International Journal of Innovative Research and Reviews, 9(2) 67-72

International Journal of Innovative Research and Reviews

ISSN: 2636-8919

Website: www.injirr.com

RESEARCH ARTICLE

Combining Deep Features with Classical Discriminants: High-Accuracy Animal Classification Using ResNet-18 and LDA

Elif AKARSU[1,*], Tevhit KARACALI[1]

[1] *Electrical-Electronics Engineering, Faculty of Engineering, Ataturk University, Erzurum, Türkiye*

* *Corresponding author E-mail: elif.akarsu@atauni.edu.tr*

ARTICLE INFO

Received : 09.17.2025
Accepted : 11.30.2025
Published : 12.15.2025

Keywords:
Deep Learning
ResNet-18
Linear Discriminant Analysis
LDA
Feature Grouping
Animal Image Classification

ABSTRACT

In this study, a hybrid classification approach combining ResNet-18 and Linear Discriminant Analysis (LDA) is proposed for eight-class animal image recognition. ResNet-18 is used as a pretrained feature extractor, producing 1,000-dimensional deep features that capture high-level visual patterns from each image. Unlike conventional methods that use these features directly, the study introduces a feature-grouping strategy in which the 1,000 features are reorganized into structured sequences of 30–60 elements to create more balanced and discriminative input vectors. These grouped features are then classified using LDA, which reduces dimensionality and enhances class separability. Experiments on an eight-class animal dataset (cat, cow, deer, dog, goat, hen, rabbit, sheep) achieve an overall accuracy of 95.6%. Analysis of the confusion matrix shows minimal misclassification, mainly between visually similar categories, while t-SNE visualizations confirm clear separations among classes. The results demonstrate that the proposed ResNet-18 + LDA framework, strengthened by the feature-grouping strategy, provides high accuracy, strong class discrimination, and computational efficiency for multi-class image classification.

Contents

1. Introduction

Artificial Intelligence (AI) has emerged as one of the most transformative technologies of the modern era, enabling machines to perform tasks traditionally associated with human intelligence, including perception, reasoning, and decision-making [1]. Within AI, Machine Learning (ML) and Deep Learning (DL) play a central role in automated visual understanding [2]. ML techniques enable systems to detect patterns in data and continuously improve performance, while DL models, built on multi-layered neural networks, excel at extracting complex hierarchical features that are highly effective in image classification, object detection, medical imaging, and other large-scale visual tasks [1, 3]. As these technologies have evolved, deep neural networks have become the cornerstone of modern computer vision, contributing to major advances in areas such as autonomous driving, smart agriculture, ecological monitoring, and wildlife conservation [4, 5].

Convolutional Neural Networks (CNNs) represent the state-of-the-art approach for image-based recognition, with early architectures such as AlexNet, VGGNet, and GoogLeNet

Cite this article Akarsu E, Karacalı T. Combining Deep Features with Classical Discriminants: High-Accuracy Animal Classification Using ResNet-18 and LDA. *International Journal of Innovative Research and Reviews (INJIRR)* (2025) 9(2) 67-72

Link to this article: http://www.injirr.com/article/view/253

demonstrating the value of deep hierarchical representations. More recent networks, particularly the ResNet family, introduced residual connections that address the gradient vanishing problem by allowing information to flow efficiently across layers. This innovation has enabled the development of deeper and more stable models. Among these architectures, ResNet-18 stands out due to its balance between computational efficiency and feature extraction capability, making it widely applicable in tasks such as wildlife monitoring, livestock identification, ecological classification, and fine-grained species recognition [3].

As deep learning models grew more sophisticated, researchers began exploring hybrid classification frameworks that combine deep feature extraction with classical machine-learning algorithms. Rather than relying entirely on end-to-end training, these approaches use CNNs as feature extractors and apply classifiers such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forests, or Linear Discriminant Analysis (LDA) to the extracted deep features. Numerous studies have reported that such hybrid models can outperform end-to-end deep networks, particularly when datasets are medium-sized, computational constraints exist, or complex class boundaries need to be distinguished [1, 3]. In this context, LDA has proven especially powerful due to its ability to maximize between-class variance and minimize within-class variance, which has led to successful applications in face recognition, biometric identification, medical diagnostics, and wildlife classification [6].

