

Article

Fundus Image Registration Technique Based on Local Feature of Retinal Vessels

Roziana Ramli ^{1,*}, Khairunnisa Hasikin ^{2,*} , Mohd Yamani Idna Idris ¹, Noor Khairiah A. Karim ³ 
and Ainuddin Wahid Abdul Wahab ¹ 

¹ Department of Computer System and Technology, Faculty of Computer Science and Information Technology, Universiti Malaya, Kuala Lumpur 50603, Malaysia; yamani@um.edu.my (M.Y.I.I.); ainuddin@um.edu.my (A.W.A.W.)

² Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia

³ Regenerative Medicine Cluster/Imaging Unit, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Kepala Batas 13200, Malaysia; drkhairiah@usm.my

* Correspondence: roziana.ramli@um.edu.my (R.R.); khairunnisa@um.edu.my (K.H.)

Abstract: Feature-based retinal fundus image registration (RIR) technique aligns fundus images according to geometrical transformations estimated between feature point correspondences. To ensure accurate registration, the feature points extracted must be from the retinal vessels and throughout the image. However, noises in the fundus image may resemble retinal vessels in local patches. Therefore, this paper introduces a feature extraction method based on a local feature of retinal vessels (CURVE) that incorporates retinal vessels and noises characteristics to accurately extract feature points on retinal vessels and throughout the fundus image. The CURVE performance is tested on CHASE, DRIVE, HRF and STARE datasets and compared with six feature extraction methods used in the existing feature-based RIR techniques. From the experiment, the feature extraction accuracy of CURVE (86.021%) significantly outperformed the existing feature extraction methods ($p \leq 0.001^*$). Then, CURVE is paired with a scale-invariant feature transform (SIFT) descriptor to test its registration capability on the fundus image registration (FIRE) dataset. Overall, CURVE-SIFT successfully registered 44.030% of the image pairs while the existing feature-based RIR techniques (GDB-ICP, Harris-PIIFD, Ghassabi's-SIFT, H-M 16, H-M 17 and D-Saddle-HOG) only registered less than 27.612% of the image pairs. The one-way ANOVA analysis showed that CURVE-SIFT significantly outperformed GDB-ICP ($p = 0.007^*$), Harris-PIIFD, Ghassabi's-SIFT, H-M 16, H-M 17 and D-Saddle-HOG ($p \leq 0.001^*$).

Keywords: image registration; fundus image; feature extraction



Citation: Ramli, R.; Hasikin, K.; Idris, M.Y.I.; A. Karim, N.K.; Wahab, A.W.A. Fundus Image Registration Technique Based on Local Feature of Retinal Vessels. *Appl. Sci.* **2021**, *11*, 11201. <https://doi.org/10.3390/app112311201>

Academic Editors:
Alessandro Stefano, Albert Comelli and Federica Vernuccio

Received: 30 September 2021
Accepted: 19 November 2021
Published: 25 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Retinal fundus image registration (RIR) is an essential tool in facilitating the diagnosis and treatment of retinal diseases [1]. RIR aligns fundus images according to geometrical transformation estimated from correspondence between fixed and moving images. Existing RIR techniques can be grouped based on the type of correspondence utilized in estimating the geometrical transformation, namely, intensity-based and feature-based.

The intensity-based RIR technique searches the similarity between the intensity patterns in fixed and moving images to estimate the geometrical transformation. The similarity between the intensity patterns is established using a similarity metric such as mutual information [2], cross-correlation [3] and phase correlation [4,5]. However, the registration performance of the intensity-based RIR technique is limited in the presence of non-uniform intensity distribution and homogenous texture [6], which is commonly observed in fundus images. Furthermore, the intensity patterns from the non-overlapping area can mislead the similarity metric in estimating inaccurate geometrical transformation.

Generally, the feature-based RIR technique is more reliable and robust in registering fundus images compared to the intensity-based RIR technique. This is because the feature-based RIR technique estimates the geometrical transformation according to the correspondence of local features such as feature points. However, the feature-based RIR technique requires the feature points to be extracted from reliable information to ensure accurate registration. Reliable information is distributed throughout an image and repeatable despite the changes in viewpoint or intensity [7]. The feature-based RIR technique is mainly comprised of feature extraction, feature descriptor, matching and estimating geometrical transformation. Feature extraction plays a crucial role in ensuring the feature points are detected and selected from reliable information by examining the local patches.

The feature extraction method in the existing feature-based RIR techniques extracts feature points from retinal vessels [8], vessel bifurcations [9], corner [10], extrema [11–13] or distinctive structure information [14]. Among this information, the retinal vessel is the most reliable because it can be found throughout the fundus image and is repeatable despite the changes in viewpoint or intensity. Additionally, the appearance of the retinal vessels within the local patches are consistent as a continuous line in 2-dimensional, and curvature shape in 3-dimensional, despite its size and contrast. However, the noises such as the retinal nerve fiber layer, underlying choroidal vessels, microaneurysm and exudates can also appear as curvature shapes in the local patches.

