Literature review on “How computer vision along with Image compression techniques are integrated for better performance and power saving in detection helping visually impaired people”

In this part of research we are talking about some researches which were made in this field which is computer vision (CV) to help visually impaired people and our focus is on how to use prefiltering algebraic matrix methods before bypassing to heavy CV model to ignore normal pre-given data (e.g. empty streets, patient’s home, etc) which the user can easily deal in this environment to save the heavy power of neural network AI models to be more accessible in wearable devices like smart glasses. This techniques depends on compressing a give images data and extracting the main information in it using matrices, so that a new image data can first be passed on this matrix and check its novelty, if found novel data, in this can it will be passed to heavy CV models to detect items in it and alert him, otherwise, ignore the normal data and do not pass to the easy CV model, and the normal data here is the environments where the person can deal easily without the help of CV model (e.g. his home). We divided our literature into three parts 1-what is the attempts of using CV at the general. 2-Compression techniques focusing on SVD. 3-what are the attempts of integrating compression techniques and CV to be more efficient and power saving.

1- Computer Vision Based Assistive Technology for Blind and Visually Impaired People

Assistive technologies for visually impaired individuals have advanced significantly in recent years, evolving from traditional tools such as the white cane, which offers only limited tactile feedback, to intelligent AI-based systems that enhance mobility and independence. Modern smart canes and wearable devices now integrate computer vision and machine learning to detect obstacles, understand surroundings, and guide users in real time. These systems typically employ Convolutional Neural Networks (CNNs) and object-detection models such as YOLO, along with Simultaneous Localization and Mapping (SLAM) techniques that fuse data from RGB-D cameras, LiDAR, and IMUs to construct spatial maps for navigation. Despite these advancements, challenges remain—particularly high power consumption and computational costs, which increase device size, weight, and price. Environmental sensitivity, poor indoor performance without GPS, and cognitive overload from complex feedback systems also limit usability. Therefore, there is a growing need for lightweight, efficient computer vision approaches to achieve affordable, real-time assistive devices suitable for widespread use [1].

A recent research has focused on enhancing obstacle detection systems to improve navigation safety for visually impaired users. Traditional assistive vision technologies, relying on ultrasonic sensors or conventional cameras, often fail in complex, crowded, or low-light environments. To address this limitation, a novel Bilateral Vision-Aided Transformer Neural Network (BViT-CKNN) has been introduced, combining bilateral filtering, Vision Transformers (ViT), and Convolution Kernel Neural Networks (CKNN). Bilateral filtering effectively removes noise while preserving image edges, the Vision Transformer provides global scene understanding by analyzing image patches, and the CKNN captures local features to detect small or partially obscured obstacles. Tested on the COCO dataset, the model achieved 93% precision, 91% recall, and a 92% F1-score, outperforming conventional architectures such as YOLOv5 and SSD MobileNetV2 in both accuracy and speed. This hybrid architecture demonstrates the potential of combining transformer-based global perception with detailed local analysis to achieve more reliable, real-time assistive vision systems for the visually impaired [2].

2- Fundamental SVD-Based Compression Techniques

Kahu and Rahate [3] presented a comprehensive analysis of image compression using SVD, exploring the fundamental mathematics and practical applications of this technique. Their work demonstrated that SVD expresses image data in terms of eigenvalues and eigenvectors, effectively exploiting psychovisual redundancies in images. The authors emphasized that an image matrix can be decomposed into three component matrices (U, Σ, and V), where Σ is a diagonal matrix containing singular values arranged in descending order. Their experimental results showed that the first singular value contains the greatest amount of information, with subsequent values containing progressively less image information. The study established that compression is achieved by retaining only the first k singular values while discarding lower-order values that contain negligible information. Using a 392×392 image, they demonstrated that when k equals the rank of the image matrix, the reconstructed image achieves a Mean Square Error (MSE) of -79.5 dB and Peak Signal-to-Noise Ratio (PSNR) of 240.77 dB, nearly identical to the original. However, they noted the critical tradeoff: smaller values of k yield higher compression ratios (up to 50.52 for k=5) but result in significant quality degradation (MSE of 35.192 dB).