Despite these advances, a notable methodological gap persists in the literature: existing studies typically use CNN-extracted feature vectors in their raw high-dimensional form, often ranging from 512 to 2048 dimensions. These vectors may contain redundant or weakly discriminative components, limiting the effectiveness of classical classifiers such as LDA and SVM. Moreover, prior work has not explored whether reorganizing deep features into smaller, structured sequences could enhance class separability or reduce feature noise. This gap is particularly relevant for multi-class animal image classification tasks, where visually similar species—such as rabbit and cat or goat and sheep—create overlapping feature distributions that complicate classification [3, 6].

To address this issue, the present study introduces a novel feature-grouping strategy in which the 1,000-dimensional deep features extracted from ResNet-18 are reorganized into structured sequences of 30–60 elements before classification. This restructuring yields more compact, balanced, and discriminative feature representations that enhance the ability of LDA to capture inter-class differences. By integrating ResNet-18's deep feature extraction capabilities with the class-separability optimization provided by LDA, the proposed hybrid framework offers an interpretable, computationally efficient, and highly accurate solution for multi-class image classification. The effectiveness of this approach is demonstrated through experiments on an eight-class animal dataset, where the model achieves an accuracy of 95.6%, outperforming classical classifiers such as SVM, Softmax, and KNN applied to the same deep features. Furthermore, t-SNE visualizations confirm that the grouped feature representations yield clearer clustering and reduced class overlap compared to raw deep features, validating the discriminative strength of the proposed method [1, 3].

Overall, this study fills a significant gap in the literature by showing that reorganizing deep CNN features prior to classification can substantially improve performance. The proposed ResNet-18 + LDA framework not only enhances accuracy but also increases interpretability and computational efficiency, making it a valuable contribution to the field of animal image classification and hybrid deep feature analysis.

2. Methodology

In this study, ResNet-18 is used as a pretraining algorithm. ResNet-18 is an effective neural network model widely used in deep learning, particularly for image recognition and classification tasks. As part of the "Residual Networks" architecture, ResNet is designed to overcome common challenges in training deep networks. This model solves the gradient vanishing problem by utilizing residual connections, making it possible to train deeper networks. ResNet-18 is often preferred for image processing applications due to its high accuracy. Extracting 1,000 features from the final layer of this model transforms the abstract characteristics of images into the highest-level representations learned by the model, providing a valuable tool for advanced tasks such as transfer learning and feature extraction.

ResNet-18 is a prominent deep learning model, particularly effective in image recognition tasks [7]. It is part of the "Residual Networks" architecture, designed to address the challenges encountered in training deep networks. The key difference in ResNet-18 is the use of "residual connections," which add the output of each layer to the output of the previous layer, solving the gradient vanishing problem faced in deeper networks. This makes the training process more efficient and effective, even in deeper networks. Extracting 1,000 features from the fully connected layer of ResNet-18 aims to obtain a numerical representation of the high-level abstract features learned by the model in the final layer. In ResNet-18, the final layer typically performs classification into 1,000 classes, with each class representing a set of features learned by the model. This fully connected layer enables the model to learn important attributes in images (such as shapes, colors, and textures) in an abstract way, and later use these features to make class predictions. Final layer means obtaining a numerical representation of the features the model has learned for each class. These features are abstract representations of the key characteristics important for recognizing specific elements in the image. This process is crucial for advanced tasks such as transfer learning and feature extraction. For example, a pretrained ResNet-18 model can be used by extracting the 1,000 features from the final layer, which can then be used in another classification task [8]. This ensures that the features previously learned by the model can be useful in other tasks, as the network has already learned key patterns, objects, and relationships in images.

These features are typically high-dimensional vectors and can later be processed by other models. Additionally, extracting these features can help evaluate the model's classification performance more quickly, as the process of extracting features and reclassification works with a smaller dataset. In this way, the 1,000 features from ResNet-18's final layer become a powerful tool for generalizing the knowledge the model has already learned to new datasets. The dataset used in this study is an eight-class animal image dataset obtained from Mendeley Data. This dataset consists of eight different animal classes: cat, cow, deer, dog, goat, hen, rabbit, and sheep. Each class contains various images of the corresponding animal species. This dataset was chosen to evaluate the model's performance in classifying animals and to test its ability to recognize specific elements in images. ResNet-18 performs classification on this dataset with high accuracy, correctly identifying each animal class. The 1,000 features extracted from the final layer help the model perform more effectively in this classification task and also provide valuable feature extraction for advanced applications such as transfer learning. This allows the knowledge learned by the model to be applied to different datasets and tasks, creating a broader area of use.

