deep Generative Models" in the field of medicine. The authors point out that while generative adversarial networks (GANs) have been successfully used for inpainting in natural images, the field of medical imaging has not been properly investigated. The study examines how three inpainting models based on deep learning—context encoders, semantic picture inpainting, and the contextual attention model—perform when applied to chest X-ray images [4]. In the field of historical preservation, the utilization of advanced technologies has become essential to safeguard and restore cultural artifacts. Artificial intelligence (AI) and deep learning techniques, in particular, have shown great potential in the realm of computer vision. Xuhui Fu's (2021) research attempted to revive traditional Chinese designs. "Research and Application of Deep Convolutional Neural Network Based Ancient Chinese Pattern Restoration."The other two approaches are tested on the same section of the Qinghai traditional embroidery image data set as the strategy that is based on the deep convolutional neural network. With a total of 2000 images, the Qinghai embroidery image data set consists of embroidery of the Tu nationality, Guinan

Tibetan embroidery, Haixi Mongol embroidery, Hehuang embroidered, and Huangzhong pile embroidery [5].

As can be seen, the development of deep learning techniques is of great significance in the historical field. Vietnam possesses an extensive collection of physical cultural heritage, comprising millions of precious artifacts, antiquities, and national treasures. These valuable items are carefully preserved and exhibited across a network of 125 museums located throughout the country. The museums are strategically distributed in various regions and are particularly concentrated in major cultural and tourist hubs [6]. However, the application of deep learning techniques to restore ancient Vietnamese patterns remains largely unexplored. Vietnamese antiquities hold a significant cultural and historical value that needs to be preserved and protected. However, the historical value of Vietnam is gradually fading both physically and mentally, especially among young people. The study titled ”Educating Traditional Cultural Values in Vietnam Universities” conducted by Nguyen Sy Trung and Vu Hong Van reveals that the education of traditional cultural values has not received significant attention in Vietnamese universities for an extended period. Furthermore, the research highlights that only a limited number of universities incorporate cultural subjects into their curriculum at the university level [7]. Many artifacts have been damaged or lost over time and many people are not aware of the value and meaning of Vietnamese antiques, making it challenging to preserve them [6]. Over the years, some experts in culture and fine arts have admitted that our fields such as movies, fine art, interior decoration, graphic design, etc. are being ”Chineseized” ”. One of the important reasons pointed out is that due to the lack of resources, Vietnam currently does not have a pattern library. ancient culture to serve as an application basis for products bearing the spirit of Vietnamese culture. Meanwhile, many other countries have thousands of patterns, gathered into libraries of vector patterns, optional online.

This study focuses on the preservation and restoration of ancient Vietnamese patterns on cobalt blue underglazed porcelains the Nguyen Dynasty. By focusing primarily on assembling a comprehensive data set of pattern images of this era from various collections and archival sources. The research will contribute to the archiving and digitization of historical image data of Vietnam. This study applies and evaluates image restoration models based on deep convolutional neural networks from previous studies. Tests performed on the collected data set will shed light on the effectiveness of these models in restoring complex patterns on porcelain artifacts.

Applying Deep Learning technologies to preserve and restore Vietnam’s cultural heritage is a modern and advanced method, attracting the attention of the following scientific, political, local and international communities. The importance of this study lies in the immense value of preserving and understanding cultural heritage. To effectively navigate the challenges and leverage the opportunities presented by Industry 4.0, it is essential for us to proactively develop its culture in the coming years [8]. By employing advanced deep learning techniques, the research will contribute to the growing body of knowledge in the field of computer vision, specifically in historical restoration. Furthermore, this is a special way to spread, educate and reach young people in the current era. The recovery and restoration of ancient patterns on cobalt blue underglazed porcelains will not only provide insights into the artistic traditions of the Nguyen Dynasty but also enhance our overall comprehension of Vietnamese history and culture. Furthermore, By doing so, it aims to enrich our understanding of Vietnamese cultural heritage and inspire further advancements in the field of historical preservation.

