Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal Vol. 5, No. 2, April 2023, pp: 89-98; eISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Manuscript received March 30, 2023; revised April 14, 2023; accepted April 28, 2023; date of publication April 30, 2023
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeemi.v5i2.294

How to cite: Rachmad Setiawan, Rayhan Akmal Fadlurahman, and Nada Fitrieyatul Hikmah, “Design of Low Vision Electronic Glasses with Image Processing Capabilities Using Raspberry Pi”, Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 5, no. 2, pp. 89–98, April 2023

Design of Low Vision Electronic Glasses with Image Processing Capabilities Using Raspberry Pi

Rachmad Setiawan , Rayhan Akmal Fadlurahman, and Nada Fitrieyatul Hikmah
Department of Biomedical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Corresponding author: Rachmad Setiawan (e-mail: rachmad@bme.its.ac.id).

ABSTRACT Poor vision is one of the most common eye health issues worldwide. Low vision patients are typically treated with optical devices or by substituting hearing or touch for visual capabilities. Head-mounted displays are currently the most promising form of low-vision assistive technology since they utilize the user's remaining natural visual capabilities. This study aims to design and evaluate a prototype of low-vision electronic glasses with image-processing capabilities using Raspberry Pi. The primary contribution of this study is a prototype head-mounted display-based low-vision tool in the form of electronic glasses utilizing evaluation of user experience in visual response. The prototype was created using a Raspberry Pi 4 B coupled with cameras to allow real-time video acquisition. The LCD on the electronic eyewear frame as the camera showed the video recording. The prototype also included software utilizing five image processing modes magnification, brightness enhancement, adaptive contrast enhancement, edge enhancement, and text detection and recognition. OpenCV was used with Python to create the software system. The performance of brightness showed a PSNR of 2.2 and bias of 80, CNR of 3.0, and bias of 100. Meanwhile, the adaptive contrast mode had the best PSNR performance with a threshold of 1.0 and grid of 16x16, and CNR with a threshold of 1.0 and grid of 2x2. Average framerate measurements of 30–40 FPS for brightness and contrast improvement modes, 20 FPS for zooming and edge enhancement modes, and 1.3 FPS for text identification modes showed that the concept of electronic spectacles was successfully implemented in this research.

INDEX TERMS Electronic glasses, Low vision, Raspberry Pi, Real-time image processing.

I. INTRODUCTION

Low vision is a condition when the visual ability of a person falls below normal limits. Compared to normal eyes, which have a visual acuity of 6/6 to 6/9, those with low vision have a visual acuity of 6/18 or greater on the Snellen scale [1]. Cataracts, glaucoma, and uncorrected refractive conditions are the main causes of low vision [2]. According to data from the Global Vision Database, there were 295 million persons with low vision in the world in 2020, and that figure is expected to rise to 474 million by the year 2050 [3]. Therefore, it is crucial to develop treatment and assistive technology that is readily available for those with impaired eyesight.

The development of low vision aids that have been carried out has various approaches. For personal computer users, there are several auxiliary software products for adjusting digital device screens to make them easier to see, such as ZoomText, JAWS, and Dolphin. Meanwhile, for non-digital media such as written media, there are already image magnifying devices in the form of monitors such as the ClearView C 24 HD and portable ones such as the Optelect Compact + [4]. However, these devices are mostly focused on reading activities and can only be used in static conditions, so they still have deficiencies in terms of mobility and flexibility.

The approach for building low-vision aids that especially focuses on mobility and navigation is mostly represented in the form of smart canes that can offer users audible feedback, such as WeWalk [5]. By using object recognition technology and acoustic feedback, other more portable and adaptable devices are developed, for instance, OrCam [6]. Devices that employ audio feedback are not always an option, despite the fact that they can be useful in some circumstances. The reason is that they can interfere with the audio information from real settings and have trouble accurately conveying complicated structures through auditory feedback.

Recently, head-mounted display technology is one of the options that has undergone significant improvement. Apart from being mobile, it will be able to offer a more natural interface and support more complicated types of activities by utilizing the original vision that its users still have for it. eSight Eyewear, NuEyes Pro Smartglasses, CyberTimez Cyber Eyez, Evergaze seeBOOST, IrisVision, and Enhanced Vision System are some new gadgets in the market [7]. Numerous studies, like those undertaken by Afinogenov and Zhao [8], have also been conducted to build alternative equivalent tools. The prototype that has been made has been successful in implementing embedded systems, but its use is still limited to basic functions including image magnification.

Adjustments and more varied options are required in image processing techniques to cater to the needs of people with limited vision who have a variety of conditions [9]. Maculopathy, a degenerative condition affecting the macular region of the retina, is one example of this. Along with a decline in distant vision, people with Maculopathy conditions, including Age-Related Macular Degeneration (ARMD) and Juvenile Macular Dystrophy (JMD) experience a decline in contrast sensitivity [10]. Hence, image processing in the form of increased contrast will significantly enhance the participants' vision who have maculopathy. Another example is people with a hazy central vision from conditions like ARMD, glaucoma, or diabetic retinopathy yet can still see clearly out of their peripheral vision [11]. Image processing, such as contour line sharpening, can aid patients with this disease by enhancing the visual information gained [12]. Zhao [13] designed a display device with more diverse image processing functions with increased contrast and borders through a prototype called ForeSee. However, its implementation has not been optimized as a mobile device because it still has to be connected to a laptop as a processor.

