4The text compares traditional computer vision techniques with deep learning. It explains that deep learning is a subset of machine learning based on artificial neural networks that can achieve more accurate results in tasks such as image classification, semantic segmentation, object detection, and Simultaneous Localization and Mapping (SLAM). Deep learning models are trained rather than programmed, and they can be re-trained for any use case. The text also explains that the traditional approach to computer vision involves feature extraction, which can be cumbersome and requires expert analysis and fine-tuning. Deep learning introduced the concept of end-to-end learning where the machine is given a dataset of images annotated with what classes of object are present in each image, and the neural networks automatically discover the underlying patterns in classes of images and work out the most descriptive and salient features. The development of convolutional neural networks (CNNs) has had a tremendous influence on the field of computer vision in recent years, and it has enabled an increase in the ability to recognize objects.

The article discusses the challenges in traditional computer vision and the benefits of using a hybrid approach that combines traditional CV algorithms with deep learning (DL) for better performance. The article highlights the challenges of DL, which require large amounts of computing resources for training and inference, and the importance of image resolution for achieving adequate performance in object classification. It also discusses the benefits of making the best use of edge computing for reducing latency, costs, cloud storage, processing requirements, and bandwidth requirements. The article suggests that using hybrid or composite approaches involving conventional CV and DL can take great advantage of the heterogeneous computing capabilities available at the edge.

The text discusses various topics related to computer vision, such as Visual SLAM, image-based localization, 360 cameras, panorama stitching, and dataset annotation and augmentation. Visual SLAM is a method that uses a vision system for registering landmarks in a scene and has advantages over LiDAR. Image-based localization involves place recognition and pose estimation, and the success of place recognition is largely attributed to the ability to extract image feature descriptors. 360 cameras have a 360-degree field of view and are important in applications such as robotics, but have challenges such as barrel distortion. Panorama stitching is a research problem in this area, and methods such as direct image alignment and correlation-based image alignment have been investigated. Finally, there is a debate about the combination of CV and DL and whether we should re-evaluate our methods from rule-based to data-driven.