II. METHODOLOGY

A. Data

Antique images are data of great value, especially in the field of artificial intelligence, big data contributes greatly to the success of models. However, to achieve this goal, very large datasets are needed to train and test AI models accurately and reliably. The effectiveness of algorithms depends heavily on the quality of the input data representation. A suitable representation leads to better performance compared to a poor representation [9]. The dataset collected in this study provides pattern images on cobalt blue underglazed porcelains in the Nguyen Dynasty of Vietnam that can enrich, diversify and expand the existing data sets.

1) Objectives: The data collected initially included images of cobalt blue underglazed porcelains under the Nguyen Dynasty in Vietnam. Pattern designs include dinner bowls, dinner plates, covered tureen, tea-set, offertory fruit-tray, altar flower vases, basin and flower pot, large flower vases, decorative pedestals. Decorative motifs include decorative motifs with animals, decorative motifs with vegetation, decorative motifs landscape, figure and story, decorative motifs with objects, and various forms of diaper patterns.

Nguyen Dynasty: The final reign of a monarch in Vietnamese history. The Nguyen Dynasty lasted 143 years, beginning with Nguyen Anh (Gia Long)’s accession to the throne in 1802 and ending with Bao Dai’s abdication in 1945.

Porcelain: The kaolin content in porcelain bone is high and the melting of porcelain bone proves that the firing temperature has exceeded 1,300oC, which is the temperature to create porcelain products in the true sense of the word.

Cobalt blue underglazed porcelains (Ordered patterned porcelains): That is sending samples to Chinese ceramic kilns to follow the orders of the orderers (the kings, mandarins and commoners of Vietnam from the Le Trung Hung period to the Nguyen Dynasty). These porcelains are made in China, but carry ”Vietnamese criteria” such as:

- Vietnamese monuments including the Thien Mu Pagoda, Thuy Van Mountain, Tam Thai Mountain, Hai Van Mountain, and Thuan Hoa Market are depicted in decorative porcelain painting.
- Poetry on porcelain, poetry written in the Vietnamese-invented Nom script, which is not used in China, poetry written in Chinese but with a Vietnamese author, such as that of Lord Nguyen Phuc Chu, Dao Duy Tu, King Thieu Tri, or King Tu Duc...
- The year marks or insignia on the porcelain that corresponds to the years of the Vietnamese mission to China include names of Vietnamese rulers including Gia Long, Minh Mang, Thieu Tri, Tu Duc, and Khai Dinh.
- These porcelains are only intended for usage by Vietnamese people and are not sold on the modern Chinese market.

2) Dataset collection: Raw data sets are collected from highly reputable sources. The time to collect and annotate the images is about 3 months, starting from March 1, 2023. The raw dataset includes 701 images of various sizes. The images are in PNG format. Images are collected and archived from a variety of data sources, including Books, private collections, Youtube videos, Blogs and Websites.

The results of synthesizing the number of raw object images containing patterns from multiple sources are shown in Table 1.

TABLE I

Source	Number of Images
"Commissioned Patterned Porcelains In the Nguyen Era" - Tran Duc Anh Son	240
Official websites of museums	103
Khanhhoathuynga's collection Blog	95
The collection of Record holder Dinh Cong Tuong	93
Dai Ngan - Youtube	87
Ngoc Tinh Phung - Youtube	49
TTR Tube - Youtube	34
Total	**701**

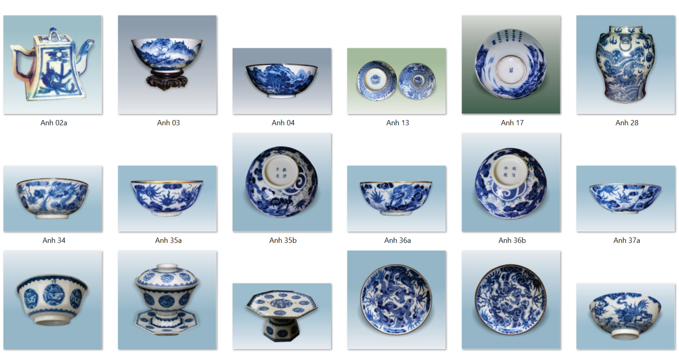

Fig. 1. Part of raw image data collected from sources.