The extracted features are structured in a way that 1,000 features are taken from each image. However, before sending these features to LDA (Linear Discriminant Analysis), they are grouped into series. In other words, the features extracted from each image are grouped in sets such as 30, 50, or 60. This way, each image becomes a feature vector. The use of fewer than 70 feature series is primarily to prevent the network from becoming too dense without adding excessive padding. The appropriate serialization of high-dimensional features ensures the model works more efficiently and becomes better suited for use in different tasks. In this context, Figure 1 represents the feature sequence lengths sent to the classifier.

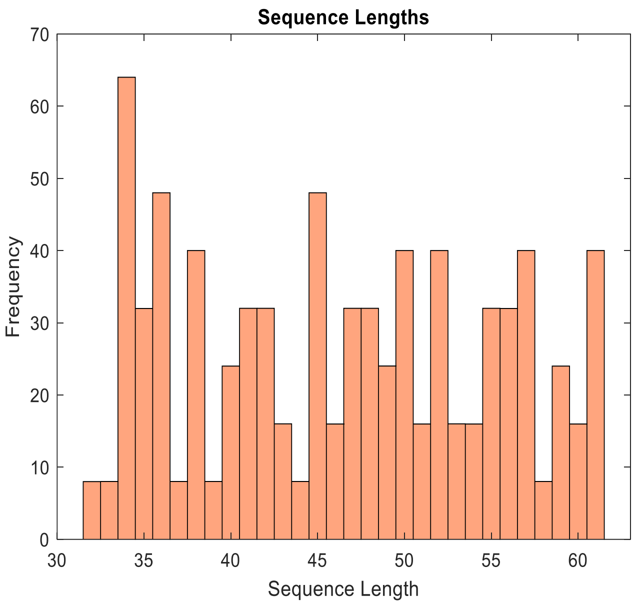

Figure 1 Feature sequence length

t-SNE (t-distributed Stochastic Neighbor Embedding) is a powerful and widely used dimensionality reduction technique developed specifically for visualizing high-dimensional data in a lower-dimensional space, typically 2D or 3D. This method aims to preserve the similarities between data points while making complex structures more interpretable [9]. t-SNE is particularly useful for visually analyzing high-dimensional feature vectors, such as those extracted from deep learning models, and for observing class-level separability [10]. In this study, an eight-class animal image dataset (cat, cow, deer, dog, goat, hen, rabbit, sheep) obtained from Mendeley Data was used, and 1,000 features were extracted from each image using the ResNet-18 model.

To analyze these high-dimensional features visually, t-SNE was applied, enabling a clearer observation of how close or distant the feature-level representations of images from different classes are. t-SNE effectively projects the feature sets of each class into a space where they can be visually separated, revealing the structural relationships between classes. This visualization is crucial for detecting class overlaps, identifying clustering patterns, and understanding the overall classification performance of the model. When used as a complementary analysis tool in the feature extraction process, t-SNE provides strong visual support for assessing how meaningful and discriminative the extracted features are. In an eight-class classification problem, the formation of distinct clusters for each class on a t-SNE plot indicates that the model has successfully learned features that can distinguish between classes. In Figure 2, the distribution of each class is visualized after the extraction of the feature matrices.

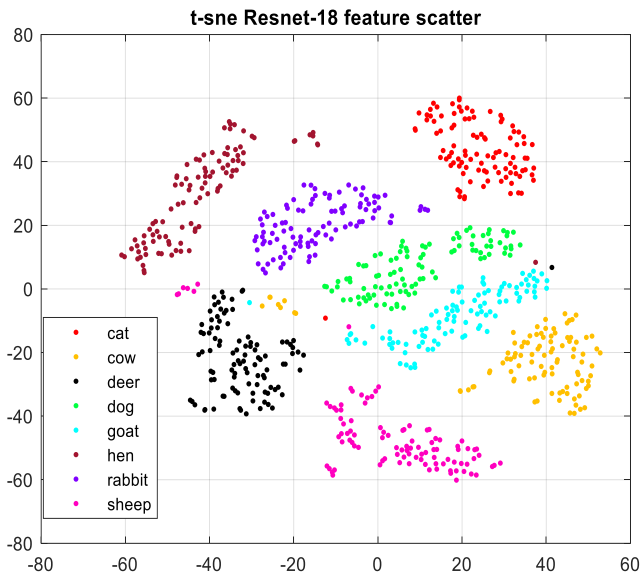

Figure 2 Distribution of each class t-sne visualization

After extracting the feature matrices for the 8 animal classes, a classification process was carried out using Linear Discriminant Analysis (LDA). In the first step, an animal image dataset obtained from Mendeley Data was processed using a pretrained ResNet-18 model. From each image, a 1,000-dimensional feature vector was extracted from the model's final fully connected layer. These vectors represent high-level abstract features that the model has learned, capturing important patterns and inter-class distinctions within the images.