Therefore, this paper introduces a new feature extraction method based on the local feature of retinal vessels (CURVE). The proposed CURVE extracts feature points throughout the fundus image with the ability to discriminate the aforementioned noises. To register the fundus images, a feature-based RIR technique framework (CURVE-SIFT) is described where CURVE is paired with the scale-invariant feature transform (SIFT) descriptor [15].

The remainder of this paper is organized as follows. Section 2 highlights and discusses the issues of the feature extraction method in the existing feature-based RIR techniques. Section 3 describes the methodology of the CURVE-SIFT technique. The experimental settings in developing and evaluating the CURVE-SIFT technique are presented in Section 4. Section 5 reports and discusses the experimental results. Finally, the conclusion and future work are given in Section 6.

2. Related Works

The majority of the existing feature-based RIR techniques [13,16–18] mainly utilized the SIFT detector [15] to extract the feature points. SIFT detector finds extrema from local patches in a hierarchical difference of Gaussian (DoG) scale-space to allow feature points to be found based on the structure of various sizes. Then, the extrema that are low contrast and on edges are rejected to ensure the final feature points are distinctive and repeatable. However, the retinal vessels exhibit inconsistent contrast levels throughout the fundus image. Therefore, Ghassabi et al. [11] utilized robust uniform SIFT (UR-SIFT) [19] to overcome this issue.

The UR-SIFT is an improvement of the SIFT detector, where the feature points are selected according to the strength of the texture surrounding the points. This enables UR-SIFT to be more efficient in extracting feature points on retinal vessels compared to the standard SIFT detector. Furthermore, UR-SIFT ensures the extracted feature points are distributed throughout the hierarchical DoG scale space. The distribution is set in reverse from the scale coefficients of the scale space. This results in more feature points being extracted in the lower part of the hierarchical DoG scale space where the images are larger and finer. Opposite to this, fewer feature points are extracted in the upper part of the hierarchical DoG scale space where the images are smaller and coarser.

Ghassabi et al. further improved their work by introducing a stability score as part of the selection criterion [8]. The stability score incorporates Frangi's vesselness measure (FVM) [20], a vessel enhancement filter that suppresses noise in the image. Incorporating FVM enabled the ability of [8] to discriminate between retinal vessels and noises.

Extracting feature points on retinal vessels from the underexposed region in the fundus image is addressed in [12], where the illumination invariant Difference of Gaussian (*iiDoG*) operator was incorporated into the hierarchical scale space [21]. *iiDoG* operator is composed of normalized difference of Gaussian (*nDoG*) and DoG operators based on a piecewise function. The combination of these operators increases the visibility of the underexposed region while leaving the correctly exposed region unchanged. This work utilized a similar approach as in SIFT detector to extract extrema from the hierarchical *iiDoG* scale space. A threshold is introduced to discard the extrema on the retinal surface before the final feature points are selected. The threshold is based on the distribution of the intensity in the local patch.

Other than SIFT, the existing feature-based RIR techniques [10,22–24] extract geometric corner [25], Harris corner [26] and speeded up robust features (SURF) [27,28]. Meanwhile, Ramli et al. [14] introduced D-Saddle to extract feature points from the low-quality region based on distinctive structural information.

There are several issues that can be outlined from the highlighted feature extraction methods. First, feature enhancement algorithms such as DoG and *iiDoG* operators are mainly incorporated in building the hierarchical scale space. These operators increase the visibility of the retinal vessels as well as the noises, which make it more challenging for the feature extraction method to discriminate between them.

Second, the feature extraction methods are mainly without a proper selection module to select feature points on retinal vessels. A proper selection module should consider both retinal vessels and noise information as they may appear similarly within a local patch. Therefore, considering both in the selection module allows for more robust discrimination between retinal vessels and noises.

3. Methodology

The CURVE-SIFT technique constitutes five main stages, as shown in Figure 1. Stage 1 converts the input images to grayscale. The proposed CURVE in Stage 2 extracts feature points from the input grayscale images, which also highlights the main contribution of this paper. Stage 3 computes the SIFT descriptor to describe the surrounding region of each CURVE feature point. From the computed descriptors, matches are established, and outliers are removed in Stage 4. Finally, Stage 5 estimates the geometrical transformation between fixed and moving images. The details of these stages are explained in the following sub-sections. The mathematical symbols and notation used in this section are listed in Appendix A (Table A1).

