JOURNAL LA MULTIAPP

VOL. 07, ISSUE 03 (608-619), 2026
DOI: 10.37899/journallamultiapp.v7i3.3256

Comparative Analysis of CNN Architectures for Clean and Non-Clean Outfit Classification in Fashion Images

Noviana Wahyu Basuki[1], Imam Yuadi[2]

[1]Master's Program in Human Resource Development, Postgraduate School, Universitas Airlangga, Indonesia
[2]Department of Information and Library Science, Faculty of Social and Political Sciences, Universitas Airlangga, Indonesia

**Corresponding Author: Noviana Wahyu Basuki*
Email: noviana.wahyu.basuki2025@pasca.unair.ac.id

Article Info

Article history:
Received 16 April 2026
Received in revised form 2 May 2026
Accepted 27 May 2026

Keywords:
Clean Outfit
Non-Clean Outfit
Image classification
Convolutional Neural Network
Machine Learning

Abstract

The purpose of this study is to create and contrast image classifiers which are able to identify clean clothing from dirty using machine learning and deep learning techniques. Our utilized dataset contains 200 images of an outfit which are obtained from different official fashion brand sites and well-known e-commerce platforms in Indonesia. The information is used from secondary data which are digital images of the two style categories. Image preprocessing (resizes, normalizations and data augmentations), feature extraction from VGG-16, VGG-19 and inception v3 is done. The extracted features are then fed into the classifiers namely Logistic Regression, Neural Network and Support Vector Machine (SVM). The evaluation of the model is performed by different metrics (e.g., AUC, accuracy, F1-score, precision, recall and MCC) and visual examination using MDS plot and Silhouette Plot. The results demonstrate that the integrated model involving VGG-16 and Logistic Regression performs best obtaining highest AUC when compared with other model combinations. The MDS and Silhouette Plot visualizations also supported that VGG-16 has the most superior feature separation between clean outfits and non-clean outfits. In a word, our study unveils that fashion style recognition accuracy can be improved significantly through CNN-based feature extraction and traditional classification model. We hope that our work will encourage the comparison of CNN feature extraction and classification algorithms, and also can lay the foundation for further research in image-based outfit guidance systems serving a range of fashion industry and service sectors where professional appearance is a criterion.

Introduction

shorthand for credibility and professionalism especially when you're in an industry where customers decide whether or not they should trust you. Previous research has found that the attire of employees affect customers and impacts their expectations of service quality and intention to purchase, indicating the significance of appearance in customer experience (Kulkarni & Harnoorkar, 2020; Lee et al., 2022; Saut & Bie, 2024; Sudirjo et al., 2023; Akram et al., 2022).

Along with the developments on image processing and deep learning, fashion image classification receives an increasing research interest in computer vision community. Convolutional Neural Networks (CNNs) have proved highly effective in extracting discriminative visual features for clothing attribute recognition and dressing style analysis 9

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Kaur & Pandey, 2023; Liu, 2025; Zhang et al., 2023; Rabbi et al., 2023). Extensive datasets and benchmark evaluations have also facilitated CNN based solutions for fashion-related tasks (Khalid & Gong, 2026; Szegedy et al., 2016). Nevertheless, most current work mainly aims at achieving higher classification accuracy or the design of new architectures; comparative works that comprehensively assess the combinations between feature extraction based on CNN and classifiers from traditional field in applied context are still very scarce (Xie et al., 2022; Kasim et al., 2025 Hossain et al., 2025).

Clean / Non-clean outfit in this work has been chosen to satisfy the need of a practical problem such as customer-service, marketing/ sales or food industry in which image representation having neat and professional attire is necessary. A clean outfit is an outfit that appears neat, simple, professional and contextually appropriate; whereas a non-clean one would come across as less tidy-looking, too casual or not for professional purpose. Image-based classification is therefore a possible and objective approach since such differences are visual in nature (Singh et al., 2023; Choudhary & Sethi, 2023; Wang et al., 2023).

Methodologically, we make use of CNN-based feature extraction due to its efficient generation of high-level visual descriptions. VGG architectures (16-19 layers) are commonly employed in transfer learning due to their capacity to generate rich and generalizable features. Inception V3, by contrast, focuses on computational efficiency with multi-scale convolutional blocks and factorized convolutions that make it ideal for problems of large image classification (Xiao et al., 2017; Elfatimi et al., 2024; Hu et al., 2026; Wang et al., 2025).

