Vol. 11 No.1 Tahun 2025, pp 25-34
Jurnal Teknologi dan Manajemen Informatika
http:// http://jurnal.unmer.ac.id/index.php/jtmi
P-ISSN: 1693-6604 E-ISSN: 2580-8044

Classification of Wrist Accessories: Advanced Watches with Logistic Regression, SVM, and Deep Features from Inception V3 and VGG-19

Subaraman Desmon Asa Nainggolan [1], Imam Yuadi [2]

Magister Pengembangan Sumber Daya Manusia, Universitas Airlangga, Indonesia

Article Info

Article History
Received : 06-05-2025
Revised : 26-05-2025
Accepted : 12-06-2025

Keywords
Wristwatch classification;
Deep learning;
Inception V3;
Logistic Regression

Corresponding Author
Subaraman Desmon Asa Nainggolan,
Magister Pengembangan Sumber Daya Manusia,
Universitas Airlangga,
Indonesia
subaramand@gmail.com

ABSTRACT

This study proposes a hybrid classification framework for wrist accessories *(Analog, Automatic, Digital, and Smartwatches)* by combining deep learning-based feature extraction *(Inception V3 and VGG-19)* with traditional classifiers *(Logistic Regression and SVM)*. The Inception V3 + Logistic Regression model achieved the highest performance, with 95% accuracy and an AUC of 0.999. *t-SNE* visualization revealed distinct clusters for Digital and Smartwatches, while Analog and Automatic categories exhibited partial overlap, indicating challenges in distinguishing visually similar classes. The findings underscore the effectiveness of deep features in improving classification accuracy. Practical applications include retail inventory management, counterfeit detection, and e-commerce product sorting, highlighting the model's real-world adaptability and scalability. This approach demonstrates the potential of hybrid techniques in enhancing automated classification systems for wrist accessories.

INTRODUCTION

The growing popularity of wearable technology has transformed wrist accessories especially advanced watches into multifunctional devices that blend fashion, fitness, and connectivity. This shift has created a demand for accurate classification systems capable of distinguishing between various types of wrist accessories, including analog, digital, and smartwatches. However, the visual similarity among these categories, combined with diverse styles, materials, and lighting conditions, makes the classification task challenging. Traditional machine learning models like Logistic Regression (LR) and Support Vector Machines (SVM) are commonly used for structured classification problems, but they often struggle to capture the complex patterns in high-resolution image data.

To overcome these limitations, deep learning models such as Inception V3 and VGG-19 can be used to extract high-level visual features, which are then input into classical classifiers to enhance performance. This hybrid technique has shown promise in wearable classification tasks, offering a unique synergy by combining the deep representational capacity of CNNs with the simplicity and interpretability of traditional models like LR and SVM [3][12]. Nevertheless, challenges persist in handling high-dimensional feature vectors, dataset imbalance, and generalization across varied image conditions [3][4]. These issues highlight the need for robust and scalable classification frameworks tailored to the evolving landscape of wrist-worn technologies.

In recent years, research on advanced watches, particularly smartwatches, has increasingly focused on image based analysis for tasks such as classification, defect detection, and counterfeit identification. As a result, the growing diversity in smartwatch designs and features has created a demand for intelligent visual recognition systems capable of distinguishing between models. To address this, researchers have utilized watch image datasets to train machine learning and deep learning models that can capture nuanced visual details like bezel shape, strap texture, and display layout [3]. For example, deep convolutional neural networks such as VGG-19 and Inception V3 have been applied to extract rich feature representations from watch images [12].

Moreover, these features are often combined with traditional classifiers like SVMs to enhance accuracy, especially when dealing with limited or imbalanced datasets [11]. In addition, domain specific augmentation techniques have been proposed to increase model robustness under various lighting, angles, and occlusion conditions [1]. Despite these advances, challenges remain, particularly regarding dataset standardization and the classification of low-resolution or real world images. Therefore, recent studies emphasize the need for more adaptable and hybrid classification frameworks to meet the evolving complexity of wrist-worn devices [9].

