Applications and Developments of Deep Learning in Image Processing

Xiang Wang[a]
School of Business & IT, James Cook University, Singapore Campus, Singapore

Keywords: Deep Learning, Image Processing, Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), Medical Image Analysis.

Abstract: Deep learning has become a game changer in image processing, and it has served as the most effective approach for most practical image processing problems like object recognition, image segmentation, and image enhancement, among others. Regarding deep learning for image processing, this paper relates these enhanced techniques with some of the applications and trends in health care, security, and self-driving cars. CNNs and GANs have become two indispensable techniques for future development in medical image analysis, facial recognition, and synthetic image synthesis. Moreover, advancements in deep learning architectures like transformer models and self-supervised learning have given further boost and generalizability. However, prospective issues like impractical computation needs, limited data availability, and issues of fairness and privacy are still not beyond the state of the art. However, following the propositions of this paper and considering recent literature, it is possible to discuss the major trends and potential developments in this area further to cover these challenges. Accordingly, while acknowledging the theoretical developments of deep learning and learning thereof from applications practiced in the literature, this paper outlines the role of deep learning in enhancing image processing technologies.

1 INTRODUCTION

Deep learning has been helpful in areas such as medical image processing, self-driving automobiles, security surveillance, and industrial robotics (Boopathi, Pandey, & Pandey, 2023). Previous approaches to computer vision rely on the programmer extracting features, which hampers generalization across different datasets and cases. This field has recently been transformed by the CNN learning models, intense learning, which allows models to learn a hierarchical representation of the image data (GeeksforGeeks, 2020).

Computer vision using deep learning methods is on an entirely different level than earlier machine learning techniques. In their paper, Tsirtsakis et al. assessed that deep learning recognition surpassed other image feature engineering techniques regarding precision and speed and produced better performances in automated vision-related tasks. Therefore, deep learning has been widely applied in many applications, and it is now one of the essential techniques for solving many related vision tasks (Tsirtsakis et al., 2025).

Deep learning has been one of the most significant advancements, especially in image processing, because of its capability to handle big data. Deep neural networks can learn rich patterns from massive datasets, which allow them to accurately detect, segment, and classify objects (Noor & Ige, 2025). ResNet, DenseNet, and Vision Transformers (ViT) architectures have emerged, which improve model efficiency and scalability to tackle existing issues, like vanishing gradients and inefficient computations (Yan et al., 2025).

The influence of deep learning goes beyond research to practical applications. It has brought early disease detection in healthcare with AI-driven radiology and pathology analysis with improved patient outcomes (Pinto-Coelho, 2023). Deep learning models play a role in real-time perception systems, which power such driving in autonomous driving, enabling safe navigation in dynamic environments (Gupta et al., 2021). In addition,

[a] https://orcid.org/0000-0000-0000-0000

Wang, X.
Applications and Developments of Deep Learning in Image Processing.
DOI: 10.5220/0013687900004670

In *Proceedings of the 2nd International Conference on Data Science and Engineering (ICDSE 2025)*, pages 277-284
ISBN: 978-989-758-765-8

industries have utilized AI-powered quality control systems to increase production efficiency and reduce defects (Finio & Downie, 2024).

However, deep learning in image processing is not without its challenges. Some research areas are still to be focused on, like high computational costs, the requirement of large labeled datasets, and issues related to model interpretability. Solving these problems will be critical to realizing more broadly accessible and reliable deep learning systems. Future research will seek to increase model efficiency, reduce dependency on big datasets, and enhance explainability to promote broader adoption in industrial scenarios.

2 DEEP LEARNING ARCHITECTURES FOR IMAGE RECOGNITION

The convolutional neural networks are undoubtedly the main ones responsible for the rise of image recognition because of the significant improvements in accuracy, speed, and scalability that deep learning architectures provide. Some effective architectures that have revolutionized the field include CNN, ResNet, ViT, and most attention-based models. They solve important issues such as feature extraction, vanishing gradients, and Long Short Term Memory for computer vision, making them inevitable in today's AI solutions.

2.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have been employed chiefly for depth image recognition learning using convolutional layers for spatial feature extraction, pooling layers for feature size reduction, and fully connected layers for classification (Zhao et al., 2024). This inevitably makes CNNs capable of extracting low and high-level features from the images for better accuracy and robustness.

