

Hybrid Swin Transformer with Capsule Attention for Ceramic Tile Surface Classification

V. S. Manjula^{1*}, Hanadi Ahmed Elnaem¹, Lwanga N. Kulsum¹

¹*Department of Computer Science, Kampala International University, Kampala, Uganda, East Africa.*

**Corresponding author: manjula.subramanian@kiu.ac.ug*

Abstract. Ceramic tile surface classification is very important for quality control since it shows how various tile types have varied textures, patterns, and surface properties. To overcome the issue, a hybrid Swin Transformer with Capsule Attention architecture is created to successfully capture both global contextual data and fine-grained spatial interactions. The architecture combines hierarchical feature extraction from the Swin Transformer with dynamic routing-based capsule attention. This makes it possible to better represent features using a dual-branch design and a feature fusion method. The VxC TSG database is used for experiments. It has several tile categories, such as Agata, Berlin, Lima, Marfil, Mediterranea, Oslo, Petra, Santiago, Somport, Vega, and Venice. Each category has three classes that show different surface conditions. The suggested model has an average accuracy of 98.83%, a precision of 98.25%, a recall of 98.25%, and an F1-score of 98.25%. This shows that it works well in all areas. Also, an AUC of 97.81% shows that the difference between things quite well. The results show that using transformer-based learning with capsule attention works well for strong and accurate classification of ceramic tile surfaces.

Keywords: Ceramic Tile Surface Classification, Swin Transformer, Capsule Attention, Feature Fusion, Texture Analysis

INTRODUCTION

Modern production and quality inspection processes need precise ceramic tile surface classification to assure product dependability and aesthetic uniformity. Automated inspection methods struggle with tile class-specific texture, pattern, and structural composition. Manual inspection procedures are too slow, subjective, and error-prone for large-scale industrial use. Thus, intelligent and automated systems that handle complex surface features with accuracy and consistency are in demand. Many traditional machine learning methods use statistical texture descriptors and frequency-domain transformations for feature extraction. These approaches can discriminate, but they typically miss complicated tile surfaces' spatial linkages and global contextual information. Deep learning has enhanced convolutional neural networks (CNNs) by automatically learning hierarchical features. CNN-based models are confined to local receptive fields and may fail to capture global context and long-range interdependence.

Transformer-based designs have become popular because they can capture global linkages using self-attention. The Swin Transformer's hierarchical structure with shifting window-based attention allows fast processing while keeping global contextual knowledge. Despite these benefits, transformer models may not represent fine-grained spatial linkages and part-whole interdependence needed to discern visually identical surface patterns. Capsule networks construct spatial hierarchies using vector-based representations and dynamic routing algorithms to overcome these restrictions. Capsule networks capture spatial details by preserving positional and structural information. Independent capsule networks may struggle to capture global contextual information in big datasets with various texture patterns.

A hybrid design that combines transformer-based models with capsule networks may work in this case. The Hybrid Swin Transformer with Capsule Attention framework uses capsule networks to describe spatial relationships and Swin Transformer global feature extraction. The dual-branch architecture lets the system learn complimentary representations and combine them to improve categorization. This integration helps the model distinguish complicated ceramic tile surface types by handling global and local feature dependencies. The suggested technique is tested on the VxC TSG database, which includes ceramic tile types Agata, Berlin, Lima, Marfil, Mediterranea, Oslo, Petra, Santiago, Somport, Vega, and Venice, each with three surface condition classes. The dataset's variety offers a thorough baseline for model resilience and generalization. The proposed framework

Received: 06.12.2025 Revised: 25.02.2026 Accepted: 11.03.2026
Licensed under a CC-BY 4.0 license | Copyright (c) by the authors

uses sophisticated learning features and hybrid architectural design to surpass current methodologies and offer a scalable industrial surface classification solution. The main contributions of this work are summarized as follows:

- A novel hybrid architecture integrating Swin Transformer and Capsule Attention for enhanced ceramic tile surface classification.
- A dual-branch feature extraction mechanism that captures both global contextual features and fine-grained spatial relationships.
- Implementation of dynamic routing-based capsule attention to preserve structural and positional information.
- Effective feature fusion strategy that combines transformers and capsule representations into a unified feature space.
- Comprehensive evaluation of the VxC TSG database with multiple tile categories and class variations.
- Demonstration of improved robustness and generalization compared to conventional machine learning and deep learning approaches.

The structure of the approach is organized as follows: Section 2 presents a review of related work in ceramic tile surface classification and deep learning-based approaches. Section 3 describes the proposed Hybrid Swin Transformer with Capsule Attention architecture in detail. Section 4 explains the experimental setup, dataset description, and evaluation methodology. Section 5 discusses the results and performance analysis of the proposed model. Section 6 concludes the paper with insights, limitations, and future research directions.

RELATED WORKS

Dual-Branch Convolutional Neural Network with Acoustic Signal Features for Tile Debonding Detection was studied in [1]. The study uses tapping acoustic signal acquisition and converts the signals into two-dimensional feature representations using MFCC and CWT. It then uses a dual-branch convolutional neural network to extract features in parallel, fuses the features by concatenation, and classifies them using a deep learning-based SoftMax classifier. Deep learning architectures for tile defect detection performance evaluation was shown in [2]. The research utilizes image acquisition and preprocessing, succeeded by data augmentation, and implements various deep learning architectures such as CNN, MobileNetV2, ResNet50, and EfficientNetB0 through a transfer learning methodology, with model evaluation conducted via k-fold cross-validation and confusion matrix-based performance assessment.

Microcrack feature extraction for chaotic silicon nitride ceramic bearing roller optical surface using multi-scale wavelet transform augmentation and enhanced Particle Swarm Optimization–Fuzzy C-means (PSO-FCM) coupling was detailed in [3]. The research employs multi-scale wavelet transform (DWT) for image decomposition and enhancement, utilizes VisuShrink-based denoising and feature fusion, and implements an optimized PSO-FCM clustering algorithm for segmentation and microcrack feature extraction from chaotic optical surface images. Classifying ancient pottery provenance using microscopic images and ensemble deep learning was suggested in [4]. The research introduces a machine vision-oriented deep learning framework that incorporates multi-image fusion, dual-branch architecture comprising a fully convolutional network and a differential Siamese network, edge-based attention extraction, feature pyramid fusion, and anchor-free multi-head object detection with non-maximum suppression for the identification of ceramic defects.

