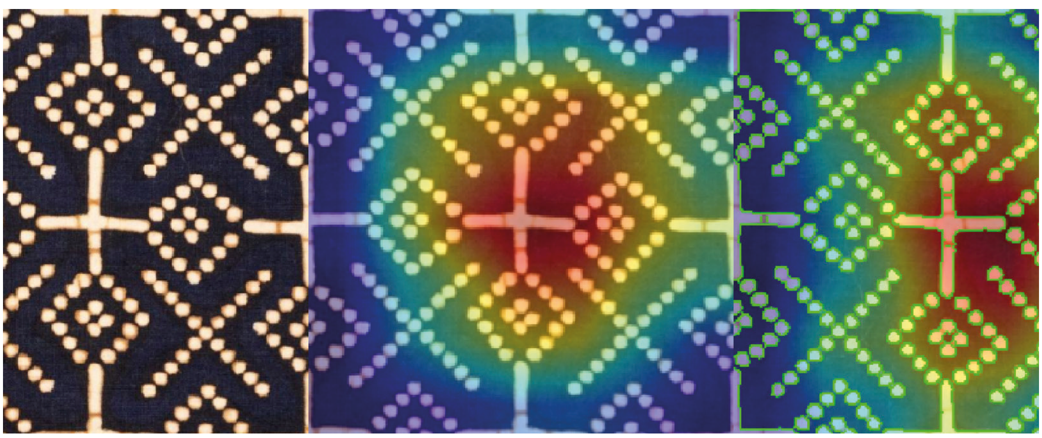
Algorithmic Representation of Batik Motifs: Visual Classification as a Form of Digital *Èkphrasis*

Giulia Flenghi

Abstract

In recent years, artificial intelligence has assumed an increasingly prominent role across various research fields, extending its scope to include the representation and study of architecture and cultural heritage. This rapidly evolving domain highlights the potential of AI technologies to provide innovative tools for the analysis, interpretation, and understanding of visual and decorative systems embedded in artistic and architectural traditions.
This paper examines the application of CNNs (Convolutional Neural Networks), with particular focus on the ResNet50 model, for the automatic classification of Batik patterns –a traditional form of visual expression with profound symbolic and cultural significance. Through the analysis of a dedicated dataset, the model training process, and the interpretation of results, the study investigates the potential of algorithmic description as a new form of *èkphrasis* –understood as a practice of translation between visual and textual languages, here mediated by computational methods.
The proposed approach integrates computer vision techniques with theoretical reflections on the relationship between digital representation and visual culture, emphasizing how artificial intelligence can contribute not only to the automation of classification processes but also to the understanding, enhancement, and preservation of artistic and decorative heritage.

Keywords
Deep Learning Classification, Grad-CAM Visualization, Algorithmic Representation, Geometric Pattern Recognition.



Grad-CAM visualization of a Batik Nitik motif. From left to right: original image, heatmap highlighting key regions for classification, and heatmap with edge detection overlay.

doi.org/10.3280/oa-1430-c899

Introduction

In recent years, artificial intelligence has increasingly expanded its role across various research domains, including the digital representation of architecture and cultural heritage. This development highlights how computational tools can support the analysis, interpretation, and enhancement of complex visual and decorative systems.
This study proposes a semi-automatic workflow for the classification and algorithmic representation of geometric patterns, employing deep learning techniques. The visual classification algorithm thus serves as a bridge between the visual language of patterns and their interpretation, mediated by computational processes. The study, representation, and manual classification of decorative motifs within the same historical or geographical context often constitute time-consuming tasks due to the quantity and variety of specimens.
The application of CNNs (Convolutional Neural Networks) to these specific tasks could accelerate this process but it is a field that remains underexplored, while recent research has primarily focused on the analysis of architectural elements, point clouds, and 3D models.
The aim of this research is to demonstrate how the automatic classification of geometric patterns through a CNN can be configured as a descriptive operation, enriching the dialogue between visual representation and digital tools. The proposed approach is structured into three main phases (fig. 1):

1. classification of decorative pattern images using deep learning;
2. evaluation of the classification results (F1-score, confusion matrices, and learning curves);
3. visualization and interpretability analysis (Grad-CAM heatmaps, edge detection, and sample predictions).

The methodology was tested on a dataset of Batik Nitik patterns, demonstrating the reliability, scalability, and reproducibility of the proposed approach.

