Digital forensics for cultural preservation: multi-device image classification of the historic Surabaya City Hall

Ulfa Meilinda Putri[1], Imam Yuadi[2]
[1]Magister Study Program of Human Resource Development, Postgraduate School, Universitas Airlangga, Surabaya, Indonesia
[2]Department of Information Science and Library, Faculty of Social and Political Science, Universitas Airlangga, Surabaya, Indonesia

Article Info

Article history:

Received Nov 12, 2025
Revised Apr 16, 2026
Accepted May 17, 2026

Keywords:

Cultural heritage
Digital forensics
Gaussian filter
Image classification
Machine learning

ABSTRACT

Cultural heritage preservation increasingly relies on digital forensics to ensure authenticity and consistency in heritage documentation. This study presents a digital forensic framework based on machine learning for classifying multi-device images of the historic Surabaya City Hall. The dataset was collected from nine smartphone devices and preprocessed through standardization, 360° rotational augmentation, and three filtering methods: gaussian, median, and laplacian. Three supervised algorithms (support vector machine (SVM), K-nearest neighbor (KNN), and logistic regression (LR)) were evaluated using accuracy, macro average, and weighted average of precision, recall, and F1-score. The results indicate that image preprocessing substantially affects model performance, with the gaussian-filtered KNN achieving the best result, reaching 92% accuracy, and balanced macro and weighted F1-scores of 0.92-0.93. Confusion-matrix analysis revealed minor misclassifications among iPhone models with similar sensor characteristics, while other devices were accurately identified. The findings confirm that gaussian filtering improves feature consistency and that KNN's distance-based classification exhibits robustness across heterogeneous image sources. However, the study is limited to a single heritage object and a restricted number of devices, which may affect generalizability. The proposed framework provides a reproducible and interpretable method that supports digital authenticity verification and aligns with UNESCO's vision for open, transparent cultural heritage preservation.

Corresponding Author:

Ulfa Meilinda Putri
Magister Study Program of Human Resource Development, Postgraduate School, Universitas Airlangga
Surabaya, Indonesia
Email: ulfa.meilinda.putri-2024@pasca.unair.ac.id

1. INTRODUCTION

Cultural heritage connects past and present societies, reflecting creativity, historical identity, and collective memory [1]-[3]. In Indonesia, colonial-era buildings like Surabaya City Hall symbolize civic governance and cultural resilience [2], [4]. However, rapid urbanization, environmental degradation, and insufficient documentation practices continue to threaten the sustainability of many heritage sites [5]-[7]. UNESCO and similar organizations advocate digital technologies for reliable, reproducible preservation [6], [7].

Advances in AI and digital forensics combine computational imaging with authentication methods to support cultural preservation [1], [8]-[11]. Although AI-based restoration and reconstruction are increasingly applied [1], [12], challenges persist due to data reliability and device heterogeneity [2], [5], [8]. Images from diverse devices vary in sensors and compression, risking authenticity and classification accuracy [13], [14]. Therefore, integrating digital forensics with machine learning for image classification is crucial for verifiable workflows [9], [14]-[16].

Digital forensics now also serves heritage documentation alongside its traditional cybercrime role [9], [17]-[19]. Machine learning enhances forensic analysis with automated detection and attribution [16], [17], [20]-[23]. Algorithms like support vector machines (SVM) [24], K-nearest neighbor (KNN) [25], and logistic regression (LR) [26] balance interpretability and accuracy for image data [27]-[30]. These have been used in camera identification [31], forgery detection [32], [33], and tampering localization [34], showing promise in heritage forensics.

However, few studies address multi-device forensic classification for heritage images, often lacking comprehensive preprocessing to handle sensor noise and distortion [35]-[41]. There is a need for reproducible pipelines combining multi-device acquisition, preprocessing, and classification to verify heritage image authenticity [42]-[44]. This study proposes such a framework focused on Surabaya City Hall images, integrating gaussian, median, and laplacian filters with SVM, KNN, and LR classifiers to evaluate preprocessing effects on accuracy and forensic reliability. In the context of cultural preservation, digital forensics plays a crucial role not only in documenting heritage assets but also in verifying the authenticity and provenance of digital images. Multi-device image classification enables the identification of image sources, which is essential for ensuring data integrity, preventing manipulation, and maintaining trustworthy digital archives. Therefore, linking device-level forensic analysis with heritage documentation provides a robust foundation for reliable and verifiable cultural preservation systems. The contributions of this study are threefold: establishing a reproducible digital forensic and machine learning workflow for heritage imagery; examining multi-device variability on model performance; and providing an interpretable framework supporting sustainable heritage preservation policies. This study contributes to bridging the gap between digital forensics and cultural heritage preservation by demonstrating how device-level classification can enhance data authenticity, traceability, and reliability in heritage documentation.