This study aims to build a working prototype of a low-vision aid in the form of electronic glasses. Because their use is not restricted to static settings, the glasses' complete hardware system, including their CPU, is designed to be worn. The system's software is designed with efficient image processing features to aid those with low eyesight in obtaining more visual data. This research is intended to yield benefits, such as assisting those with low vision in their daily activities, particularly those involving reading from close and medium distances. Several contributions of this research are:

a) The designed prototype can function as a tool that helps improve the user's visualization abilities. By increasing visual abilities, users can use tools for various needs without being stuck with only one specific activity.
b) Providing options for various visual conditions. The image processing patterns used are techniques that have been widely accepted by experts in the field of low vision and have been tested in other studies.
c) Low-cost and easy to manufacture. The prototype is expected to be an alternative tool that is more accessible to most people with low vision. The hardware design is made of widely used materials, so it can be more easily repaired or replaced when there is damage.
d) Supports hands-free interaction. The prototype is designed to be flexible and comfortable for daily needs without limiting the user's physical activity. The setting pattern on the glasses is also designed to be easy to control.

II. RELATED WORK

The design of low vision aids in the form of head-mounted displays is developed because it has a more natural interface by utilizing the user's vision which can still be used. Many of these tools are designed to be portable or wearable. The following describes research developments related to assistive devices in terms of image processing and head-mounted displays.

A. IMAGE PROCESSING FOR LOW-VISION PEOPLE

In the category of low-vision aids, there is a distinction between electronic devices and more traditional optical equipment. In contrast to optical technologies, where the magnification process depends on the strength and form of the lens, electronic devices allow for digital magnification [14]. Electronic devices' image processing also makes it feasible to offer additional adjustments that can benefit a larger range of viewing circumstances. According to a study by Moshtael, people with impaired vision favor several image processing techniques, including contrast augmentation, outlines/contour identification and enhancement, background attenuation and simplification, and remapping and retargeting [9]. This study also revealed the requirement for an image processing method that was specifically made for a certain patient.

The benefits of image enlargement for people with low vision can be seen in the research conducted by Christen and Haji [15]. The results of research on low vision simulations and real subjects showed a significant increase in reading ability when given an enlarged image. A correlation between contrast sensitivity and reading speed was discovered in participants with macular degeneration in a study by Brussee 10]. Meanwhile, Bowers conducted a study to assess the influence of brightness levels on people with macular illness, where subjects with limited vision require at least a high lighting level to read optimally [16].

Kwon studied the effectiveness of image processing by sharpening contour lines for those with limited vision and discovered that edge enhancement significantly increases search speed for older age groups [12]. A study by Wolfssohn discovered that people with impaired vision considerably favored edge-enhanced views over unfiltered displays [17]. Huang's research aimed to use text detection algorithms to aid low-vision users in navigation. Subjective evaluation reveals that participants who used a text detection application said the app helped them navigate, while walk-pattern analysis reveals that they may take shorter routes.

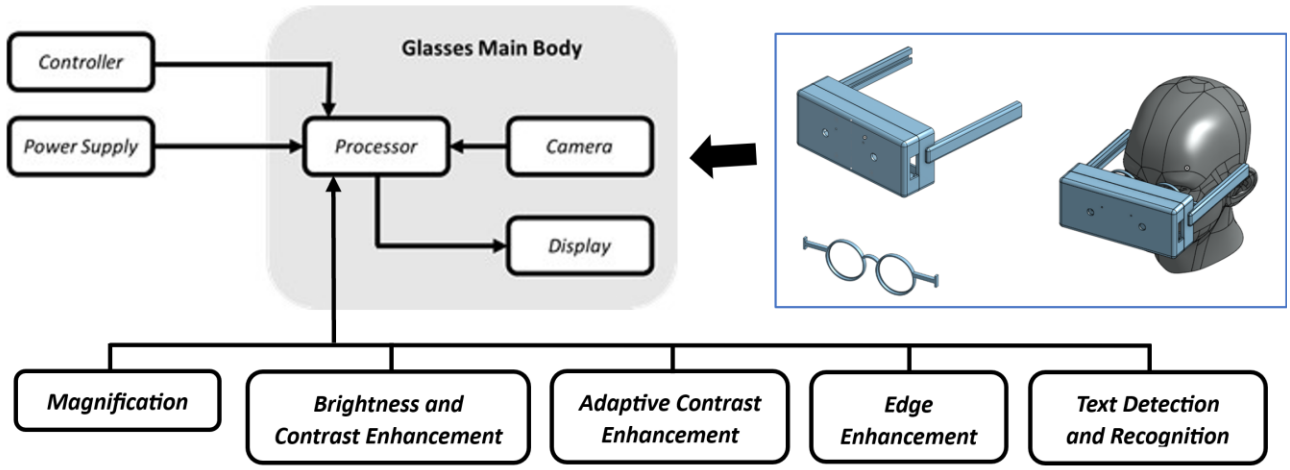

FIGURE 1. **A Prototype for Hardware and Software Design.**

B. HEAD-MOUNTED DISPLAY DESIGN

The head-mounted display device design that inspired this research was taken from two main sources: the prototype made by Afinogenov and Zhao. Afinogenov researched to create prototypes of visual aids for blind people, particularly for students in developing nations with little financial resources [8]. As a result, the prototype created was inexpensive yet provided basic image-processing functions to aid learning. Six elements make up the prototype system: a camera, LCD, lens, image processing, battery, and packaging. A Raspberry Pi camera with a resolution of 1920x1080 pixels was used. The monitor utilized a 2-inch Raspberry Pi LCD with a 320x240 pixel resolution. The lens utilized a Fresnel lens that was parallel to the pupil. Raspberry Pi 2 with 1 GB RAM was selected as the CPU for image processing. The API from the Python Picamera package was used to perform image enlargement. While battery capacity of 7200 mAh was the minimum requirement for the system to run for longer hours.

