IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

Characterizing and Correcting Capture-Specific Defects in Images Taken by Visually Impaired Users to Improve OCR-Based Text Recognition by Comparing OCR Engines.

Iqra Sayyed[1] and Ms. Snigdha Shukla[2]
[1]Department of Computer Applications,
[2]Assistant Professor,
Sadhu Vaswani Institute of Management Studies for Girls, Pune, India
[1]iqra.sayyed@svims-pune.edu.in and [2]snigdhashukla@svims-pune.edu.in

Abstract: *OCR, optical character recognition, has proven to be effective in reading text from images. However it can be a non-trivial task, especially from the point of view of visually impaired users, due to possible defects in the image. We explore whether OCR preprocessing should be tailored to the actual defect present in an image or whether it is always optimal to apply the same preprocessing, regardless of image defects. We analyze 2000 images from the VizWiz corpus, each with four possible types of defects: blur, tilt, low illumination, and overexposure. We compare three conditions: the raw image, a standard preprocessing, and a defect-specific preprocessing. We evaluate these using two OCR engines, Tesseract and EasyOCR. The error rates in terms of CER and WER are computed on a quality-screened set of 177 images with reference text. With different results for each OCR engine, the Tesseract's best CER and WER were achieved with the defect-specific preprocessing, whereas for EasyOCR the best results were obtained with the standard preprocessing. Therefore, it is not possible to identify a single best preprocessing for all images and OCR engines.*

Keywords: Optical Character Recognition, OCR Preprocessing, VizWiz, Visually Impaired Users, Image Defects, Defect-Aware Preprocessing, Tesseract, EasyOCR, CER, WER

I. INTRODUCTION

OCR is used today to read text from photographs. It can be useful for visually impaired users since text on labels, on packages, in documents, signs, and screens can be inaccessible to them. However, OCR is dependent on image quality, and a photograph can be blurry (due to camera movement), have an incorrect orientation, incorrect lighting, or be overexposed, thus impacting character recognition negatively. Images captured by visually impaired people may have additional problems, as the photographer cannot see if the camera was focused, if the camera was tilted, if the lighting was adequate, or if the object was captured in full. The images in the VizWiz corpus were captured by blind or visually impaired users and thus may help evaluate such cases. Typical preprocessing steps that may be applied to images in order to improve OCR results are converting the image to grayscale, increasing the contrast, thresholding, sharpening, correcting uneven illumination, and rotation. However, each of these operations may not be necessary or applicable to a particular defect, e.g., a blurry image may need sharpening, but an image with insufficient illumination may need more contrast. The motivation for this research is to evaluate whether OCR accuracy can be improved upon by selecting an appropriate preprocessing depending on the type of defect present in the image, and whether similar results can be obtained for different OCR engines which can later be integrated into a software .

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40476

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

Literature Review

There is a body of literature exploring the image quality issues that arise when photographs are taken by vision-impaired users. Chiu, Zhao, and Gurari [1] introduced the VizWiz Image Quality Issues dataset and explored potential image defects such as blur, under- and overexposure, framing, obstructions, and rotation. They conclude that images taken by blind users can exhibit quality issues that are not present in conventional image corpora. Chen, Anjum, and Gurari [2] investigate the nature of visual questions asked by blind users, finding that often useful information is present in images but is challenging to detect for such users. Mandal et al. [3] analyze image-quality feedback for photographs taken by vision-impaired users, identifying blur, exposure problems, and noise. Atitallah et al. [4] propose a text recognition system targeting color images for visually impaired users. Their approach includes image processing, text localization, enhancement, and OCR. Bafghi and Gurari [5] examine how image-classification models perform on images taken by blind people and find that such models are generally not robust to variations in image composition. Reynolds, Nagesh, and Gurari [6] investigate salient object detection in the context of images taken by people with low vision. They also find that standard computer vision tasks are impacted by the image quality and non-standard composition of such images.

More recent works continue the exploration of vision-related issues in the context of low vision and blindness. Pal et al. [7] describe an approach to visual question-answering using smartphone videos and preprocessing, OCR, and pretrained models. Gao et al. [8] introduce the VI-OCR method to evaluate text accessibility under simulated reductions in visual acuity and contrast sensitivity. These works contribute to the understanding of the challenges involved but do not directly address the selection of preprocessing steps for images taken by such users.