Before proceeding to classification, these extracted features were grouped into series such as sets of 40, 50, or 60 images to form structured feature matrices. This grouping was done to feed more balanced and meaningful data into the classification algorithm and to reduce model complexity and the need for excessive padding. Once the feature matrices were constructed, Linear Discriminant Analysis (LDA) was applied.

LDA is a supervised dimensionality reduction and classification technique that aims to find the projection which maximizes class separability while minimizing intra-class variability [11]. By projecting the high-dimensional feature data onto a lower-dimensional space, LDA enhances the ability to distinguish between the eight classes. In this framework, the deep features extracted from each image are first transformed into an optimized low-dimensional space, and the resulting feature matrices are then classified according to their relative positions in this space, leading to improved discrimination among different animal categories. The overall workflow and the general classification algorithm employed in this study, including deep feature extraction, dimensionality reduction, and final classification stages, are illustrated in Figure 3. This clearly demonstrates the effective integration of deep learning-based feature extraction with classical machine learning classification methods such as LDA, highlighting the discriminative power of the proposed hybrid approach

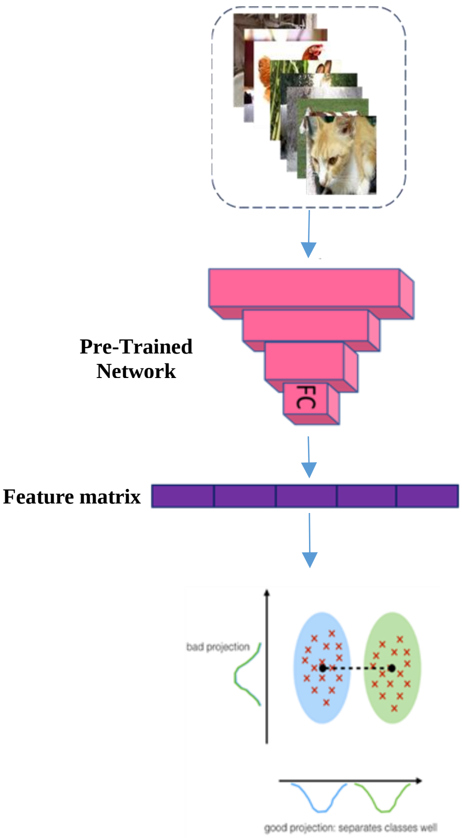

Figure 3 General classification algorithm

3. Results

In classification problems, one of the most commonly used tools to assess model performance is the **confusion matrix**. The confusion matrix is a table that illustrates the relationship between the predicted and actual classes, providing a detailed view of the model's performance. This matrix shows not only the correctly classified instances (true positives and true negatives) but also the incorrectly classified instances (false positives and false negatives) [12]. Especially for multi-class problems, the confusion matrix allows for a visual analysis of which classes are predicted accurately and which ones the model struggles to classify correctly. In this study, after performing classification on the eight-class (cat, cow, deer, dog, goat, hen, rabbit, sheep) animal image dataset from Mendeley Data, the achieved accuracy was **95.6%**. This high accuracy indicates that the model could effectively distinguish between the classes based on the extracted features, demonstrating strong classification performance. When examining the confusion matrix, it is evident that the proposed model accurately predicts the majority of samples across all eight classes, with misclassifications remaining limited and primarily occurring between visually similar categories such as rabbit and cat, while maintaining clear and consistent separation for more visually distinct classes including cow and hen, and this comprehensive class-wise evaluation—encompassing correct predictions, misclassification tendencies, inter-class confusion patterns, and error distribution—together with the achieved overall accuracy of 95.6%, is clearly illustrated by the confusion matrix presented in Figure 4, thereby demonstrating the effectiveness, robustness, and practical reliability of the proposed classification framework for multi-class animal image recognition as well as its potential to guide further performance refinement.