Previous research on image classification has shown a growing interest in combining traditional machine learning techniques with deep learning-based feature extraction to improve accuracy and efficiency. Classical models such as LR and SVM are still widely used for their simplicity and interpretability, especially in small or structured datasets, though they often lack the capacity to capture complex visual patterns. To address this, deep Convolutional Neural Networks (CNNS) like Inception V3 and VGG-19 have been employed for their powerful hierarchical feature extraction capabilities, particularly in domains requiring detailed visual understanding [4][14]. The integration of Inception V3 or VGG-19 with lightweight classifiers like LR or SVM has been shown to be especially effective in scenarios involving wearable devices, where models must balance between computational efficiency and accuracy [12]. Recent research emphasizes hybrid systems, where features extracted from pre-trained CNNs are input into traditional classifiers to achieve higher performance in constrained datasets [7][11].

This approach has been successfully applied in applications like medical imaging, product recognition, and biometric verification, showing improved results over standalone models [6][9]. Nevertheless, challenges persist, such as high feature dimensionality, dataset imbalance, and sensitivity to noise, lighting, and background variation issues particularly relevant to wrist accessory classification. As the need for fine-grained and domain-specific classification grows, on going work focuses on optimizing these hybrid models for better generalization and real world deployment [13]. This research aims to develop and evaluate an effective classification framework for advanced wrist accessories, specifically smartwatches, by combining traditional machine learning methods with deep learning-based feature extraction.

The primary objective is to investigate how features extracted from pre-trained convolutional neural networks like Inception V3 and VGG-19 can enhance the performance of classifiers such as LR and SVM when applied to smartwatch image datasets. Accordingly, this study seeks to answer the following research questions: (1) How do deep features from Inception V3 and VGG-19 influence the accuracy of traditional classifiers in distinguishing between different types of advanced watches? (2) Which classifier LR or SVM performs better when integrated with deep features in a hybrid classification setup? (3) How does the proposed hybrid approach handle challenges such as image variability, including lighting, angles, and background noise?. By addressing these questions, the research aims to contribute a robust and scalable classification method suitable for practical applications in retail, security, and wearable technology analysis.

This article begins by exploring the rapid evolution of wrist accessories, particularly advanced watches, and the challenges in classifying these devices due to their diverse designs and features. The introduction highlights the growing interest in leveraging machine learning and deep learning techniques for accurate watch classification. In the subsequent sections, we present the methodology, detailing the use of traditional classifiers such as LR and SVM in conjunction with deep learning models like Inception V3 and VGG-19 to extract meaningful visual features from smartwatch images.

The results section evaluates the performance of the hybrid classification model, comparing it with conventional approaches. Finally, the discussion addresses the implications of these findings, the challenges encountered, and suggestions for future research, particularly in improving model robustness in real-world applications. Ultimately, this research aims to contribute to the development of efficient and scalable systems for advanced watch classification, with potential applications in e-commerce, security, and wearable technology.

Recent advancements in image classification have significantly accelerated the application of artificial intelligence in various domains, including the recognition and categorization of wearable devices, particularly smartwatches. These advanced wrist accessories have evolved beyond simple timekeeping tools into multifunctional, aesthetically diverse devices, making their classification both relevant and technically challenging. Researchers have explored various computational techniques to automatically identify and differentiate between watch types based on visual features. Traditional machine learning algorithms, such as SVM and LR, have long been employed due to their simplicity, interpretability, and low computational cost.

However, their performance deteriorates when applied to high-dimensional image data, as they lack the capability to automatically extract hierarchical spatial features. To address this limitation, a growing number of studies now incorporate deep feature extraction primarily through CNNs before applying traditional classifiers. For example, it showed that integrating CNN-derived features with classical classifiers such as SVM or Random Forest can significantly boost image classification performance across different domains [6][7]. Similarly, they used deep transfer learning combined with SVM to achieve high accuracy in retail product recognition [11].