Research Gap

The main research question is whether OCR preprocessing should be tailored to the actual defect present in an image or whether it is always optimal to apply the same preprocessing, regardless of image defects. We explore whether OCR accuracy can be improved upon by selecting an appropriate preprocessing depending on the type of defect present in the image, and whether similar results can be obtained for different OCR engines.

II. METHODOLOGY

Dataset and Defect Description

We analyze a subset of 2000 images from the VizWiz corpus for defect detection, and a quality-screened subset of 177 images for OCR evaluation. For the latter, we selected images with acceptable reference transcriptions. We classify each image as having one of the four defects or being free of defects. We describe how we detect these defects in terms of automatic processing steps.

The four defects we consider are blur, tilt, low illumination, and overexposure. We classify the image as being affected by blur using the variance of the Laplacian: $B = \mathrm{Var}(L(I))$. A lower value of B indicates a higher probability of being blurry. Specifically, images with a value of B below 100 were classified as blurry. We classify tilt using the Canny edge detection and Hough transform procedures. We obtain an approximate tilt angle and correct it if needed. We estimate the mean pixel intensity and classify images with a mean intensity value below 70 as having insufficient illumination and images with a mean intensity value above 220 as being overexposed.

Preprocessing and OCR

We implement two approaches to image preprocessing: a standard preprocessing and a defect-specific preprocessing. For the standard preprocessing we applied grayscale conversion, histogram equalization, and Otsu thresholding. For the defect-specific preprocessing we applied sharpening, rotation correction, contrast enhancement, and illumination correction, depending on the detected image defect(s). We applied the OCR twice: with the default Tesseract settings and with pretrained EasyOCR models. We compared the OCR results for three types of input images: raw, standard-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

preprocessed, and defect-specific preprocessed. We did not train a new OCR model for this study since our aim was to compare different preprocessing methods rather than improving the OCR model itself.

Algorithm: Defect-Aware OCR Processing

The proposed algorithm starts by taking an image from the VizWiz corpus and initially analyzing its qualities. Blur is detected using the variance of the Laplacian, tilt is detected using edge detection and line analysis, and illumination problems are detected using pixel brightness. Based on the detected defect, the defect-specific preprocessing steps are applied: images with blur are sharpened, tilted images are rotated, low-light images have their contrast enhanced, and overexposed images have their illumination and contrast adjusted. The resulting image is then processed with Tesseract or EasyOCR, with the resulting OCR text being compared to the reference text using CER and WER.

Ground-Truth Preparation and Quality Screening

We needed reference transcriptions to evaluate OCR accuracy. Some of the available transcriptions were short and did not appear to correspond to the text on the image. First, we automatically screened the images to identify potential issues, and then we manually reviewed the suspicious images. Finally, we obtained a list of 177 images that could be used for quality assessment. Note that this is a quality-screened subset and should not be confused with fully manually-vetted corpora.

Experimental Conditions

OCR Engine	Preprocessing	CER	WER
Tesseract	Raw	1.4640	1.4168
Tesseract	Standard	2.0241	1.5565
Tesseract	**Defect-specific**	**1.4236**	**1.3728**
EasyOCR	Raw	3.3258	3.3308
EasyOCR	**Standard**	**2.7988**	**2.8754**
EasyOCR	Defect-specific	3.5573	3.3717

Figure 1. Comparison conditions used in the OCR experiment.

Each image in the evaluation set was analyzed under three conditions: (1) raw image, (2) standard preprocessing, and (3) defect-specific preprocessing. We applied these preprocessing steps using Tesseract and EasyOCR engines. Thus, we compared 3 × 2 = 6 sets of OCR results. We evaluated OCR accuracy for each condition relative to the reference text. We analyzed the results for all images in the subset, comparing raw, standard-preprocessed, and defect-specific preprocessed image with respect to one another.

Evaluation Metrics

The results were presented in terms of two standard OCR evaluation metrics, namely Character Error Rate (CER) and Word Error Rate (WER). The said metrics were used to compare the OCR results with the reference text and both accounted for substitution, deletion, and insertion errors. A lower value indicates better OCR performance. We compute CER as

$$CER = (S + D + I) / N,$$

where S is the number of substituted characters, D is the number of deleted characters, I is the number of inserted characters, and N is the number of characters in the reference text. The WER metric follows the same principle but for words.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40476

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

Experimental Workflow and Comparative Evaluation

The final OCR evaluation used 177 quality-screened images with available reference transcriptions. The screening was done automatically first, followed by manual checking of suspicious cases. The 177 images should therefore be described as a quality-screened evaluation set, not as a completely manually verified dataset. For every evaluation image, three versions were tested: (1) raw/original image, (2) standard-preprocessed image, and (3) defect-specific-preprocessed image. All three versions were passed through both Tesseract and EasyOCR. The OCR output was compared with the reference text. Character Error Rate (CER) measures character-level errors. Word Error Rate (WER) measures word-level errors. Both consider substitutions, deletions, and insertions. Lower CER and WER mean better OCR recognition. The comparison was done in two ways: first, overall results were compared across all evaluation images; second, results were checked separately for blurry, tilted, low-light, overexposed, and clean images.