Attention-based convolutional neural networks for ceramic tile defect classification are efficient was proposed in [5]. A deep learning framework for classifying defects in ceramic tiles that uses a CNN with the CBAM attention mechanism is provided, as well as image preprocessing and defect masking. The approach improves feature extraction by using channel and spatial attention modules to concentrate on the parts of the image that have defects. Cross-Interaction Transformer for ceramic tile surface flaw detection was detailed in [6]. A Cross Interaction Transformer (CIFormer) is presented, integrating coarse–fine granularity feature interaction and adaptive cross-scale attention techniques. The model uses a transformer-based architecture to find multi-scale contextual dependencies and use cross-attention-based feature fusion methodologies to find small errors. Finding surface defects in complex-textured ceramic tiles was presented in [7]. The development of a machine vision-based defect detection system involves image capture, preprocessing, saliency detection, and sub-block analysis, culminating in defect determination modelling. For finding flaws in complicated textured ceramic tiles, the method uses edge detection, segmentation, illumination correction, and feature-based classification. A comparison of old building surface fracture detection utilizing deep learning-based YOLO algorithms was briefed in [8]. For

identifying surface cracks, we use a deep learning-based object identification system that is based on YOLO architectures (YOLOv5–YOLOv10). The approach uses single-stage identification with grid-based bounding box prediction and tests many YOLO versions to find cracks quickly on structural surfaces.

A manufacturing surface fault inspection survey using generative models was shown in [9]. The study provides an overview of surface defect inspection utilizing generative models, classifying methodologies into Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Diffusion Models (DMs), and hybrid models. It emphasizes synthetic defect generation, feature learning, and data augmentation to enhance deep learning-based inspection systems. New microwave sensor for indoor residual spraying quality assurance was detailed in [10]. A microwave-based sensing system is created using a microstrip split-ring antenna, along with signal acquisition and machine learning classifiers (SVM, Random Forest, KNN) to find out the properties of materials by analyzing their frequency-domain electromagnetic response and classifying the patterns of the signals that were measured. Using active thermography and automated flaw identification to analyze building façade ceramic tile specimens was analyzed in [11]. The study utilizes active infrared thermography (IRT) alongside automated image processing and divide methods for defect detection, employing thermal excitation, heat flux analysis, and thermal image decomposition with spatio-temporal filtering to pinpoint and localize subsurface defects in ceramic tiles. Microstructure and characteristics were comparing flash sintering with industrial firing of ceramic wall tiles was briefed by [12]. The research examines flash sintering (FS) as an alternative processing method, employing electric field-assisted rapid sintering, and juxtaposes it with traditional firing techniques, evaluating microstructural development, phase formation, and material characteristics under regulated electrical and thermal conditions.

Comparing the Bayesian network and logistic regression models for ceramic tile defect prevention was detailed in [13]. The study uses Bayesian Networks and Logistic Regression to model the occurrence of defects through conditional probability relationships and statistical classification. Bayesian networks use directed acyclic graphs (DAGs) and conditional probability tables (CPTs) to show how different production variables affect ceramic tile defects. A novel method for characterizing ceramic tile glaze tribomechanical properties was described in [14]. The study suggests a glaze characterization protocol (GCP) that uses both analytical methods (thermal analysis, SEM, XRD, microhardness, profilometry) and technological tests (impact, abrasion, wear resistance based on ISO standards) to look at surface properties and help make ceramic tile glazes that work better in terms of tribology. The lightweight CTDD-YOLO algorithm identifies minor tile flaws was studied in [15]. A lightweight deep learning detection model (CTDD-YOLO) that combines context-anchor attention modules was suggested, dynamic convolution (ODConv), multi-scale feature fusion (CGRFPN), and optimized loss functions. This model improves feature extraction and the detection of tiny flaws in ceramic tile images by using an improved YOLO-based architecture. Evaluation of ceramic tile wear resistance using various techniques was detailed in [16]. The study utilizes a multi-attribute wear resistance evaluation, incorporating abrasion testing (ISO 10545-7), gloss variation measurement, scratch resistance analysis, microhardness testing, and SEM-based microstructural characterization to examine the tribological properties of glazed ceramic tiles through integrated analytical and standardized testing methodologies.

A study of its engobe and glaze formulas of Frit-free sun reflecting porcelain stoneware ceramic tiles made from recycled granite waste was proposed in [17]. The research suggests a sustainable ceramic tile formulation method utilizing recycled granite waste in the engobe and glaze layers, employing Design of Experiments (simplex-centroid mixture design) alongside thermal analysis, XRD, SEM, and solar reflectance assessment to create frit-free, solar-reflective porcelain stoneware tiles. Evaluation of substrate layer, adhesive mortar, ceramic tile product effects on ceramic covering performance was suggested in [18]. The study utilizes an experimental evaluation framework for ceramic tile systems, examining substrate adhesive tile interactions through mechanical testing (shear strength, flexibility), thermal and moisture cycling, and standardized adhesion tests to evaluate performance degradation and bonding behavior under varying environmental and material conditions. Automatic ceramic motif recognition and digital reconstruction using deep learning was proposed in [19]. A deep learning-based framework that uses CNNs, ViT, GANs, and diffusion models to automatically recognize and digitally recreate ceramic motifs was suggested in this study. This framework combines techniques for feature extraction, pattern classification, and generative reconstruction to help preserve cultural heritage. Automatic tile panel debonding evaluation using convolutional neural network was analyzed in [20]. The research creates a structural health monitoring system based on CNNs that uses wavelet-based scalogram modification of vibration data and image-based classification to find and measure interfacial debonding flaws in ceramic tile systems using automated pattern recognition.