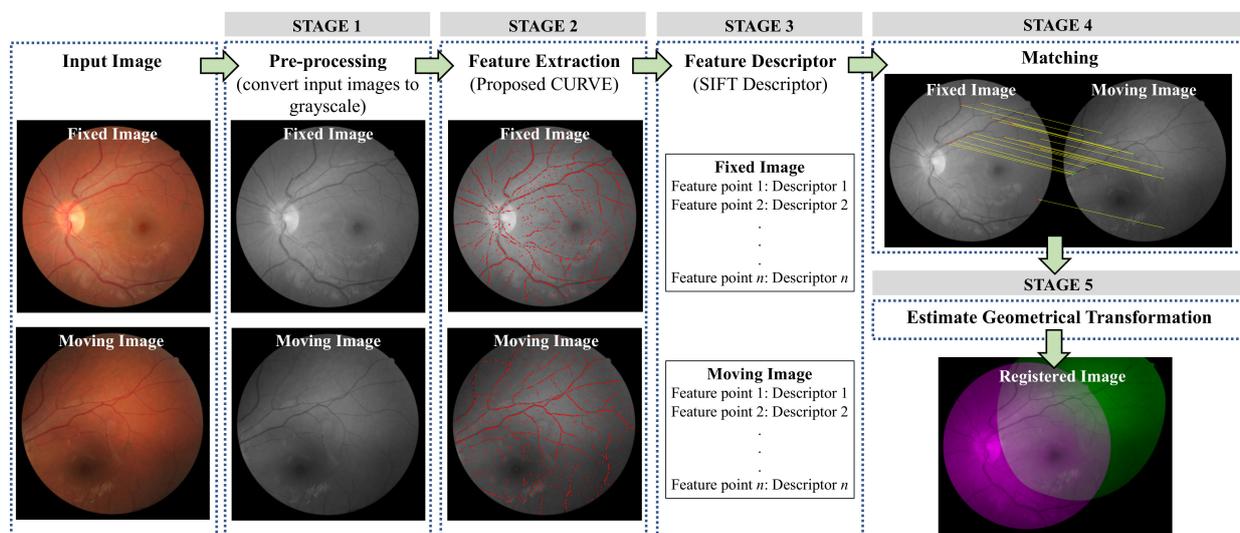


Figure 1. A general framework of the CURVE-SIFT technique.

3.1. STAGE 1: Pre-Processing

The conversion of the input images from color to grayscale follows the calculation of luminance in Recommendation ITU-R BT.601-7 [29] given below:

$$I = 0.2989\mathcal{R} + 0.587\mathcal{G} + 0.1140\mathcal{B} \quad (1)$$

where, I is the input image in grayscale, \mathcal{R} is the red channel, \mathcal{G} is the green channel and \mathcal{B} is the blue channel. The grayscale conversion based on luminance was chosen for this study because it has been shown to be superior to other grayscale conversions in terms of highlighting texture visibility [30] and trade-off between accuracy and processing cost.

3.2. STAGE 2: Feature Extraction

This sub-section describes the proposed CURVE to extract feature points on retinal vessels. CURVE is composed of feature detection and feature selection modules. The feature detection module detects candidate feature points according to the curvature shape of the retinal vessels. The curvature shape of the retinal vessels is observed when its grayscale image is depicted in 3D (see Table A2 in Appendix B). However, the detected candidate feature points are located on retinal vessels as well as noises. Therefore, the feature selection module removes the detected candidate feature points associated with noises by considering the unique characteristics of both retinal vessels and noises in intensity profiles. Then, the final feature points are chosen based on the strength of the retinal vessels. The steps in the feature detection and feature selection modules are summarized in Figure 2.

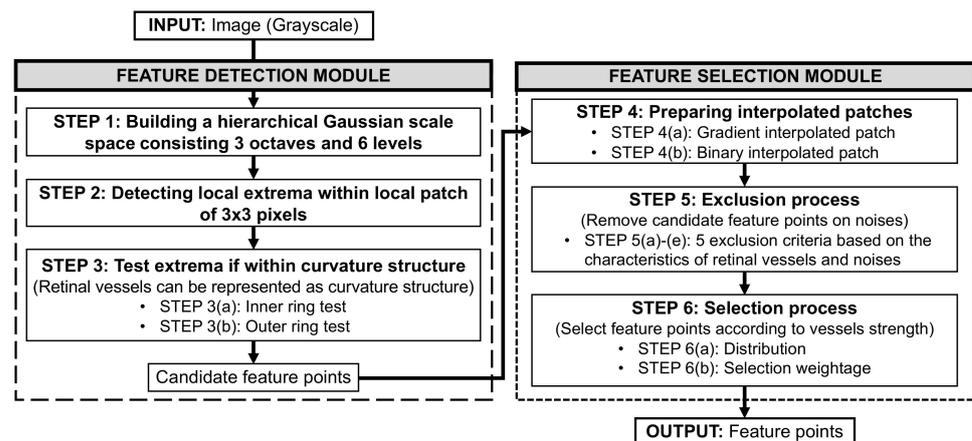


Figure 2. Overview of CURVE feature extraction in Stage 2. CURVE is composed of a feature detection module and feature selection module.

3.2.1. Feature Detection Module

The feature detection module examines local patches in the images of the hierarchical Gaussian scale space to detect extrema within the curvature shape of various sizes. This module involves three main steps, as explained below.

(a) STEP 1: Building a hierarchical Gaussian scale space

The initial step of the feature detection module is to build a hierarchical Gaussian scale space. The hierarchical Gaussian scale space enables the detection of the candidate feature points on various sizes of retinal vessels at the lower octave as the images are larger and finer with detailed information. At the higher octave, the candidate feature points are detected on thicker retinal vessels as the images are smaller and coarser with prominent information.

Building the hierarchical Gaussian scale space (G) involves generating three octaves ($P = 3 \mid p = 0, \dots, P - 1$) and six levels ($Q = 6 \mid q = -1, \dots, Q - 2$) per octave, as in [15,31].