Model Comparison

The OCR performance was compared under three different image-processing conditions to determine whether preprocessing affected text recognition. Raw Image is the original image was directly provided to the OCR engine without applying any preprocessing. Standard Preprocessing is a common preprocessing pipeline was applied to all images. This consisted of grayscale conversion, histogram equalization, and Otsu thresholding. Defect-Specific Preprocessing is were the preprocessing operation was selected according to the defect identified during the defect-detection stage. Sharpening was applied to blurry images, rotation correction was applied to tilted images, contrast enhancement was applied to low-light images, and illumination/contrast correction was applied to overexposed images.
The three conditions were evaluated using both Tesseract and EasyOCR. The comparison was performed using CER and WER to determine whether defect-specific preprocessing provided better OCR performance than applying the same standard preprocessing to every image.The comparison was conducted at two levels. First, the overall performance of the three conditions was compared across the complete evaluation subset. Second, the results were analyzed separately for each detected image condition. This allowed the study to determine whether the effectiveness of preprocessing changed according to the type of capture defect and the OCR engine being used.

Proposed Defect – Aware OCR Preprocessing Framework

The proposed system is a defect-aware OCR preprocessing framework for images captured by visually impaired users. Its purpose is to identify selected image-quality defects before OCR and apply a suitable correction instead of using one fixed preprocessing method for every image.Based on the detected condition, a corresponding preprocessing operation is selected. The processed image is then provided to the OCR engine for text recognition.

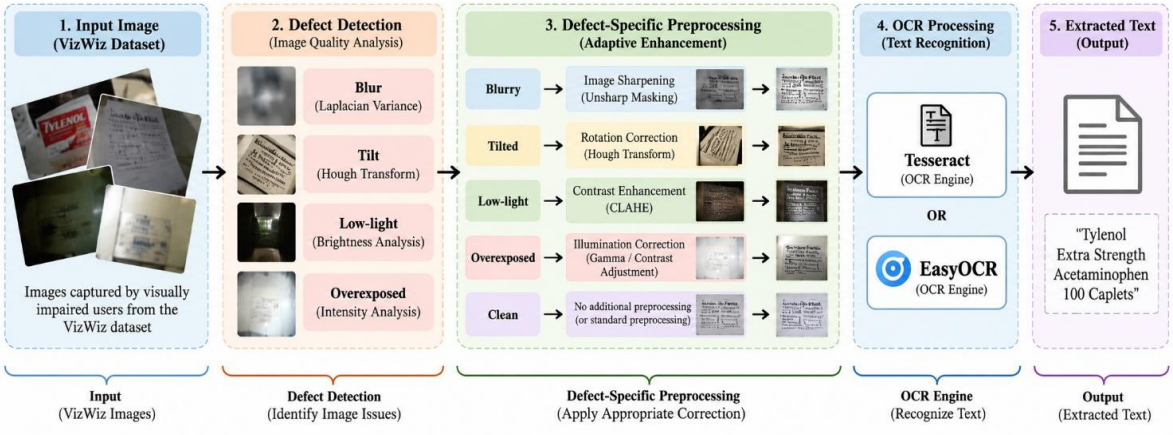


Figure 2. Proposed defect-aware OCR preprocessing framework.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

The defect-aware processing stage uses the output of the defect-detection module to determine which correction should be applied. For blurry images, sharpening is applied to enhance character edges. For tilted images, rotation correction is performed. Low-light images undergo contrast enhancement using CLAHE, while overexposed images are processed using illumination and contrast correction. When multiple defects are detected, the corresponding preprocessing operations are applied sequentially.

After preprocessing, the resulting image is passed to the OCR engine for text recognition. Both Tesseract and EasyOCR were used in the study. The extracted text is subsequently evaluated against the available reference transcription using CER and WER.