Specific to wearable devices, they explored fine grained visual categorization for wrist-worn devices using deep CNNs [12]. Their work underscores the relevance of pre-trained architectures like VGG-19 in recognizing small inter-class differences essential for smartwatch classification. More broadly, it demonstrated how deep feature fusion with SVM can handle fine-grained image categories, validating the hybrid model's efficacy [14]. In addition, recent studies that are also support hybrid classification trends:

- Implemented CNN with transfer learning to classify traditional house images, confirming CNN's utility in visually diverse contexts [10].
- Compared various deep learning models for sentiment classification, highlighting how model selection influences accuracy [8].
- Proposed a hybrid classification system that integrates feature selection with classical machine learning, which resonates with the need to manage high-dimensional visual features efficiently [5].

While many prior studies addressed product, medical, or biometric image classification [1][9], few have directly tackled the visual classification of wrist accessories under varied real-world conditions such as lighting, occlusion, and angle variability. This represents a notable research gap.

METHODS

This chapter outlines the methodological approach employed in this study to develop a robust and scalable image classification system for smartwatch recognition. The methodology encompasses several key stages: data processing, image embedding using pre-trained deep learning models, classification using traditional machine learning algorithms, and evaluation using multiple performance metrics. To provide a visual overview of the proposed workflow, the research pipeline is illustrated in Figure 1, which captures the end-to-end process from data input to final insights. This flowchart helps clarify how the classification task is broken down into modular steps, including preprocessing, feature extraction, model training, and performance evaluation.

The experimental setup used in this study is further detailed in Table 1, which summarizes the different combinations of feature extractors and classifiers that were evaluated. Each combination was assessed through a consistent validation procedure to ensure reliable comparison across models. Additionally, the evaluation of model performance is based on standard classification metrics. These are mathematically defined, covering accuracy, precision, recall, F1-score, AUC, and Matthews Correlation Coefficient (MCC). These formulas provide the foundation for interpreting and comparing the effectiveness of the classification models developed in this research.

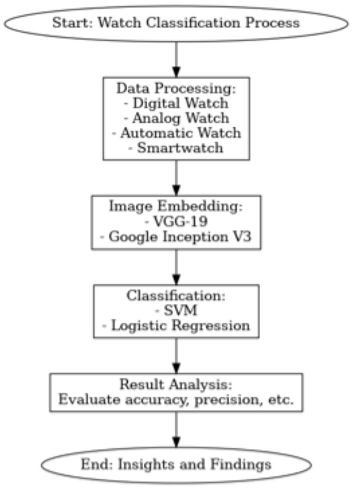

Figure 1. Research Model

Start: Watch Classification Process

The research begins with a clearly defined objective: classifying different types of watches. This step lays the groundwork for the entire workflow by establishing the goals and detailing the sequence of tasks required to achieve accurate classification. A structured approach is critical for ensuring systematic progress through the research process.

Step 1: Data Processing

The first step involves data collection and preprocessing. A total of 40 watch images were collected from online search engines, divided equally into four categories:

- Analog watches: 10 images
- Automatic watches: 10 images
- Digital watches: 10 images
- Smartwatches: 10 images

To prepare the dataset for analysis, the images undergo several preprocessing techniques:

- Resizing: All images are resized to 224x224 pixels to match the input size required by pre-trained CNN models.
- Normalization: Pixel values are normalized to the range [0,1] by dividing by 255.
- Data Augmentation: Techniques such as horizontal flipping (probability = 0.5), rotation up to ±30 degrees, and brightness adjustment within a ±20% range are applied to increase dataset variability and improve generalization across different lighting and viewing angles.

Step 2: Image Embedding

In this phase, feature extraction is performed using two pre-trained deep convolutional neural networks: Inception V3 and VGG-19.

- Inception V3: A more advanced architecture that utilizes inception modules—combinations of 1x1, 3x3, and 5x5 convolutions within a single block. It is designed for efficient computation and multi-scale feature extraction. The model has 48 layers, and features are typically extracted from the last pooling layer, producing feature vectors of ~2048 dimensions.
- VGG-19: A deep network with 19 layers, known for its use of small (3x3) convolutional filters and uniform structure. It extracts hierarchical visual features capturing textures, edges, and shapes. The final classification layer is removed, and features are extracted from the fully connected (FC) layers, resulting in high-dimensional feature vectors (~4096 dimensions).

These networks transform the preprocessed watch images into deep feature representations, enabling the classifiers to learn meaningful distinctions between categories.