MATERIALS AND METHODS

The proposed system introduces a hybrid deep learning framework that integrates the Swin Transformer and Capsule Attention to effectively classify ceramic tile surfaces. The architecture is designed to capture both global contextual information and fine-grained spatial relationships, which are essential for distinguishing visually similar tile categories. The architecture has a dual-branch design that blends hierarchical global feature extraction with spatial relationship modelling to improve how well it classifies things. Let the input image be represented as:

$$X \in \mathbb{R}^{224 \times 224 \times 3} \quad (1)$$

The preprocessing stage includes resizing and normalization:

$$X' = \frac{X - \mu}{\sigma} \quad (2)$$

where μ is mean, σ is standard deviation. This ensures uniform input distribution for both branches. The image is divided into non-overlapping patches:

$$X_p \in \mathbb{R}^{N \times (P^2 \cdot C)} \quad (3)$$

where P is 4 (patch size), N refers to number of patches. Each patch is projected into a feature space:

$$Z_0 = X_p W_e \quad (4)$$

where W_e is the embedding matrix. The architecture from Figure 1 has a dual-branch design that blends hierarchical global feature extraction with spatial relationship modeling to improve how well it classifies things. The Swin Transformer branch gradually decreases the number of spatial dimensions while increasing the number of features from 96 to 768. This allows it to successfully capture multi-scale contextual representations. The capsule branch uses convolutional feature extraction, principal capsules, and dynamic routing-based attention to maintain part-whole relationships through 16-dimensional capsule vectors at the same time. By combining features from both branches, a rich 944-dimensional representation that lets the model use both global and local information can be studied. Before being put into one of 11 classes, this composite representation is improved even further by a completely linked layer. The architecture strikes a good mix between attention-based learning and capsule-driven feature encoding. This makes the model better at telling the difference between complicated surface patterns in ceramic tiles.

Figure 2 shows how data moves through the proposed hybrid architecture, showing both simultaneous feature extraction and dimensional adjustments. The procedure starts with the tile image dataset, where unprocessed images are pre-processed to make them the same size and shape. Then, the processed data is sent to two parallel branches at the same time to extract features. The Swin Transformer branch makes global features, which makes a 768-dimensional feature vector. The capsule-based branch, on the other hand, extracts spatial features, which makes a 176-dimensional capsule feature vector ($11 \times 16D$). These features are saved and then put together using a technique called fusion to produce a 944-dimensional unified feature vector ($768D + 176D$). The classification step gets the fused representation. There, it is improved by a fully connected layer and then mapped to output classes using a SoftMax function. It clearly shows the input flows in several dimensions, how parallel processing works well, and features are integrated in a systematic way to get the final classification. As shown in Data Flow Diagram, the system follows a structured pipeline consisting of preprocessing, parallel feature extraction, feature fusion, and classification. The dual-branch design enables complementary feature learning, where transformer-based attention captures global dependencies, and capsule networks model spatial hierarchies.