Considering this, the objective of this work is establishing a classification model for clean and non-clean outfits, which also involves guidance in terms of analysis through comparative assessment between CNN-based methods of feature extraction (VGG-16, VGG-19 and InceptionV3) along their combination with several others classification algorithms including Logistic Regression, Support Vector Machine (SVM) and Neural Network.

Related Work

Fashion analysis has been widely adopted by computer vision as target tasks such as category/attribute/style classification of apparels. Figure 2(a) Figure 2(b)Benchmark datasets have already been proposed for the evaluation of machine learning models on sets of fashion images and as references standard. Deep learning in fashion reviews also affirm CNNs as the preferred method for classification, recommendation and aesthetic analysis tasks (Shao et al., 2004; Deldjoo et al., 2023; Buradagunta & Balakrishna, 2025; Shushi & Abdulazeez, 2024).

In addition to the technical aspects, consumer behavior literature suggest that customer perception is influenced by employees? es look. Research on dress in service-contact positions, has shown that appropriate attire affects customers perceptions of the quality of services and purchase intentions. Focus group research has further suggested that the style of one's clothing is associated with perceptions of trust, professionalism, and ethicality, again balancing the importance of outfit type in an organizational setting.

CNNs are widely employed as feature extractors (image embeddings) in a transfer learning paradigm because they abstract the high-level visual features that are hard to engineer by hand. The VGG architecture shows that increasing network depth does lead to better and transferable representation with small convolution filters (3×3). However, the Inception architecture has been developed for efficient computation and multi-scale feature extraction to learn more-representative features with minimal computational burden (Simonyan & Zisserman, 2014; Khan & Iqbal, 2025; Lv et al., 2024).

Some of the previous works have presented classification pipelines of CNN-based embeddings and conventional classifiers (e.g. Logistic Regression, SVM), particularly when

ISSN: 2716-3865 (Print), 2721-1290 (Online)

dataset is limited or when one goal of a work was to investigate feature representation quality instead of end-to-end learning. Visual techniques like Multidimensional Scaling (MDS) and Silhouette analysis are extremely popular for evaluating embedding quality and cluster separability (Chae, 2022; Mudgal, 2026; Moujahid & Dornaika, 2025; Gómez-Talal et al., 2026).

Even though there are significant studies on fashion image classification, most of them focus on product category or low-level attributes such as color and texture. Research that specifically poses a clean vs. non-clean clothing classification task and compares VGG versus Inception embeddings employed with various classifiers as extensive as ours is still limited. This gap is filled by means of the following analysis.

Methods

Research Design and Methodology Framework

This is quantitative experimental study in machine learning and computer vision. The methodology is framed for producing a effective image classification technique with comparative performance evaluation of CNN based feature extraction models namely VGG and Inception architectures.

The research process includes the steps of image preprocessing, character feature extraction, classification, and performance evaluation. Image embeddings are produced by pre-trained CNN models, and classifications are conducted with traditional machine learning classifiers. This hybrid model has the ability to learn both linear and nonlinear decision surfaces (Nassar et al., 2026).

To validate model performances, we employed different quantitative metrics including AUC, classification accuracy, F1-score (which is the Harmonic means between precision and recall), Precision (positive predictive value), Recall(sensitivity), MCC. Feature separability and clustering quality is also visualized with MDS and Silhouette Plot analysis (Panpaeng et al., 2023). The methodology is illustrated in the following diagram :

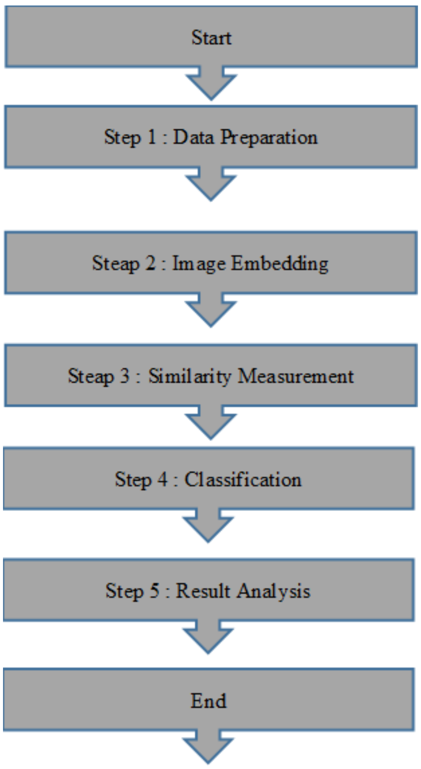

Figure 1. Research Process

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Dataset and Data Preprocessing

The dataset used in this study consists of 200 outfit images categorized into two classes: clean outfit and non-clean outfit. All images are secondary data collected from official fashion brand websites and leading e-commerce platforms in Indonesia, such as Zalora, Zara, and Shopee. These sources were selected because they provide high-quality and visually consistent product images representative of real-world usage.