Yann LeCun also created LeNet-5 for digit recognition, wanting to showcase that convolution layers were suitable for feature extraction, as presented in the Convolutional Neural Network (CNN) Architectures (GeeksforGeeks, 2023). These features allow AlexNet to be a ReLU activation, data augmentation, and GPU acceleration to make deep networks easier to process large-scale datasets of images (Li et al., 2021).

Henceforth, some more progression occurred, including the VGGNet created by researchers at Oxford University in 2014. Small 3x3 convolutional kernels and deeply layered structures corresponded to another benefit of VGGNet, where it was observed that while accuracy and efficiency decreased, computation power was reduced immensely (Boesch, 2024). The improvements in CNN architectures have contributed significantly to enhancing image recognition. They are indispensable for medical imaging, computer vision, robotic vision, and Face ID.

2.2 Residual Networks (ResNet)

Deep learning models have also become deep and have experienced problems such as the vanishing/exploding gradient problem, which made training difficult. To overcome this, Residual Networks (ResNet) employ skip connections or residual blocks to keep the gradient flow efficient even with intense networks (Mahaur et al., 2022). Their deep architecture led to ResNet-50 and ResNet-101 models performing state-of-the-art accuracy on large-scale image classification benchmarks.

On ResNet's foundation, DenseNet was created in 2017 and improved the information flow by connecting every layer to all previous layers, hence reducing the redundancy and improving the gradient propagation (Zhou et al., 2022). This denser connectivity resulted in higher efficiency, facilitating deep networks to leverage fewer parameters efficiently and achieve good performance. DenseNet's feature reuse mechanism is proven highly effective in medical imaging applications where deep feature extraction is essential (Wang & Zhang, 2020).

2.3 Vision Transformer (ViT)

However, traditional CNNs rely on local receptive fields, restricting their ability to learn the global dependency within images. The Vision Transformer (ViT) brought the self-attention mechanism from natural language understanding (NLP) to image understanding, which allows models to learn long-range dependencies in visual data (Wang et al., 2025). Unlike CNNs, where convolutional filters are applied for spatial regions, ViT segments images into patches and operates on them using self-attention mechanisms.

Rein Bugnot noted that ViT models outperform CNNs on large datasets with pretraining on extensive labeled data (Bugnot, 2025). Unfortunately, ViTs have a significant computational cost, making them hard to deploy in resource-constrained environments. Researchers are investigating hybrid models combining CNNs for feature extraction with transformer-based self-attention to balance efficiency and performance.

2.4 Attention-Based Networks

Modern deep-learning architectures incorporate attention mechanisms that allow models to focus on key image features while ignoring distractions. Another primary attention model is the Squeeze-and-Excitation (SE) Network (Ghaffarian et al., 2021), which injects a channel-wise attention mechanism to recalibrate features dynamically. SE modules improve classification accuracy by applying global pooling followed by fully connected layers that enhance important features and suppress less relevant ones.

The Non-local Network is another robust attention-based architecture that extends self-attention to model long-range image dependencies. However, unlike CNNs, non-local networks compute relationships between all pixels of the image using a global receptive field and have been proven more effective in video analysis and action recognition tasks. Applications of these models have seen a marked improvement in security surveillance, medical imaging, and real-time image segmentation.

Deep learning architectures have been game changers for solving crucial problems with feature extraction, gradient optimization, and global dependency modeling in the image recognition domain. Although ResNet and DenseNet have allowed us to build more profound and efficient networks, CNN remains a fundamental image-processing tool. Self-attention mechanisms in Vision Transformers help model long-range dependency, and attention-based networks improve feature selection and interpretability. With continued research in deep learning, future advancements will probably come from hybrid architectures that fuse the advantages of CNNs and transformers while maximizing efficiency in computation. These advancements will motivate further innovation in AI-powered image recognition for applications ranging from healthcare to autonomous systems and industrial automation.

3 TRAINING OPTIMIZATION STRATEGIES

Deep learning optimization is critical to improving accuracy, overcoming overfitting, and increasing computational efficiency. Various training techniques have been created to improve model performance, such as data augmentation, transfer learning, hyperparameter optimization, regularization, and self-supervised learning.

3.1 Data Augmentation

Artificial expansion of training datasets through data augmentation is commonly used to improve model generalization. It is essential in applications where data collection is complex, like medical imaging. For example, image transformations in radiology have allowed models to detect diseases in the same image under different conditions (Islam et al., 2024). Moreover, other sophisticated techniques like GANs have further enriched the dataset, thus creating synthetic examples required for training. According to Chen et al., it has been proved that GAN-based augmentation performs much better for medical imaging when the labeled data are limited in the datasets (Chen et al., 2022).