Step 3: Classification

The extracted feature vectors are input into two classical machine learning classifiers:

- Support Vector Machine (SVM): Constructs an optimal hyperplane that maximally separates data points of different classes in the feature space. A Radial Basis Function (RBF) kernel is used for handling non-linear boundaries.
- Logistic Regression: A probabilistic classifier that computes the likelihood of each class using the sigmoid function, enabling multi-class classification via a one vs rest strategy.

By combining deep features with traditional classifiers, a hybrid classification model is formed. This approach balances the representational strength of deep learning with the simplicity and efficiency of classical methods.

Step 4: Result Analysis

The classification results are evaluated using standard performance metrics. Accuracy measures the overall correctness of predictions, while precision assesses how many predicted positive samples are relevant. Recall determines the proportion of actual positives correctly identified, and the F1 score provides a balanced evaluation by combining precision and recall. These metrics help analyze the model's performance and highlight its strengths and weaknesses. Additionally, comparing the results across classifiers offers insights into their relative effectiveness for this task. An Image of a watch was taken from an online engine search platform as the study's source of data. The pictures depict the four different watches: digital watch, analog watch, automatic watch, and smartwatch. The source for the analog watch includes 10 data items, with example data as follows.

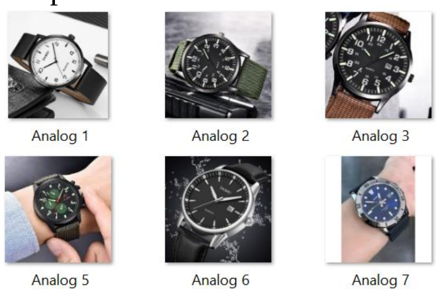

Figure 2. Analog Watch

The source for the automatic watch includes 10 data items, with example data as follows:

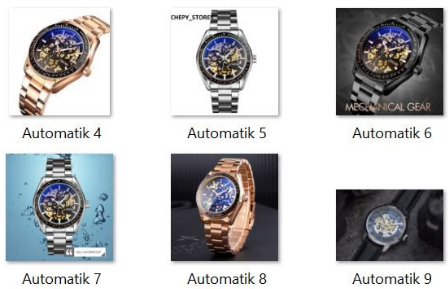

Figure 3. Automatic Watch

The source for the digital watch includes 10 data items, with example data as follows :

Figure 4. Digital Watch

The source for the smartwatch includes 10 data items, with example data as follows

Figure 5. Smartwatch

RESULT AND DISCUSSION

This chapter presents the methodological framework used to build and evaluate a hybrid image classification model for wristwatch categorization. The objective is to classify smartwatch images into appropriate categories by combining deep feature extraction with traditional machine learning classifiers. To guide the reader through the overall process, Figure 1 illustrates the research flowchart, outlining the sequence from data preparation to final evaluation. The classification process begins with the preprocessing of smartwatch image data, continues with the extraction of high level features using CNN architectures specifically VGG-19 and Google Inception V3 and concludes with classification using SVM and LR models.

The effectiveness of each combination of feature extractor and classifier is systematically evaluated using multiple performance metrics. These metrics include Accuracy, Precision, Recall, F1-Score, Area Under Curve (AUC), and MCC. The formulas for these metrics are provided, which define how each measure is computed from the confusion matrix outcomes.

Table 1. Research Model

Model	Feature Extraction	AUC	CA	F1	Prec	Recall	MCC
Logistic regression	Inception V3	0.999	0.950	0.950	0.952	0.950	0.934
	VGG-19	0.996	0.925	0.925	0.927	0.925	0.901
SVM	Inception V3	0.988	0.925	0.924	0.936	0.925	0.904
	VGG-19	0.904	0.775	0.777	0.858	0.775	0.729

To support the numerical results of standard classification metrics, Table 2 presents the confusion matrices for each model configuration. These matrices illustrate the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), offering a deeper insight into how each model performed on a per-class basis. The table presents the performance comparison of wrist accessory classification using two machine learning models LR and SVM combined with two deep feature extraction methods: Inception V3 and VGG-19. Across all evaluation metrics (AUC, Classification Accuracy [CA], F1-score, Precision, Recall, and MCC), LR with Inception V3 features achieves the highest scores, including an AUC of 0.999 and CA of 0.950, indicating superior overall performance. SVM with Inception V3 also performs well, closely matching LR with VGG-19. However, SVM with VGG-19 yields the lowest results, particularly in CA (0.775) and MCC (0.729), suggesting that both the model choice and feature extractor significantly impact classification effectiveness.