Swathi et al. [6] further explored SVD-based compression with a focus on color image processing. Their approach involved transforming RGB images into the YCbCr color space before applying SVD, recognizing that RGB representation contains redundant information that is suboptimal for compression. By decomposing the Y (luminance), Cb, and Cr (chrominance) components separately and applying frequency domain thresholding to U and V matrices, they achieved improved compression while maintaining visual quality. Their implementation in R programming demonstrated compression ratios varying from 1.09 to 174.55 depending on the value of k, with image sizes ranging from 106.6 KB to 215.8 KB for a 700 KB original image.

Quality Assessment Metrics for SVD Compression

A significant challenge identified in the literature is the inadequacy of traditional quality metrics for SVD-compressed images. Razafindradina et al. [4] critically analyzed this limitation, noting that conventional metrics such as PSNR and Structural Similarity Index (SSIM) are not suitable for evaluating SVD compression performance. They proposed a novel quality metric based on energy ratio, calculated as the ratio between the energy of the compressed image and the original image. Their comprehensive testing on 512×512 pixel images revealed three distinct quality zones: (1) poor quality with 99% to 99.85% energy restoration (k=8 to 32, PSNR 27-34 dB), (2) good quality with 99.9% to 99.98% energy restoration (k=40 to 120, PSNR 35-42 dB), and (3) very good quality with 99.99% to 100% energy restoration (k≥128, PSNR 43-98 dB). The authors demonstrated that for rank k=40, corresponding to SSIM=0.94 or PSNR=35 dB, approximately 99.9% of the image energy is restored. This energy-based metric provides higher precision compared to PSNR and SSIM, with the difference between minimum and maximum energy ratios being less than 0.1.

- Hybrid Compression Approaches

Recognizing the limitations of standalone SVD compression, researchers have explored hybrid techniques combining SVD with other transform methods. Prabakar Joshua et al. [5] proposed a novel approach integrating SVD with Discrete Wavelet Transform (DWT) to leverage the complementary strengths of both techniques. Their work addressed a fundamental limitation: while SVD offers excellent PSNR values, it typically achieves low compression ratios, whereas DWT provides high compression ratios. The proposed cascading approach first compresses images using SVD, then applies DWT compression to the reconstructed image. Testing on standard images (Lena, Peppers, Boats, Baboon, and Barbara) demonstrated superior performance compared to individual techniques, with compression ratios reaching 48.21 for Lena and 50.02 for Peppers. The overall compression percentage was calculated as the product of SVD-based compression and DWT-based compression, effectively multiplying the benefits of both approaches.

- Comparative Analysis and Performance Considerations

The literature reveals several consistent patterns regarding SVD compression performance. All studies confirm that the choice of k (number of retained singular values) is critical and application-dependent [3], [4], [5], [6]. The theoretical upper limit for k is established by the condition $m \times n > k \times (m+n+1)$, where m and n are the image dimensions [1]. When storage space is prioritized over quality, lower k values are selected; conversely, when image fidelity is paramount, higher k values are necessary.

Kahu and Rahate [3] concluded that optimal compression results are obtained when MSE is less than or equal to 30 dB, corresponding to k=128 in their experiments. This finding contrasts somewhat with Razafindradina et al. [4], who suggested that PSNR values of 30 dB, while typical for other compression techniques, represent suboptimal quality for SVD compression, recommending 35 dB as a more appropriate threshold.

- **Advantages and Limitations**

The literature consistently identifies several advantages of SVD-based compression: (1) lower computational complexity compared to other compression techniques [2], (3); (2) applicability to arbitrary, square, reversible, and non-reversible matrices of any m×n size [3], [6]; (3) inherent energy compaction properties that concentrate information in initial singular values [4]; and (4) versatility beyond compression, including applications in noise reduction, face recognition, and watermarking [28].