2. RESEARCH METHOD

The digital forensic workflow of this study is summarized in Figure 1, comprising the following stages: selection of the research object, multi-device data collection, dataset compilation and standardization, image filtering, data augmentation, classification, and performance evaluation.

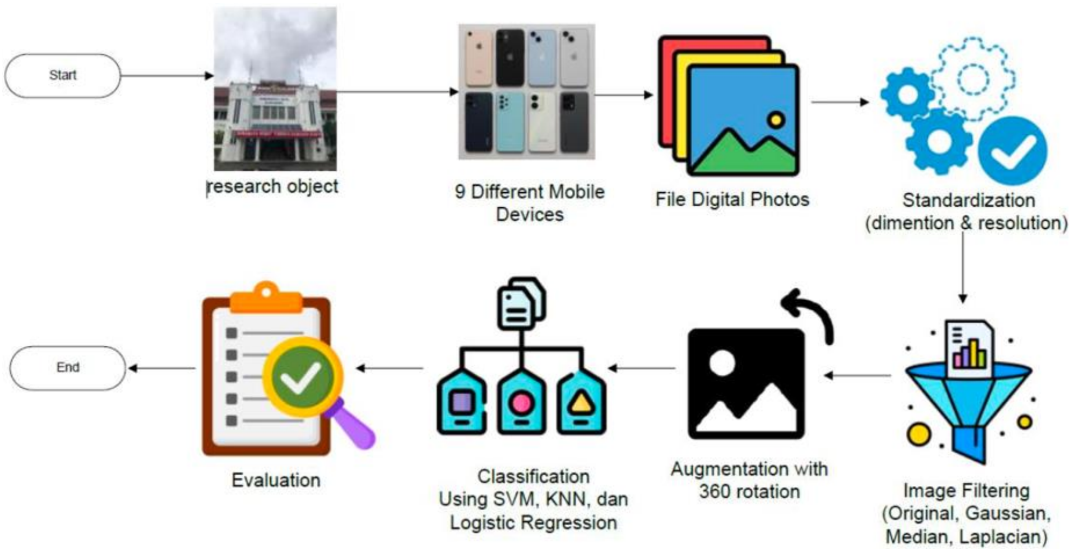

Figure 1. Workflow of the digital forensics classification process

2.1. Research object

The historic Surabaya City Hall was selected as the primary research object. This landmark represents dutch colonial architecture with its distinctive facades and symmetrical design, presenting complex visual features that make it suitable for digital forensic image classification consistent with UNESCO’s cultural-heritage documentation principles [6], [7].

2.2. Data collection

The dataset was collected by photographing Surabaya City Hall using nine different smartphone models spanning multiple brands, including iPhone, Oppo, Vivo, Samsung, and Xiaomi, with a total of 3,240 original images. Each device captured approximately 360 images at different elevations and angles, ensuring 360° coverage around the building. Metadata such as EXIF information, timestamps, and geolocation were preserved to maintain evidential integrity, consistent with forensic data-collection standards [14], [30], [31].

2.3. Dataset compilation and standardization

To eliminate inconsistencies caused by device variation, all collected images underwent strict standardization. All images were converted to a common format (JPEG/PNG) and normalizing resolution, resizing every image to 128×128 pixels. This step prevented classification biases based on dimensional disparities. Images were systematically labeled and organized into device-specific folders to preserve provenance traceability. This preprocessing step ensured that subsequent classification performance reflected algorithmic effectiveness rather than input inconsistencies [32], [39], [45].

2.4. Image filtering and preprocessing

To improve image quality and reduce noise prior to classification, four dataset variants were prepared: one unfiltered (original) and three filtered using gaussian, median, and laplacian techniques. These filters adjust image attributes such as smoothness, contrast, and edge definition. These filters are widely recognized in digital forensics and image processing research [32], [33], [39]-[41]. Filters were applied post-standardization to maintain consistent intensity distribution and facilitate comparative evaluation of their impact on classification accuracy.

Figure 2 presents representative samples from the four datasets. Figure 2(a), Figure 2(b), Figure 2(c), and Figure 2(d) show the original, gaussian-filtered, median-filtered, and laplacian-filtered images, respectively. Gaussian and median filters yield cleaner, more uniform inputs suitable for pattern-based classifiers, whereas the laplacian filter emphasizes geometry and contrast for edge-focused feature extraction. These variations are expected to influence the classification performance discussed in section 3, where their quantitative effects are compared.