Class labeling was performed manually based on visual criteria. Clean outfits are defined as neat, simple, and professional appearances, while non-clean outfits represent less tidy or overly casual styles. To ensure consistency and validity, labeling was conducted carefully according to these operational definitions.

Before model training, all images underwent preprocessing steps, including background removal**,** resizing to 224 × 224 pixel**s**, and format conversion to JPG, following common practices in CNN-based image classification (Bharadwaj et al., 2021; Dillon & Muhammad, 2024). To mitigate overfitting due to the limited dataset size, data augmentation techniques such as rotation, horizontal flipping, and random cropping were applied.

Results and Discussion

Supervised Classification Result

Table 1. Vgg 16, Vgg 19, & Inception V3 Test and Score Result

Model Image Embedding	Machine Learning Model	AUC	CA	F1	Prec	Rec	MCC
VGG 16	SVM	0.991	0.955	0.955	0.955	0.955	0.910
	Neural Network	0.988	0.965	0.965	0.967	0.965	0.932
	Logistic Regression	0.994	0.975	0.975	0.975	0.975	0.950
VGG 19	SVM	0.988	0.955	0.955	0.957	0.955	0.912
	Neural Network	0.989	0.955	0.955	0.955	0.955	0.910
	Logistic Regression	0.985	0.960	0.960	0.961	0.960	0.921
Inception V3	SVM	0.987	0.950	0.950	0.952	0.950	0.902
	Neural Network	0.987	0.940	0.940	0.940	0.940	0.880
	Logistic Regression	0.990	0.960	0.960	0.960	0.960	0.920

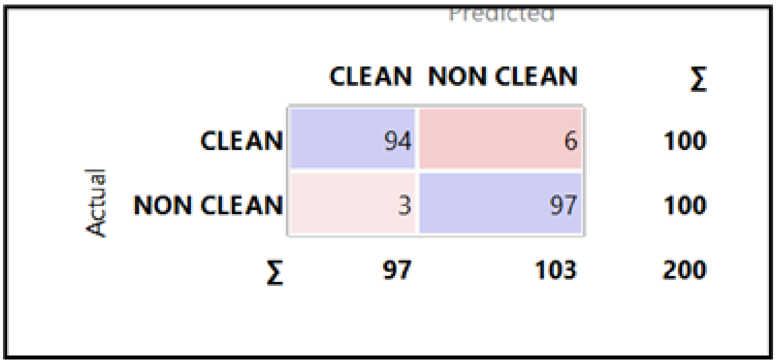

Figure 2. SVM – VGG 16

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Actual \ Predicted	CLEAN	NON CLEAN	Σ
CLEAN	100	0	100
NON CLEAN	7	93	100
Σ	107	93	200

Figure 3. Neural Network – VGG 16

Actual \ Predicted	CLEAN	NON CLEAN	Σ
CLEAN	97	3	100
NON CLEAN	6	94	100
Σ	103	97	200

Figure 4.. Neural Network – VGG 19

Actual \ Predicted	CLEAN	NON CLEAN	Σ
CLEAN	99	1	100
NON CLEAN	8	92	100
Σ	107	93	200

Figure 5. SVM – VGG 19

Actual \ Predicted	CLEAN	NON CLEAN	Σ
CLEAN	98	2	100
NON CLEAN	6	94	100
Σ	104	96	200

Figure 6. Logistic Regression – VGG 19

Actual \ Predicted	CLEAN	NON CLEAN	Σ
CLEAN	95	5	100
NON CLEAN	7	93	100
Σ	102	98	200

Figure 7. Neural Network – Inception V3

ISSN: 2716-3865 (Print), 2721-1290 (Online)