3) Preprocessing: To process the raw data set, a number of tasks were carried out including: Checking and deleting duplicate images, poor quality images. All images were then cropped to square sizes and removed unused pattern-free areas from the image. Areas that do not qualify are also ignored. Finally, the number of processed pattern images is 3106.

Fig. 2. The image simulates the processing of raw data into pattern image data.

4) Ethical considerations: Researchers are interested that Vietnamese antiques, especially those of the Nguyen Dynasty, are very rare and not easy to find and collect. Therefore, the researcher makes sure to respect the copyright of the images and videos and has also asked the owner's permission before using it. For each data source used in the study, we contacted the owner via email or cell phone and asked for permission to use the image, video, or photograph (with real artifacts). We guarantee to have all documents related to copyright issues. Furthermore, we have clarified that this is applied research on artifact images and does not make any physical changes to the artifact.

B. Model

Several computer vision jobs require the use of the image inpainting technique, which is used to replace missing or damaged portions of an incomplete image [10]. Several deep learning algorithms have been successfully used and openly reported for the inpainting job. Some popular techniques include Convolutional Neural Networks (CNNs), which are neural network architectures capable of learning the structural features of an image through the application of convolutional filters and pooling layers [11]. Encoder-Decoder Networks are another commonly used architecture, where the encoder extracts information from the original image, while the decoder network is used to reconstruct the missing regions [12]. Generative Adversarial Networks (GANs) are a type of adversarial neural network architecture, consisting of a generator network used to predict and generate pixel values for the inpainted regions, and a discriminator network used to distinguish between the original and completed images [13]. Attention Mechanisms incorporate attention mechanisms that allow the network to focus on important regions of the image and prioritize the completion of these regions first [14]. In this study, we will test three different models on the data set.

1) Text Preprocessing: Performed the pre-training steps of the data, including image format conversion, resizing, normalizing pixel values, and other processing suitable for the inpainting problem. These steps have ensured that the data is correctly prepared and suitable for model training and testing. In the process of addressing the inpainting problem, the division of the image dataset into training and testing sets is an important step to ensure fairness and effective evaluation of the model. In this study, we utilized a dataset comprising 3106 images and performed the data split following a 9:1 ratio, resulting in 2806 images for training and 300 images for testing. The random selection of images for each set was carried out with considerations for randomness and fairness, ensuring that the generated datasets are representative of the original data.

2) Generative Image Inpainting with Contextual Attention: We build on earlier work by Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang in this paper. They created a coarse-to-fine network architecture in their paper titled "Generative Image Inpainting with Contextual Attention" to carry out the process of adding missing pixels to an image, also known as image inpainting [5]. Here are the specific details of this model:

The model has an image-to-image structure, where one picture serves as both the input and the output. Two stages make up the model's architecture: a coarse network and a refinement network. The Refinement Network uses the initial coarse prediction from the Coarse Network as input and produces the final detailed result.

The distorted image is fed into the coarse network, which produces a rough forecast of the missing region. To gather multi-scale contextual data, it employs a conventional encoder-decoder architecture with skip connections and dilated convolutions. 8 convolutional layers with a stride of 2 and a kernel size of 4 make up the encoder, which is followed by a fully connected layer. The decoder begins with a convolutional layer with a kernel size of 1, then 8 deconvolutional layers with a stride of 2 and a kernel size of 4. Between the encoder and decoder layers with the same spatial resolution, skip connections are inserted. With dilation rates of 2, 4, 8, and 16, the dilated convolutions are used in the middle layers of the encoder and decoder.