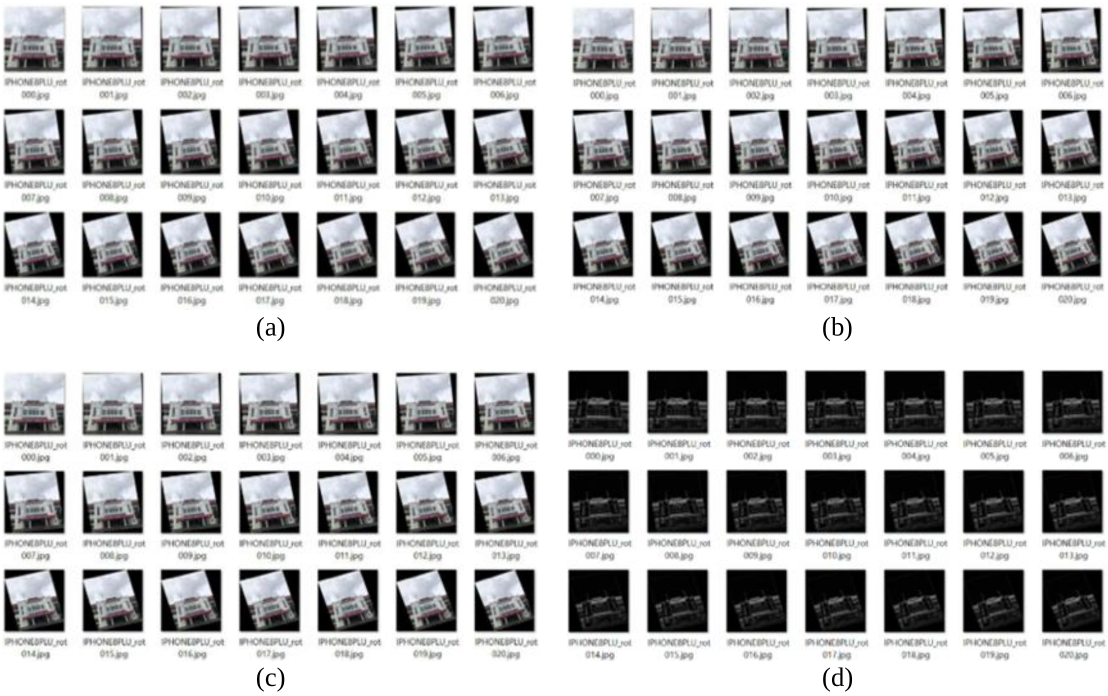

Figure 2. Sample images illustrating the effects of preprocessing methods: (a) original, (b) gaussian-filtered, (c) median-filtered, and (d) laplacian-filtered images

2.5. Original dataset
The original dataset contains standardized images without filtering, preserving authentic visual features, including sensor-specific noise, color tones, and compression artifacts. This dataset served as the baseline for evaluating the impact of filtering techniques, as these subtle characteristics are essential for device-level forensic attribution, where authenticity relies on subtle pixel-level differences [30], [31].

2.6. Gaussian filter
The gaussian filter smooths the image by averaging neighboring pixels weighted based on their distance from the center pixel. It effectively reducing high-frequency noise and illumination irregularities while maintaining architectural structure. This technique is commonly used in forensic imaging to stabilize intensity variance before feature extraction [28], [39], [41]. In this study, gaussian filtering produced balanced brightness and reduced speckle noise across device sources. The image appears smoother with reduced random noise and consistent tone across surfaces.

2.7. Median filter
The median filter applies a non-linear approach replacing each pixel with the median value of its neighborhood, effectively removing “salt-and-pepper” noise while preserving sharp edges and details that critical for architectural features like window frames, ornamentation and wall patterns without excessive blurring [33], [39]. This filter balances noise reduction with edge preservation, and random bright/dark pixel artifacts are eliminated, that suited to heritage imagery.

2.8. Laplacian filter
The laplacian filter, a second-derivative operator, enhances edges by highlighting regions with rapid intensity changes, accentuating structural contours such as rooflines and façades [32], [40]. Due to high noise sensitivity, it is typically applied after normalization or mild gaussian smoothing to avoid false edges [39], [41]. This filter aids forensic analysis by sharpening degraded images and emphasizing architectural boundaries [32], [33].