The initial Gaussian image $G_{p,q}$ at $p = 0$ and $q = -1$ is created through convolution of input image I with width of relative Gaussian kernel $\check{\sigma}_{p,q}$ at $p = 0$ and $q = -1$ as follows:

$$G_{0,-1} = I * \check{\sigma}_{0,-1} \tag{2}$$

with, $\check{\sigma}_{0,-1}$ is denoted by:

$$\check{\sigma}_{0,-1} = \sqrt{\sigma_{0,-1}^2 - \sigma_s^2} \tag{3}$$

The width of the relative Gaussian kernel $\check{\sigma}_{p,q}$ assumes the input image I is pre-filtered with a sampling Gaussian kernel $\sigma_s \geq 0.5$ [15]. Thus, $\sigma_{0,-1}$ can be expressed as in [15,31]:

$$\sigma_{0,-1} = \sigma_0 \cdot 2^{-1/Q-3} \tag{4}$$

where, $\sigma_0 = 1.6$ is the base width of the Gaussian kernel.

$$\check{\sigma}_{0,-1} = \sqrt{\sigma_{0,-1}^2 - \sigma_s^2} \tag{5}$$

To obtain $G_{p,-1}$ at higher octave $p \in [1, \dots, P-1]$, $G_{p-1,2}$ is downsampled by half. The subsequent $G_{p,q}$ at $p \in [0, \dots, P-1]$ and $q \in [0, \dots, Q-2]$ can be obtained from the convolution between $G_{p,-1}$ in the respective octave with the relative Gaussian kernel of width $\check{\sigma}_q$ given by:

$$\check{\sigma}_q = \sigma_0 \cdot \sqrt{2^{2q/Q-3} - 1} \tag{6}$$

(b) STEP 2: Detecting local extrema

The feature detection module continues with the detection of extrema within the local patches of 3×3 pixels. An extremum represents the maximum or minimum intensity value of the center pixel compared to the eight immediate neighbors in the local patch. The local patches across the image are overlapped by $1/3$ of its size. The extrema found near the border of the field of view (FOV) are excluded from further processing using a mask image.

(c) STEP 3: Test extrema if within curvature structure

The retinal vessels generally exhibit curvature shape in 3-dimensions. Therefore, the extrema are tested if they are within the curvature structure by performing two tests as reported in [32]. These tests are the inner ring test and outer ring test.

- STEP 3(a): Inner ring test

The inner ring test considers eight pixels surrounding an extremum ($a_j \mid j \in [1, \dots, 8]$), as depicted in Figure 3a. Four out of eight pixels are tested at a time for patterns \times and $+$, as shown in Figure 3b–e. These patterns are formed when the intensities of two opposing pixels are brighter (dark green dot) than the other two opposing pixels in orthogonal (pink dot). The extrema can pass this test with one or two patterns. Then, the central intensity value β is estimated by taking the median value of four pixels if the extremum passes with one pattern, and eight pixels if it passes with two patterns. The extrema that failed the inner ring test are eliminated.

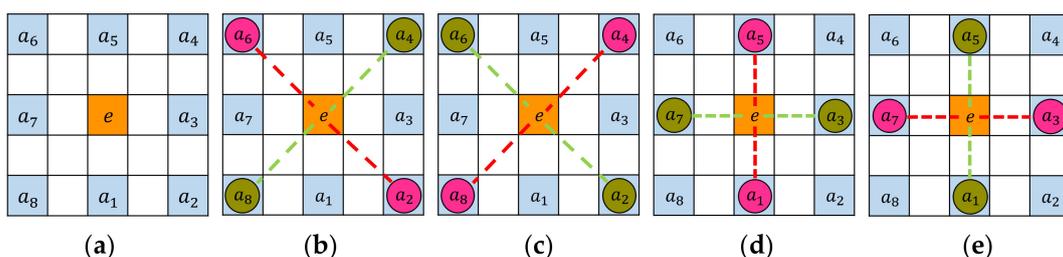


Figure 3. Inner ring test. (a) Eight pixels denoted by $a_j, j \in [1, \dots, 8]$ surrounding an extremum, e . (b,c) Patterns in the shape of \times . (d,e) Patterns in the shape of $+$. Pixels with a dark green dot have higher intensity values than pixels with a pink dot.

- STEP 3(b): Outer ring test

A circumference of 16 pixels surrounding an extremum that passes the inner ring test forms the outer ring pixels ($b_l \mid l \in [1, \dots, 16]$) as shown in Figure 4a. These pixels are divided into groups of *low*, *middle* and *high* as defined below:

$$\begin{aligned} \text{Group low (red dot)} &: I_{b_l} < \beta - \varepsilon \\ \text{Group middle (purple dot)} &: \beta - \varepsilon \leq I_{b_l} \leq \beta + \varepsilon \\ \text{Group high (green dot)} &: I_{b_l} > \beta + \varepsilon \end{aligned} \tag{7}$$

where, I_{b_l} is the intensity of the outer ring pixels and ε is the offset. The offset ε is set to 0.0010 as the intensity value of the pixels is in the range of $[0, 1]$ [14].