User controls were included in a belt pack containing zoom controls and a power button. The prototype frame was made using 3D printing with plastic material. This prototype was evaluated by students with low vision at MIT using the Snellen table. The zoom control on the prototype allows the user to read 4 additional lines from 0.46 m and 3 additional lines from 6 meters, compared to without using the prototype.

Meanwhile, Zhao's study looked at the efficacy of several image-enhancing approaches for those with low vision (13). As a result, ForeSee, an augmented reality-based see-through head-mounted display video tool for those with low vision, was developed. The developed prototype included five commonly used standard image processing techniques: magnification, contrast enhancement, edge sharpening, and black-and-white inversion. In addition, there were two display options available: full display, which showed the image in its entirety, and window display, which only extracted the region that the user had chosen. The Oculus Rift DK2 was used to create the prototype, with a WebCam F100 camera for video input and a laptop for processing. The entire system was created using Unity and OpenCV for real-time video processing. The input design, meanwhile, was created utilizing a smartwatch and a voice command.

The test results showed that all participants could read and recognize images at near and far distances using ForeSee, but required adjustments to different image processing methods and display modes. While on the test using different inputs, the results showed that most participants preferred to use voice command input rather than a smartwatch.

III. MATERIALS AND METHODS

In this research, a prototype assistive device for people with low vision was made as a head-mounted display in the form of glasses. Before starting prototyping, a hardware design was made as a first step to describe the design requirements that were expected to be realized in the research. The hardware design would be a reference for selecting software specifications.

A. HARDWARE DESIGN

This prototype consists of two parts: a head-mounted display (HMD) in the form of glasses and a controller connected to Bluetooth. The HMD has a camera, display, and processor, while the power supply and controller are separate from the main part of the HMD. The hardware scheme can be seen in Fig. 1.

The camera utilized an 8-megapixel Raspberry Pi V2 module that enables video qualities up to 1080p while being tiny and weighing only 3 grams. The camera is mountedon the eyeglass frame, so its diminutive size and light weight are crucial. Meanwhile, a 3.2-inch LCD witha 480x800 pixel resolution is employed for the display. This size was chosen to provide adequate pixel resolution whilestill being comparable to the size of a typical eyeglass lens.A Fresnel lens was utilized since it may be moved between the display and the eye to change the focus because the display will be positioned at a distance of around 3–4 cm directly in front of the eyes.

The lens, display, and camera were all combined into a single 3D-printed eyeglass frame. The web-based design tool Onshape was used to create 3D eyeglass frames. The camera is mounted on the front of the glasses to maintain a line of sight with the wearer's eyes. A Raspberry Pi 4 model B computer with 4 GB of memory served as the processor. The Raspberry Pi was chosen partly due to its compact size (10x7x3 cm), which is comparable to modern computers in terms of computational power, and partly due to howsimple it is to utilize Python and OpenCV to implement video processing. Two mini HDMI connections are also included in Generation 4, making it simpler to display video streams.

The BLE-M3 controller, which is compact (50x30x10.5 mm) and can be connected to the Raspberry Pi using Bluetooth, is used to collect user input. The power supply was chosen in accordance with suggestions for the Raspberry Pi, which has a voltage of 5 V, an output current of 3 A, and a USB-C connection. A power bank with a capacity of 20,000 mAh is used to extend usage time.

B. SOFTWARE DESIGN

Python and OpenCV were used in software development. Because of its reasonably comprehensive library for image and video processing, OpenCV was chosen. Five image processing techniques magnification, global brightness and contrast improvement, adaptive contrast enhancement, edge enhancement, and text detection are included in the software designed to give image enhancement that can aid people with low vision.

1) MAGNIFICATION

In digital image processing, bilinear interpolation is one of the widely used algorithms to perform transformations on spatial dimensions [18]. For a two-dimensional square shape, bilinear interpolation is obtained by performing two linear interpolations on the horizontal and vertical axes The mathematical formula is given in (1),

$$P = \frac{1}{(x_2-x_1)(y_2-y_1)}\begin{bmatrix} x_2 - x & x - x_1 \end{bmatrix}\begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}\begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \quad (1)$$

where (x_n,y_n) is the pixel position in the initial image array, (x,y) is the pixel position in the output image array, Q_{mn} is the neighboring pixel value in the initial image, and P is the pixel value in the output image.

2) BRIGHTNESS AND CONTRAST ENHANCEMENT

Modifying pixels with the point operator makes it possible to increase the brightness and contrast of the image. A constant enters the pixel value into the addition or multiplicationarithmetic function. Equation (2) outlines this function mathematically [18],

$$g(x) = af(x) + b, \qquad a > 0 \quad (2)$$

where $f(x)$ is the value of the input pixel and $g(x)$ is the outcome. The constant value a is commonly known as the gain parameter and influences the amount of contrast, while constant value b is the bias parameter and influences the brightness level.

3) ADAPTIVE CONTRAST ENHANCEMENT

The contrast or brightness of the image can be changed by executing a histogram equalization, which entails stretching the intensity range of thepixels on the histogram [18]. The cumulative distribution function in (3) can be used to equalize histograms,

$$c(I) = \frac{1}{N}\sum_{i=0}^{I} h(i) = c(I-1) + \frac{1}{N}h(I) \quad (3)$$

where N denotes the number of pixels in the image, $h(I)$ denotes the initial histogram, and $c(I)$ is the cumulative distribution value. Histogram equalization must be performed on each color channel in color images (RGB).