2.9. Data augmentation
To improve model generalization and simulate variations in viewing angles, a 360° rotational augmentation was implemented. Each image was rotated at fixed intervals, 15°, 30°, 45°, to increase dataset diversity and prevent model overfitting. This augmentation is particularly relevant for architectural heritage objects where orientation invariance enhances generalization [13], [28], [29]. In a digital forensics context, rotational augmentation is particularly relevant because it ensures that device-specific features such as sensor noise patterns and compression artifacts remain invariant under geometric transformations. This allows the model to focus on intrinsic device characteristics rather than orientation-dependent features, improving robustness in real-world forensic scenarios where images may be captured from various angles.

2.10. Machine learning classification
Three supervised machine learning algorithms were used to classify images according to the device source. First, SVM: utilizes kernel-based optimization to separate classes in high-dimensional feature spaces and is effective for non-linear boundaries [24], [27], [36]. Second, KNN: a distance-based instance classifier that assigns labels based on similarity, offering robustness in smaller datasets [25], [30]. Last, LR: provides interpretable probabilistic predictions through weighted feature coefficients, facilitating transparent forensic evidence interpretation [26], [46]. Each model was trained using identical training-validation splits, and hyperparameters were optimized via grid-based cross-validation to ensure fair and rigorous comparison.

2.11. Evaluation
Model performance was quantitatively evaluated using accuracy, macro average, weighted average metrics [27]. Confusion matrix were generated to identify common misclassifications and assess robustness across device classes. Cross-validation mitigated risks of biased results due to data partitioning.

3. RESULTS AND DISCUSSION
This study addresses the gap in previous research related to multi-device forensic classification for cultural heritage images, particularly in handling device heterogeneity and preprocessing variability. The experimental evaluation compared the classification performance of three supervised machine learning algorithms, SVM, KNN, and LR across four preprocessing conditions: original, gaussian-filtered, median-filtered, and laplacian-filtered datasets. The comparison results are presented in Table 1, which summarizes the accuracy, macro-averaged, and weighted-averaged values for precision, recall, and F1-score.

Table 1. The comparison of performance measurement results from all models

Dataset	Model	Accuracy	Macro Avg precision	Macro Avg recall	Macro Avg F1-score	Weighted Avg precision	Weighted Avg recall	Weighted Avg F1-score
Original	SVM	0.79	0.82	0.79	0.79	0.82	0.79	0.79
	KNN	0.92	0.93	0.92	0.92	0.93	0.92	0.92
	LR	0.84	0.84	0.84	0.84	0.84	0.84	0.84
Gaussian-filtered	SVM	0.79	0.82	0.79	0.79	0.82	0.79	0.79
	KNN	0.92	0.93	0.92	0.92	0.93	0.92	0.92
	LR	0.84	0.84	0.84	0.84	0.84	0.84	0.84
Median-filtered	SVM	0.80	0.81	0.80	0.80	0.81	0.80	0.80
	KNN	0.91	0.91	0.90	0.91	0.91	0.90	0.91
	LR	0.79	0.80	0.79	0.79	0.80	0.79	0.79
Laplacian-filtered	SVM	0.78	0.79	0.77	0.78	0.79	0.77	0.78
	KNN	0.86	0.87	0.86	0.86	0.87	0.86	0.86
	LR	0.76	0.77	0.76	0.76	0.77	0.76	0.76

Overall, the KNN model consistently outperformed the other classifiers in all preprocessing variations. In particular, the gaussian-filtered KNN model achieved the highest overall performance, with an accuracy of 0.92, macro average precision of 0.93, recall of 0.92, and F1-score of 0.92. This demonstrates the superior generalization capability and robustness of the gaussian-KNN combination in handling heterogeneous multi-device image data. This behavior is consistent with the nature of sensor pattern noise (SPN), where images captured by the same device tend to form compact clusters in feature space, making distance-based classifiers such as KNN particularly effective.

In contrast, SVM achieved lower performance, with accuracy ranging between 0.78 and 0.80. This limitation is attributed to SVM's sensitivity to overlapping features and its dependence on kernel optimization for nonlinear data separation. LR, on the other hand, delivered moderate results with accuracy values between 0.76 and 0.84. While it provided clear probabilistic interpretations, its linear nature made it less capable of capturing the complex, nonlinear patterns inherent in architectural textures. Among the preprocessing methods, gaussian filtering consistently produced the most balanced results, followed by median filtering, which achieved slightly lower accuracy (0.91). The laplacian filter yielded the weakest performance (0.76–0.86), as it tends to amplify edge noise and reduce tonal stability, confirming previous findings that edge-only representations are insufficient for robust classification [32], [40].