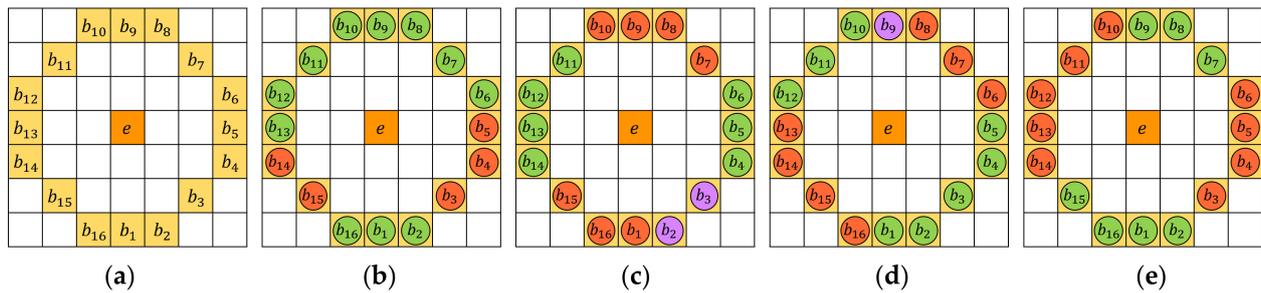


Figure 4. Outer ring test. (a) Sixteen pixels denoted by $b_l, l \in [1, \dots, 16]$ surrounding an extremum, e . (b–e) Examples of outer ring patterns. Pixels with red dot are from group *low*, pixels with purple dot are from group *medium* and pixels with green dot are from group *high*.

Then, the extrema are tested for the outer ring patterns consisting of consecutive and alternating arcs from groups *low* and *high*. The length of each arc can be between 2 to 8 pixels. These arcs can also be separated by pixels from group *middle* up to two pixels. Examples of the outer ring patterns are depicted in Figure 4b–e. The extrema that pass the outer ring test are the extrema found within the curvature structure, as shown in Figure 5. These extrema are assigned as candidate feature points and included in the feature selection module.

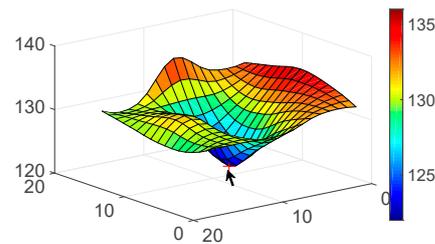


Figure 5. Example of candidate feature point (pointed by black arrow) from feature detection module. The candidate feature point is an extremum within a curvature structure.

3.2.2. Feature Selection Module

The feature selection module includes exclusion and selection processes. The exclusion process discards the candidate feature points associated with noises while the selection process selects the final feature points according to the strength of the retinal vessels. These processes require gradient and binary interpolated patches as input.

- (a) STEP 4: Preparing gradient and binary interpolated patches

The initial step of the feature selection module is to extract a square patch with the size of $s_p \times s_p$ pixels for each candidate feature point from the respective $G_{p,q}$. The size of the patch is varied depending on the octave position (p) of the candidate feature point to

ensure the retinal vessel can be captured within the patch despite the image size of $G_{p,q}$. The side length (s_p) of the patch is an odd number computed as follows:

$$s_p = s_{initial} - 4(p + 1) \tag{8}$$

where, $s_{initial}$ is the initial side length. There are three possible values for $s_{initial}$ as defined in (9). $s_{initial}$ is set by referring to the size of the initial Gaussian image $G_{0,-1}$. These values are determined by observing the retinal vessels with the thickest width on the fundus images from five datasets; CHASE_DB1 [33,34], DRIVE [35,36], HRF [37,38], STARE [39,40] and Fundus Image Registration (FIRE) dataset [41]. Furthermore, by considering scale or zoom less than 1.5 [8]. The $s_{initial}$ is suitable for input images larger than that of the largest image used to determine $s_{initial}$ (10 megapixels). This is because hierarchical Gaussian scale space down-sampled the input image by half as the level increased and reduced the image details as the octave increased, allowing the vessels of varying sizes to fit in the square patch even for input images larger than 10 megapixels.

$$s_{initial} \begin{cases} 35 \text{ pixels} & \text{if } G_{0,-1} > 1000 \times 1000 \text{ pixels} \\ 25 \text{ pixels} & \text{if } G_{0,-1} \leq 1000 \times 1000 \text{ pixels} > 600 \times 600 \text{ pixels} \\ 21 \text{ pixels} & \text{if } G_{0,-1} \leq 600 \times 600 \text{ pixels} \end{cases} \tag{9}$$

The extracted gradient patch is up-sampled using cubic interpolation with a refinement factor of two to smooth the region around the vessel edges. Then, this interpolated patch is converted to a binary image as depicted in Figure 6(a*ii*,b*ii*). These patches are used as input for exclusion and selection processes.