		Predicted		
		CLEAN	NON CLEAN	Σ
Actual	CLEAN	98	2	100
	NON CLEAN	8	92	100
	Σ	106	94	200

Figure 8. SVM – Inception V3

		Predicted		
		CLEAN	NON CLEAN	Σ
Actual	CLEAN	97	3	100
	NON CLEAN	5	95	100
	Σ	102	98	200

Figure 9. Logistic Regression – Inception V3

Under the confusion matrix section, we can see two classes are created: clean outfit and non-clean outfit, which helps further assess a model's power to classify images into these categories by looking at the overall distribution of true positive, false positive, true negative and false negative predictions. The true positives (TP), false negatives (FN) and false positives for that set are listed along with the number of true negatives (TN).

From the confusion matrix analysis we can see that there is no other setting in which VGG-16 behaves as well as with logistic regression, so that it is possible to confirm that VGG-16 + Logistic Regression indeed the best configuration for clean outfit vs. non-clean outfit classification. This model not only exhibits better quantitative behaviour but also ends up being more stable and balanced in terms of class-agnostic classification for applications like outfit recommendation systems at workplaces. VGG-16 results in the most balanced prediction distribution as well as having the lowest error rates among classifiers.

MDS Analysis

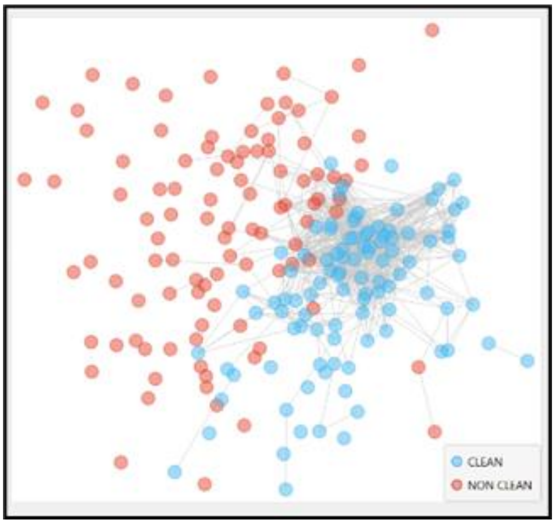

Figure 10. MDS – VGG 16

ISSN: 2716-3865 (Print), 2721-1290 (Online)

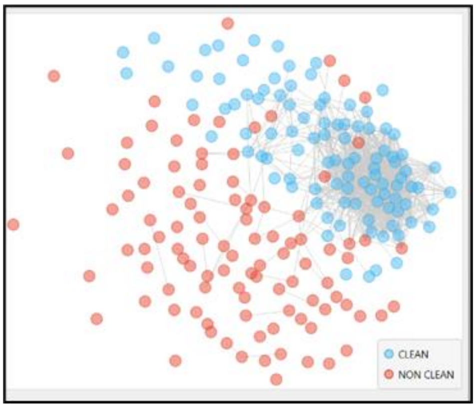

Figure 11. MDS – VGG 19

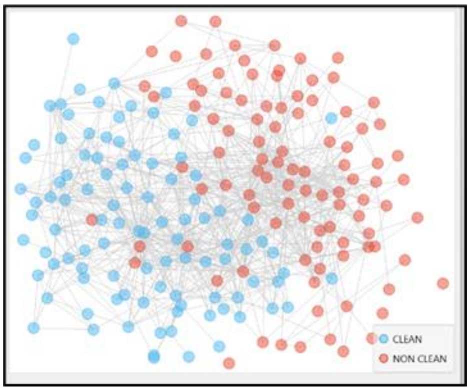

Figure 12.MDS – Inception V3

Given the Multidimensional Scaling (MDS) visualisation we can see the 3 CNN architectures provide different amount of class separability in learning features of clean vs non-clean outfits. The separation between clusters is the most visually obvious for VGG-16, each main group can be identified more as different layers of overlap as in ResNet. This suggests that the features learnt by VGG-16 are more discriminative between the visual appearances of the two classes. While cluster separation is still clear in VGG-19, the area of overlap between clusters is larger than that from the VGG-16. On the other hand, Inception V3 has high overlap ranging from 0.01 to 1, with less clean and non-clean outfit points observed. These facts are aligned with the quantitative performance and confirm CNN architecture affects greatly to embedding quality and classification accuracy in fashion image classification (Krizhevsky et al., 2012; Pan & Yang, 2010).