These outcomes highlight that image preprocessing plays a critical role in enhancing classification performance, particularly when dealing with data collected from multiple devices. Gaussian filtering effectively suppresses random intensity fluctuations while preserving structural details, providing cleaner input for the KNN algorithm to compute accurate distance-based similarity measures. This explains why the gaussian-KNN configuration consistently produced the best results across all performance metrics. Furthermore, this finding aligns with earlier studies in digital forensics, which report that preprocessed images lead to higher attribution accuracy and improved forensic reliability [16], [36], [37]. The strong F1-score balance between precision and recall further indicates that the model maintained consistent recognition capability without overfitting to specific device characteristics.

To further validate the superior performance of the Gaussian-filtered dataset, Figure 3 presents the confusion matrices for all three models (SVM, KNN, and LR) under Gaussian preprocessing. The SVM model shows that iPhone cluster (12, 13, 15, 8+) exhibited strong overlap, while non-iPhone devices (e.g., Oppo_A60, VivoY36, Xiaomi14T) were more separable. This indicates that SVM struggled mainly with visually similar iPhone models, while excelling on devices with more distinct features. The KNN model shows near-perfect classification for several devices, including iPhone16 and VivoY36, with only minor overlap observed among similar iPhone models (12, 13, 8+) and between Xiaomi14T and Samsung_A52s. KNN's instance-based approach proved effective for most models, outperforming SVM by handling intra-class similarities more robustly. LR shows that misclassifications were concentrated in the iPhone series (12, 13, 8+) and occasional overlap between Oppo_A60 and Samsung_A52s. While less robust than KNN, LR maintained a strong balance between interpretability and performance.

From a digital forensic and cultural heritage perspective, these findings demonstrate that machine-learning-driven authentication can substantially improve the reliability and reproducibility of heritage image documentation. Specifically, the gaussian-KNN model offers a practical, interpretable framework that can be incorporated into automated systems for verifying image provenance across diverse acquisition devices. Such capabilities align closely with UNESCO's recommendations for digital archiving and authenticity verification in heritage preservation [6], [7]. Future research should further investigate inter-device misclassification patterns, particularly among visually similar smartphone models, to enhance forensic discrimination capability.

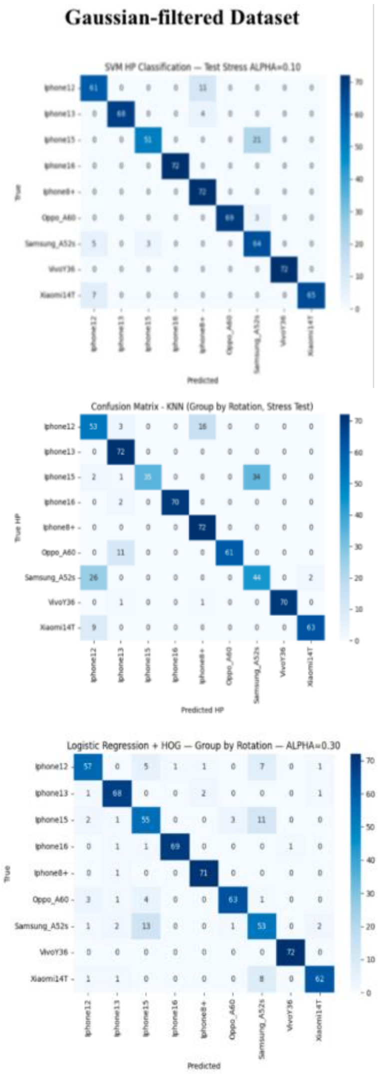

Figure 3. Confusion matrices of SVM, KNN, and LR models on gaussian-filtered dataset

4. CONCLUSION

This study proposed a digital forensics-based framework for cultural heritage preservation through multi-device image classification of the historic Surabaya City Hall. The framework integrates standardized preprocessing, noise filtering, and supervised machine learning algorithms to evaluate the authenticity and consistency of heritage imagery captured by nine different smartphones. The results demonstrate that preprocessing plays a pivotal role in enhancing model performance when dealing with heterogeneous image sources. Among the three models (SVM, KNN, and LR), the KNN model combined with gaussian filtering achieved the highest accuracy of 92%, with balanced precision, recall, and F1-score values between 0.92 and 0.93. The gaussian smoothing process effectively reduced sensor-related noise and intensity variation, enabling KNN to robustly identify similarity patterns across devices. In contrast, SVM and LR performed less effectively due to their sensitivity to overlapping features and linear decision boundaries.

Confusion matrix analysis further confirmed the superior device-level discrimination capability of the gaussian-filtered KNN approach, with misclassifications primarily confined to visually similar iPhone models. These findings suggest that the synergy of gaussian filtering and KNN offers a practical and interpretable solution for multi-device digital forensic classification, supporting enhanced data authenticity and reproducibility in cultural heritage documentation. This approach aligns with UNESCO's vision for transparent and reliable digital preservation of heritage assets.