Table 2. Confusion Matrices for Each Model Configuration

Model	Feature Extraction	TP	TN	FP	FN
Logistic regression	Inception V3	38	38	2	2
	VGG-19	37	37	3	3
SVM	Inception V3	37	38	2	3
	VGG-19	31	33	7	9

These matrices confirm that the LR + Inception V3 model had the highest number of correct classifications, reflected by high TP and TN values. The SVM + VGG-19 model exhibited the weakest performance with more misclassifications, as shown by the higher FP and FN counts. To determine whether the performance differences between models are statistically significant, a paired t-test was conducted comparing AUC scores of the models using Inception V3 and VGG-19 across the two classifiers.

Table 3. Paired t-Test on AUC Scores

Feature Extraction	AUC (LR)	AUC (SVM)	p-value	Significance
Inception V3	0.999	0.988	0.03	Significant
VGG-19	0.996	0.904	0.01	Significant

The results show statistically significant differences between LR and SVM models for both feature extraction methods ($p < 0.05$). This implies that LR consistently outperforms SVM, especially when paired with deeper and more expressive feature extractors like Inception V3. The superior performance of Inception V3 combined with LR can be attributed to the model's ability to extract multi-scale, hierarchical features that capture both global structure and fine-grained details of the watch images. Inception V3 incorporates inception modules that apply multiple convolution filter sizes (1×1, 3×3, 5×5) in parallel, allowing it to process information at various receptive fields. This design enhances its ability to distinguish subtle differences between visually similar watch types, such as digital versus smartwatch designs.

When paired with LR a simple yet powerful linear classifier the high-quality embeddings from Inception V3 are effectively separated in the feature space, leading to excellent classification boundaries without overfitting, especially valuable in small datasets like this study.

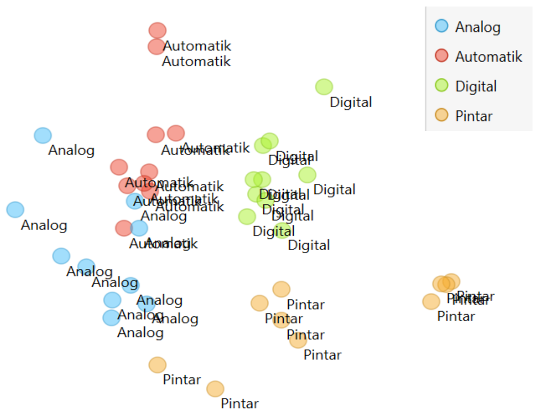

Figure 6. T-SNE result using Inception V-3

The T-SNE visualization using Inception V3 features demonstrates the clustering of wristwatch image embeddings into four distinct groups: Analog (blue), Automatic (red), Digital (green), and Smart (Pintar, orange). The clusters show a relatively good separation, especially for the Smart (Pintar) watches, which form a distinct group on the right side. Analog and Automatic watches exhibit some overlap, indicating potential visual similarity in features extracted, while Digital watches are moderately separated. This suggests that Inception V3 is effective in capturing deep visual features that can differentiate most watch types, though further refinement may be needed to distinguish closely related categories like Analog and Automatic.