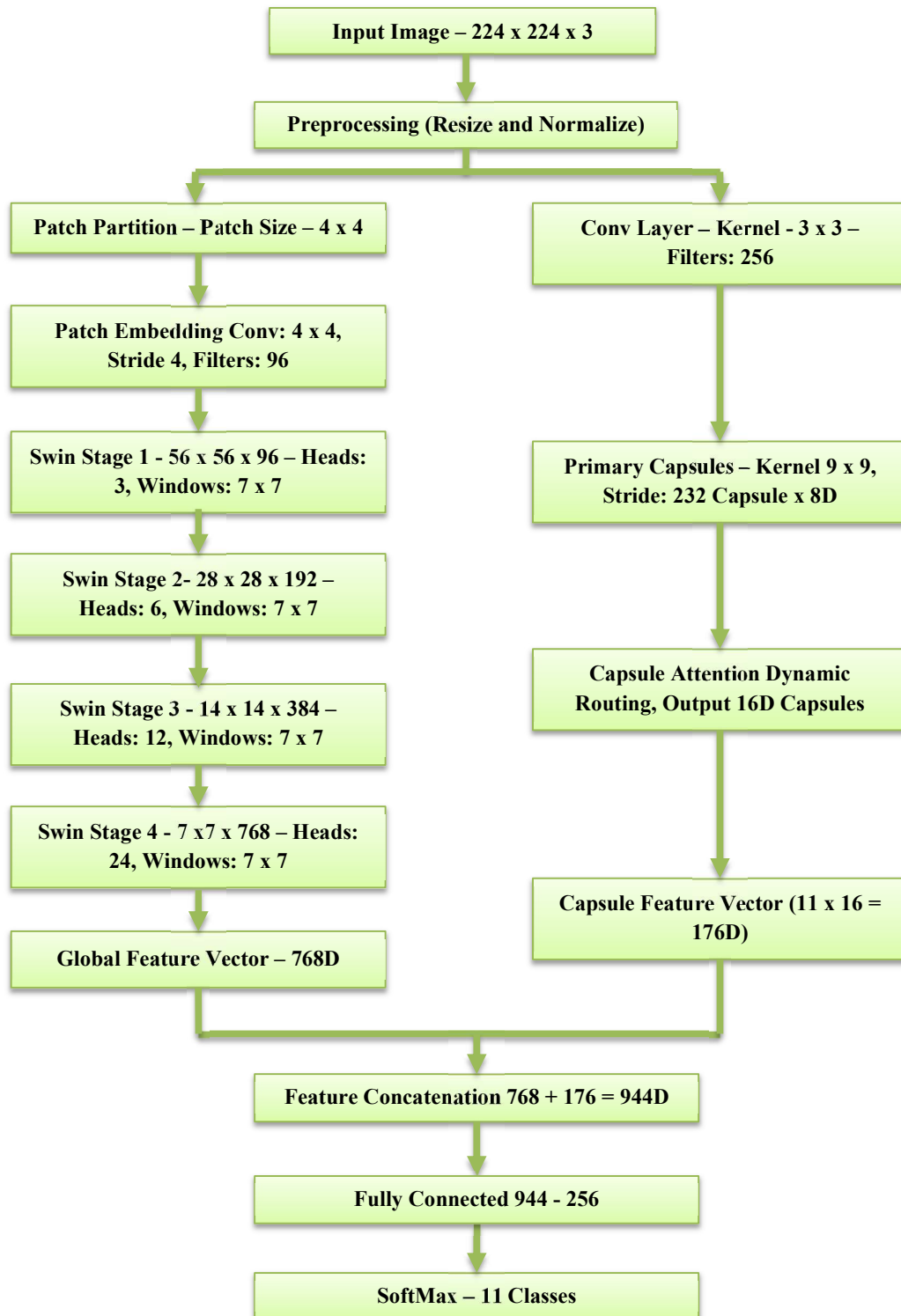


FIGURE 1. Hybrid Swin Transformer with Capsule Attention for Ceramic Tile Surface Classification

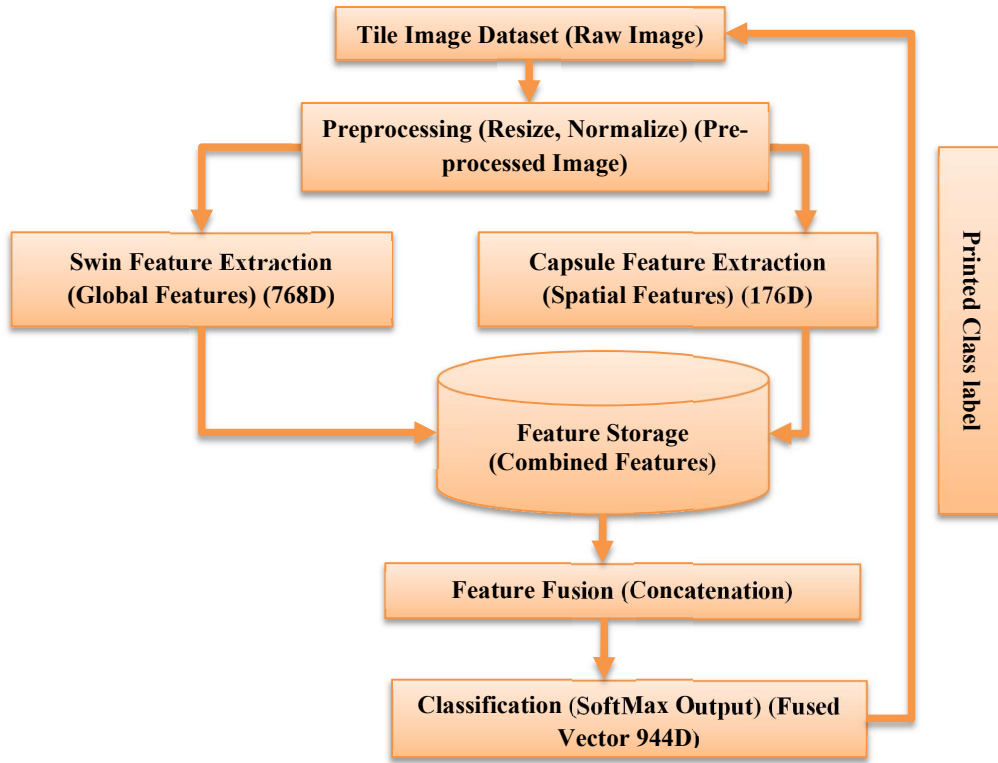


FIGURE 2. Data Flow Diagram of Hybrid Swin Transformer with Capsule Attention for Ceramic Tile Classification

The Swin Transformer applies Window-based Multi-Head Self-Attention (W-MSA):

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (5)$$

The shifted window mechanism enhances cross-window interaction. The hierarchical feature maps evolve as:

$$56^2 \rightarrow 28^2 \rightarrow 14^2 \rightarrow 7^2 \quad (6)$$

with feature dimensions:

$$96 \rightarrow 192 \rightarrow 384 \rightarrow 768 \quad (7)$$

Global Average Pooling (GAP):

$$F_{swin} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W Z_{ij} \quad (8)$$

$$F_{swin} \in \mathbb{R}^{768} \quad (9)$$

Initial feature extraction:

$$F_c = Conv(X') \quad (10)$$

RESULTS AND DISCUSSIONS

The VxC TSG database is a comprehensive dataset designed for tile surface classification and defect analysis, consisting of 956 images with 42 surface grades and 14 different ceramic models across 11 classes. Each class corresponds to a specific tile type, such as Venice, Vega, Marfil, Agata, Berlin, Lima, Oslo, Somport, Santiago, Petra, and Mediterranea, representing a diverse range of surface textures. The dataset includes variations in texture

patterns such as fixed, pseudorandom, and random structures, which are characteristic of different tile materials like marble and granite. The number of images per class is not strictly uniform, creating a slightly imbalanced dataset, which is beneficial for evaluating the robustness and generalization ability of machine learning and deep learning models. Since there is imbalance in classes data augmentation was implemented. A fivefold manifestation technique was adopted for solving the class imbalance. The data was split into a ratio of 70 % for training (4800 images), 15 % for validation (720 images), 15% for testing (720 images) With 956 total images, the dataset provides sufficient variability in texture, lighting, and structural appearance, making it suitable for tasks such as surface classification, texture recognition, detection. Figure 3 shows images of ceramic tile surfaces from different classes, showing how the texture, color distribution, and structural patterns in the dataset differ.