Silhouette Plot Analysis

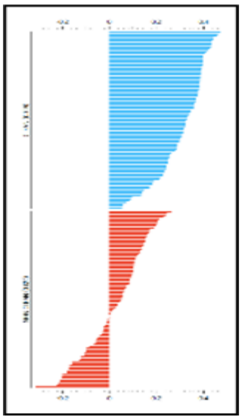

Figure 13. Silhouette Plot of VGG 16

ISSN: 2716-3865 (Print), 2721-1290 (Online)

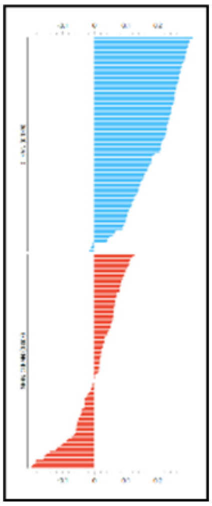

Figure 14. Silhouette Plot of VGG 19

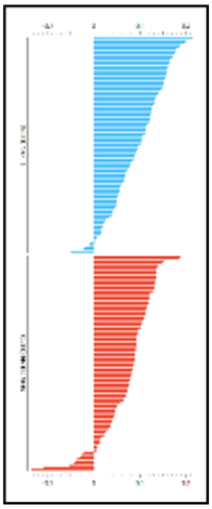

Figure 15. Silhouette Plot of Inception V3

In general the silhouette Plot results corroborate that VGG-16 produces superior quality embeddings, followed by VGG-19 andLastly Inception V3. This observation is consistent with the experimental results obtained from MDS visualization and quantitative verification pattern, and it demonstrates that classification performance of recognising clean versus non-clean outfits depends much on which CNN architecture used for extracting feature representation.

The experiment results indicate the selection of CNN architecture is critical in featurization quality and classification performance. The VGG-16 and Logistic Regression has achieved the best performance in various evaluation metrics, implying that the features are well discriminative and linearly separable.

VGG-19, on the other hand despite of its very high number of layers seems to deliver marginally worse accuracy which supports that scaling up the model does not necessarily enahance class separation. The greatest feature separations are induced by Inception V3, as recovered in MDS and Silhouette analysis, which means more dispersed features and more overlaps across class (Arandjelovic et al., 2016; Bengio et al., 2013; Rousseeuw, 1987). The correspondence between quantitative evaluations and visual inspections suggests VGG-16 achieves the best feature representation for both clean or non-clean outfits recognition in this study.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Conclusion

In this work, we comparative study popular CNN-style feature extractors paired with traditional classifiers for fashion images classification. Results reveal that VGG-16 architecture with Logistic Regression performs best overall.

The results of experiments show that the model combined with VGG-16 and Logistic Regression has better performance than other models in terms of AUC values and several other evaluation indexes. VGG-19 has an acceptable performance, but not better than VGG-16 and Inception V3 records the lowest class separability. The results in Table 3 reveal that the choice of CNN architecture plays a crucial role, both for generating high quality embeddings and obtaining good classification results.

Future works involve adding more products to enlarge the dataset, retraining CNN structure on fashion-specific training data, integrating other visual properties like content and texture, developing end-to-end technique for deploying in real-world.

References

Akram, M. W., Abbas, A., & Khan, I. A. (2022). Effects of perceived value, service quality and customer trust in home delivery service staff on customer satisfaction: evidence from Pakistan. *International Journal of Management Research and Emerging Science*, *12*(4). https://doi.org/10.56536/ijmres.v12i4.351

Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2016). NetVLAD: CNN Architecture for Weakly Supervised Place Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5297–5307. https://doi.org/10.1109/CVPR.2016.572

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bharadwaj, Prakash, K. B., & Kanagachidambaresan, G. R. (2021). *Pattern Recognition and Machine Learning* (pp. 105–144). https://doi.org/10.1007/978-3-030-57077-4_11

Buradagunta, S., & Balakrishna, S. (2025). A two-stage deep learning approach for optimizing fashion product recommendations. *SN Computer Science*, *6*(4), 345. https://doi.org/10.1007/s42979-025-03909-2

Chae, M. (2022). Gender Differences in Adaptive Clothing: Applying Functional, Expressive, and Aesthetic (FEA) Needs of People with Movement Impairments. *International Journal of Fashion Design, Technology and Education*, *15*(3), 360–370. https://doi.org/10.1080/17543266.2022.2071468