The distorted image and the coarse forecast are inputs into the refinement network, which produces a refined prediction of the missing region. It makes explicit use of a contextual attention module to copy or borrow data from far-off spatial regions. A softmax operation is used by the attention module to give weights to each patch after computing a similarity map between patches in the missing region and patches in the surrounding region. The missing section is then recreated using the weighted patches. Following an attention module, the refinement network has 4 convolutional layers with stride 1 and kernel size 3, then 4 more convolutional layers with stride 1 and kernel size 3.

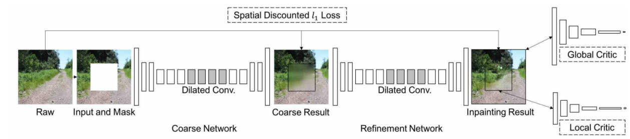

Fig. 3. Network structure of Contextual Attention [5]

The model is trained end-to-end with an adversarial loss and a reconstruction loss. The adversarial loss encourages the model to generate realistic images that can fool a discriminator network. The discriminator network has 5 convolutional layers with stride 2 and kernel size 4, followed by a fully connected layer. The reconstruction loss measures the pixel-wise difference between the generated image and the ground truth image. It consists of two terms: an L1 loss that penalizes large errors and a perceptual loss that penalizes semantic errors.

3) Generative Image Inpainting with Context-Encoder GAN: An encoder and a decoder make up the generator, which is a U-Net model. Eight convolutional layers with leaky ReLU activation and stride 2 make up the encoder. Eight deconvolutional layers with stride 2 and ReLU activation make up the decoder. In order to maintain spatial information, the decoder also has skip connections from the encoder. The decoder's output layer uses a tanh activation to generate a picture in the [-1, 1] range.

A PatchGAN model with five convolutional layers activated with stride 2 and leaky ReLU serves as the discriminator. A sigmoid activation on the discriminator's output layer creates a probability map of the input image's actual and false patches.

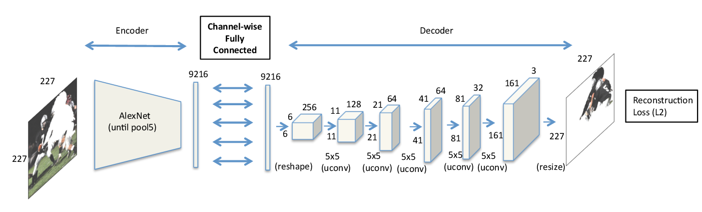

Fig. 4. Network structure of Context-Encoder GAN [15]

To train the generator, the model combines adversarial loss and reconstruction loss. The binary cross-entropy between the discriminator's predictions and the ground truth labels is the adversarial loss. The L2 difference between the generated image and the original image in the missing region is the reconstruction loss.

4) Generative Image Inpainting with Image Fine-grained Inpainting: We build on earlier research done by Zheng Hui, Jie Li, Xiumei Wang, and Xinbo Gao in their paper titled "Image Fine-grained Inpainting" in this study. [16]. In this paper, a one-stage generative adversarial network (GAN) model for image inpainting is proposed. Earlier inpainting methods frequently created finished works with a strange structure or blurriness. The suggested model uses dense combinations of dilated convolutions to achieve larger and more effective receptive fields, addressing this issue and making it simpler to recover vast parts in an incomplete image.

The authors present a novel self-guided regression loss that concentrates on ambiguous regions and improves semantic details in order to successfully train the generator. A geometrical alignment constraint is created to align the feature center coordinates between predicted features and ground-truth ones, correcting for the pixel-based distance in addition to the widely utilized VGG feature matching loss.

To guarantee consistency in local-global contents, the model has a discriminator with local and global branches. Additionally, a discriminator feature matching technique is added to the local branch to dynamically minimize the similarity of intermediate features between artificial and real-world patches, enhancing the caliber of the images that are generated.

The suggested method's performance is evaluated on numerous open datasets and measured against cutting-edge approaches, proving its superiority.