III. FINDINGS AND RESULTS

The defect distribution across the 2000 VizWiz images is shown below. Overall, we find a significant presence of blurry images (35.7%) and images with low-light conditions (18.7%). Tilt and overexposure were relatively rare with 2.5% and 4.3%, respectively. Orientation was undetectable by Tesseract OSD (Orientation Detection) for 77.9% of the images, suggesting significant variation in rotation.

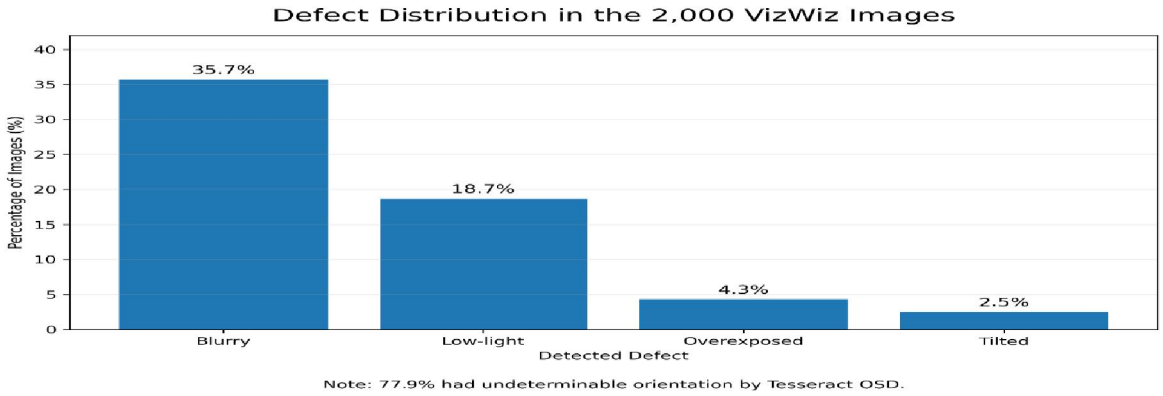


Figure 3. Defect distribution in the 2,000 VizWiz images.

The overall OCR performance was compared using Character Error Rate (CER) and Word Error Rate (WER) under three conditions: raw images, standard preprocessing, and defect-specific preprocessing. . For Tesseract, defect-specific preprocessing gave the lowest overall CER (1.4236) and WER (1.3728). Raw images were slightly worse, while standard preprocessing produced the highest CER.

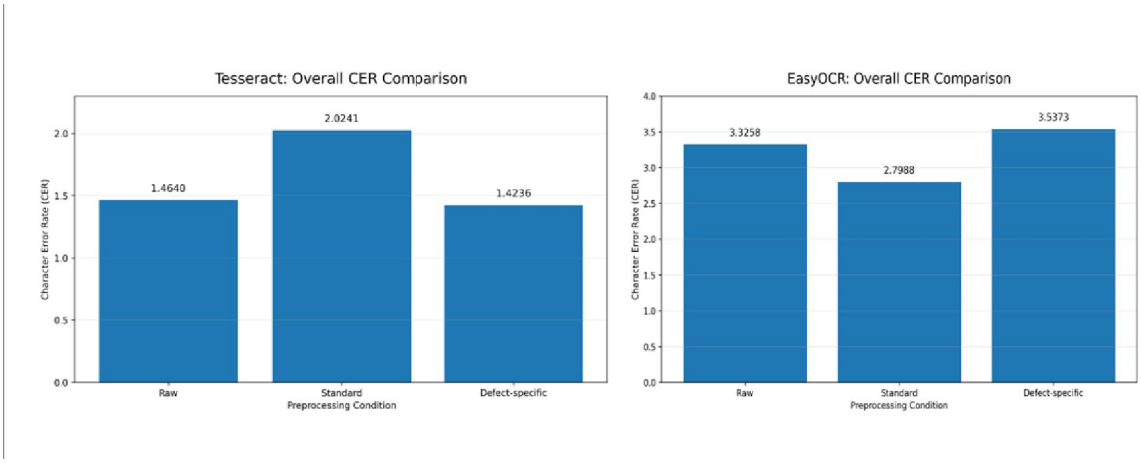


Figure 4. Overall CER comparison for Tesseract and EasyOCR.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40476

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

For EasyOCR, the pattern was different. Standard preprocessing produced the lowest overall CER (2.7988) and WER (2.8754), while defect-specific preprocessing produced higher errors than both standard and raw processing. These results show that preprocessing does not have the same effect on different OCR engines. In this study, Tesseract benefited more from the defect-specific approach, whereas EasyOCR performed better with the standard preprocessing method.