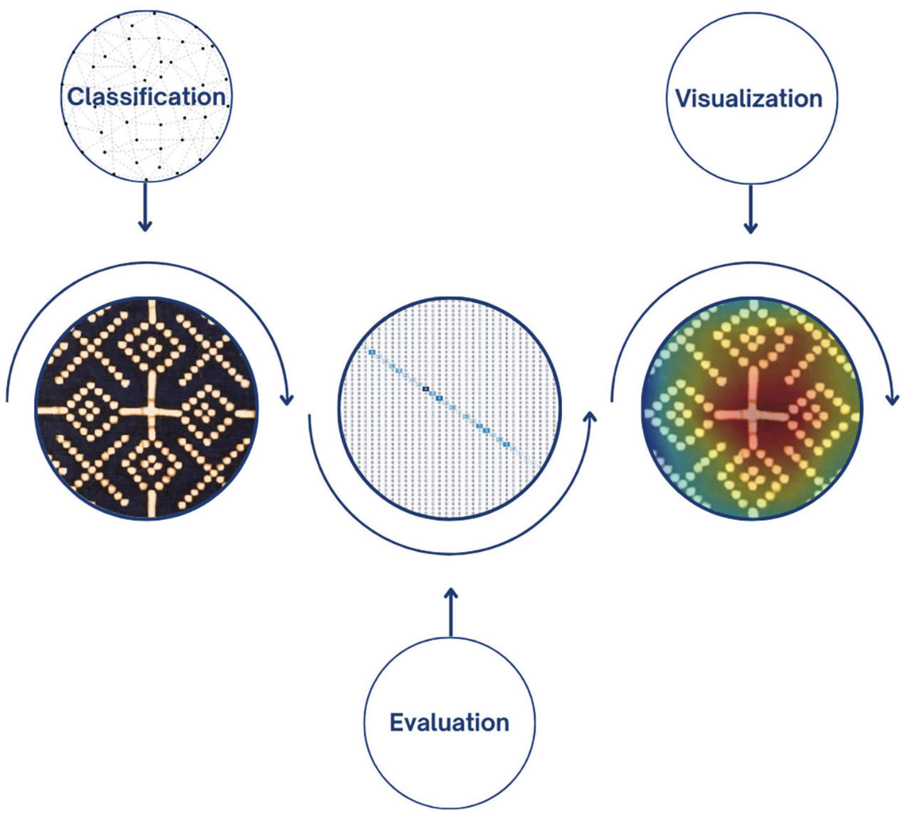


Fig. 1. The three main phases of the proposed methodology.

State of art

The application of computer science and image processing to Cultural Heritage has grown significantly in recent years [Bordoni, Mele, Sorgente 2016; Andrić *et al.* 2021; Bellavia *et al.* 2022; Fiorucci *et al.* 2020; Flenghi, Proietti 2024; Croce, Cera 2024]. However, research on the recognition, classification, and interpretation of geometric patterns in 2D data remains limited, with a notable absence of approaches combining automatic pattern classification with visualization tools to explain neural network decisions, supporting heritage analysis and conservation.
CNNs (Convolutional Neural Networks), introduced by LeCun *et al.* [1998] and later refined by LeCun, Bengio, and Hinton [2015], are now the leading technique for image recognition and classification. Their ability to capture spatial information makes them particularly effective for analysing repetitive decorative patterns and geometric structures. CNNs have been applied to cultural heritage [Bernasconi, Cetinić, Impett 2023; Billi *et al.* 2023; Castellano, Lella, Vessio 2021]; for instance, Ghosh *et al.* [2021] developed a deep learning framework for classifying geometric shapes in mosaics.
To address the opacity of CNNs, Grad-CAM (Gradient-weighted Class Activation Mapping), introduced by Selvaraju *et al.* [2019], generates heatmaps highlighting the image regions most influential to model predictions, aiding interpretability. In cultural heritage, Federico Milani [2021] used Grad-CAM for iconographic analysis, while Mahdi Bahrami and Amir Albadvi [2024] applied it to identify Iranian historical buildings in need of conservation. Despite these applications, Grad-CAM remains underexplored for geometric and ornamental patterns. The combined use of CNNs and Grad-CAM in this domain is an emerging research area, offering new tools to analyse subtle pattern variations and enhance documentation, analysis, and conservation.
The novelty of this approach lies in integrating CNNs and Grad-CAM to investigate the classification process of repetitive geometric patterns, improving both algorithmic transparency and the documentation of decorative heritage.

The Batik Nitik 960 dataset

The dataset used for this study is the Batik Nitik 960 [Minarno, Nugroho, Soesanti 2023b], a collection of 960 digital images documenting the 60 traditional patterns of Batik Nitik, a refined variant of Indonesian Batik (fig. 2).
Batik is an ancient textile art from Java, Indonesia, characterized by geometric and symbolic motifs that reflect tradition, aesthetics, and cultural identity. Batik Nitik, originating in Yogyakarta, is among the oldest and most intricate styles. It features geometric patterns composed of dots and lines, creating a woven fabric-like visual effect. This style was influenced by Indian Patola textiles, known for their double ikat technique [Zuhro *et al.* 2020]. Batik Nitik was formally recognized on February 19, 1940, when Bendara Raden Ayu Brongtodiningrat cataloged the first 56 motifs. Later, Haryani Winotosastro expanded the corpus to 60 patterns [Minarno, Nugroho, Soesanti 2023b].
The dataset was created to facilitate the application of artificial intelligence in the automatic classification of textile patterns [Minarno, Nugroho, Soesanti 2023a]. It initially comprised four images per motif, later increased to 16 through data augmentation, resulting in 960 .jpeg images (512×512 pixels). This augmentation enhanced the clarity of patterns and supported the training of machine learning models.
The dataset was selected for its open-access availability and the diversity of its geometric motifs, making it a robust benchmark for classification tasks.

Methodology

The research adopts a methodological approach that integrates automatic classification techniques with visualization and result interpretation tools, aimed at studying geometric decorative patterns.