One of the most widely used algorithms is the variant that uses a contrast gain limit to avoid noise amplification, namely contrast limited adaptive histogram equalization (CLAHE). This algorithm is carried out by calculating the image histogram in $M \times M$ blocks, and then interpolating it when moving to the next block [19]. AHE can be done with the bilinear blending function in (4),

$$f_{s,t}(I) = (1-s)(1-t)f_{00}(I) + s(1-t)f_{10}(I) + (1-s)tf_{01}(I) + stf_{11}(I) \quad (4)$$

where $f(I)$ is the value of pixels, and (s,t) is the coordinate of the pixels.

4) EDGE ENHANCEMENT

Canny edge detection is one of the most used techniques for edge detection [20]. This algorithm is carried out in stages, smoothing using a Gaussian filter to lessen noise. The two-dimensional Gaussian mathematical equation can be seen in (5),

$$G(x,y) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{x^2+y^2}{2\pi\sigma^2}\right) \quad (5)$$

where x and y are pixel coordinates, and σ is the standard deviation. Next, the process was finding an edge on a gradient with a high magnitude, namely non-maximum suppression, which involves designating the local maximum in the pixel neighborhood as an edge and removing other pixels (the value becomes zero); and finally, optimization. Hysteresis removed linked edges with sure edges by eliminating unconnected edges with potential edges.

5) TEXT DETECTION AND RECOGNITION

The text detection and recognition program algorithm relied on an optical character recognition (OCR) system with the Tesseract engine. Tesseract is an open-source OCR engine commonly used in various text recognition applications [21]. The first pre-processing stage was applied to identify the blur level in the frame. Since the video stream consists of a collection of image frames, selecting frames with low blur levels is important to proceed to the next process and eliminate frames with high blur levels. Besides providing better-quality input, this can also reduce the computational burden. For detecting blur, several methods can be used, one of which is convolution using the Laplacian operator [22]. Laplacian operator is a derivative operator for generating edges in an image. However, images with low and high blur can be estimated by calculating the variance value in the output image. The magnitude of the variance is known using (6) [23],

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \quad (6)$$

where S^2 is the variance value, n is the number of pixels, X_i is the pixel value, dan $\bar{X}_n$ is the average pixel value.

C. EXPERIMENTAL METHOD AND EVALUATION

The camera took image processing, and an assessment was carried out using three different methods. For image magnification mode, the variance of the Laplacian value was used to determine the increase in blurring at each enlargement stage. For the brightness, contrast, and edge sharpening modes, the peak signal-to-noise ratio (PSNR) value was used as a measurement. PSNR is a value that is widely used to measure image quality [24]. By using PSNR, it is possible to compare the initial image with the processed image by calculating the noise/error that appears in the process. The calculation of PSNR value is given in (7),

$$PSNR = 10\log_{10}(peakval^2)/MSE \quad (7)$$

where the peakval value is 255 for the 8-bit unsigned integer data type. Meanwhile, the *MSE* value or mean squared error for an image with dimensions $M\times N$ with $g(m,n)$ as the initial image and $\hat{g}(n,m)$ for the resulting image is obtained through (8).

$$MSE = \frac{1}{MN}\sum_{n=0}^{M}\sum_{m=1}^{N}[\hat{g}(n,m) - g(m,n)]^2 \quad (8)$$

where $g(m,n)$ is the original image with size *mxn* columns and rows, and $\hat{g}(n,m)$ is the filtered image.

Specifically for adaptive brightness and contrast adjustment modes, the contrast-to-noise ratio (CNR) was used. Using CNR, it can be seen the comparison of the strength of the contrast between the two images based on the mean and variance. The CNR value can be calculated based on the mathematical equation shown in (9) [25],

$$CNR = \frac{|\mu_i - \mu_o|}{\sqrt{\sigma_i^2 + \sigma_i^2}} \quad (9)$$

where μ_i and μ_o is the average pixel value of two images, while σ_i^2 and σ_o^2 is the variance value.

IV. RESULTS

Software testing is carried out to evaluate the image processing program created. Testing is carried out by connecting the camera to the Raspberry Pi and then connecting it to a laptop via a Virtual Network Computing (VNC) connection. The displayed image resolution has a size of 1024x768 pixels. The image results obtained from the five implemented modes are discussed further in the following sections.

A. ZOOM MODE

Magnification mode was evaluated from three distances: 30 cm close distance, 1 m medium distance, and 2 m long distance. Story reading text was used from close range, while pangram phrase text was used from a medium and far distance. At close and far distances, it was necessary to adjust the focus in advance on the camera because the focus distance was fixed at a distance of 1 meter.

The results of taking close-up data showed that the program requires magnification of up to 2.4x from the initial image so that the entire text can fill the display. Meanwhile, magnification of up to 4.4x was required to fill the display with the entire text when capturing medium-range data for long distances, up to 11x magnification was required.

The enlarged image results were evaluated by calculating the variance of Laplacian. The process was done to determine the blurring level in the image in each enlargement stage. The results of calculating the variance value for each measuring distance are presented in TABLE 1.

TABLE 1
Variance of Laplacian in Zoom Mode

From 30 cm		From 1 m		From 3 m	
Scale	Variance	Scale	Variance	Scale	Variance
1,0x	190,0	1,0x	193,3	1,0x	179,6
1,2x	44,5	0,5x	25,1	2,0x	11,0
1,4x	25,8	2,0x	11,3	3,0x	7,1
1,6x	17,1	2,5x	10,5	4,0x	3,3
1,8x	22,7	3,0x	9,4	5,0x	3,1
2,0x	11,9	3,5x	6,9	6,0x	2,2
2,2x	7,1	4,0x	4,9	7,0x	2,0
2,4x	7,4	4,5x	5,0	8,0x	1,7
2,6x	6,3	5,0x	4,4	9,0x	1,7
2,8x	5,5	5,5x	3,9	10,0x	1,6
3,0x	4,8	6,0x	3,1	11,0x	1,4

B. BRIGHTNESS AND CONTRAST MODE

Global brightness and contrast setting modes were evaluated using pangram phrase text from 1 m away. Peak signal-to-noise ratio (PSNR) and contrast-to-noise ratio (CNR) values were calculated to measure the effect of increased brightness and contrast on the image. The brightness and contrast are affected by the bias and gain parameters at the point operator. The results of the increase in the gain parameter value with changes in PSNR and CNR are shown on the left of Table 2. Meanwhile, the increase in the value of the bias parameter with changes in PSNR and CNR is shown on the right of TABLE 2, and the results of brightness mode can be seen in FIGURE 2(a).