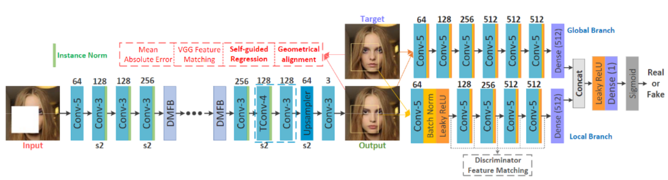

Fig. 5. Network structure of Image Fine-grained Inpainting [16]

III. Model application results and discussion

A. Experimental Environment and Data Set

The models are experimentally conducted on the image data set of patterns on cobalt blue underglazed porcelains under the Nguyen Dynasty. Pre-processing steps are fully conducted before entering the model to ensure accuracy and fairness. Shown in Fig 6

B. Evaluation Index

In order to more comprehensively evaluate the results of different models in experimenting on the above data set. We use the following 4 metrics:

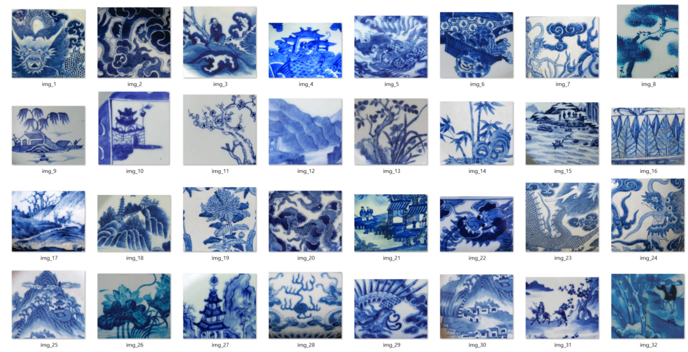

Fig. 6. Part of the the data set of patterns on cobalt blue underglazed porcelains under the Nguyen Dynasty.

L1 Loss: Mean absolute error (MAE), measures the average absolute difference between the predicted and ground truth pixel values. It provides a measure of how much the model's predictions deviate from the true values.

$$L1Loss = \frac{\sum |x^f - x^r|}{\sum_{i\in\omega} x_i^f} \tag{1}$$

L2 Loss: The average squared difference between the predicted and actual pixel values is measured by the mean squared error (MSE). Greater errors are amplified by it than by L1 loss.

$$L2Loss = \frac{\sum (x^f - x^r)^2}{\sum_{i\in\omega} (x_i^f)^2} \tag{2}$$

Peak Signal-to-Noise Ratio, or PSNR, is a widely used statistic to rate the quality of an image. It calculates the difference between the maximum power of a signal and the strength of the noise that degrades the signal's fidelity. Better image quality is indicated by higher PSNR values.

$$PSNR = 20 * \log_{10}\left(\frac{MAX_1}{\sqrt{MSE}}\right) \tag{3}$$

$$MSE = \frac{1}{m*n}\sum_{i=1}^{m-1}\sum_{j=1}^{n-1}(x_{i,j}^f - x_{i,j}^r)^2 \tag{4}$$

Total Variation Loss (TV Loss): TV loss is a regularization term that encourages smoothness in the image by penalizing rapid changes in pixel values. It helps in reducing noise and producing visually pleasing results. Lower TV loss indicates smoother image reconstruction.

$$R_{V^\beta}(f) = \int_\Omega \left(\left(\frac{\partial f}{\partial u}(u,v)\right)^2 + \left(\frac{\partial f}{\partial v}(u,v)\right)^2\right)^{\frac{\beta}{2}} dx \tag{5}$$

$$R_{V^\beta}(x) = \sum_{i,j}\left((x_{i,j+1} - x_{i,j})^2 + (x_{i+1,j} - x_{i,j})^2\right)^{\frac{\beta}{2}} \tag{6}$$

By considering these evaluation indicators together, the research aims to provide a comprehensive assessment of the performance differences between the models. Each indicator contributes a different aspect of image quality, such as absolute difference, squared difference, image fidelity, and smoothness.

C. Experimental Results and Conclusions

Table 2 displays the comparison outcomes of image restoration on portions of data image patterns on cobalt blue underglazed porcelains produced during the Nguyen Dynasty.