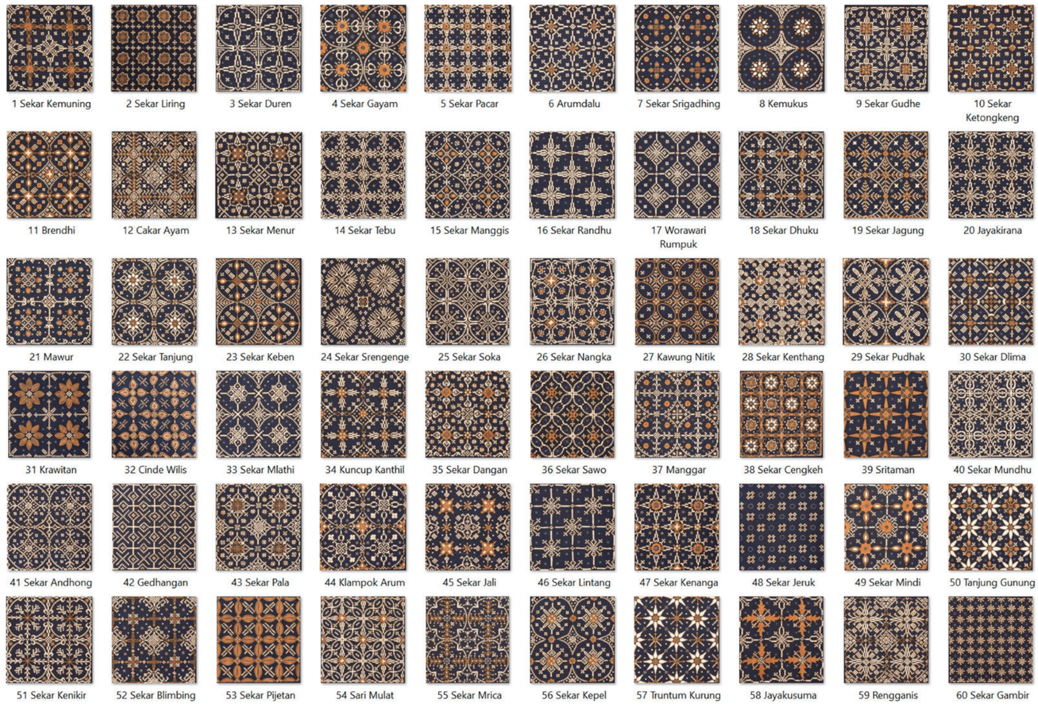


Fig. 2. The 60 categories within the dataset Batik Nitik 960 from Minarno *et al.*, 2023b.

This approach seeks to expedite the typological subdivision of decorations, facilitate their visual analysis, and support the study and preservation of artistic artifacts.
The proposed approach supports scalability and facilitates the transferability of the methodology to other decorative heritage studies.

Dataset Organization
The dataset was automatically organized into folders corresponding to the respective classes using a Python script called *folders.py*.
This structure facilitated the subdivision of data into three subsets:
- training set: used for model training;
- validation set: used to monitor model performance and prevent overfitting;
- test set: used for the final evaluation of model performance on unseen data.

The Batik Nitik 960 dataset already included images processed with offline data augmentation techniques, increasing the number of samples to 16 per class.
In addition to this, online data augmentation was applied dynamically during training. Specifically, transformations such as random resized cropping and horizontal flipping were used to introduce further variability in the training data.
This two-tiered approach –offline and online augmentation– was aimed at enhancing model robustness despite the limited dataset size.

Model Architecture
For the classification phase, the ResNet50 model was selected, a deep CNN widely recognized for its high performance in image classification and object detection tasks. ResNet50 is a pre-trained model on ImageNet, a large manually annotated image database, whose use allows leveraging features learned from a general domain for specific applications through a technique known as Transfer Learning.
This approach offers several advantages:
- reduction of training time;
- performance improvement even on small datasets.

In the adopted configuration, the initial convolutional layers of the model were 'frozen', meaning they were made non-trainable to preserve the features learned on ImageNet. The final layers were fine-tuned on the Batik Nitik dataset.

Training

The training process was conducted using the PyTorch library, with the following hyperparameters:

- batch size: 32;
- number of epochs: 20;
- early stopping with a patience of 5 epochs (interrupting training if no improvements are observed for 5 consecutive epochs);
- optimizer: SGD (Stochastic Gradient Descent), with:
  - learning rate: 0.001;
  - momentum: 0.9;
  - weight decay: 1e-4;
- scheduler: StepLR with step_size=7 and gamma=0.1 (reducing the learning rate by a factor of 0.1 every 7 epochs);
- loss function: CrossEntropyLoss, suitable for multi-class classification problems.

At the end of the process, the model successfully classified the images in the dataset (fig. 3).

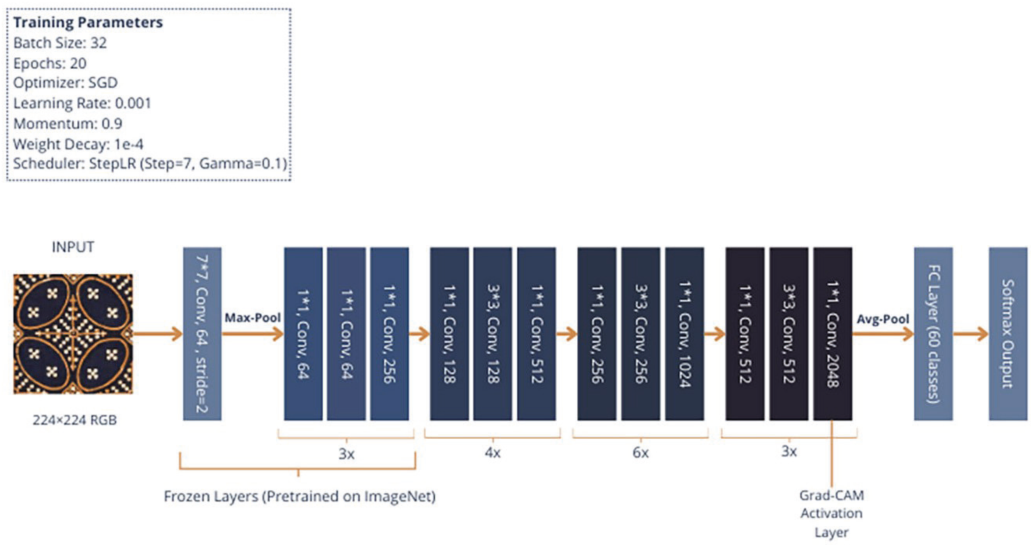


Fig. 3. ResNet-50 architecture for Batik Nitik classification. Frozen layers (pretrained on ImageNet) process 224×224 RGB images, with Grad-CAM applied to the final convolutional block. The fully connected layer outputs 60 classes. Training parameters are shown in the top-left.