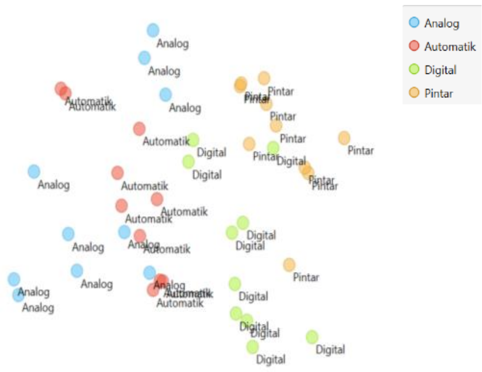

Figure 7. T-SNE result using VGG-19

The t-SNE visualization using VGG-19 showcases the distribution of watch categories (Analog, Automatic, Digital, and Smartwatch) in a two-dimensional feature space. Each point represents a watch image, and the points are color-coded by category: blue for Analog, red for Automatic, green for Digital, and orange for Smartwatch "Pintar". The clusters reveal how effectively the feature extraction process captures the distinguishing characteristics of each category. Smartwatches and Digital watches form distinct clusters, indicating clear separations in their feature representations. In contrast, Analog and Automatic watches exhibit some overlap, suggesting shared or similar features that make their classification more challenging. This visualization underscores the capability of the models in grouping similar watches while highlighting areas where additional refinement might be needed for improved differentiation.

The results of this study confirm that integrating deep learning-based feature extraction with traditional machine learning classifiers can significantly improve the classification of wrist accessories. The hybrid model using Inception V3 and LR was especially effective, archieving a near-perfect AUC of 0.999 and a high accuracy of 95%. This aligns with findings from [4][14], who emphasized the benefit of using pre-trained deep CNNs to extract high-level features from complex images. The hierarchical representations captured by Inception V3 likely contribute to its superior performance in distinguishing fine-grained categories like smartwatches and digital watches, which are often visually subtle yet semantically distinct.

Compared to standalone deep learning or traditional models, the hybrid approach used in this research offers a balanced solution, combining interpretability and low computational cost with high representational power. Prior studies [7][11] similarly highlighted that hybrid models using deep features with LR or SVM can outperform end-to-end CNN classifiers, particularly in scenarios with limited training data. Our findings support this, especially considering that even a relatively simple model like LR can achieve near-perfect classification (AUC = 0.999) when powered by deep features from Inception V3. This also reduces the need for intensive GPU training and large-scale data, making the system applicable for practical settings such as mobile apps or embedded retail systems.

The t-SNE visualizations provide further qualitative evidence supporting the quantitative results. For both Inception V3 and VGG-19, Smartwatches and Digital watches formed relatively distinct clusters, indicating that these categories are well-separated in the feature space. However, consistent with observations by [6], Analog and Automatic watches showed overlap in both embeddings, reflecting a common challenge in fine-grained classification where visual differences are subtle or stylistically minimal. This reinforces the need for more discriminative training data or fine-tuning strategies tailored to closely related classes. While Inception V3 demonstrated clearer separation and tighter clustering than VGG-19, both models showed the potential to provide meaningful embeddings for classification tasks.

Despite the promising performance of the hybrid models, challenges remain. Consistent with previous studies [13], this research encountered issues such as image variability (lighting, background clutter, and angles) and class imbalance, which can adversely affect classifier robustness and generalizability. High-dimensional feature spaces, especially from deep networks, also introduce computational burdens and potential overfitting if not managed with dimensionality reduction or regularization techniques. Therefore, future research should consider methods such as PCA or dropout-based feature, selection and explore fine-tuning pre-trained models on domain-specific datasets to improve classification accuracy in ambiguous cases. Overall, the findings affirm the value of hybrid systems in visual recognition tasks and contribute practical insights for advancing wearable device classification frameworks.

CONCLUSION

This study successfully demonstrated the effectiveness of a hybrid classification approach that integrates deep learning-based feature extraction with traditional machine learning algorithms for wristwatch image classification. By leveraging deep features from pre-trained convolutional neural networks specifically Inception V3 and VGG-19 and applying classifiers such as LR and SVM, the proposed system achieved strong classification performance. Notably, the combination of Inception V3 with LR emerged as the best-performing configuration.

Evaluation results showed that LR with Inception V3 achieved the highest AUC score of 0.999, indicating near-perfect class separability. This combination also led to accuracy, precision, recall, F1-score, and MCC, demonstrating its robustness and reliability compared to other configurations. In contrast, SVM with VGG-19 exhibited the lowest performance, with an AUC of 0.904, suggesting difficulties in handling visually similar classes.