TABLE II
EVALUATION INDEX OF THREE DIFFERENT METHODS

Method	**L1 loss (%)**	**L2 loss (%)**	**PSNR**	**TV loss (%)**
Method 1	11.16	7.28	17.59	19.22
Method 2	40.73	30.59	11.16	11.83
Method 3	3.76	1.12	20.05	90.35

According to Table 2, the first technique in this study work performs better than the other two methods in terms of a number of different indices, including L1 loss, L2 loss, PSNR, and VT loss. The Nguyen Dynasty pattern design dataset's conventional pattern images were restored using all three procedures, namely methods 1, 2, and 3. The results that followed were attained.

In Figures 7-10, (a) displays the original designs on cobalt blue underglazed porcelains from the Nguyen Dynasty, (b) the processed image. Pattern image (c) is the result of the masking process. Pattern image (d) is the result of method 1 repair. Pattern image (e) is the result of method 2 repair. Pattern image (f) is the result of method 3 repair.

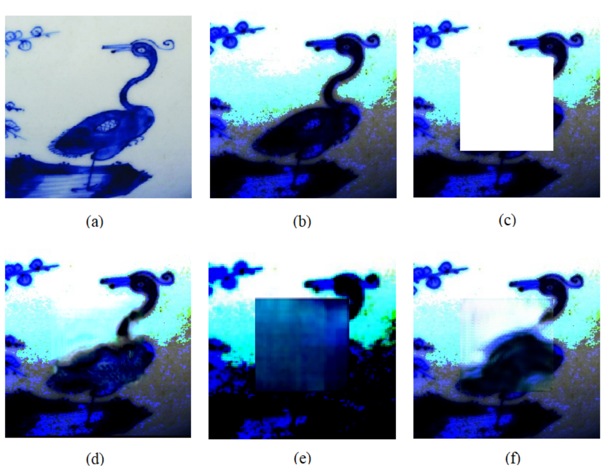

Fig. 7. Results of several mending techniques are compared. The original picture (a). Image input (b). c) A disguised image. (d) The outcome of Repair Method 1. (e) The outcome of Method 2 repair. The repair outcome using Method 3 is (f).

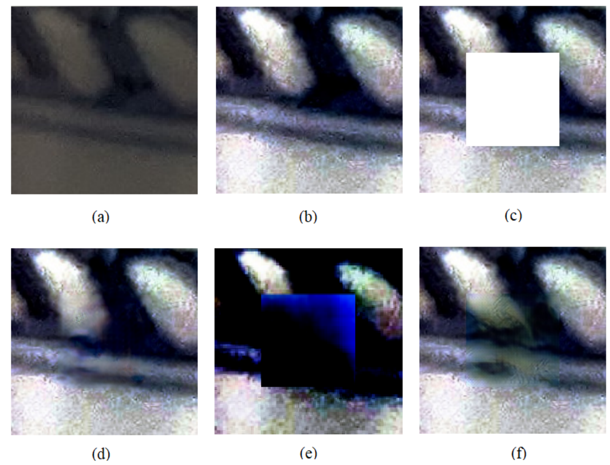

Fig. 8. Results of several mending techniques are compared. The original picture (a). Image input (b). c) A disguised image. (d) The outcome of Repair Method 1. (e) The outcome of Method 2 repair. The repair outcome using Method 3 is (f).

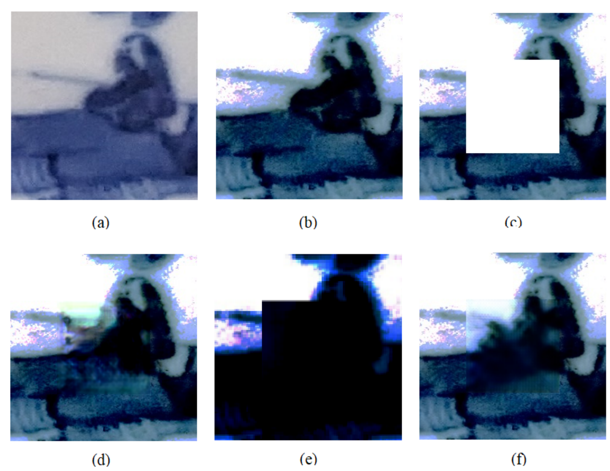

Fig. 9. Results of several mending techniques are compared. The original picture (a). Image input (b). c) A disguised image. (d) The outcome of Repair Method 1. (e) The outcome of Method 2 repair. The repair outcome using Method 3 is (f).