Evaluation Metrics

The primary evaluation metric used was the F1 score, which balances precision and recall in multi-class classification tasks. These are defined as follows:

precision = TP / (TP + FP);

recall = TP / (TP + FN);

F1 Score = 2 × (precision × recall) / (precision + recall).

These metrics provides a reliable assessment of model performance, especially in cases of class imbalance.

Visualization

Various tools were employed to visualize the results: confusion matrices, learning curves, representative samples, and Grad-CAM heatmaps:

- Confusion Matrices: generated for both the validation (fig. 4) and test (fig. 5) phases, these matrices report the number of correct and incorrect predictions for each class. The matrix rows represent actual classes, while the columns indicate predicted classes. Elements

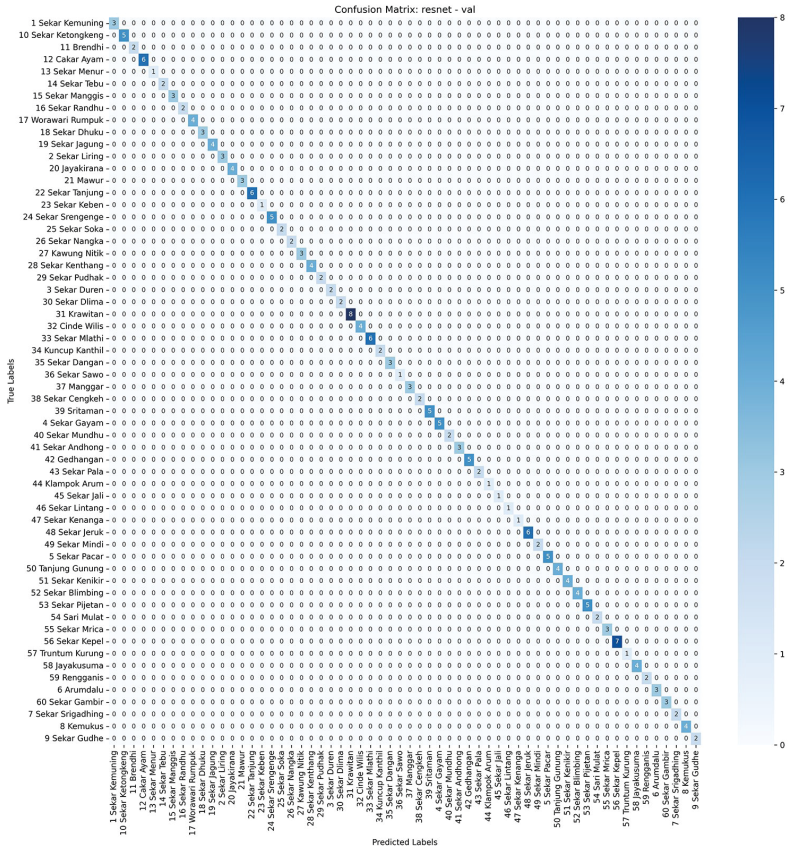


Fig. 4.Confusion matrix of the validation phase.

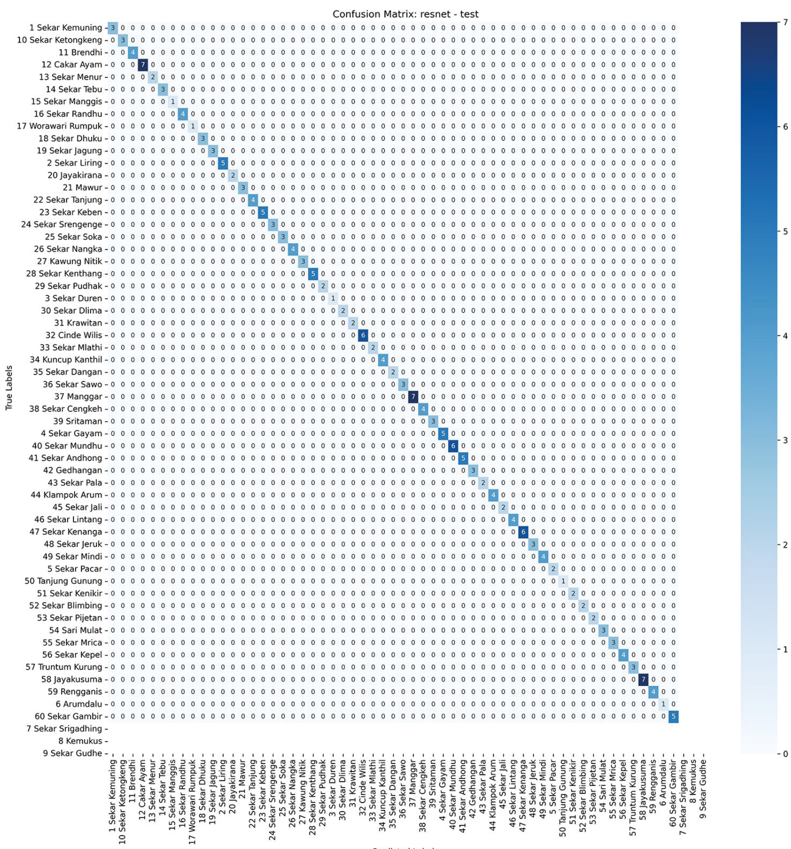


Fig. 5. Confusion matrix of the test phase.