Despite promising outcomes, several limitations merit attention. The dataset was restricted to nine smartphone devices and focused on a single heritage site, limiting generalizability to broader contexts. The study employed two-dimensional imagery and traditional filtering methods, without exploring advanced deep feature extraction or multimodal forensic cues. Future work should expand dataset diversity, integrate deep learning techniques for hierarchical feature analysis, and incorporate Explainable AI (XAI) frameworks to improve transparency and interpretability. Additionally, extending the framework to incorporate 3D reconstruction and multispectral imaging modalities could further enhance authenticity verification and advance sustainable, open-access digital heritage preservation. Overall, this study provides practical evidence that integrating digital forensics and machine learning can significantly enhance the reliability and trustworthiness of digital cultural heritage documentation.

ACKNOWLEDGMENTS

The author would like to express sincere gratitude to Dr. Imam Yuadi, S.Sos., M.MT., Ph.D., for his valuable guidance and support throughout this research. Appreciation is also extended to all lecturers and staff of the Master Program of Human Resource Development (PSDM), Faculty of Social and Political Sciences, Universitas Airlangga, for their assistance during the study. Special thanks are addressed to the Government of Surabaya City for providing permission, data, and valuable information essential to this research. Finally, the author is deeply thankful to her family and friends for their encouragement and moral support during this academic journey.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ulfa Meilinda Putri	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Imam Yuadi		✓					✓			✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, U.M.P., upon reasonable request.

REFERENCES

[1] N. Beridse, “Artificial intelligence in cultural preservation: reviving heritage through deep learning and virtual reconstruction,” 2025.

[2] A. Lukman, G. Assilmi, and I. N. Imandiharja, “Cultural heritage digitization in Indonesia: a new perspective on preserving depok colonial heritage,” *Kapata Arkeologi*, vol. 15, no. 1, pp. 15–24, Jun. 2020, doi: 10.24832/kapata.v15i1.15-24.

[3] R. Ristawati, R. Salman, S. A. Fitriany, and S. Taskesen, “Constitutional Protection of Cultural Heritage in Indonesia: the role of museums in preserving national identity and public welfare,” *Sriwijaya Law Review*, pp. 49-70, Jan. 2025, doi: 10.28946/slrev.Vol9.Iss1.3348.pp49-70.

[4] F. Ribera, A. Nesticò, P. Cucco, and G. Maselli, “A multicriteria approach to identify the highest and best use for historical buildings,” *Journal of Cultural Heritage*, vol. 41, pp. 166–177, Jan. 2020, doi: 10.1016/j.culher.2019.06.004.

[5] I. Robertus and R. Rasmita, “Preserving the cultural heritage of Indonesian society through digital preservation in libraries,” *Knowledge Garden*, vol. 3, no. 1, pp. 48–63, Apr. 2025, doi: 10.21776/ub.knowledgegarden.2025.3.1.32.

[6] UNESCO, “Operational guidelines for the implementation of the world heritage convention,” 2025. [Online]. Available: https://whc.unesco.org/fr/orientations

[7] UNESCO, “UNESCO and cultural heritage preservation 2017,” 2017. [Online]. Available: http://www.unesco.org/open-access/terms-use-ccbysa-en