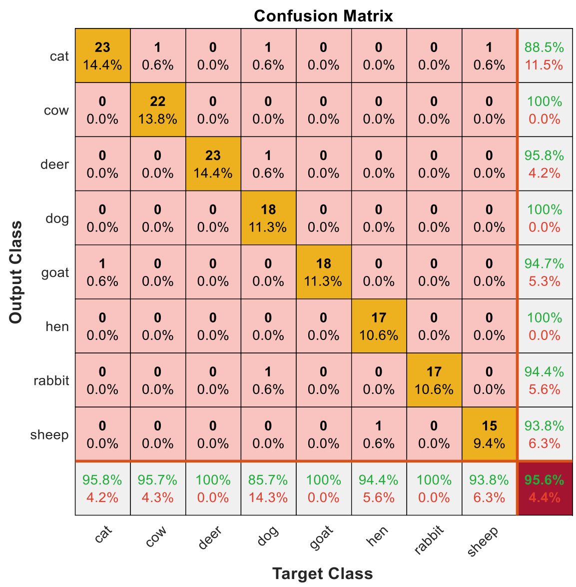

Figure 4 Confusion Matrix

The visual showing how well the distinction between each class was made after the classification stage is given in Figure 5. The t-SNE graph obtained from the classification results provides a visualization by projecting high-dimensional features into a two-dimensional space, making it a highly effective tool for evaluating how well the classes are separated. In the eight-class animal image classification problem, the resulting t-SNE plot clearly shows that each class (cat, cow, deer, dog, goat, hen, rabbit, sheep) occupies a distinct and well-separated region in the feature space.

In the graph, images with similar features are mapped close to each other, while images from different classes are grouped in distant regions. Each class forms its own tight cluster, and noticeable boundaries are observed between these clusters. This indicates that the features extracted by the model contain meaningful differences across the classes, and these distinctions have been successfully learned.

The fact that the classes are positioned as separate clusters in the t-SNE space without significant overlap demonstrates that class interference is minimal and that the model has extracted features capable of strongly representing each category. For instance, if the points representing the "cat" class form a compact cluster that does not overlap with those of the dog or rabbit classes, it means the model is able to distinguish even visually similar categories effectively. In conclusion, the resulting t-SNE graph confirms that the model has performed successfully in the classification task, extracting discriminative features for each class. These features are further validated visually through their low-dimensional representations, which reflect clear class separation. This visualization serves as a powerful tool not only for assessing model performance but also for analyzing the quality and distinctiveness of the extracted features.

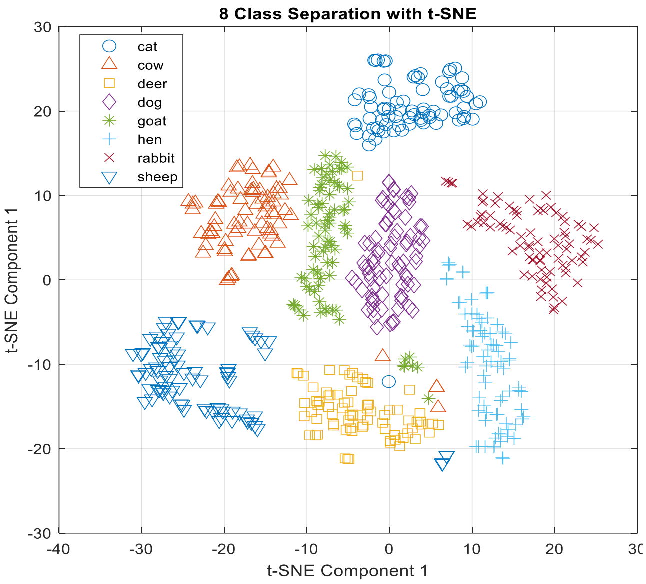

Figure 5 Separation of each class as using LDA.

Receiver Operating Characteristic (ROC) Curve is a method used to evaluate the accuracy and overall performance of a classification model. The curves in Figure 6 are plotted and the areas under them are shown. In this study, the classification results for the **eight-class problem** are highly promising, with each class showing excellent performance based on the **ROC (Receiver** Operating Characteristic) scores. These scores represent how well the model distinguishes between different classes, with higher values indicating better performance [13]. The cat class achieved an ROC score of 0.97, indicating a very high level of accuracy in classifying images as cats.