Another functional augmentation approach is Mixup. It combines two images and their associated labels to generate new training samples and improve robustness (Yang & Xiang, 2024). Augmentation methods improve the recognition of faces with different lighting and angles and help ensure the reliability of deep learning models in real-world applications.

3.2 Transfer Learning

Transfer learning employs pre-trained deep learning models on massive datasets to solve other tasks better. It significantly reduces the training time by enabling models to use the previous features learned, making it ideal for applications with small labeled data (Transfer Learning, 2022). Popular models, including ResNet, VGG, and Vision Transformers (ViT), have been widely adopted for transfer learning in medical imaging and autonomous driving.

Already, CNNs pre-trained on ImageNet have been utilized successfully for X-ray and MRI classification tasks in medical imaging, mitigating the need for large numbers of labeled datasets (Li et al., 2023). Fine-tuning pre-trained models for domain-specific tasks effectively achieves high diagnostic accuracy and efficiency. Transfer learning has enabled the rapid deployment of AI-based defect detection systems in industrial automation.

Cross-modal transfer learning is another growing approach, where models developed on images are adapted for video analysis. Crossmodal transfer learning has furthered real-time video analytics for surveillance systems, teaching AI based on still images for moving frames.

3.3 Hyperparameter Optimization

Hyperparameter tuning is an important part of deep learning model optimization. The 'right' learning rate, batch size, and optimizer can significantly affect a model's performance and convergence speed (Raiaan et al., 2024). Standard hyperparameter tuning methods are Grid Search, Random Search, and Bayesian Optimization. AutoML techniques have induced popularity, automating hyperparameter selection to increase model performance.

Another crucial factor in optimization is learning rate scheduling. Adaptive learning rate strategies, such as cyclical learning rates and warm restarts, can accelerate training convergence and improve the final model accuracy. Furthermore, we also see how batch normalization works to standardize activations during training and how this technique helps stabilize optimization and reduce sensitivity to initialization. Also, the "use of hyperparameter optimization in deep reinforcement learning helped to enhance autonomous vehicle navigation by developing more effective decision-making policies.

3.4 Regularization Techniques

Regularization techniques are harnessed to prevent overfitting and improve model generalization. L2 regularization (weight decay) is a common technique in penalizing large weight values to prevent overfitting. Other effective regularization techniques include dropout, which acts by randomly deactivating neurons during training so that the model develops redundant representations for increased robustness.

In residual networks, a newer approach called Stochastic Depth allows models to skip layers during training, thereby cutting costs and maintaining accuracy. Stochastic depth has shown great promise in facilitating deep networks to be efficient for resource-constrained environments such as mobile devices. According to Ioffe & Szegedy, Batch Normalization normalizes the inputs to each layer and allows for faster training and stabler deep networks (Ioffe & Szegedy, 2015).

3.5 Self-Supervised Learning

Deep learning models can learn from unlabeled data with the help of pretext tasks through self-supervised learning (SSL). SSL methods like contrastive learning and predictive modeling have shown tremendous success in reducing the reliance on massive labeled datasets. Contrastive learning-based approaches like SimCLR and MoCo have been shown to outperform supervised learning on many vision tasks with self-supervised feature learning.

In medical imaging, self-supervised learning played a crucial role in training models with fewer labeled medical scans, thus improving the diagnostic resolution. SSL techniques such as jigsaw puzzle solving and rotation prediction have enhanced feature extraction functionality, enabling deep learning models to do more with less data.

SSL has been applied in industrial environments, such as predictive maintenance systems, where deep learning models can analyze sensor data to explore equipment's early failure symptoms. Self-supervised learning-based anomaly detection in industrial automation has dramatically lowered unexpected downtimes, resulting in significant cost savings.

The role of training optimization strategies in improving deep learning models for image processing is crucial. Data augmentation, transfer learning, hyperparameter tuning, and regularization all boost model performance, while self‑supervised learning eliminates the need for labeled data. With the ongoing advancements in deep learning, future optimization methods will also lead to even more significant improvements in computational efficiency and model generalization. Future research will likely focus on the AutoML hyperparameter selection, expanding on SSL techniques, and even developing energy-efficient training methods—the following section details real-world use cases and challenges of implementing deep learning models for image processing.