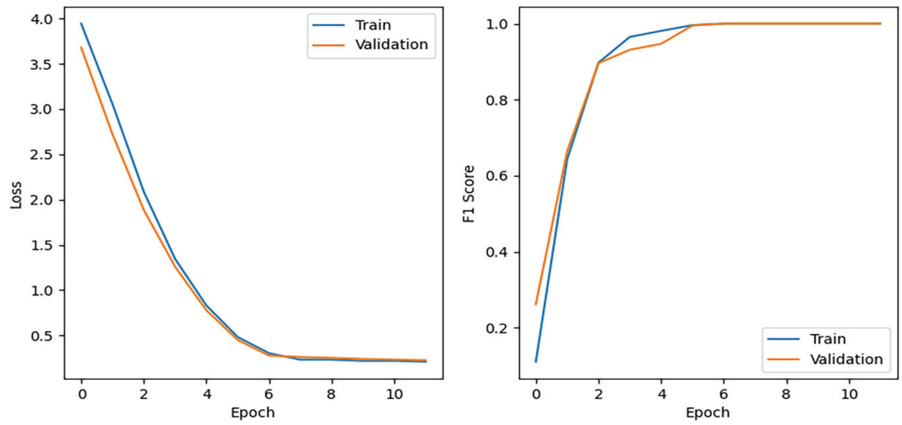


Fig. 6. Learning Curves. On the left, the loss computed on training and validation data; on the right, the F1 score trends.

along the diagonal represent correctly classified instances TP (True Positives). Off-diagonal elements highlight classification errors, which can be FP (False Positives) or FN (False Negatives);

- Learning Curves (fig. 6): these plots show the performance metrics over the epochs, such as the loss function and F1 score:

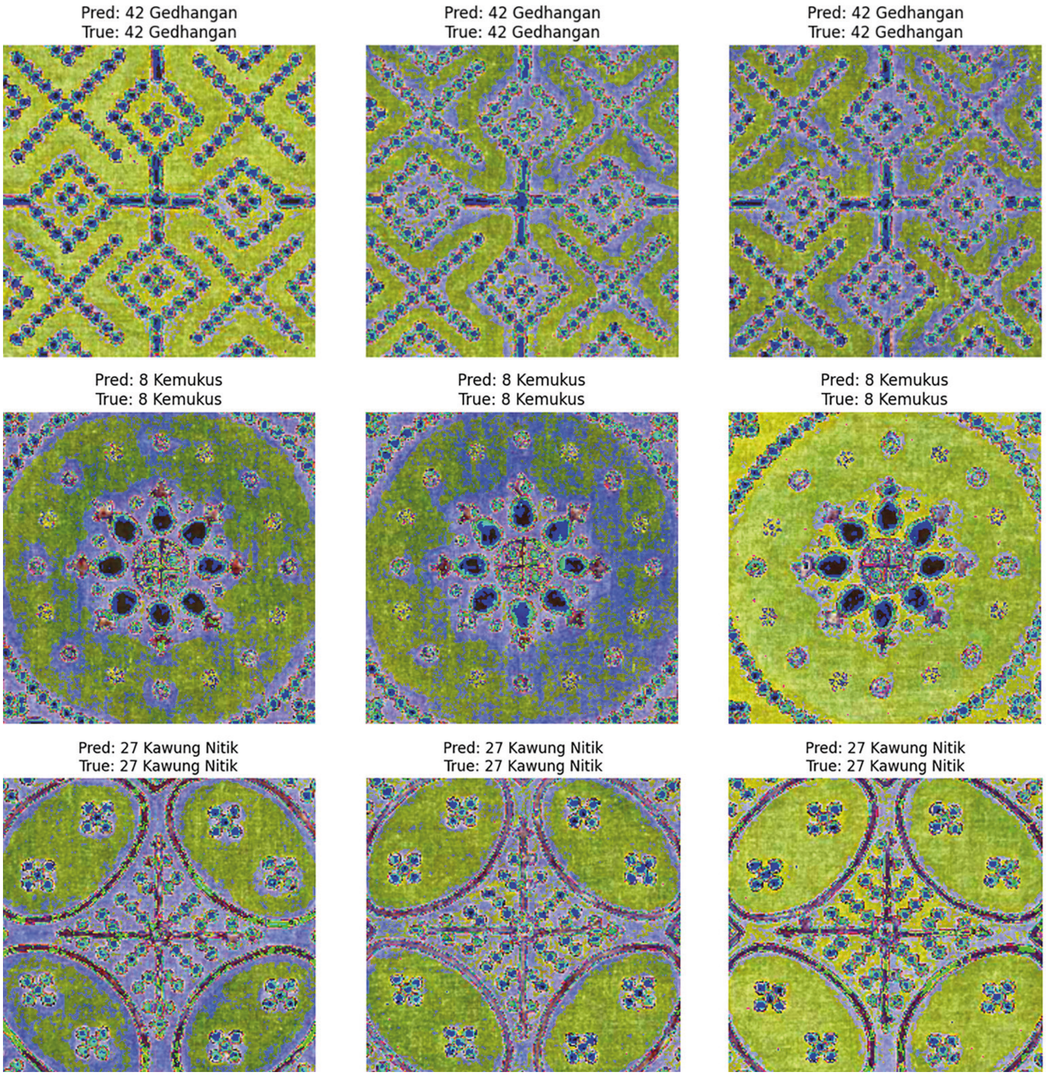


Fig. 7. Samples of correctly classified images

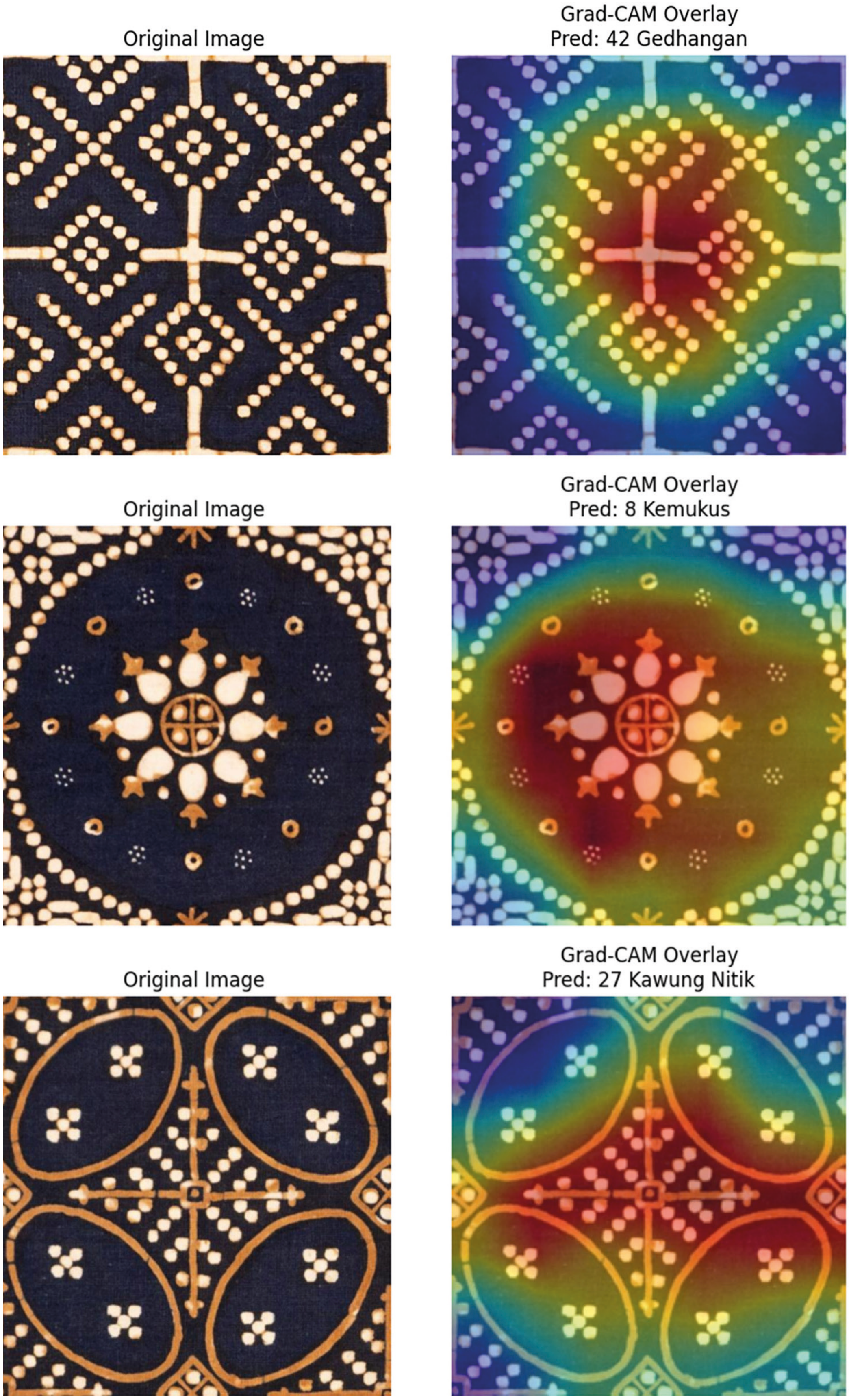


Fig. 8. The Grad-CAM heatmaps overlaid on the original images.

- left graph displays the loss computed on training and validation data –y-axis indicates the loss value; x-axis: represents the number of epochs (training cycles); blue line (train) shows the loss computed on the training set; orange line (validation) shows the loss computed on the validation set;
- right graph displays the F1 score trends –y-axis represents the F1 score, indicating the model's predictive quality. Higher values suggest better performance; x-axis represents the number of epochs; blue line (train) represents the F1 score on the training set; orange line (validation) represents the F1 score on the validation set;
- Sample Predictions: samples of correctly classified images and misclassified cases were selected and visualized. Each image was accompanied by its predicted and actual class labels to identify any visual ambiguities or recurring features within the decorative pattern classes (fig. 7);

- Grad-CAM Analysis: the model's interpretability analysis was conducted using Grad-CAM (Gradient-weighted Class Activation Mapping), a technique that highlights the salient areas of images that most influenced the model's decision.

This technique generates heatmaps overlaid on the original images, indicating regions with the highest activation (fig. 8):

- red and yellow areas denote regions with the greatest influence on the classification decision;
- blue areas represent less relevant regions.

To enhance visual clarity and facilitate the reading of heatmaps, Canny edge detection was applied, outlining the contours of the geometric shapes in the decorative patterns.

The superimposition of Grad-CAM heatmaps with detected edges allowed correlating the network's activation regions with the formal characteristics of the patterns, providing a more detailed understanding of the automatic recognition processes (fig. 9).

The analysis of Grad-CAM outputs enabled the verification of whether the model focused on geometrically significant elements for motif identification or, conversely, on marginal details. Visualization revealed that the model, in cases of correct classification, emphasized specific structural elements of the decorative patterns. In misclassification cases, the heatmaps indicated either attention dispersion or focus on irrelevant areas for typological recognition.

This visual analysis thus contributed not only to assessing the model's quality but also to gaining deeper insight into the geometric abstraction mechanisms adopted by the algorithm.