TABLE 2
PSNR and CNR of Brightness and Contrast Adjustment

Contrast Parameter			Brightness Parameter		
Gain	PSNR	CNR	Bias	PSNR	CNR
1.0	inf	0.0	0	inf	0.00
1.2	23.0	0.6	10	23.36	0.25
1.4	23.6	1.1	20	21.78	0.51
1.6	22.8	1.5	30	22.15	0.76
1.8	22.7	1.9	40	25.30	1.01
2.0	23.5	2.2	50	20.44	1.27
2.2	23.9	2.4	60	31.32	1.52
2.4	23.7	2.5	70	27.80	1.78
2.6	23.6	2.6	80	48.80	2.02
2.8	23.6	2.6	90	21.23	2.28
3.0	23.6	2.7	100	27.20	2.57

C. ADAPTIVE CONTRAST MODE

As with the global brightness and contrast tests, adaptive contrast tests were also carried out using pangram text from a distance of 1 m and evaluated by calculating the PSNR and CNR. As explained in the previous section, the adaptive contrast algorithm used was CLAHE. Two parameters can be changed to adjust the contrast level: the threshold value for contrast limiting on the histogram and the number of local histogram counting regions.

D. CONTOUR LINE SHARPENING MODE

The sharpening of contour lines with the Canny edge detection algorithm was evaluated using pangram text from a distance of 1 m. This algorithm uses the maximum and minimum thresholds to determine the sure edge and potential edge. Therefore, an experiment was carried out at this testing stage by changing the thresholding parameters. Three thresholding distances were used: narrow with a threshold distance of 20, medium with a threshold distance of 50, and wide with a gradient distance of 100. Three minimum thresholds were started at a value of 100. The measurement of quality of the contour sharpening results is measured using PSNR.

Qualitatively, it can also be observed that each threshold distance has a different sensitivity to fulfill the edge of the phrase 'the lazy dog' in the test text. Narrow thresholding can stably meet edges starting at -20 level, medium thresholding meets edges starting at -40 level, and wide thresholding only meets edges at -60 level. The results of sharpening contour lines can be seen in Fig. 2(b).

E. DETECTION AND RECOGNITION TEXT MODE

Text detection and recognition modes were tested using pangram text in the form of digital prints from a distance of 1 m in dark and bright conditions. Testing digital printed text in bright conditions gave the best results with all the words in the sentence being detected and recognized correctly. Meanwhile, digital printing in dark conditions gives results where the phrases 'the quick brown fox' and 'jumps over' can be recognized correctly. A comparison of text detection results can be seen in Fig. 2(c) and 2(d).

F. HARDWARE TESTING

The hardware consisted of eyeglass frames as a casing and a series of electronics consisting of a camera, Raspberry pi, and LCD. As a mechanism to control parameter changes in the application, a Bluetooth remote mini was used. The system was to make it easier for users to set their needs without needing a large control panel like a keyboard. The power supply used can be in the form of a power bank to enable mobile use. The magnitude of the frame processing speed for each mode used can be seen in TABLE 3. The program speed was calculated using the average FPS (frame per second) observed for one minute. Therefore, the possibility of applying an electronic glasses system for people with low vision without the risk of cybersickness due to high delay can be measured.

TABLE 3
Calculation of FPS Each Mode for A Minute

Mode	Average FPS
Zoom	22,2
Brightness and Global Contrast	40,8
Adaptive Contrast	30,8
Edge Enhancement	19,3
Text Detection	1,3

V. DISCUSSIONS

Magnification, done with the focal point of the magnification in the center and not at the edges of the image, gives a stable magnification result on an object of interest in the middle of the camera's field of view, even though several magnification stages are carried out. However, this can also be a drawback if the focus point of the zoom is expected to be off-center. For example, in the case of low vision, the scotoma is in the middle, so it can only rely on peripheral vision. In cases like this, the advantage for the user can be higher if the zoom focus point can be positioned. In detail, it follows the position of the object of interest so that the object of attention is not blurred or lost from view when the user enlarges the image.

To read story text from a distance of 30 cm, it is necessary to magnify up to 2.4x until the entire text fills the field of view. Blurring will occur in the image at each enlargement stage because the footage is taken from the same image dimensions as the initial image. If seen from the results of calculating the variance, then at magnification with a scale of 2.4x it has reduced the variance value from 190.0 on a scale of 1x to 7.4. The result shows a significant increase in blurring, even up to more than 90%. Changes in the number of edges can be caused by the object of interest experiencing blurring or because the background changes when enlarged. Even though the variance value has decreased drastically, the quality of the enlargement results can still be read clearly. However, the results of image observations with a variance value of more than 5.0 still provide images with quite good results.