4 APPLICATIONS AND CHALLENGES

Deep learning and its state-of-the-art solutions have revolutionized image-processing tasks. This section explore key application areas of deep learning in medical imaging, autonomous driving, security surveillance, and industrial automation, as well as the challenges, including data requirements, computational costs, and model interpretability.

4.1 Applications

4.1.1 Medical Imaging

The advancements in medical imaging via deep learning have dramatically enhanced diagnostic accuracy and speed. Deep learning models are capable of automatically detecting tumors while reducing false favorable rates and promoting early diagnosis. Convolutional Neural Networks (CNNs) and Vision Transformers are increasingly used in cancer cell detection, disease classification, and medical image segmentation (Jiang et al., 2023). Based on these models, it has been possible to design models that will assist in analyzing MRI and CT scans for tumor detection and also help radiologists diagnose diseases with much precision.

Deep learning-powered automated systems are helpful in enhancing the sensitivity for detecting any abnormalities in radiology images and are not a burden for radiologists. The well-known CNN architecture that is widely used is U-Net, which has significantly improved the accuracy of the organ and lesions' segmentation. This means that with these advancements, one would be able to diagnose faster and more effectively.

Federated learning is gradually becoming one of the critical instruments in medical AI. Several deep learning models are trained across multiple institutions across datasets without patients' information being transferred across institutions. As explained, this method makes deep learning-based radiology models more generalizable to different patient populations while also considering ethics and privacy issues. In medical AI, federated learning helps improve the identification of disease and diagnosis without compromising data protection.

The utilization of Deep Learning in Medical Imaging is set to expand as the field develops with an expectation of producing advanced and accurate diagnostic techniques. This could also pave the way for further expansion of the use of AI in live clinical decision support systems to improve the metrics on such areas from the viewpoint of medical professionals and patients. Deep learning and federated learning are two approaches to enhancing medical imaging and are very promising because they improve disease detection without violating patients' privacy and ethical issues.

4.1.2 Autonomous Driving

Autonomous driving uses deep learning extensively to achieve real-time object recognition, lane detection, and collision evasion pedestrian, Vehicle, and Traffic sign recognition models using CNNs and LiDAR-based deep learning. Tesla and Waymo use deep learning-driven vision systems to interpret complex driving environments. One popular deep learning model for object detection, YOLO (You Only Look Once), has significantly advanced real-time object detection, permitting autonomous vehicles to decide instantly while driving in real-time (Gheorghe et al., 2024).

In addition to object detection, deep learning models based on reinforcement learning allow self-driving systems to make real-time adaptive driving decisions. Deep networks processing raw sensor inputs to predict vehicle controls are explored to enhance decision-making capabilities using end-to-end learning approaches. One challenge in ensuring deep learning-based driving models' safety and integrity is rigorously testing them under various environmental conditions.

4.1.3 Security and Surveillance

Facial recognition, behavior analysis, and threat detection have been built based on deep learning, which enhances security and surveillance. Governments and private organizations track public spaces using AI-powered surveillance systems to detect suspicious actions. FaceNet is a deep-learning model that performs accurate facial recognition for identity verification. For example, recurrent neural networks (RNNs) are used to analyze behavior and detect anomalies in surveillance footage, preventing real-time security threats.

Through the use of deep learning, it is reasonably possible to process hundreds of video streams simultaneously and detect the presence of unusual activities and potential threats. Nevertheless, ethical issues concerning mass surveillance and privacy violations call for the development of fairness-aware AI models that aim to address particular biases in facial recognition systems (Shabbir et al., 2024).

4.1.4 Industrial Automation

Some utilization areas of deep learning include manufacturing and quality control for defect detection, product classification, and preventive maintenance. CNNs and transformer models enable effective analysis of product images on assembly lines to identify defects, reducing reliance on manual inspection (Thomas et al., 2023). Using artificial intelligent controlled systems decreases the prospect of errors and increases production capacity. Specifically, it encourages robotic automation because industrial robots are capable of analyzing data and making precise time decisions about a particular manufacturing process.

Through the application of deep learning in the industrial automation processes, the number of produced defects has been reduced by 30 percent, generally enhancing predictive maintenance. Some of the loads from digitized manufacturing include predictive maintenance for equipment supported by advanced IT, such as artificial intelligence, in that it is used to identify possible machinery breakdowns before they occur, hence avoiding downtime and costly maintenance. Thus, wear and tear may be predicted based on the data from sensors and observations of the machine's previous performance to avoid potential breakdown.