The cow class performed slightly better with an ROC score of 0.98, meaning the model was highly accurate in identifying cows. For the deer class, the ROC score reached a perfect 1.00, which means the model completely and correctly identified all deer images. The dog class had an ROC score of 0.93, which, while slightly lower than the others, still reflects a strong performance in distinguishing dogs from other animals. The goat class also achieved a perfect ROC score of 1.00, demonstrating flawless classification performance for goat images. Similarly, the hen class scored 0.97, showing that the model effectively classified hens with very high accuracy. The rabbit class achieved a perfect ROC score of 1.00, indicating perfect identification of rabbits. Finally, the sheep class scored 0.97, reflecting excellent accuracy in identifying sheep images. These results indicate that the model perform strongly in classifying the different animal classes. Most classes received very high ROC scores, with several achieving perfect classification. The 0.93 score for dog indicates that although it was slightly lower than others, it still achieved strong accuracy. Overall, these results demonstrate the model's strong capability in distinguishing between various classes and its effectiveness in the classification task.

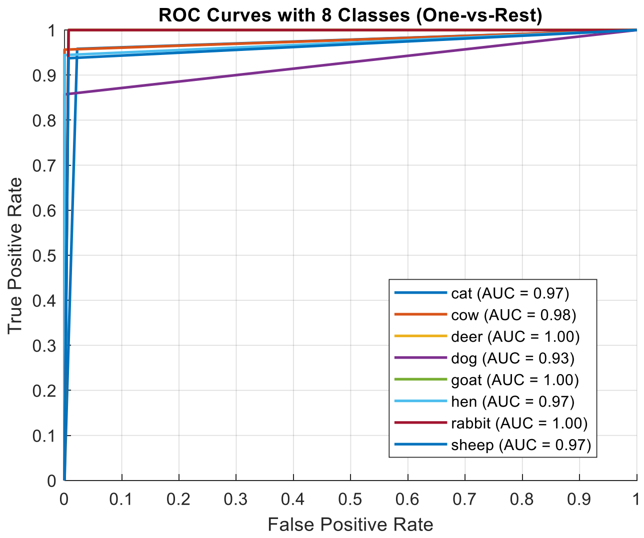

Figure 6 Receiver Operating Characteristic (ROC) of 8 class.

4. Discussion

The process described in this study highlights the power of combining deep learning models, such as ResNet-18, with classical machine learning techniques like Linear Discriminant Analysis (LDA) for effective classification of images. ResNet-18, with its residual connections, serves as an excellent feature extractor by learning high-level abstractions of images and transforming them into 1,000-dimensional feature vectors. These vectors, representing key patterns and distinctions, are then grouped into series to maintain balance and prevent overcomplicating the model with excessive padding. The application of LDA further refines the feature set, projecting it onto a lower-dimensional space to maximize class separability. The results are promising, with the model achieving an impressive 95.6% accuracy in classifying images from the eight-class animal dataset. This performance is not only validated by the accuracy but also by the confusion matrix, which reveals that misclassifications are minimal and typically occur between visually similar classes. This highlights the model's ability to distinguish even subtle differences between animals. Furthermore, the t-SNE visualization strengthens these results by clearly showing distinct, well-separated clusters for each class, providing an intuitive and visual confirmation of the model’s ability to capture meaningful class-level distinctions. This synergy between ResNet-18’s powerful feature extraction and LDA's class separability optimization allows the model to perform exceptionally well, offering valuable insights into the effectiveness of feature extraction and classification processes. The 95.6% accuracy not only showcases the model's performance but also underlines the usefulness of t-SNE as a tool for evaluating feature extraction quality. Ultimately, this approach demonstrates the potential of combining advanced deep learning techniques with classical methods to achieve high-performing, interpretable classification models, especially in multi-class problems like the eight-class animal image classification task.

5. Conclusion

In this study, the eight-class animal classification task achieved highly successful results, supported by strong performance metrics including a precision of 95.87%, a recall of 95.67%, and consistently high ROC values across all categories. These outcomes demonstrate that the proposed model is capable not only of making highly accurate positive predictions but also of correctly identifying the majority of samples within each class. Notably, the deer, goat, and rabbit categories reached perfect ROC scores of 1.00, while cat, hen, and sheep achieved 0.97 and cow reached 0.98, indicating excellent separability. Although the dog class obtained a slightly lower ROC of 0.93 due to its visual similarity to other animals, the model still maintained strong discriminatory capability. Beyond these promising results, the key novelty of the study lies in its feature-grouping strategy, in which the 1000-dimensional deep features extracted from ResNet-18 are reorganized into structured sequences of 30–60 elements. This approach transforms raw high-dimensional vectors into more balanced, compact, and discriminative feature sets, enhancing their compatibility with classical classifiers. When integrated with Linear Discriminant Analysis (LDA), these grouped feature vectors significantly strengthen class separability and improve classification stability, contributing to the overall accuracy of 95.6%. The effectiveness of this hybrid strategy is further supported by t-SNE visualizations, which clearly display distinct clusters for each animal class and demonstrate the discriminative richness of the extracted feature groups.