[8] P. Westenberger and D. Farmaki, "Artificial intelligence for cultural heritage research: the challenges in UK copyright law and policy," *European Journal of Cultural Management and Policy*, vol. 15, Sep. 2025, doi: 10.3389/ejcmp.2025.14009.
[9] M. C. Stamm, Min Wu, and K. J. R. Liu, "Information forensics: an overview of the first decade," *IEEE Access*, vol. 1, pp. 167–200, 2013, doi: 10.1109/ACCESS.2013.2260814.
[10] F. Dedouit *et al.*, "The current state of forensic imaging – perspectives," Nov. 01, 2025, Springer Science and Business Media Deutschland GmbH. doi: 10.1007/s00414-025-03466-6.
[11] F. Colace, R. Gaeta, A. Lorusso, M. Pellegrino, and D. Santaniello, "New AI challenges for cultural heritage protection: a general overview," *Journal of Cultural Heritage*, vol. 75, pp. 168–193, Sep. 2025, doi: 10.1016/j.culher.2025.07.019.
[12] C. Gao *et al.*, "Applying optimized YOLOv8 for heritage conservation: enhanced object detection in Jiangnan traditional private gardens," Herit Sci, vol. 12, no. 1, p. 31, Jan. 2024, doi: 10.1186/s40494-024-01144-1.
[13] J. L. Lerma, M. Cabrelles, S. Navarro, and S. Fabado, "From digital photography to photogrammetry for cultural heritage documentation and dissemination," 2013, doi: https://doi.org/10.6092/issn.1828-5961/3850.
[14] D. Quick and K.-K. R. Choo, "Big forensic data reduction: digital forensic images and electronic evidence," *Cluster Computing*, vol. 19, no. 2, pp. 723–740, Jun. 2016, doi: 10.1007/s10586-016-0553-1.
[15] V. Srivastava, S. Omar, V. Shrivastava, and A. Professor, "Digital forensics investigation tool using machine learning for person detection in images," 2023.
[16] Foluke Ekundayo, "Big data and machine learning in digital forensics: predictive technology for proactive crime prevention," *World Journal of Advanced Research and Reviews*, vol. 24, no. 2, pp. 2692–2709, Nov. 2024, doi: 10.30574/wjarr.2024.24.2.3659.
[17] D. Dunsin, M. C. Ghanem, K. Ouazzane, and V. Vassilev, "A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response," *Forensic Science International: Digital Investigation*, vol. 48, Mar. 2024, doi: 10.1016/j.fsidi.2023.301675.
[18] M. Arjamand *et al.*, "The role of artificial intelligence in forensic science: transforming investigations through technology," *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, vol. 7, no. 5, pp. 67-70, 2024.
[19] C. S. R. S. Chandra, S. Palit, and C. Sinha Roy, "Machine Learning in Digital Forensic Investigation," *International Journal of Advances in Computer Science and Cloud Computing*, no. 10, pp. 2321–4392, 2022, [Online]. Available: http://iraj.in
[20] A. J. Jones, "Machine learning in digital forensic analysis," in *Digital Forensics in the Age of Ai*, 2024, pp. 219-246. doi: 10.4018/979-8-3373-0857-9.ch009.
[21] L. Tageldin and H. Venter, "Machine-learning forensics: state of the art in the use of machine-learning techniques for digital forensic investigations within smart environments," *Applied Sciences*, vol. 13, no. 18, p. 10169, Sep. 2023, doi: 10.3390/app131810169.
[22] T. Nayerifard, H. Amintoosi, A. G. Bafghi, and A. Dehghantanha, "Machine learning in digital forensics: a systematic literature review," *Arxiv*, 2023, [Online]. Available: http://arxiv.org/abs/2306.04965
[23] A. M. Qadir and A. Varol, "The role of machine learning in digital forensics," in *2020 8th International Symposium on Digital Forensics and Security (ISDFS)*, IEEE, Jun. 2020, pp. 1–5. doi: 10.1109/ISDFS49300.2020.9116298.
[24] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.
[25] O. Kramer, *Dimensionality reduction with unsupervised nearest neighbors*, vol. 51. in Intelligent Systems Reference Library, vol. 51. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-38652-7.
[26] D. W. Hosmer, S. Lemeshow, and R. X. *Sturdivant, Applied logistic regression*, Wiley, 2013.
[27] S. Islam *et al.*, "Image processing and support vector machine (SVM) for classifying environmental stress symptoms of pepper seedlings grown in a plant factory," *Agronomy*, vol. 14, no. 9, p. 2043, Sep. 2024, doi: 10.3390/agronomy14092043.
[28] E. Adamopoulos, "Learning-based classification of multispectral images for deterioration mapping of historic structures," *Journal of Building Pathology and Rehabilitation*, vol. 6, no. 1, Dec. 2021, doi: 10.1007/s41024-021-00136-z.
[29] R. Janković, "Machine learning models for cultural heritage image classification: comparison based on attribute selection," *Information*, vol. 11, no. 1, p. 12, Dec. 2019, doi: 10.3390/info11010012.
[30] S. Ferreira, M. Antunes, and M. E. Correia, "A dataset of photos and videos for digital forensics analysis using machine learning processing," *Data (Basel)*, vol. 6, no. 8, p. 87, Aug. 2021, doi: 10.3390/data6080087.
[31] J. Luka, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006, doi: 10.1109/TIFS.2006.873602.
[32] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, Mar. 2009, doi: 10.1109/TIFS.2008.2012215.
[33] Monika and A. Passi, "Digital image forensic based on machine learning approach for forgery detection and localization," *Journal of Physics: Conference Series*, vol. 1950, no. 1, p. 012035, Aug. 2021, doi: 10.1088/1742-6596/1950/1/012035.
[34] P. Chitti, K. Prabhushetty, and S. Allagi, "A deep learning and machine learning approach for image classification of tempered images in digital forensic analysis," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 10, 2022, doi: 10.14569/IJACSA.2022.0131069.
[35] Y. Al Balushi, H. Shaker, and B. Kumar, "The use of machine learning in digital forensics: review paper," in *Proceedings of the 1st International Conference on Innovation in Information Technology and Business (ICIITB 2022)*, Atlantis Press International BV, 2023, pp. 96–113. doi: 10.2991/978-94-6463-110-4_9.
[36] C. Shi, L. Chen, C. Wang, X. Zhou, and Z. Qin, "Review of image forensic techniques based on deep learning," *Mathematics*, vol. 11, no. 14, p. 3134, Jul. 2023, doi: 10.3390/math11143134.
[37] J. R. Del Mar-Raave, H. Bahşi, L. Mršić, and K. Hausknecht, "A machine learning-based forensic tool for image classification - a design science approach," *Forensic Science International: Digital Investigation*, vol. 38, p. 301265, Sep. 2021, doi: 10.1016/j.fsidi.2021.301265.
[38] J. Lee and J. M. Yu, "Automatic surface damage classification developed based on deep learning for wooden architectural heritage," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-M-1–2023, pp. 151–157, Jun. 2023, doi: 10.5194/isprs-annals-X-M-1-2023-151-2023.
[39] R. C. Gonzalez, *Digital image processing*, Pearson, 2009.
[40] D. C. Marr and E. C. Hildreth, "Theory of edge detection," in *Proceedings of the Royal Society of London - Biological Sciences*, vol. 207, no. 1167, pp. 187–217, Feb. 1980, doi: 10.1098/rspb.1980.0020.
[41] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*, MIT Press, 2006.
[42] A. M. Qadir and A. Varol, "The Role of Machine Learning in Digital Forensics," in *8th International Symposium on Digital Forensics and Security, ISDFS 2020*, 2020, doi: 10.1109/ISDFS49300.2020.9116298.