Choudhary, G., & Sethi, D. (2023). From conventional approach to machine learning and deep learning approach: an experimental and comprehensive review of image fusion techniques. *Archives of Computational Methods in Engineering*, *30*(2), 1267-1304. https://doi.org/10.1007/s11831-022-09833-5

Deldjoo, Y., Nazary, F., Ramisa, A., Mcauley, J., Pellegrini, G., Bellogin, A., & Noia, T. D. (2023). A review of modern fashion recommender systems. *ACM Computing Surveys*, *56*(4), 1-37. https://doi.org/10.1145/3624733

Dillon, E., & Muhammad, A. (2024). The Power of Professional Dress: Competence, Confidence, and Generational Shifts. *The Lookout: Journal of Undergraduate Research & Creative Activity Journal at ECU*, 26–33. https://lookout.ecu.edu/index.php/lookout/article/view/31

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Elfatimi, E., Eryiğit, R., & Elfatimi, L. (2024). Deep multi-scale convolutional neural networks for automated classification of multi-class leaf diseases in tomatoes. *Neural Computing and Applications*, *36*(2), 803-822. https://doi.org/10.1007/s00521-023-09062-2

Gómez-Talal, I., Bote-Curiel, L., Lozano-Paredes, D., Feijoo-Martínez, J. R., & Rojo-Álvarez, J. L. (2026). Evaluating Manifold Learning Techniques for Dimensionality Reduction on Industrial IoT Cybersecurity Data. *IEEE Access*. https://doi.org/10.1109/ACCESS.2026.3662155

Hossain, M. S., Basak, N., Mollah, M. A., Nahiduzzaman, M., Ahsan, M., & Haider, J. (2025). Ensemble-based multiclass lung cancer classification using hybrid CNN-SVD feature extraction and selection method. *PLoS One*, *20*(3), e0318219. https://doi.org/10.1371/journal.pone.0318219

Hu, Y., Wang, P., Jia, M., Zhang, Y., Hong, J., Wan, G., ... & Li, T. (2026). EMFFTrans: Efficient Multi-Scale Feature Fusion Transformer for Road Scene Semantic Segmentation. *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2026.3679203

Kasim, S., Malek, S., Tang, J., Kiew, X. N., Cheen, S., Liew, B., ... & Shariff, R. (2025). Multiclass leukemia cell classification using hybrid deep learning and machine learning with CNN-based feature extraction. *Scientific reports*, *15*(1), 23782. https://doi.org/10.1038/s41598-025-05585-x

Kaur, N., & Pandey, S. (2023). Predicting clothing attributes with CNN and SURF based classification model. *Multimedia Tools and Applications*, *82*(7), 10681-10701. https://doi.org/10.1007/s11042-022-13714-1

Khalid, L., & Gong, W. (2026). Exploring AI in Fashion: A Review of Aesthetics, Personalization, Virtual Try-On, and Forecasting. *Multimedia Systems*, *32*(3), 1–10. https://doi.org/10.1007/s00530-026-02232-x

Khan, S. H., & Iqbal, R. (2025). A comprehensive survey on architectural advances in deep CNNs: challenges, applications, and emerging research directions. *arXiv preprint arXiv:2503.16546*. https://doi.org/10.48550/arXiv.2503.16546

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, *25*, 1–10. http://code.google.com/p/cuda-convnet/

Kulkarni, S., & Harnoorkar, S. (2020). Comparative Analysis of CNN Architectures. *International Research Journal of Engineering and Technology*, 1–6. www.irjet.net

Lee, V., Park, S., & Lee, D. (2022). The effect of e-commerce service quality factors on customer satisfaction, purchase intention, and actual purchase in Uzbekistan. *Global Business & Finance Review (GBFR)*, *27*(3), 56-74. https://doi.org/10.17549/gbfr.2022.27.3.56

Liu, X. (2025, April). A convolutional neural network-based platform for clothing style feature extraction and new style generation: integrating optical feature extraction for enhanced design. In *Fifth International Conference on Telecommunications, Optics, and Computer Science (TOCS 2024)* (Vol. 13629, pp. 95-102). SPIE. https://doi.org/10.1117/12.3067441