All traditional pattern images requiring restoration in this study have their central image pattern concealed, resulting in the majority of image information being masked. The restoration findings show that procedure 2 has the least effective restoration. The typical pattern image has illogical geometric elements and texture information because the repaired area is dark and highly impacted by irrelevant backdrop colors. Method 3 performs better, with the restored image's geometric information aligning more closely with human perception of traditional patterns. However, method 3 struggles to produce satisfactory texture effects. On the other hand, the first method proposed in this paper achieves superior

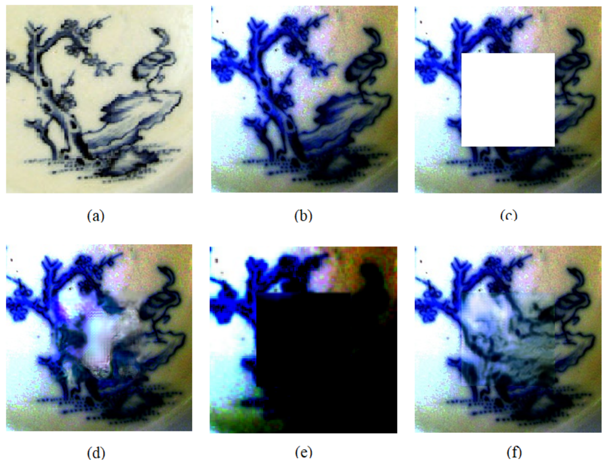

Fig. 10. Results of several mending techniques are compared. The original picture (a). Image input (b). c) A disguised image. (d) The outcome of Repair Method 1. (e) The outcome of Method 2 repair. The repair outcome using Method 3 is (f).

restoration results compared to the other two methods. In order to do a more logical picture restoration based on the surrounding information of the missing area, it successfully learns the geometric structure and texture information of the pattern.

The results of the experiments conducted on the Nguyen Dynasty pattern design dataset demonstrate the superiority of the approach described in this paper for assessing the efficacy of restoring conventional pattern pictures. It is difficult to recover the geometric and texture details of the missing sections using technique 2 because background pixel information can interfere with traditional needlework images' sophisticated information structures. Method 3 is less effective at recovering texture information even though it can better restore geometric structure information. In comparison, the restoration outcomes from the first method show more logical geometric structures and pattern information. In order to successfully restore information, the first model proposed in this paper generates a logical conventional pattern picture structure and texture information for the missing areas based on the information qualities surrounding those areas.

IV. Conclusion

This research aims to contribute to the archiving and digitization of historical image data of Vietnam by assembling a comprehensive data set of pattern images from various collections and sources. By applying and evaluating image restoration models based on deep convolutional neural networks, the study seeks to restore complex patterns on porcelain artifacts and shed light on the effectiveness of these models.

Although the application of deep learning techniques in historical conservation has shown significant progress, the restoration of ancient Vietnamese patterns remains largely unexplored. The recovery results of this study have not yet reached the expected results. In the future, it is necessary to expand the research on more complete data sets and improve the model structure to bring the best possible results. Besides focusing on pattern cobalt blue underglazed porcelains under the Nguyen Dynasty, the study can serve as a guide for the development of other image datasets.

The importance of this study lies in the preservation and understanding of Vietnam's cultural heritage. The restoration of ancient patterns on cobalt blue underglazed porcelains will provide insights into the artistic traditions of the Nguyen Dynasty and enhance our overall comprehension of Vietnamese history and culture. The application and bringing of new technologies such as artificial intelligence to solving historical problems brings a fresh approach to history for young people. Ultimately, this research aims to enrich our understanding of Vietnamese cultural heritage and inspire further advancements in the field of historical preservation.