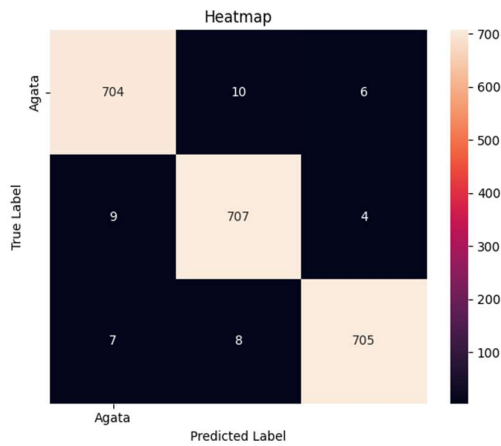
FIGURE 3. Ceramic tile surface images of different classes.

. The surfaces include a lot of different features, such as smooth, even areas and rough, uneven patterns. This is because of changes in the materials used. The first row has textures that are mostly the same, with just a few differences. This means that the surface attributes are the same. The second row shows mild changes in texture, with clear granular forms and changes in tone. The third row has textures that are more complicated and uneven, with greater roughness and irregularity. Table 1 details the performance assessment measures utilized to evaluate the proposed classification system. Accuracy, precision, recall, and F1-score.

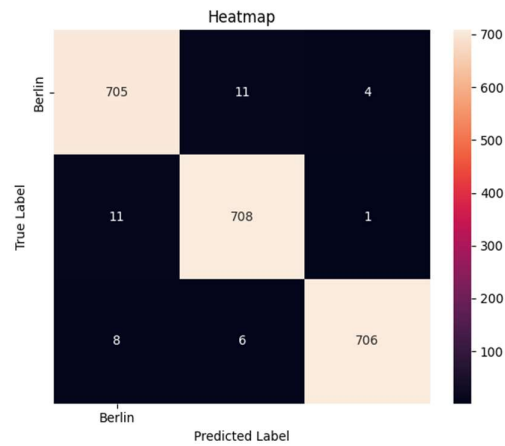
Table I. Performance evaluation metrics for ceramic tile surface classification

Metric	Definition	Equation	Description (Based on Your System)
Accuracy	Overall correctness of classification	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$	Measures total correctly classified ceramic tile images across all classes
Precision	Correct positive predictions	$Precision = \frac{TP}{TP + FP}$	Indicates how many predicted tile classes are actually correct
Recall (Sensitivity)	Ability to detect actual positives	$Recall = \frac{TP}{TP + FN}$	Measures how well the model identifies each tile class
F1-Score	Harmonic mean of precision and recall	$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$	Balances precision and recall for multi-class classification

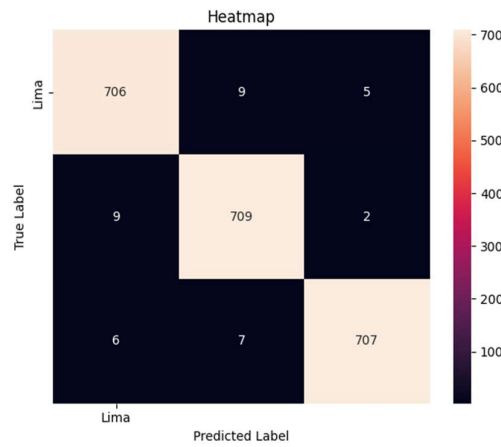
Figure 4 (a)–(k) shows the confusion matrix heatmaps for several types of ceramic tiles. These heatmaps show how well the proposed model works across various classifications. In Figure 4(a) Agata, Figure 4(b) Berlin, and Figure 4(c) Lima, there are large true positive values (over 700) and very few misclassifications.