The global brightness and contrast settings in the program are adjusted by changing the bias and gain parameter values. In conditions of poor room lighting, this setting mode can help users see objects more clearly up to a certain lighting

(a) (b) (c) (d)

FIGURE 2. **Image processing results from camera view: (a) Brightness mode, (b) Edge enhancement using thresholding of 10 and level of -50, (c) Text detection in light condition, and (d) Text detection in dark condition.**

level. However, due to its global nature for all parts of the image, increasing the brightness at the highest bias and gain values only causes the display to be dominated by white and can eliminate the information that should be obtained.

Evaluation using PSNR shows an irregular pattern, both for changes in the higher gain value, as well as the higher bias value. In changing the gain value, the highest PSNR was recorded at a gain value of 2.2, which is 23.9. PSNR then did not experience much change starting from the gain value of 2.6 and above. Meanwhile, for changes in the bias value, the highest PSNR was recorded at a bias value of 80, which was 48.8. Changes in fluctuating PSNR values indicate that the increase in gain and bias values does not always positively impact the output image results. However, there are specific values where the gain and bias are used to provide the highest signal information. Therefore, the range of gain and bias values in the program does not need to be specifically widened, but the sensitivity can be increased by adjusting the value changes shorter. That way, users can get more leverage in adjusting the brightness level they need, even though it will take more time to set it in a trade-off.

The adaptive contrast setting has a test model that is almost the same as the brightness setting, but with different parameters, which are in accordance with the CLAHE parameters. The parameters are the limiting threshold value for the histogram and the number of image parts or regions used to calculate the local histogram. Unlike the brightness enhancement mode which is able to improve the overall visibility of the image in poor room lighting conditions, the adaptive contrast setting mode does not directly improve visibility. This mode works better for the need to sharpen the foreground part of the image. Changes in the PSNR value due to an increase in the threshold and the number of regions have an irregular pattern. where the highest PSNR value for the threshold parameter of 1.0 is 24.0. As for the number of regions parameter, the highest PSNR value is 27.9 on the grid parameter of 16x16. Therefore, the increase in the threshold value and the number of regions is not linearly related to the PSNR value. It is necessary to adjust these two parameters according to the needs of the conditions experienced by the user so that the results of adaptive contrast settings can provide an ideal display output.

Meanwhile, the evaluation using the CNR value gives different results for the threshold parameter and the number of regions. Increasing the threshold value gives a gradual decrease in the CNR value. The largest CNR value of 0.18 was recorded when the threshold value was one and then decreased to 0.06 when the threshold value was three. This is probably due to an increase in threshold causing higher noise, where the increase in noise is higher than the increase in contrast. Meanwhile, the increase in the number of regions causes fluctuating changes in the CNR value. The largest CNR value of 0.19 is when the number of regions is 2x2, while the smallest value is 0.14 in the 6x6 and 16x16 regions. Changes in fluctuating CNR values can be understood because the number of regions determines the increasing number of equalization histograms that need to be calculated. The different equalization results in each region will affect the contrast level during cumulative calculations so that the change in contrast level in the output image, which is influenced by the histogram of each region can fluctuate according to the division of the region. Thus, to get an image with a good CNR value, it is necessary to pay attention to a small threshold level, combined with the number of regions according to the conditions.

The results of the contour line sharpening test obtained data on changes in PSNR values for level changes at three threshold distances. The threshold distance in question is the maximum threshold value minus the minimum threshold in the hysteresis thresholding process in the Canny edge detection algorithm. In comparison, the level is the arithmetic value of the increase or decrease of the two thresholds. In general, decreasing the level will also reduce the PSNR value. This is because the results of selecting edge detection with a low magnitude threshold will reveal more edges obtained from noise as well. As for evaluating further the best threshold distance, three kinds of threshold distances are used. The three threshold distances used, namely narrow (50), medium (100), and wide (200) will affect changes in PSNR values differently for each level change. This can be observed through data on changes in PSNR values at each level.

For the narrow threshold, the decrease in the PSNR value occurs linearly from level 100 to level 0, which is around 0.5 dB. It shows a significant decrease from level 0 to -100, with the biggest drop from level -80 to level -100 of 6.7 dB. Meanwhile, for the medium threshold, the decrease in PSNR is linear from level 100 to level -40. Then there was a drastic drop from -40 to -100 level with the biggest drop at -80 to -100 level of 5.3 dB. Meanwhile, for the wide threshold, it was found that the PSNR decreased linearly up to -60 level, with the biggest drastic decrease at -80 to -100 level, which was 1.8 dB. Based on three thresholds, it can be seen that the narrow threshold is the most sensitive to edges while the wide threshold is the most non-sensitive. As a result, at levels below 0, the narrow threshold will detect more edges and even edges in noise, while the wide threshold is more stable because it does not detect many edges in noise. However, a narrow threshold can provide more accurate contour sharpening results on an object of interest with fewer step levels because it is more sensitive to edges. So, there is a trade-off between the amount of noise likely to be detected with accurate edge-sharpening results. The medium threshold can be chosen to maximize this trade-off because the sensitivity is between the narrow and wide threshold.

The results of text detection and recognition were obtained using digitally printed writing and handwriting results in light and dark conditions. The recognition results using the Tesseract engine depend a lot on the quality of the image pre-processing before being inserted into the engine. In the program created, pre-processing was realized through the frame selection stage using the variance of Laplacian, denoising with a Gaussian filter, converting color to grayscale, and finally adaptive binarization. Text detection results for digital print writing in bright lighting conditions give good results where all the words in a sentence can be detected and recognized. Meanwhile, in dark lighting conditions, there is an error in the detection of the phrase 'the lazy dog' because it has a thinner font than the other two phrases. To conclude, it takes good lighting conditions to be able to provide the correct text detection results. Compared to previous research, the text detection and recognition program algorithm relies on an optical character recognition (OCR) system with the Tesseract engine. Tesseract is an open-source OCR engine commonly used in various text recognition applications [21]. Even though Tesseract is capable of performing fairly good text recognition on images, for input conditions in the form of video streams such as the system created it is necessary to carry out several pre-processing stages. Therefore, the text detection results obtained in this study are more accurate, even with image quality and background scenes that vary from video frame.