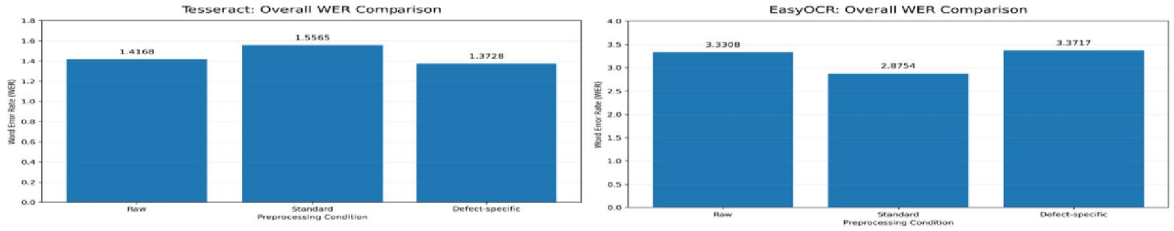


Figure 5. Overall WER comparison for Tesseract and EasyOCR.

The results were also examined separately for blurry, tilted, low-light, overexposed, and clean images. This helps show whether the best preprocessing method changes with the type of image defect.

For Tesseract, defect-specific preprocessing slightly improved CER for blurry and tilted images compared with raw images. For low-light images, it performed better than standard preprocessing, although raw images still had a slightly lower CER. For overexposed images, standard preprocessing gave the lowest error. Clean images showed very little difference between raw and defect-specific processing.

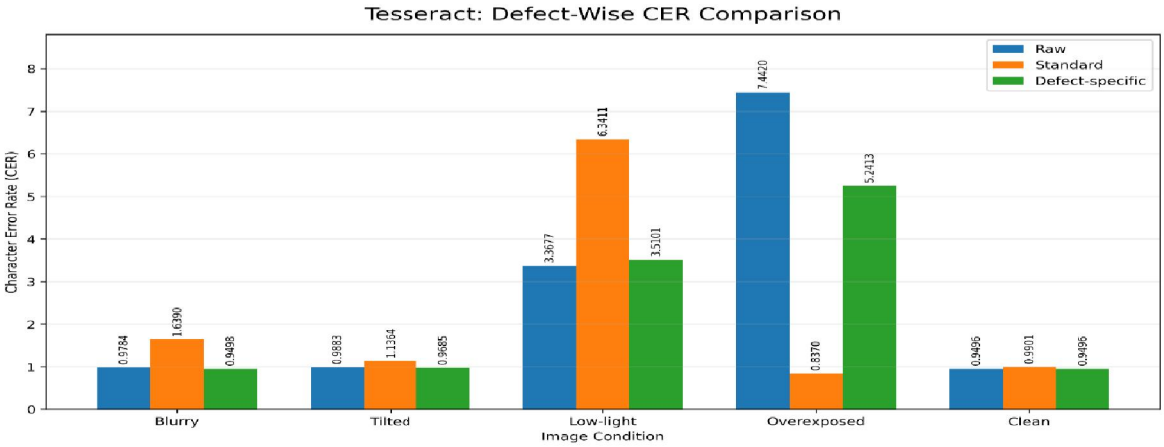


Figure 6. Tesseract defect-wise CER comparison.

EasyOCR showed a different pattern. Standard preprocessing produced the lowest CER for blurry, tilted, low-light, and clean images. The overexposed images were particularly difficult for EasyOCR, with considerably higher CER values under all three preprocessing conditions. The results, therefore, suggest that defect-specific preprocessing did not provide EasyOCR with a significant advantage for the considered categories of defects.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