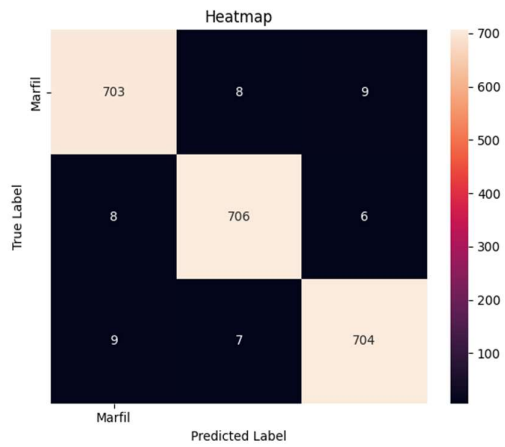
(a) Class: Agata



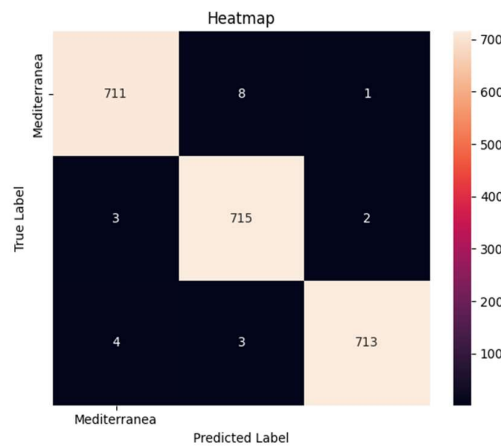
(b) Class: Berlin



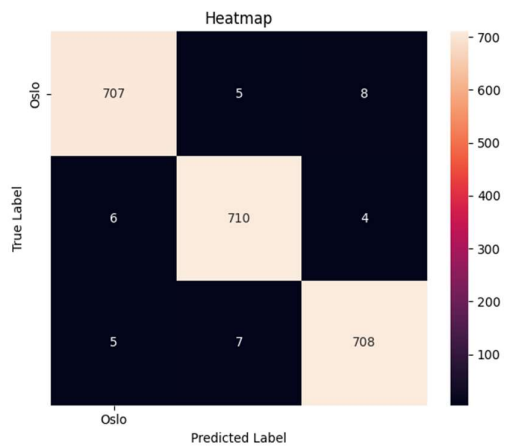
(c) Class: Lima



(d) Class: Marfil



(e) Class: Mediterranea



(f) Class: Oslo

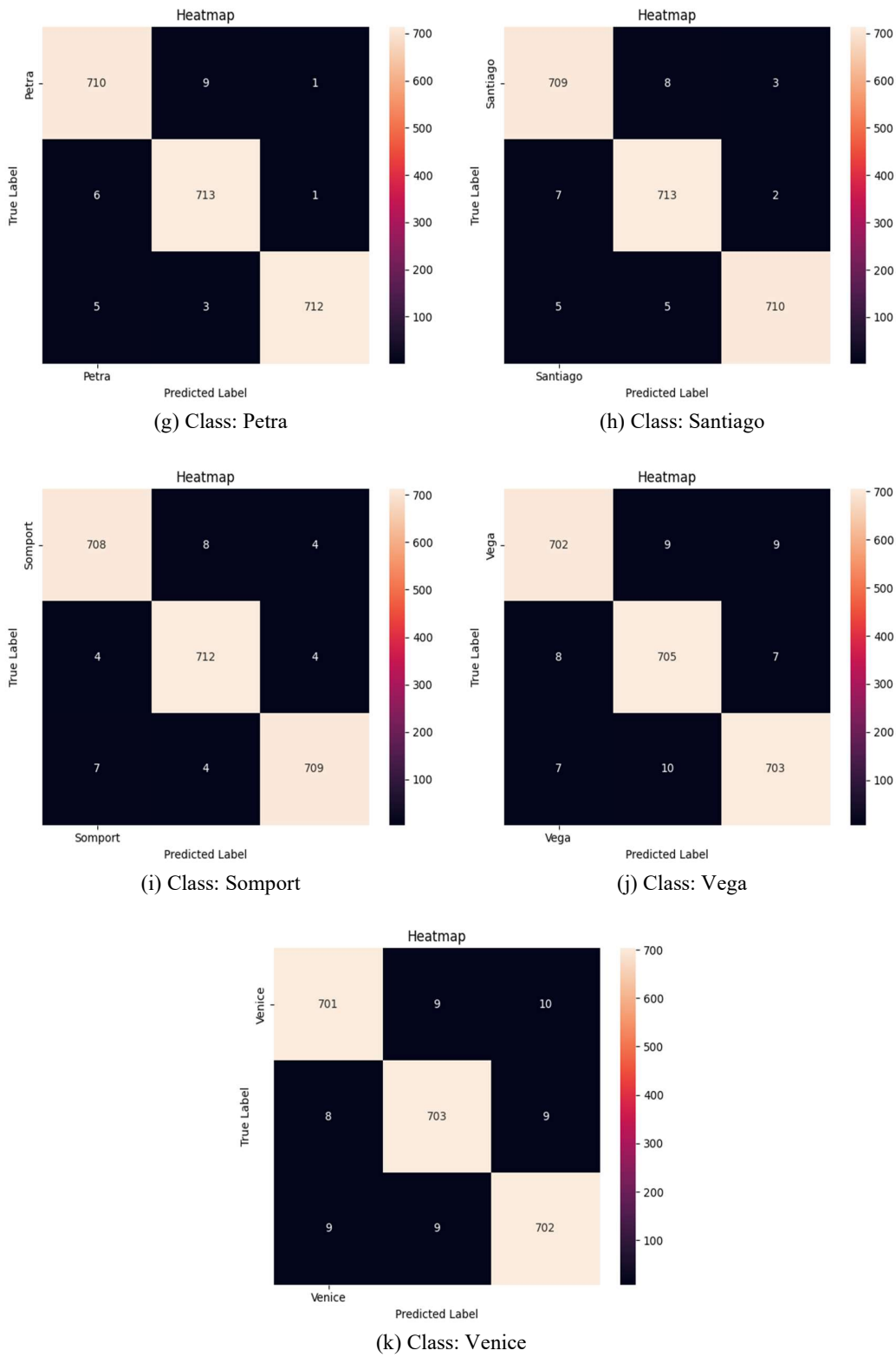


FIGURE 4 (a)–(k). Heatmaps for ceramic tile surface classification

Accuracy measures classification accuracy by including properly predicted positive and negative examples. Precision indicates prediction reliability by measuring the percentage of accurately predicted classes among all predicted positives. Recall measures the model's sensitivity to each class by identifying all relevant occurrences. The F1-score balances accuracy and recall assessing classification ability, particularly in multi-class situations. Specificity indicates the capacity to accurately reject non-relevant classes, whereas FPR and TPR are critical for ROC analysis, showing the trade-off between false alarms and correct detections. The ROC curve shows this link, while the AUC measures the model's discrimination. This means that the texture patterns are being correctly identified. Figure 4(d) Marfil and Figure 4(e) Mediterranea likewise has great diagonal dominance, with Mediterranea having extremely high accuracy and very low confusion, which shows that the features are easier to separate. Figure 4(f) Oslo and Figure 4(g) Petra's performance stays constant, with just a few off-diagonal numbers. This suggests that there isn't much overlap across classes. Figure 4(h) Santiago and Figure 4(i) Santiago Somport has consistent classification performance with low misclassification counts, which shows that it can generalize well.

Figures 4(j) Vega and 4(k) Venice show that the off-diagonal values are a little higher than those of other classes, which means that they are more comparable to classes next to them. The diagonal values are still much higher, which confirms that the classification is reliable. Overall, the heatmaps show that the hybrid Swin Transformer with Capsule Attention structure works well since it has excellent predictive power and little confusion. Figure 5 shows the Receiver Operating Characteristic (ROC) curves for various types of ceramic tiles. This shows how well the suggested model works for different types of tiles. Most classes have a high true positive rate and a low false positive rate, which is seen by the curves being near to the top-left corner. The Area Under the Curve (AUC) values for most classes are between 0.98 and 0.99, which shows that they are quite good at telling the difference between things. Petra (0.9913) and Mediterranea (0.9927) are two classes that perform the best, which means they are the most distinct. Venice, Vega, Marfil, Agata, Berlin, Lima, Oslo, and Somport are some of the other classes that also have consistently high AUC values, which shows that they are stable and dependable for categorization. Santiago has a lower AUC (0.8926), which means that it is more like other classes. In a multi-class context, it still works well enough. The dashed diagonal line shows random classification, and all the curves are much above this baseline, which shows that the suggested hybrid Swin Transformer with Capsule Attention model works well. In general, the image shows that the model can generalize well and classify well over a wide range of ceramic tile surface classes. Figure 6 compares accuracy, precision, recall, and F1-score of different methods.