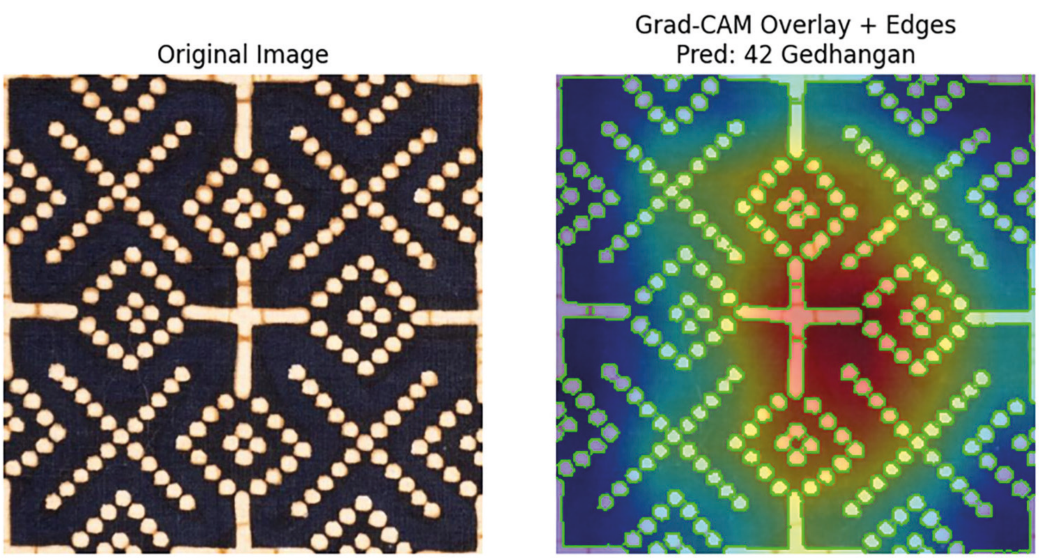


Fig. 9. The superimposition of Grad-CAM heatmaps with detected edges on the original images.

Results

The model demonstrated strong classification performance, achieving the following results:
- accuracy on the validation set: 100%;
- F1-score on the validation set: 1.00;
- accuracy on the test set: 100%;
- F1-score on the test set: 1.00.

The learning curves indicated a rapid convergence toward high performance, while the confusion matrices revealed no classification errors. Grad-CAM visualizations showed that the model focused on specific details of the patterns to determine class membership. Representative samples of images assigned to specific classes further confirmed the consistency between model predictions and actual labels. The results also indicate a strong internal coherence of the model, which was able to establish consistent and repeatable boundaries between motif classes. Notably, correctly classified samples often corresponded to patterns with strong geometric centrality or regular symmetries.

While no classification errors were observed in the final test set, earlier validation phases suggested that errors tended to occur in patterns exhibiting internal noise or subtle intra-class variations. This observation offers valuable insight into the model's behaviour and potential limitations.

The attainment of perfect metrics suggests the presence of overfitting raising concerns about the reliability of the validation process. This phenomenon is likely attributable to two main factors:
- the limited size of the dataset (16 images per class);
- the potential presence of highly similar images across the training, validation, and test subsets.

Although both offline and online data augmentation techniques were employed to increase the diversity of training samples, the limited variety of original data may have still constrained the model's generalization capacity. The augmentations, while useful for introducing small geometric or positional variations, were not sufficient to simulate broader variations in motif style or texture. As such, their contribution was effective in stabilizing training but insufficient to fully mitigate overfitting when similar samples existed across data splits.

It is also important to note that the dataset was randomly split into training, validation, and test sets without grouping by source pattern. Given that offline augmented versions of the same original images may have been distributed across different subsets, this approach may have led to a form of data leakage, potentially inflating performance metrics.

Moreover, although additional online augmentations (such as flipping and cropping) were applied during training, they did not prevent the possibility that highly similar images appeared in both training and evaluation sets. This limitation is acknowledged and motivates future work toward stricter dataset partitioning strategies that better evaluate generalization capabilities.

To mitigate overfitting and enhance the model's generalization capability, future research developments will focus on:
- expanding the dataset with additional images;
- increasing the use of data augmentation techniques to enhance sample variability;
- testing the model on images acquired under different conditions to evaluate its robustness and adaptability to contexts not represented in the training data.

Despite these limitations, the results highlight the potential of convolutional neural networks in the automatic analysis of geometric patterns.

Beyond simple classification, the model appears capable of capturing visually significant structures, even across subtle formal differences. The heatmaps generated with Grad-CAM provided insights into which areas of each pattern most influenced the model's decisions –typically revealing consistent attention to compositional axes, repetitive modules, or key contrast points.

These visualizations open new opportunities for applications beyond automation, including interpretive support for researchers, comparative stylistic studies, and computational typology in design and heritage contexts.

Conclusions

This research demonstrates the effectiveness of an AI-based approach for the automatic classification and algorithmic visualization of decorative geometric patterns. The integration of Grad-CAM with deep learning models has proven particularly valuable, not only for enhancing classification accuracy but also for revealing the visual features driving the model's decisions, fostering a more transparent interpretation of artistic artifacts.

The proposed method shows potential for broader applications within cultural heritage, serving as a tool for supporting study, cataloguing, and conservation practices. Grad-CAM visualizations, in particular, offer a novel means of computational description, facilitating the documentation and comparative analysis of ornamental details.

Nevertheless, some limitations persist –primarily the risk of overfitting– highlighting the need for future research to:

- develop more advanced models capable of addressing complex classification tasks;
- employ synthetic pattern generation techniques to enrich and diversify datasets;
- further investigate algorithmic visual tools to enhance digital documentation and the understanding of decorative heritage.