This research can be a development of science in the field of assistive technology and image processing. From the field of assistive technology, namely as a tool for people with low vision in acquiring visual information, and from the field of image processing as the implementation of real-time image processing in embedded systems. Then the research contribution from a practical perspective is a prototype of a visual aid that can help people with low vision in reading activities from near and medium distances.

Further research is expected to develop these electronic glasses from at least three aspects: software, hardware, and eyewear design. In this study, there are still deficiencies in terms of the long processing time used for image processing, especially for text detection algorithms. Thus, the software that is expected in further research can provide faster computations and minimize delay. Further research can also provide variations of the algorithm for each mode, so that image processing options can better suit user needs. Because high-level low-vision subjects were not used in this study's testing, it will be necessary to test image processing implementation again in future studies using real low-vision subjects or low-vision simulations. This will allow for a more accurate assessment of the benefits of image processing.

Hardware-wise, this study only used a single camera and a single display, thus in order to more accurately imitate the vision of average people, a system with two cameras and two screens must be implemented. Development by selecting the type of camera with better specifications to improve focus and increase image resolution can also be done. Finally, it is envisaged that a spectacles design with a smaller frame size, or the size of glasses in general, can be produced.

VI. CONCLUSION

This research proposes an electronic eyewear system equipped with several image processing modes to help improve the visual abilities of people with low vision. Various image processing modes are needed to provide options for adjusting to the needs of people with different low vision conditions. The image processing modes designed in this study are magnification mode, global brightness and contrast enhancement, adaptive contrast adjustment, contour line sharpening, and text detection and recognition. The findings of data collection in brightness mode revealed a changing PSNR value with a maximum gain of 2.2 and a bias of 80. The CNR value, meanwhile, had a gain of 3.0 and a bias of 100. The results of data gathering indicated a greater PSNR value for the adaptive contrast setting mode, oscillating with the maximum value at a 16x16 grid with a 1.0 threshold setting. The CNR value changed at each increase in grid size, peaking at the 2x2 grid, whereas the CNR value declined with each rise in threshold value, reaching a maximum at the threshold of 1.0. In the contour line sharpening mode, data retrieval findings for three threshold distances—narrow by 50 magnitudes, medium by 100 magnitudes, and wide by 200 magnitudes show that the PSNR values decline differently for each distance as the level is lowered. The PSNR decreases linearly from level 100 to level -40 and then sharply from level -40 to level -100, indicating the medium threshold's presence, which yields a good sensitivity rating. The implementation of software made into hardware in the form of a series of electronic glasses gives satisfactory results. Raspberry pi can perform computational image processing modes in real time. Images captured by the camera and then processed by the program can be displayed on the LCD screen without delay. Conceptually, the electronic glasses system has been successfully actualized. Although it still requires some development to improve performance. It is recommended that future work be undertaken in the following areas: minimizing the delay impact of text detection, adding the modes of image processing features, implementing a more compact design frame for glasses, and realizing the prototype using two cameras and two displays.