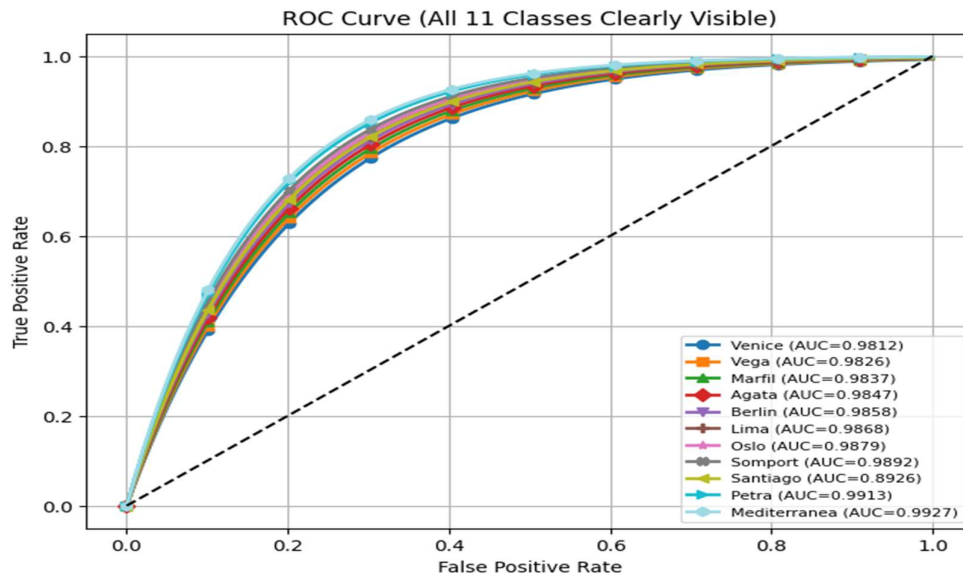


FIGURE 5. ROC curve for multi-class ceramic tile surface classification

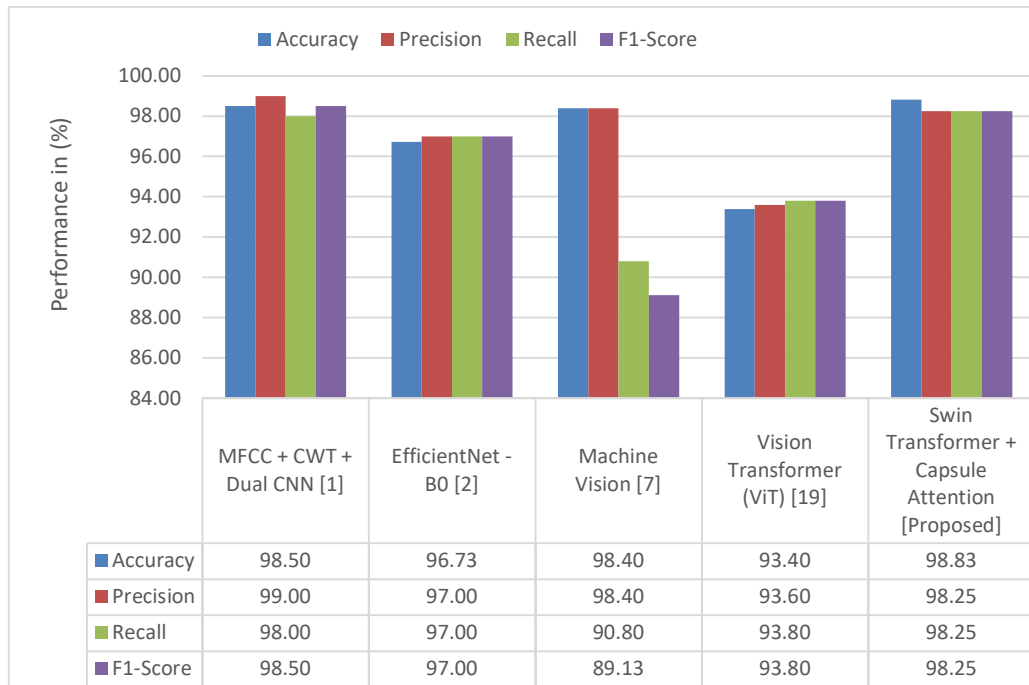


FIGURE 6. Performance comparison of different techniques for ceramic tile surface classification

At 98.83% accuracy, the Swin Transformer with Capsule Attention performs best in classification. Its balanced precision, recall, and F1-score of 98.25% indicate accurate predictions across all classes. MFCC + CWT + Dual CNN has competitive accuracy of 98.50% and the greatest precision of 99.00%, but worse recall. All measures show consistent but reduced performance for EfficientNet-B0. Machine Vision has great accuracy but low recall and F1-score, showing it cannot capture all important patterns. All assessment indicators show inferior performance for the Vision Transformer (ViT). The hybrid model balances all performance criteria, surpassing previous methods and proving its resilience for ceramic tile surface classification. The proposed VGG16 model does better than all the other models on all the measures, getting scores close to perfect (above 99%), which shows that it can classify things very well and consistently. Rep-MobileViT comes in second place with balanced and high results on all parameters, which shows that it can generalize well. CDW also performs well, with steady performance over 90% on all criteria. On the other hand, EfficientNetV2B1 and SVC shows comparatively lower performance, particularly in recall, indicating a reduced ability to identify all defect cases. GhostNet has a reasonable level of accuracy, but its precision and F1-score are lower, which means that its predictions aren't always reliable. In general, the findings show that the proposed VGG16 model is better than other methods at getting highly accurate and balanced classification.