In particular, deep learning provides quality control through real-time monitoring and anomaly detection. Deep learning leads to advanced vision systems capable of detecting minute defects in products that a human inspector may miss, resulting in higher product reliability. AI-driven automation also facilitates supply chain management by optimizing inventory tracking, demand forecasting, and logistics.

However, as deep learning remains a growing field, the applications in the manufacturing industry will grow along with it, resulting in increasingly intelligent and adaptive production systems. The future could include self-learning robots integrated with Process adjustments based on real-time feedback and more precise defect detection models. Integrating deep learning with industrial automation allows manufacturers to be more efficient, reliable, and cost-effective, making industrial production more innovative and resilient in an increasingly competitive market.

4.2 Challenges

4.2.1 Large-Scale Data Requirements

One major issue with deep learning in image processing is that in order to achieve good results, large amounts of labeled training data are required. According to (Khan et al., 2022), data annotation costs remain a significant bottleneck, especially in domains that require expert labeling, such as medical imaging and autonomous driving. Millions of labeled images are required by supervised learning approaches to achieve high accuracy. However, assembling and annotating such large datasets is tedious and costly. In order to address this problem, self-suppled learning and data-efficient training techniques are researched to limit data dependency

4.2.2 High Computational Costs

CNNs and transformers, both deep learning models, require a lot of computational resources. Training deep networks is expensive in computation and infrastructure (cost of GPUs, TPUs). Additionally, efficient real-time models that can process images for applications such as autonomous driving or surveillance are needed. Pruning and quantization techniques have been developed to reduce model size while keeping accuracy intact so that the model can be deployed at the edge devices. These techniques are being investigated to optimize deep learning architectures for deployment on low-power hardware.

4.2.3 Model Interpretability Issues

Although deep learning models are very accurate, they are often considered a black box because it can be challenging to understand their decision-making process. In medical imaging or autonomous driving, understanding how a model arrives at its conclusion is critical for reliability and trust (Boopathi, Pandey, & Pandey, 2023): model transparency and interpretability increase by developing explainable AI (XAI) methods like Grad-CAM and SHAP. AI techniques are crucial in high-stakes scenarios where profound learning decisions are interpretable by human experts.

5 CONCLUSIONS

Deep learning has revolutionized image processing, improving accuracy, efficiency, and automation. Deeper learning models have reached new heights in medical imaging, eliminating diagnosis errors and extending early detection capabilities." As a result, AI-powered diagnostic tools that help radiologists, improve pathology analysis, and allow predictive solutions in healthcare have become available. This also impacts autonomous driving, where object recognition and real-time decision-making rely heavily on deep learning. Deep learning models such as YOLO have significantly improved vehicle perception, making self-driving systems safer and more responsive.

In future research, there is likely a push for more efficient, interpretable, and accessible deep-learning models for image processing. Self-supervised and federated learning are key solutions for reducing the reliance on labeled data so that AI can be scaled to various applications." Further, adversarial robustness to these attacks and vulnerabilities will require significant efforts to protect AI systems safely and robustly in cybersecurity and fraud detection.

Finally, deep learning has altered the possibilities of image processing and has the potential to revolutionize applications in many industries. However, these methods pose difficulties regarding computational costs, data needs, and model interpretability, which must be overcome for wider adoption and longevity. Deep learning will remain one of the cornerstones of artificial intelligence by further advancing model efficiency, ethical AI practices, and interpretability, delivering more reliable and accessible solutions for practical applications.

REFERENCES

Boesch, G. 2024. Very Deep Convolutional Networks (VGG) Essential Guide. viso.ai. Retrieved from https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks.

Boopathi, S., Pandey, B. K., & Pandey, D. 2023. Advances in artificial intelligence for image processing. In Advances in computational intelligence and robotics book series (pp. 73–95).

Bugnot, R. 2025. Vision Transformers (ViT) explained: Are they better than CNNs? Towards Data Science. Retrieved from https://towardsdatascience.com/vision-transformers-vit-explained-are-they-better-than-cnns/

Chen, Y., et al. 2022. Generative adversarial networks in medical image augmentation: A review. Computers in Biology and Medicine, 144, 105382.