REFERENCES

[1] S. J. Leat, G. E. Legge, and M. A. Bullimore, "What is low vision? A re-evaluation of definitions," *Optom. Vis. Sci.*, vol. 76, no. 4, pp. 198–211, 1999, doi: 10.1097/00006324-199904000-00023.
[2] GBD 2019 Blindness and Vision Impairment Collaborators, "Causes of blindness and vision impairment in 2020 and trends over 30 years , and prevalence of avoidable blindness in relation to VISION 2020 : the Right to Sight : an analysis for the Global Burden of Disease Study," *Lancet Glob Health*, vol. 9, no.2, pp. e144-e160, 2021, doi: 10.1016/S2214-109X(20)30489-7.
[3] GBD 2019 Blindness and Vision Impairment Collaborators, "Trends in prevalence of blindness and distance and near vision impairment over 30 years : an analysis for the Global Burden of Disease Study," *Lancet Glob Health*, vol. 9, no.2, pp. e130-e143, 2021. doi: 10.1016/S2214-109X(20)30425-3.
[4] A. Blackman, "An introduction to electronic low vision aids," *Dispensing Optics*, November, pp. 18–22, 2016. [Online]. Available: https://www.abdo.org.uk/wp-content/uploads/2016/11/DO-November-2016.pdf
[5] J. H. Morris, L. Irvine, T. Tooman, S. U. Dombrowski, B. McCormack, F. V. Wijck, and M. Lawrence, "WeWalk: walking with a buddy after stroke-a pilot study evaluating feasibility and acceptability of a person-centred dyadic behaviour change intervention," *Pilot Feasibility Stud.*, vol. 9, no. 1, 2023, doi: 10.1186/s40814-022-01227-5.
[6] X. Nguyen, J. Koopman, M. M. Genderen, H. L. M. Stam, and C. J. F. Boon, "Artificial vision: the effectiveness of the OrCam in patients with advanced inherited retinal dystrophies," *Acta Ophthalmologica*, vol. 100, no. 4, 2022, doi: 10.1111/aos.15001.
[7] A. D. Deemer *et al.*, "Low Vision Enhancement with Head-mounted Video Display Systems: Are We There Yet?," *Optom. Vis. Sci.*, vol. 95, no. 9, pp. 694–703, 2018, doi: 10.1097/OPX.0000000000001278.
[8] B. I. Afinogenov *et al.*, "Wearable optical-digital assistive device for low vision students," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2016, pp. 2083–2086, doi: 10.1109/EMBC.2016.7591138.
[9] H. Moshtael, T. Aslam, I. Underwood, and B. Dhillon, "High Tech Aids Low Vision : A Review of Image Processing for the Visually Impaired," *Transl. Vis. Sci. Technol.*, vol. 4, no. 4, 2015, doi: 10.1167/tvst.4.4.6.
[10] T. Brussee, T. J. T. P. Van Den Berg, R. M. A. Van Nispen, I. De Boer, and G. H. M. B. Van Rens, "Association between Contrast Sensitivity and Reading with Macular Pathology," *Optom. Vis. Sci.*, vol. 95, no. 3, pp. 183–192, 2018, doi: 10.1097/OPX.0000000000001178.
[11] P. T. Harvey, "Common eye diseases of elderly people: Identifying and treating causes of vision loss," *Gerontology*, vol. 49, no. 1, pp. 1–11, 2003, doi: 10.1159/000066507.
[12] M. Kwon, C. Ramachandra, P. Satgunam, B. W. Mel, E. Peli, and B. S. Tjan, "Contour Enhancement Benefits Older Adults with Simulated Central Field Loss," *Optom Vis. Sci.*, vol. 89, no. 9, pp. 1374–1384, 2012, doi: 10.1097/OPX.0b013e3182678e52.
[13] Y. Zhao, S. Szpiro, L. Shi, and S. Azenkot, "Designing and evaluating a customizable head-mounted vision enhancement system for people with low vision," *ACM Trans. Access. Comput.*, vol. 12, no. 4, 2019, doi: 10.1145/3361866.
[14] J. R. Ehrlich *et al.*, "Head-Mounted Display Technology for Low-Vision Rehabilitation and Vision Enhancement," *Am. J. Ophthalmol.*, vol. 176, pp. 26–32, 2017, doi: 10.1016/j.ajo.2016.12.021.
[15] M. Christen and M. Abegg, "The effect of magnification and contrast on reading performance in different types of simulated low vision," *J. Eye Mov. Res.*, vol. 10, no. 2, 2017, doi: 10.16910/jemr.10.2.5.
[16] A. R. Bowers, C. Meek, and N. Stewart, "Illumination and reading performance in age-related macular degeneration," *Clin. Exp. Optom.*, vol. 84, no. 3, pp. 139–147, 2001, doi: 10.1111/j.1444-0938.2001.tb04957.x.
[17] J. S. Wolffsohn, D. Mukhopadhyay, and M. Rubinstein, "Image Enhancement of Real-Time Television to Benefit the Visually Impaired," *Am. J. Ophthalmol.*, vol. 144, no. 3, 2007, doi: 10.1016/j.ajo.2007.05.031.
[18] R. Szeliski, *Computer Vision: Algorithms and Applications*, vol. 142. Washington, USA: Springer, 2011.
[19] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 38, no. 1. p. 99, 1987, doi: 10.1016/s0734-189x(87)80156-1.
[20] A. Eshaghzadeh and R. S. Kalantari, "Canny Edge Detection Algorithm Application for Analysis of the Potential Field Map," *Earth Sci. India*, vol. 10, no. 3, 2017, doi: 10.31870/esi.10.3.2017.5.
[21] R. Smith, "An Overview of the Tesseract OCR Engine BT - ICDAR 07: Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2," *Lect. Google Code. Google Inc*, pp. 629–633, 2007, [Online] Available: http://portal.acm.org/citation.cfm?id=1304846%5Cn/home/kermorvant/refbase_files/2007/smith/1152_smith2007.pdf.
[22] R. Bansal, G. Raj, and T. Choudhury, "Blur image detection using Laplacian operator and Open-CV," *Proc. 5th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2016*, pp. 63–67, 2017, doi: 10.1109/SYSMART.2016.7894491.
[23] L. A. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*. 2004.
[24] U. Sara, M. Akter, and M. S. Uddin, "Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study," *J. Comput. Commun.*, vol. 07, no. 03, pp. 8–18, 2019, doi: 10.4236/jcc.2019.73002.
[25] A. Rodriguez-molares, O. Marius, H. Rindal, D. Jan, A. Austeng, and H. Torp, "The Generalized Contrast-to-Noise Ratio: A Formal Definition for Lesion Detectability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 4, pp. 745-759, 2020.

BIOGRAPHY

Rachmad Setiawan earned his bachelor degree in Electronic Engineering from Institut Teknologi Sepuluh Nopember Surabaya (ITS) in 1989, his master degree in Instrumentation from Bandung Institute of Technlogy (ITB) in 1996, and his doctoral degree in Electronic Engineering from ITS in 2010. He started giving Electronic lectures in ITS in 1995, and joined ITS Biomedical Engineering Department in 2015. His main research fields are embedded microcontroller dan instrumentation.

Rayhan Akmal Fadlurahman is currently taking his bachelor degree in Biomedical Engineering from ITS. He started his study in 2018 and took his interest in Virtual Reality and Augmented Reality research fields.

Nada Fitrieyatul Hikmah earned a bachelor's degree in Biomedical Engineering at Universitas Airlangga, Indonesia in 2012. She continued her Masters's degree at Electrical Engineering, Institut Teknologi Sepuluh Nopember (ITS), and received a master's degree in 2016. From 2014-2016, she had conducted research on Multimodal Analysis of Heart Signals (ECG, PCG and Carotid Pulse) for Classification of Normal and Abnormal Hearts. Her current activity is to become a lecturer at Department of Biomedical Engineering, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember (ITS), Indonesia.