CONCLUSIONS

The research offers an efficient framework for the classification of ceramic tile surfaces by the integration of a hybrid Swin Transformer with Capsule Attention. The system can get both contextual and structural information from tile surfaces because it uses hierarchical global feature extraction and capsule-based spatial relationship modeling. The dual-branch design, together with feature fusion, makes the model better at telling apart visually comparable categories while keeping things the same across diverse surface conditions. The suggested method works well and may be used in many other situations, as shown by testing it on the VxC TSG database, which has three classes of each tile type. But the model is more complicated to compute since it combines transformer and capsule parts, which might make it harder to deploy in real time and in contexts with limited resources. Also, performance may change when there are very uneven or hidden patterns. Future work might concentrate on model optimization, streamlined architectural design, and implementation in edge-based systems. Future enhancements might include integrating bigger and more varied datasets and investigating sophisticated attention mechanisms to increase scalability and generalization.

REFERENCES

- [1]. D. Wang, and B. Kang, 2026, "Tile debonding detection based on acoustic signal features and a dual-branch convolutional neural network," *Buildings*, 16(4), Article. 870.
- [2]. A. Demir, and H. K. Örnek, 2025, "Performance evaluation of deep learning architectures for tile defect detection," *Electrical Engineering and Energy*, 4(3), pp. 118-127.
- [3]. Y. Jiang, M. Tang, W. Dong, G. Tang, W. Hu, and J. Li, 2025, "Microcrack feature extraction method for chaotic optical surface of silicon nitride ceramic bearing roller based on multi-scale wavelet transform enhancement and optimized PSO-FCM coupling," *AIP Advances*, 15(1), Article. 015214.
- [4]. Q. Wang, X. Xiao, and Z. Liu, 2024, "Using microscopic imaging and ensemble deep learning to classify the provenance of archaeological ceramics," *Scientific Reports*, 14(1), Article. 32024.
- [5]. S. Abbas, H. Moumen, and F. Abbas, 2023, "Efficient method using attention based convolutional neural networks for ceramic tiles defect classification," *Revue d'Intelligence Artificielle*, 37(1), pp. 53-62.
- [6]. Zhang, X. Wang, Z. Sun, and Q. Zhang, 2025, "CIFormer: Cross interaction transformer for surface defect detection in ceramic tile," *Array*, 28, Article. 100553.
- [7]. H. Zhang, L. Peng, S. Yu and W. Qu, 2021, "Detection of surface defects in ceramic tiles with complex texture," *IEEE Access*, 9, pp. 92788-92797.
- [8]. H.A. Akyürek, H.İ. Kozan, Ş. Taşdemir, 2024, "Surface crack detection in historical buildings with deep learning based YOLO algorithms: A comparative study," *Computational Research Progress in Applied Science & Engineering, CRPASE: Transactions of Electrical, Electronic and Computer Engineering*, vol. 10, Article. 2904.
- [9]. Y. He, S. Li, X. Wen, and J. Xu, 2024, "A survey on surface defect inspection based on generative models in manufacturing," *Applied Sciences*, 14(15), Article. 6774.
- [10]. P. Kot, M. Muradov, A. Shaw, G. Mohi-Ud-Din, V. Karunamoothei, R. Deb, A. Ali, S. Sharma, P. K. Mishra, B. Sinhad, and R. P. Singh, 2025, "Novel microwave sensor for quality assurance of indoor residual spraying," *IEEE Sensors Letters*, 9(9), Article. 3504504.
- [11]. I. Garrido, E. Barreira, R. M. Almeida, and S. Lagüela, 2022, "Introduction of active thermography and automatic defect segmentation in the thermographic inspection of specimens of ceramic tiling for building façades," *Infrared Physics & Technology*, 120, pp. 1-20.
- [12]. Z. Çetinkaya, B. Yıldız, Y. Köse, E. Suvacı, and R. Raj, 2026, "Comparison of flash sintering and industrial firing of ceramic wall tiles: Microstructure and properties," *International Journal of Applied Ceramic Technology*, 23(1), Article. 104012.
- [13]. V. Sevinç, and M. M. Kirca, 2024, "A comparison between the Bayesian network model and the logistic regression model in prevention of the defects on ceramic tiles," *Journal of Experimental & Theoretical Artificial Intelligence*, 36(7), pp. 1385-1401.
- [14]. R. Fabris, G. Masi, D. Mazzini, L. Sanseverino, and M. C. Bignozzi, 2024, "Tribomechanical properties of glazes for ceramic tiles: A novel protocol for their characterization," *Materials*, 18(1), Article. 60.
- [15]. D. Wang, J. Peng, S. Lan, W. Fan, 2024, "CTDD-YOLO: A lightweight detection algorithm for tiny defects on tile surfaces," *Electronics*, 13(19), Article. 3931.
- [16]. R. Fabris, G. Masia, D. Mazzinib, L. Sanseverinoc, and M. C. Bignozzia, 2024, "Application of different test methods to assess the wear resistance of ceramic tiles," *Ceramic tile manufacture*, pp. 1-19.
- [17]. L. Varghese, E. I. Cedillo-Gonzalez, A. Cattini, M. Vacchi, and C. Siligardi, 2024, "Frit-free solar reflective porcelain stoneware ceramic tiles using recycled granite waste: an investigation on its engobe and glaze formulations," *Energy and Buildings*, 311, Article. 114129.
- [18]. A. P. Margarido, J. C. Carvalho, M. M. Resende, R. P. de Andrade, F. L. Maranhão, and A. V. Chaves, 2024, "Evaluation of the influence of the substrate layer/adhesive mortar/ceramic tile product on the performance of the ceramic covering," *Qualicer*, pp. 1-19.
- [19]. B. Song, and W. Yang, 2025, "Deep learning for automated recognition and digital reconstruction of ceramic motifs," *International Conference on Generative AI and Digital Media Arts*, pp. 122-127.
- [20]. D. Shen, Y. Lu, and L. Hua, 2025, "Automated debonding assessment of tile panels using convolutional neural network," *Structural Health Monitoring*, pp. 1-13.