[43] A. Kuzin, A. Fattakhov, I. Kibardin, V. I. Iglovikov, and R. Dautov, "Camera model identification using convolutional neural networks," in *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2018, pp. 3107–3110. doi: 10.1109/BigData.2018.8622031.

[44] W. A. Baroto, "Advancing digital forensic through machine learning: an integrated framework for fraud investigation," *Asia Pacific Fraud Journal*, vol. 9, no. 1, pp. 1–16, Jun. 2024, doi: 10.21532/apfjournal.v9i1.346.

[45] E. Casey, *Digital evidence and computer crime: Forensic science, computers, and the internet*, Academic press, 2011.

[46] M. Jordan, J. Kleinberg, and B. Schölkopf, "Pattern recognition and machine learning," 2006.

BIOGRAPHIES OF AUTHORS

Ulfa Meilinda Putri Ulfa Meilinda Putri was born in Bojonegoro, Indonesia, on May 19, 1994. She earned her Bachelor's degree in Statistics from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, in 2016, with an undergraduate thesis titled "Risk Investment Analysis of Islamic Stocks Using Value at Risk by Approach of Bayesian Mixture Normal Autoregressive" supervised by Prof. Drs. Nur Iriawan, M.Ikom., Ph.D. She has professional experience in a management consulting firm, focusing on data analysis, organizational performance, and process improvement. She currently works at the Organization Division of the Surabaya City Government, where she is engaged in institutional analysis, public service evaluation, and bureaucratic reform initiatives. Her academic and research interests include data analytics, organizational performance, human resource development, and public sector management. She can be contacted at email: ulfameilinda@gmail.com.

Imam Yuadi He is a senior lecturer at the Department of Information and Library Science, Faculty of Social and Political Sciences, Universitas Airlangga, Surabaya, Indonesia. Born in Indonesia, he earned a Diploma in Library Technology from Universitas Airlangga (best graduate, 1996), a Bachelor's in Library and Information Science from Universitas Padjadjaran (1999), an M.MT. in Management of Technology from Institut Teknologi Sepuluh Nopember (2009), and a Ph.D. in Information Management from National Chiao Tung University, Taiwan (2017), where he received the Outstanding Contribution Award. His teaching and research cover digital libraries, digital forensics, image processing, and information systems. Noted for an interdisciplinary approach connecting information science, digital forensics, and computational analysis to fields such as anthropology, linguistics, and data preservation, he has authored numerous national and international publications and leads collaborative academic projects. He can be contacted at email: imam.yuadi@fisip.unair.ac.id.