Ultimately, the combination of AI-based analysis and visual interpretation techniques opens new perspectives for the study and safeguarding of cultural artifacts, promoting interdisciplinary dialogue betw\en computational approaches and art-historical research.

Acknowledgments

The author sincerely thanks Kévin Réby for his essential contribution in selecting the model and developing the initial code for the experimentation.

Reference List

Andrić, V., Gajić-Kvaščev, M., Crkvenjakov, D. K., Marić-Stojanović, M., Gadžurić, S. (2021). Evaluation of pattern recognition techniques for the attribution of cultural heritage objects based on the qualitative XRF data. In *Microchemical Journal*, n. 167, Article 106267. https://doi.org/10.1016/j.microc.2021.106267.

Bahrami, M., Albadvi, A. (2024). Deep learning for identifying Iran's cultural heritage buildings in need of conservation using image classification and Grad-CAM. In *ACM Journal on Computing and Cultural Heritage*, n. 17(1), Article 16, pp. 1-20. https://doi.org/10.1145/3631130.

Bellavia, F., Colombo, C., Morelli, L., Remondino, F. (2022). Challenges in image matching for cultural heritage: An overview and perspective. In P. L. Mazzeo, E. Frontoni, S. Sclaroff, C. Distante (Eds.). *Image Analysis and Processing. ICIAP 2022 Workshops* ICIAP International Workshops, Lecce, May 23-27, 2022, vol. 1, pp. 210-222. Cham: Springer. https://doi.org/10.1007/978-3-031-13321-3_19.

Bernasconi, V., Cetinić, E., Impett, L. (2023). A computational approach to hand pose recognition in early modern paintings. In *Journal of Imaging*, n. 9(6), Article 120. https://doi.org/10.3390/jimaging9060120.

Billi, D., Croce, V., Bevilacqua, M. G., Caroti, G., Pasqualetti, A., Piemonte, A., Russo, M. (2023). Machine learning and deep learning for the built heritage analysis: Laser scanning and UAV-based surveying applications on a complex spatial grid structure. In *Remote Sensing*, n. 15(8), Article 1961. https://doi.org/10.3390/rs15081961.

Bordoni, L., Mele, F., Sorgente, A. (Eds.). (2016). *Artificial intelligence for cultural heritage*. Cambridge: Cambridge Scholars Publishing.

Castellano, G., Lella, E., Vessio, G. (2021). Visual link retrieval and knowledge discovery in painting datasets. In *Multimedia Tools and Applications*, n. 80(5), pp. 6599-6616. https://doi.org/10.1007/s11042-020-09995-z.

Croce, V., Cera, V. (2024). Evaluation of annotation ambiguity in common supervised machine learning classification approaches for cultural heritage. In A. Giordano, M. Russo, R. Spallone (a cura di). *Advances in Representation. Digital Innovations in Architecture, Engineering and Construction*. Cham: Springer. https://doi.org/10.1007/978-3-031-62963-1_30.

Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., James, S. (2020). Machine learning for cultural heritage: A survey. In *Pattern Recognition Letters*, n. 133, pp. 102-108.

Flenghi, G., Proietti, M. (2024). The recognizability of a place through generative representation of intangible qualities. In A. Giordano, M. Russo, R. Spallone (Eds.). *Advances in Representation. Digital Innovations in Architecture, Engineering and Construction*. Cham: Springer. https://doi.org/10.1007/978-3-031-62963-1_20.

Ghosh, M., Obaidullah, S. M., Gherardini, F., Zdimalova, M. (2021). Classification of geometric forms in mosaics using deep neural network. In *Journal of Imaging*, n. 7(8), Article 149. https://doi.org/10.3390/jimaging7080149.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, n. 86(11), pp. 2278-2324. https://doi.org/10.1109/5.726791.

LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. In *Nature*, n. 521(7553), pp. 436–444. https://doi.org/10.1038/nature14539.

Milani, F. (2022). *Analysis of cultural heritage data for complex iconography studies*. Tesi di dottorato in Information Technology, relatore P. Fraternali, tutor N. Gatti. Politecnico di Milano. https://www.politesi.polimi.it/handle/10589/196870.

Minarno, A. E., Nugroho, H. A., Soesanti, I. (2023a). Batik Nitik 960 dataset for classification, retrieval, and generator. In *Data*, n. 8(4), Article 63. https://doi.org/10.3390/data8040063.

Minarno, A. E., Nugroho, H. A., Soesanti, I. (2023b). *Batik Nitik 960* (Mendeley Data; Version V3); *Dataset*. https://doi.org/10.17632/sgh484jxzy.3.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D. (2019). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *International Journal of Computer Vision*, n. 128(2), pp. 336-359. https://doi.org/10.1007/s11263-019-01228-7.

Zuhro, A., Sunarya, I., Nugraheni, W. (2020). Batik Nitik's existence in the postmodern era. In *Proceedings of the 3rd International Conference on Arts and Arts Education (ICAAE 2019)*, pp. 1-6. https://doi.org/10.2991/assehr.k.200703.001.

Author
Giulia Flenghi, Sapienza Università di Roma, giulia.flenghi@uniroma1.it

To cite this chapter: Giulia Flenghi (2025). Algorithmic Representation of Batik Motifs: Visual Classification as a Form of Digital *Èkphrasis*. In L. Carlevaris *et al.* (Eds.). *èkphrasis. Descrizioni nello spazio della rappresentazione/èkphrasis. Descriptions in the space of representation*. Proceedings of the 46th International Conference of Representation Disciplines Teachers. Milano: FrancoAngeli, pp. 2765-2776. DOI: 10.3280/oa-1430-c899.

 ISBN 9788835182412