However, limitations are also acknowledged. SVD compression alone typically achieves modest compression ratios unless quality is significantly sacrificed [3], [5]. The sequential nature of SVD computation presents scalability challenges, though this can be addressed through parallelization strategies [6]. Additionally, the irreversible nature of the compression makes it a lossy technique, which may be unsuitable for applications requiring perfect reconstruction [5]..

3- Attempts of integrating compression techniques and Computer Vision (CV) model

There are many papers use different techniques for prefiltering the input images before bypassing to a heavily resource and power consuming CV models.

"*Filtering Empty Video Frames for Efficient Real-Time Object Detection*"[7] While this paper uses a hybrid time series analysis method called L-filter instead of SVD, its methodology and evaluation are highly relevant to our research. The L-filter method significantly enhances the efficiency and scalability of real-time object detection in computer vision (CV) systems by functioning as a lightweight, L-filter unit predicts and drops empty video frames. The pros of L-filter include dramatically improved performance metrices, it boosts the frame processing rate of models like YOLOv5, SSD, and EfficientDet by 31-47% compared to using them standalone. Although this is slightly different from our research using SVD, this emphasis the powerfulness of using prefiltering techniques before heavy CV models.

"SVD-GAN for Real-Time Unsupervised Video Anomaly Detection"[8]. In this paper the researchers propose a lightweight generative-adversarial network (GAN) architecture for real-time, unsupervised video anomaly detection. By using Singular Matrix Decomposition (SVD) they are able to train their model by unsupervised way, which means no labels for training; as SVD detection principle is binary (Normal or Novel), so SVD can do that by taking only the data for the required images or video frames normal state and then any different state in the images or video frames that diverge significantly (by calculating reconstruction error) will be considered anomal. Although the method is lightweight in terms of GAN standards, it is still a generative neural network with many parameters, which might be heavier than a purely linear SVD-based filter. Despite its

efficiency, the approach still requires training a dual-network GAN, which can be computationally demanding and sensitive to hyperparameter tuning.

"Novelty Detection Via Blurring"[9] This paper introduces SVD-RND, a novel and efficient Out-of-Distribution (OOD) detector designed to address the critical vulnerability of conventional novelty detection schemes, such as variational autoencoders (VAE) and Random Network Distillation (RND), which tend to assign high confidence or low uncertainty to OOD data and are particularly susceptible to blurred images. SVD-RND is based on the RND framework but utilizes self-generated blurred images as adversarial examples during training to prevent the model from overfitting to low resolution. The blurring process employs Singular Value Decomposition (SVD), where the bottom non-zero singular values of each data channel are zeroed out to construct the blurred image. The model trains a predictor network to discriminate between the original data and these low-rank blurred datasets. Empirically, SVD-RND achieves significant performance gains, dramatically improving detection accuracy (e.g., from approximately 50% to over 90%) in difficult OOD domains like CIFAR-10 : SVHN, surpassing conventional baselines, and is shown to learn a better representation of the target distribution. Although this approache empowers the SVD to be more accurate detecting small anomalities, our approach of ordinary SVD make it faster and serves our main goal of using the least possible power to make the AI model applicable for small and wearable devices, on the other hand the paper approach has a powerful point of blur images detection.