V. Acknowledgments

We would like to sincerely thank the people and organizations who contributed significantly to the success of this project. First and foremost, we would like to express our sincere gratitude to Nguyen Quoc Trung, our supervisor, for his important advice and persistent support throughout this research project. Their knowledge and intelligent criticism were crucial in determining the course and ensuring the caliber of this work.

We would like to extend our special thanks to author Tran Duc Anh Son for their kind permission to use the images from their book "Commissioned Patterned Porcelains In the Nguyen Era". The inclusion of these images has significantly enhanced the completeness and visual appeal of our dataset.

We would like to express our special thanks to record holder Dinh Cong Tuong and collector Pham Huynh Trong Hieu for allowing direct access to this precious private collection in order to collect the best quality and accurate images. Furthermore, their historical information and knowledge are also very valuable and important to our research.

We would like to thank all individuals and groups for allowing us to use their images to complete the dataset. It is an integral part of this project.

We would like to thank FPT University for holding the competition that served as the inspiration for this project. Additionally, we would like to express our profound gratitude to the staff and teachers who planned and led the seminars during the competition.

We are truly honored and grateful for all those mentioned above. Without their contributions, this project would not have been possible.

REFERENCES

[1] O. Elharrouss, N. Almaadeed, S. Al-Maadeed, and Y. Akbari, “Image inpainting: A review,” *Neural Processing Letters*, vol. 51, no. 2, pp. 2007–2028, dec 2019. [Online]. Available: https://doi.org/10.1007%2Fs11063-019-10163-0

[2] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, “Generative image inpainting with contextual attention,” 2018.

[3] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, “Edgeconnect: Generative image inpainting with adversarial edge learning,” 10 2019.

[4] E. Sogancioglu, S. Hu, D. Belli, and B. van Ginneken, “Chest x-ray inpainting with deep generative models,” 2018.

[5] X. Fu, “Research and application of ancient chinese pattern restoration based on deep convolutional neural network,” *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–15, 12 2021.

[6] B. Truong, “Bo v và phát huy giá tr kho tàng di sn văn hóa phc v s nghip phát trin bn vng vit nam,” *VNU Journal of Science: Social Sciences and Humanities*, vol. 32, no. 1S, 2016. [Online]. Available: https://js.vnu.edu.vn/SSH/article/view/4405

[7] N. Trung and V. Van, “Educating traditional cultural values in vietnam universities,” *South Asian Research Journal of Humanities and Social Sciences*, vol. 2, p. 5, 06 2020.

[8] N. M. Tri and N. T. Dung, “Impact of the industrial revolution 4.0 on vietnamese cultural development,” *Linguistics and Culture Review*, vol. 6, no. 1, pp. 43–54, 2022.

[9] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Q. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, “Review of deep learning: concepts, cnn architectures, challenges, applications, future directions,” *Journal of Big Data*, vol. 8, 2021.

[10] Z. Qin, Q. Zeng, Y. Zong, and F. Xu, “Image inpainting based on deep learning: A review,” *Displays*, vol. 69, p. 102028, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0141938221000391

[11] H. Li and J. Huang, “Localization of deep inpainting using high-pass fully convolutional network,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[12] H. Liu, B. Jiang, Y. Song, W. Huang, and C. Yang, “Rethinking image inpainting via a mutual encoder-decoder with feature equalizations,” 2020.

[13] U. Demir and G. Unal, “Patch-based image inpainting with generative adversarial networks,” 2018.

[14] C. Ran, X. Li, and F. Yang, “Multi-step structure image inpainting model with attention mechanism,” *Sensors*, vol. 23, no. 4, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/4/2316

[15] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, and A. A. Efros, “Context encoders: Feature learning by inpainting,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2536–2544, 2016.

[16] Z. Hui, J. Li, X. Wang, and X. Gao, “Image fine-grained inpainting,” *ArXiv*, vol. abs/2002.02609, 2020.