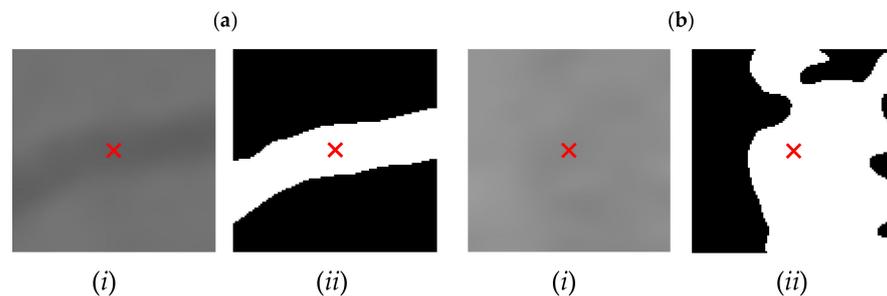


Figure 6. Examples of the (i) gradient and (ii) binary interpolated patches extracted from (a) retinal vessel and (b) noise. Red ‘x’ represents the position of the candidate feature point on the patch.

(b) STEP 5: Exclusion process

The curvature structure in the local patch represents retinal vessels of various sizes as well as noises such as the retinal nerve fiber layer, underlying choroidal vessels, microaneurysm and exudates. Therefore, five exclusion criteria specifying the characteristics of the retinal vessels and noises on the sum of intensity profiles are presented to discard candidate feature points on noises.

The intensity profile is the intensity value of the pixels extracted from a cross-sectional line running through the patch. In this study, the intensity profiles extracted from multiple cross-sectional lines are summed to distinctively highlight the characteristics of the retinal vessels and noises in the interpolated patch. The intensity profiles are extracted from a total of L_{total} cross-sectional lines that parallel each other with $L_{distance}$ distance between the lines. These cross-sectional lines are positioned either along or perpendicular to the main orientation of the interpolated patch. The main orientation is the angle between the x -axis and major axis of the ellipse on the prominently connected region of the binary interpolated patch.

The L_{total} , $L_{distance}$ and orientation of the cross-sectional lines are set according to the exclusion criteria as summarized in Table 1. The length of the cross-sectional lines in the

pixel can be determined from L_{total} and $L_{distance}$ to ensure the lines do not exceed the size of the interpolated patch in any orientation as follows:

$$L_{length} = s_{bin} - (L_{distance} \cdot L_{total}) \tag{10}$$

where, L_{length} is the length of the cross-sectional lines, s_{bin} is the side length of the binary interpolated patch, $L_{distance}$ is the distance between the parallel cross-sectional lines and L_{total} is the total cross-sectional lines.

Table 1. Settings and details of exclusion criteria in STEP 5.

		STEP 5(a): Exclusion Criterion 1	STEP 5(b): Exclusion Criterion 2	STEP 5(c): Exclusion Criterion 3	STEP 5(d): Exclusion Criterion 4	STEP 5(e): Exclusion Criterion 5
Settings to extract the sum of intensity profiles from interpolated patches						
Interpolated Patch		Binary	Gradient	–	–	Binary and gradient
Cross-sectional lines	L_{total}	5	7	–	–	7
	$L_{distance}$	3 pixels	5 pixels	–	–	5 pixels
	Orientation	Along main orientation	Perpendicular to main orientation	–	–	Perpendicular to main orientation
Details of exclusion criteria						
Input		Sum of intensity profiles from binary interpolated patch	Sum of intensity profiles from gradient interpolated patch	Valley with maximum depth from STEP 5(b)	Valley with maximum depth and global minimum from STEP 5(c)	Sums of intensity profiles from binary and gradient interpolated patches
Candidate feature point	On Vessels	A horizontal line. Figure 7(aii)	With at least a valley. Figure 8(aii)	Is global minimum. Figure 9(aii,bii)	At-axis. Figure 10a,b	Intersected when overlaid. Figure 11(aiii)
	On Noise	With at least a peak. Figure 7(bii)	Without valley. Figure 8(bii)	Is local minimum. Figure 9(cii)	At 1st or 4th section on x -axis. Figure 10c	Apart from each other when overlaid. Figure 11(biii)

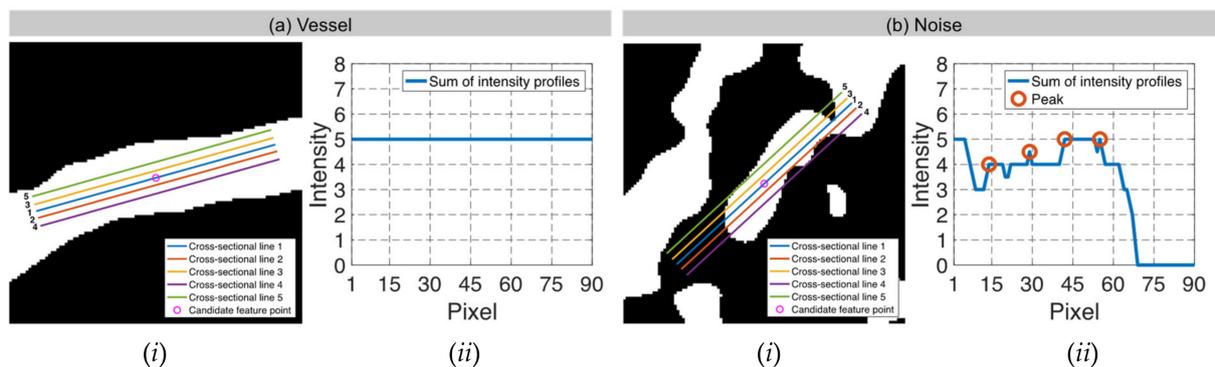


Figure 7. Exclusion criterion 1. (i) Cross-sectional lines and (ii) sum of intensity profiles for binary interpolated patch with (a) retinal vessel and (b) noise. A candidate feature point is discarded if any peak is found on the sum of intensity profiles from binary interpolated patch as in (b)(ii).