Research Gap what we will try to work on

Through our literature we discovered that there are many attempts trying to make the object detection (OB) for helping visually impaired people more efficient and accurate, but still the main obstacle for these approaches is performance in terms of computational power and battery, although techniques for compression like SVD, L-filter and other are used to enhance efficiency, they are still embedded within the heavy CV models which still consume high power hindering it from being widely adopted and applied for wearable devices with limitations in battery and computational power. So the whole idea is a trade-off and balancing between accuracy, power consuming and computational power limitations. From our literature we found there is no personal environment adaptation for each user approach adopted to enhance this point, so our idea of research is to make a gate-keeper-like technique using prefiltering techniques used, the idea is not only compressing the passed data, moreover it is about which data is to pass. Our idea depends on that the blind people always can deal easily with their familiar and day-to-day environments like their homes, their street, empty corridors in universities and schools, low-crowd walking streets, etc. So why not to use a pure mathematical technique like Singular value decomposition (SVD) which is able to take image data, compress it in a matrix and then it will be able to algebraically check whether a new given image is novel from the pregiven data (for the patient familiar environment) or not, if novel only in this case the image will be passed to heavy CV models, as the SVD alone can detect whether it is novel or not, but can not say what is novel or detect what are in the image. And if the data is not novel then that means that the user can easily deal here so no need to run CV model to alert the user else run the CV model, detect and alert the user if needed. Of coarse continuous detection is better in terms of accuracy and accessibility, but we try to make a balance to allow the CV models to be easily adopted on wearable devices with their limitations in power and resources.

Citations

[1] A. Singh, P. K. Chouhan, and M. K. Bansal, "Artificial Intelligence-Based Assistive Technologies for Visually Impaired Individuals: Advances, Challenges, and Future Directions," *SN Computer Science*, vol. 6, no. 2, p. 234, 2025. doi: 10.1007/s44443-025-00234-9.
Available: https://link.springer.com/article/10.1007/s44443-025-00234-9

[2] P. S. Bisen, V. Bisen, and R. Bisen, "Enhanced Obstacle Detection Using Bilateral Vision-Aided Transformer Neural Network for Visually Impaired Persons," *Cluster Computing*, 2024. doi: 10.1007/s10586-024-04710-1.
Available: https://link.springer.com/article/10.1007/s10586-025-05740-z

[3] S. Kahu and R. Rahate, "Image Compression using Singular Value Decomposition," *International Journal of Advancements in Research & Technology*, vol. 2, no. 8, pp. 244-248, August 2013.
Available:https://www.academia.edu/4578584/Image_Compression_using_Singular_Value_Decomposition

[4] H. R. Swathi, S. Sohini, Surbhi, and G. Gopichand, "Image compression using singular value decomposition," *IOP Conference Series: Materials Science and Engineering*, vol. 263, 042082, 2017.
Available: https://iopscience.iop.org/article/10.1088/1757-899X/263/4/042082/pdf

[5] H. B. Razafindradina, P. A. Randriamitantsoa, and N. R. Razafindrakoto, "Image Compression with SVD: A New Quality Metric Based On Energy Ratio," *IJCSN International Journal of Computer Science and Network*, vol. 5, no. 6, pp. 960-965, December 2016.
Available: https://ieeexplore.ieee.org/abstract/document/4426357

[6] T. P. Joshua, M. Arrivukannamma, and J. G. R. Sathiaseelan, "Lossy Image Compression Using Singular Value Decomposition and Discrete Wavelet Transform," *I J C T A*, vol. 9, no. 27, pp. 569-574, 2016.
Available: https://serialsjournals.com/abstract/42761_74-173_p.pdf

[7] Y. Liu and K. D. Kang, "*Filtering empty video frames for efficient real-time object detection*" *Sensors*, vol. 24, no. 10, p. 3025, May 2024.
Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11125236/

[8] D. J. Samuel and F. Cuzzolin, "*SVD-GAN for real-time unsupervised video anomaly detection*" in Proceedings of the British Machine Vision Conference (BMVC), 2021.
Available: https://www.researchgate.net/publication/355905349_SVD-GAN_for_Real-Time_Unsupervised_Video_Anomaly_Detection

[9] J. Zhao, Y. Guo, and M. Gong, "SVD-RND: An Efficient Out-of-Distribution Detector," *Proceedings of the International Conference on Learning Representations (ICLR)*, Apr. 2024. [Online].
Available: https://openreview.net/forum?id=ByeNra4FDB