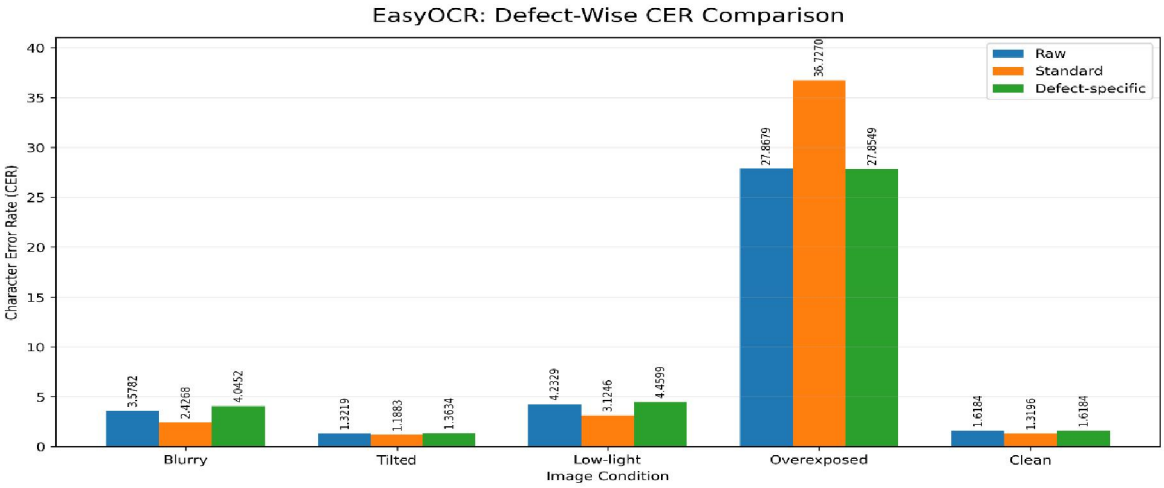


Figure 7. EasyOCR defect-wise CER comparison

As a result, it was found that for each particular category, there was no particular method that showed better accuracy in OCR tasks. Defect-specific preprocessing gave the best overall result for Tesseract, while standard preprocessing performed better for EasyOCR. The defect-wise analysis also showed that the effect of preprocessing changed with the image condition. Therefore, the results support the use of a defect-aware approach as an alternative to applying the same preprocessing operation to every image, while also showing that its effectiveness depends on the OCR engine and the type of defect present.

Major Findings

The study found that image-quality problems are common in VizWiz images, with blur and low-light conditions being the most frequently observed defects. The OCR experiments also showed that preprocessing does not have the same effect in every situation. Tesseract performed best overall with defect-specific preprocessing, while EasyOCR performed better with standard preprocessing. The defect-wise analysis further showed that the usefulness of preprocessing depends on the type of defect and the OCR engine. Therefore, the results support the idea that there is no single preprocessing method that works best for every image, and that considering the condition of the input image can be useful when designing an OCR pipeline.

Limitations

The study has certain limitations. The defect analysis was performed on a subset of 2,000 VizWiz images, while the final OCR evaluation used a quality-screened subset of 177 images. Some defect categories, particularly tilted and overexposed images, contained relatively few samples. In addition, the quality of the available VizWiz reference transcriptions was not uniform, and some images had orientation that could not be reliably determined. The study also focused on four defects and two OCR engines, so the findings may not apply to every type of image defect or OCR system.

IV. CONCLUSION

This study examined whether OCR preprocessing should be selected according to the capture-specific defects present in images taken by visually impaired users. The results showed that preprocessing effectiveness varied across both defect types and OCR engines. Tesseract achieved its best overall performance with defect-specific preprocessing, while EasyOCR performed better with standard preprocessing. These findings suggest that defect-aware preprocessing

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 7, October 2026
Impact Factor: 8.2

can be useful, but it should not be assumed to improve OCR in every situation. The study provides a basis for further work using larger, better-validated datasets and additional image defects and OCR systems.

REFERENCES

1. Tai-Yin Chiu, Yinan Zhao and Danna Gurari (2020): Assessing Image Quality Issues for Real-World Problems
2. Chongyan Chen, Samreen Anjum and Danna Gurari (2022): Grounding Answers for Visual Questions Asked by Visually Impaired People
3. Maniratnam Mandal, Deepti Ghadiyaram, Danna Gurari and Alan C. Bovik (2023): Helping Visually Impaired People Take Better Quality Pictures
4. A. B. Atitallah et al. (2023): An Efficient Text Recognition System from Complex Color Image for Helping the Visually Impaired Persons
5. Reza Akbarian Bafghi and Danna Gurari (2023): A New Dataset Based on Images Taken by Blind People for Testing the Robustness of Image Classification Models Trained for ImageNet Categories
6. Jarek Reynolds, Chandra Kanth Nagesh and Danna Gurari (2024): Salient Object Detection for Images Taken by People With Vision Impairments
7. Ratnabali Pal, Samarjit Kar, Dilip K. Prasad and Arif Ahmed Sekh (2025): Improving Visual Question Answering for Visually Impaired People Leveraging Pre-trained Models
8. Q. Gao, R. Manduchi, P. Y. Ramulu, G. E. Legge and Y. Xiong (2025): VI-OCR: “Visually Impaired” Optical Character Recognition Pipeline for Text Accessibility Assessment