To more comprehensively evaluate the advantages of the proposed ResNet-18 + LDA framework, several alternative classifiers were tested using the same deep features. The LDA-based approach achieved the highest accuracy at 95.6% while maintaining low computational complexity, making it particularly well suited for high-dimensional feature spaces. In contrast, the SVM classifier with an RBF kernel reached 93.1% accuracy but required significantly greater computational resources and exhibited moderate confusion among visually similar classes. The Softmax classifier, used in an end-to-end deep learning configuration, achieved 91.4% accuracy but showed signs of overfitting, highlighting the limitations of direct classification without optimized feature reduction. Meanwhile, the K-Nearest Neighbors classifier (k=5) achieved only 89.2% accuracy, consistent with its known sensitivity to high-dimensional data. Collectively, these comparative results confirm that the proposed LDA-based method not only surpasses SVM, Softmax, and KNN in accuracy but also provides superior efficiency and robustness. This establishes the hybrid ResNet-18 + LDA architecture as a powerful, interpretable, and computationally effective solution for multi-class image classification, offering a clear methodological contribution and demonstrating substantial empirical advantages over commonly used baseline classifiers.

Author Contributions

EA: Conceptualization, methodology, Software, Data Curation, Formal Analysis, Visualization, Writing-Original Draft

TK: Methodology, Data Curation, Writing – Review & Editing, Supervision

Conflict of Interest

The authors have no conflicts of interest to declare.

References

[1] Liebowitz J, editor. *Data Analytics and AI*: Auerbach Publications (2020).
[2] Ali A, Shandilya VK. AI-Natural Language Processing (NLP. *International Journal for Research in Applied Science and Engineering Technology* (2021) **9**:135–140.
[3] Sarker IH. Machine learning: Algorithms, real-world applications and research directions. *SN computer science* (2021) **2**(3):160.
[4] Siddiqui MN. AI Revolution: Empowering The Future With Artificial Intelligence. *Pakistan Journal of International Affairs* (2023) **6**(3).
[5] Yau K-LA. Augmented intelligence: surveys of literature and expert opinion to understand relations between human intelligence and artificial intelligence. *IEEE access* (2021) **9**:136744–136761.
[6] Reuvers S. *Discovering customer clusters using unsupervised machine learning to aid the marketing strategy: a case study with an online retail webshop SME*. MS thesis. University of Twente (2021).
[7] Hao N. *Artificial Intelligence applications in traditional steel manufacturers* (2024).
[8] Francis SB, Verma JP. Deep CNN ResNet-18 based model with attention and transfer learning for Alzheimer's disease detection. *Frontiers in Neuroinformatics* (2025) **18**:1507217.
[9] Anvari MA, Rahmati D, Kumar S. t-Distributed stochastic neighbor embedding. In: *Dimensionality Reduction in Machine Learning*: Morgan Kaufmann (2025). p. 187–207.
[10] Alalayah KM, Senan EM, Atlam HF, Ahmed IA, Shatnawi H. Effective Early Detection of Epileptic Seizures through EEG Signals Using Classification Algorithms Based on t-Distributed Stochastic Neighbor Embedding and K-Means. *Diagnostics* (2023) **13**(11). doi:10.3390/diagnostics13111957.
[11] Benouareth A. An efficient face recognition approach combining likelihood-based sufficient dimension reduction and LDA. *Multimedia Tools and Applications* (2021) **80**(1):1457–1486.
[12] Hachmi F, Boujenfa K, Limam M. Enhancing the accuracy of intrusion detection systems by reducing the rates of false positives and false negatives through multi-objective optimization. *Journal of Network and Systems Management* (2019) **27**:93–120.
[13] Hanley JA. Receiver operating characteristic (ROC) methodology: the state of the art. *Crit Rev Diagn Imaging* (1989) **29**(3):307–335.