Finio, M., & Downie, A. 2024. AI in manufacturing. IBM.com. Retrieved from https://www.ibm.com/think/topics/ai-in-manufacturing

GeeksforGeeks. 2020. Residual Networks (ResNet) - Deep Learning. GeeksforGeeks. Retrieved from https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/

GeeksforGeeks. 2023. Convolutional Neural Network (CNN) architectures. GeeksforGeeks. Retrieved from https://www.geeksforgeeks.org/convolutional-neural-network-cnn-architectures/

Ghaffarian, S., Valente, J., Van Der Voort, M., & Tekinerdogan, B. 2021. Effect of attention mechanism in deep learning-based remote sensing image processing: A systematic literature review. Remote Sensing, 13(15), 2965.

Gheorghe, C., Duguleana, M., Boboc, R. G., & Postelnicu, C. C. 2024. Analyzing real-time object detection with YOLO algorithm in automotive applications: A review. Computer Modeling in Engineering & Sciences, 141(3), 1939–1981.

Gupta, A., Anpalagan, A., Guan, L., & Khwaja, A. S. 2021. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. Array, 10, 100057.

Ioffe, S., & Szegedy, C. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv.org.

Islam, T., Hafiz, M. S., Jim, J. R., Kabir, M. M., & Mridha, M. F. 2024. A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions. Healthcare Analytics, 5, 100340.

Jiang, X., Hu, Z., Wang, S., & Zhang, Y. 2023. Deep learning for medical image-based cancer diagnosis. Cancers, 15(14), 3608.

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. 2022. Transformers in vision: A survey. ACM Computing Surveys, 54(10s), 1–41.

Li, M., Jiang, Y., Zhang, Y., & Zhu, H. 2023. Medical image analysis using deep learning algorithms. Frontiers in Public Health, 11.

Li, S., Wang, L., Li, J., & Yao, Y. 2021. Image classification algorithm based on improved AlexNet. Journal of Physics Conference Series, 1813(1), 012051.

Mahaur, B., Mishra, K. K., & Singh, N. 2022. Improved residual network based on norm-preservation for visual recognition. Neural Networks, 157, 305–322.

Mohd Noor, M. H., & Ige, A. O. 2025. A survey on state-of-the-art deep learning applications and challenges.

Pinto-Coelho, L. 2023. How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications. Bioengineering, 10(12), 1435.

Raiaan, M. A. K., et al. 2024. A systematic review of hyperparameter optimization techniques in convolutional neural networks. Decision Analytics Journal, 11, 100470.

Shabbir, A., Arshad, N., Rahman, S., Sayem, M. A., & Chowdhury, F. 2024. Analyzing surveillance videos in real-time using AI-powered deep learning techniques. International Journal of Research in Information Technology and Computer Communication. Retrieved from http://www.ijritcc.org

Thomas, J. B., Chaudhari, S. G., K. V., S., & Verma, N. K. 2023. CNN-based transformer model for fault detection in power system networks. IEEE Transactions on Instrumentation and Measurement, 72, 1–10.

Transfer Learning. 2022. Transfer learning - an overview. ScienceDirect Topics. Retrieved from https://www.sciencedirect.com/topics/computer-science/transfer-learning

Tsirtsakis, P., Zacharis, G., Maraslidis, G. S., & Fragulis, G. F. 2025. Deep learning for object recognition: A comprehensive review of models and algorithms. International Journal of Cognitive Computing in Engineering.

Wang, S.-H., & Zhang, Y.-D. 2020. DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification.

ACM Transactions on Multimedia Computing Communications and Applications, 16(2s), 1–19.
Wang, Y., Deng, Y., Zheng, Y., Chattopadhyay, P., & Wang, L. 2025. Vision Transformers for image classification: A comparative survey. Technologies, 13(1), 32.
Yan, H., Mubonanyikuzo, V., Komolafe, T. E., Zhou, L., Wu, T., & Wang, N. 2025. Hybrid-RViT: Hybridizing ResNet-50 and Vision Transformer for enhanced Alzheimer's disease detection. PLoS ONE, 20(2), e0318998.
Yang, L., & Xiang, Y. 2024. AMPLIFY: Attention-based mixup for performance improvement and label smoothing in transformer. PeerJ Computer Science, 10, e2011.
Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. 2024. A review of convolutional neural networks in computer vision. Artificial Intelligence Review, 57(4).
Zhou, T., Ye, X., Lu, H., Zheng, X., Qiu, S., & Liu, Y. 2022. Dense convolutional network and its application in medical image analysis. BioMed Research International, 2022, 1–22.