Lv, P., Xu, H., Zhang, Y., Zhang, Q., Pan, Q., Qin, Y., ... & Chen, C. (2024). An improved multi-scale feature extraction network for rice disease and pest recognition. *Insects*, *15*(11), 827. https://doi.org/10.3390/insects15110827

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Moujahid, A., & Dornaika, F. (2025). Advanced unsupervised learning: a comprehensive overview of multi-view clustering techniques. *Artificial Intelligence Review*, *58*(8), 234. https://doi.org/10.1007/s10462-025-11240-8

Mudgal, P. (2026). Clustering of temporal and visual data: Recent advancements. *Data*, *11*, 7. https://doi.org/10.3390/data11010007

Nassar, S., Hedjam, R., & Belhaouari, S. B. (2026). Novel Loss Functions for Improved Data Visualization in t-SNE. *Machine Learning and Knowledge Extraction*, *8*(2), 47. https://doi.org/10.3390/make8020047

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Panpaeng, S., Phanphaeng, P., Kumnuanta, J., Yommakit, P., Kocento, K., & Wongchompoo, P. (2023). The Application of Data Mining Techniques for Predicting Education to New Undergraduate Students at Chiang Mai Rajabhat University. *2023 IEEE International Conference on Cybernetics and Innovations (ICCI)*, 1–6. https://doi.org/10.1109/ICCI57424.2023.10112233

Rabbi, M. F., Sultan, M. N., Hasan, M., & Islam, M. Z. (2023). Tribal Dress Identification using Convolutional Neural Network. *J. Inf. Hiding Multim. Signal Process.*, *14*(3), 72-80.

Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to The Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Saut, M., & Bie, S. (2024). Impact of service expectation, experiential quality, and perceived value on hotel customer satisfaction. *Journal of Quality Assurance in Hospitality & Tourism*, *25*(4), 781-809. https://doi.org/10.1080/1528008X.2022.2141414

Shao, C. Y., Baker, J. A., & Wagner, J. (2004). The Effects of Appropriateness of Service Contact Personnel Dress on Customer Expectations of Service Quality and Purchase Intention. *Journal of Business Research*, *57*(10), 1164–1176. https://doi.org/10.1016/S0148-2963(02)00326-0

Shushi, A., & Abdulazeez, A. M. (2024). Fashion design classification based on machine learning and deep learning algorithms: a review. *The Indonesian Journal of Computer Science*, *13*(3). https://doi.org/10.33022/ijcs.v13i3.3980

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–12. https://arxiv.org/pdf/1409.1556

Singh, L. K., Khanna, M., Thawkar, S., & Singh, R. (2023). Nature-inspired computing and machine learning based classification approach for glaucoma in retinal fundus images. *Multimedia Tools and Applications*, *82*(27), 42851-42899. https://doi.org/10.1007/s11042-023-15175-6

Sudirjo, F., Yuliana, A., Novilia, F., Kalalo, R. R., & Belani, S. W. (2023). The Influence Of Service Quality And Consumer Experience On Consumer Repurchase Intention. *Innovative: Journal Of Social Science Research*, *3*(6), 3965-3973.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818–2826). https://doi.org/10.1109/CVPR.2016.308

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Wang, B., Li, Y., Zhou, M., Han, Y., Zhang, M., Gao, Z., ... & Liu, B. F. (2023). Smartphone-based platforms implementing microfluidic detection with image-based artificial intelligence. *Nature Communications*, *14*(1), 1341. https://doi.org/10.1038/s41467-023-36017-x

Wang, S., Shi, B., Wang, N., Zhang, Y., & Zhu, Y. (2025). Convformer: Multi-Scale Masked Hybrid Convolution-Transformer Network for Hyperspectral Image Super-Resolution. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. https://doi.org/10.1109/JSTARS.2025.3569410

Xiao, H., Rasul, K., & Vollgraf, R. (2017, August 25). Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms. *Computer Science*. https://arxiv.org/pdf/1708.07747

Xie, W., Li, Z., Xu, Y., Gardoni, P., & Li, W. (2022). Evaluation of different bearing fault classifiers in utilizing CNN feature extraction ability. *Sensors*, *22*(9), 3314. https://doi.org/10.3390/s22093314

Zhang, Y., He, K., & Song, R. (2023). Image multi-feature fusion for clothing style classification. *IEEE Access*, *11*, 107843-107854. https://doi.org/10.1109/ACCESS.2023.3320270

ISSN: 2716-3865 (Print), 2721-1290 (Online)