

Survey research on the new generation of deep learning in image processing

Tran Duy Tung

Abstract: This study investigates the capabilities and performance of next-generation deep learning architectures in the field of image processing. As traditional convolutional neural networks face limitations in capturing long-range dependencies and global context, emerging models such as Vision Transformers, MLP-Mixers, and state-space models offer promising alternatives. This research conducts a systematic survey and comparative analysis of these novel approaches, evaluating their effectiveness across core tasks including image classification, object detection, and semantic segmentation. Key performance metrics such as accuracy, computational efficiency, and scalability are assessed using standard benchmark datasets. The findings reveal that while next-generation models often achieve superior accuracy, particularly in tasks requiring global reasoning, they also present challenges related to data requirements and computational cost. The study further explores hybrid architectures that integrate the strengths of both convolutional and attention-based mechanisms. This survey provides a comprehensive overview of the current landscape, highlighting the transformative potential of these advanced deep learning methods for future image processing applications.

Keywords: Deep learning, image processing, computer vision, convolutional neural networks