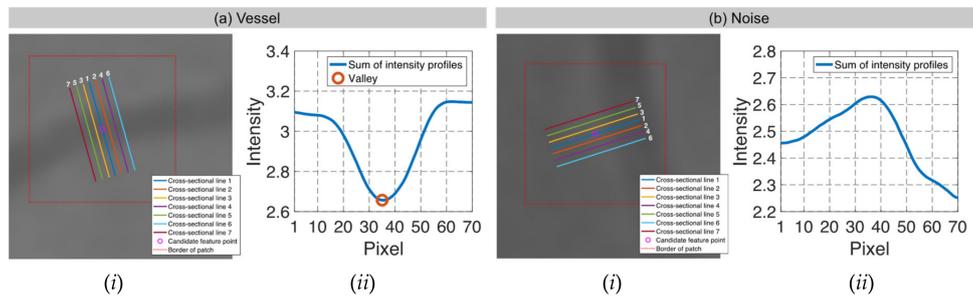


Figure 8. Exclusion criterion 2. (i) Cross-sectional lines and (ii) sum of intensity profiles for gradient interpolated patch with (a) retinal vessel and (b) noise. A candidate feature point is discarded if the sum of intensity profiles from gradient interpolated patch is without any valley as in (b)(ii).

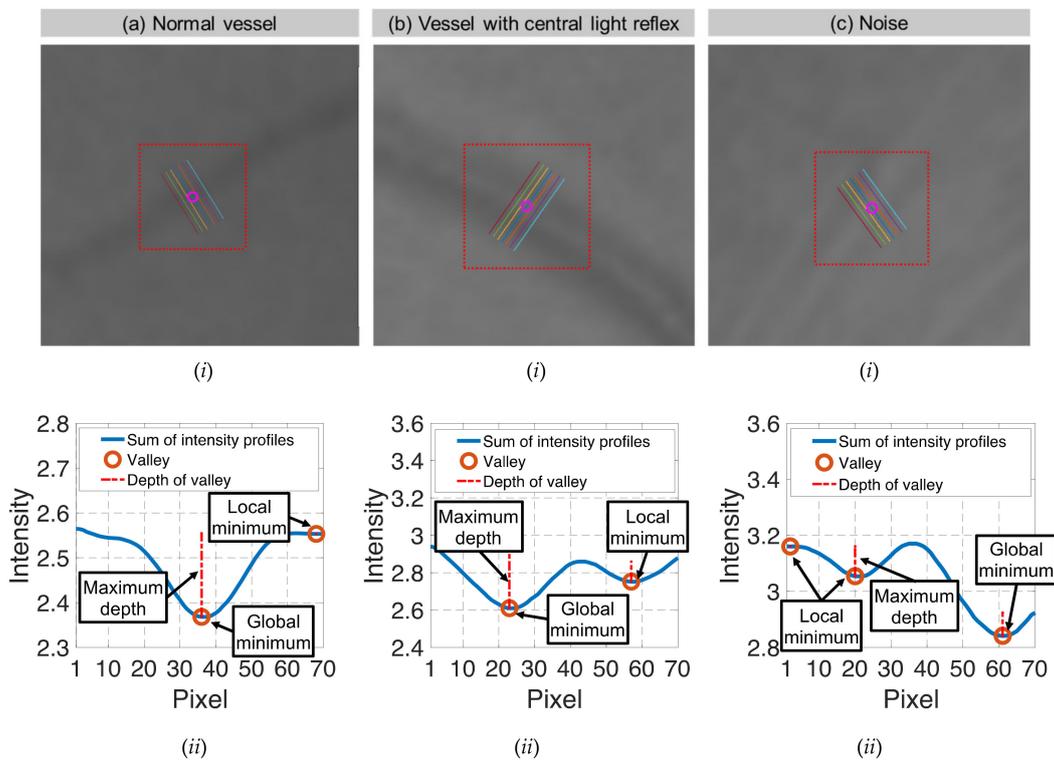


Figure 9. Exclusion criterion 3. (i) Cross-sectional lines and (ii) sum of intensity profiles for gradient interpolated patch with (a) normal retinal vessel, (b) retinal vessel with central light reflex and (c) noise. A candidate feature point is discarded when the valley with the maximum depth is a local minimum as in (c)(ii).