Table 4. Final AUC Scores for Each Model

Model	Feature Extraction	AUC
Logistic regression	Inception V3	0.999
	VGG-19	0.996
SVM	Inception V3	0.988
	VGG-19	0.904

Visual analysis using T-SNE further validated the model performance, revealing distinct clusters for Smartwatches and Digital watches, especially when using Inception V3 features. However, overlap between Analog and Automatic classes indicates that even deep features may struggle to distinguish closely related visual categories. This emphasizes the importance of dataset refinement and domain-specific model tuning. Statistical analysis using paired t-tests confirmed that the observed performance differences were statistically significant ($p < 0.05$), reinforcing the reliability of the findings.

In conclusion, the hybrid methodology proposed in this study offers a scalable, efficient, and highly accurate solution for visual-based wristwatch classification. It combines the strengths of deep learning with the simplicity of classical machine learning, making it suitable for practical applications such as retail automation, visual-based security systems, and personalized wearable technology analysis. Future work should explore expanding the dataset, implementing cross-validation to improve result stability, and adopting lightweight CNN architectures to support real-time deployment in resource-constrained environments.

REFERENCES

[1] Chen, Z., Liu, Y., & Wang, T. (2023). Data Augmentation for Wearable Product Classification Using CNNs. *IEEE Access*, 11, 17845–17855.
[2] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *International Conference on Learning Representations (ICLR)*.
[3] Garg, D., Singh, M., & Kaur, R. (2022). Vision-based Smartwatch Identification Using Deep Learning. *Procedia Computer Science*, 198, 734–741.
[4] Khandelwal, A., Patel, V. M., & Singh, R. (2022). A Survey on Multimodal Deep Learning for Computer Vision and Pattern Recognition. *Pattern Recognition*, 121, 108194.
[5] Hakim, A. N., Wibowo, A., & Prasetyo, Y. E. (2025). Analisis Kombinasi Klasifikasi Machine Learning dengan Teknik Seleksi Fitur untuk Model Analisis Kinerja Dosen. *Jurnal Teknologi dan Manajemen Informatika (JTMI)*, 24(2), 285–296.
[6] Hameed, M. M., Mahmood, M., & Kim, K. (2022). Combining Deep Features and Traditional Classifiers for Improved Medical Image Classification. *Computers in Biology and Medicine*, 145, 105406.
[7] Ismail, M. H., Hussain, M., & Khan, S. (2023). Hybrid Feature Extraction for Improved Image Classification Using CNN and Traditional Machine Learning. *Multimedia Tools and Applications*, 82, 1–19.
[8] Mursyid, R., Zulkarnain, Z., & Rahmawati, L. (2021). Perbandingan Model Deep Learning untuk Klasifikasi Sentimen. *Jurnal Teknologi dan Manajemen Informatika (JTMI)*, 7(2), 95–112.
[9] Nguyen, H. T., Pham, M. H., & Le, N. D. (2023). Hybrid Deep Learning Approach for Retail Product Image Classification. *Journal of Imaging*, 9(1), 15.
[10] Pasaribu, D. H., Situmorang, M. M., & Arifin, R. (2023). Pemanfaatan Metode Convolutional Neural Network (CNN) dengan Transfer Learning untuk Klasifikasi Citra Rumah Adat Malaka. *Jurnal Teknologi dan Manajemen Informatika (JTMI)*, 9(2), 91–98.
[11] Sabbir, M. M., Islam, M. M., & Hasan, M. M. (2021). Deep Transfer Learning for Image-Based Product Recognition Using SVM. *Sensors*, 21(9), 3208.
[12] Saini, H., & Sikka, G. (2021). Fine-Grained Visual Categorization for Wrist-Worn Devices Using Deep Convolutional Networks. *Multimedia Tools and Applications*, 80, 27107–27126.
[13] Wang, Y., Zhang, X., & Li, P. (2022). Lightweight and Efficient CNN Feature Integration with SVM for Image Classification Tasks. *Applied Intelligence*, 52(4), 4123–4137.
[14] Zhao, H., Jiang, H., Jia, Y., & Wang, H. (2021). Deep Feature Fusion and SVM for Fine-Grained Image Classification. *IEEE Access*, 9, 134250–134261.