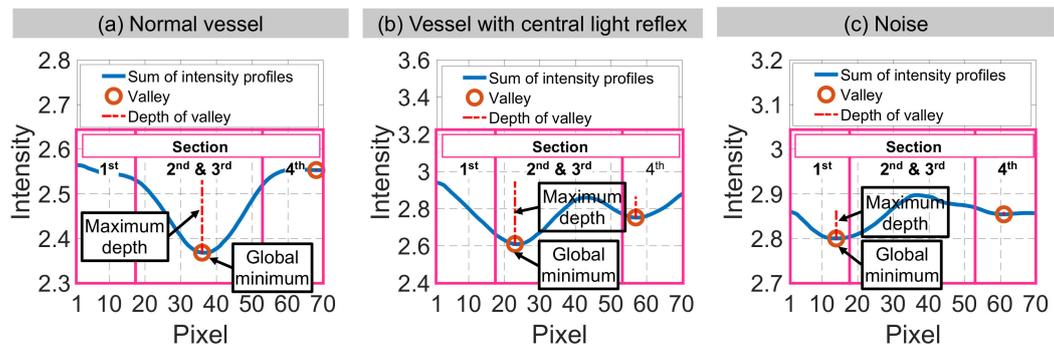


Figure 10. Exclusion criterion 4. (a,b) The valley with the maximum depth is on the 2nd or 3rd section for retinal vessels. (c) The valley with the maximum depth is on the 1st or 4th section for noise. A candidate feature point is discarded when the valley with the maximum depth is at the 1st or 4th section, as in (c).

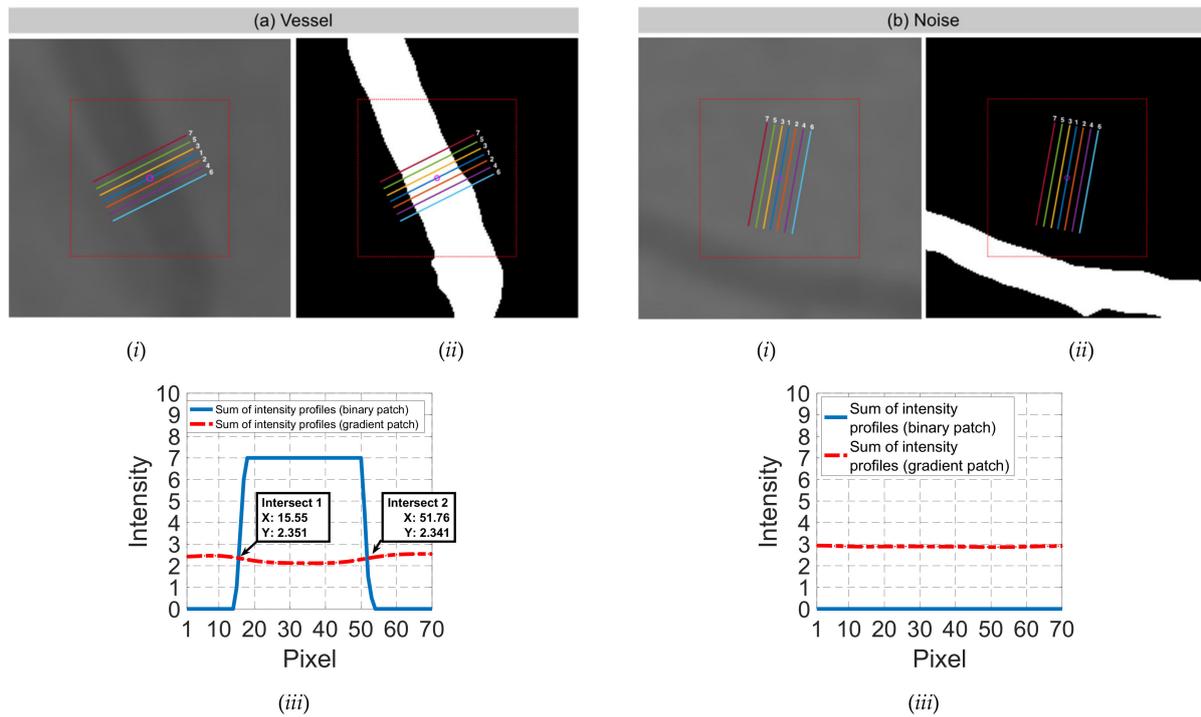


Figure 11. Exclusion criterion 5. Cross-sectional lines on (i) gradient and (ii) binary interpolated patches for (a) retinal vessel and (b) noise. (iii) The intersection between sums of intensity profiles from (i) and (ii). A candidate feature point is discarded when the overlaid sums of intensity profiles are apart from each other, as in (b)(iii).

$$L_{total}L_{distance}Atx$$

- STEP 5(a): Exclusion criterion 1

Retinal vessel in a binary interpolated patch forms a nearly straight and wide connected region, as depicted in Figure 7(ai). This characteristic can be represented by the sum of the intensity profiles extracted from five cross-sectional lines positioned along the main orientation of the patch. The L_{total} , $L_{distance}$ and orientation for these cross-sectional lines are chosen to best express the retinal vessel of various sizes in the patch. For retinal vessels, the sum of the intensity profiles appears as a horizontal line, as depicted in Figure 7(aii). Contrarily, the noise comprises an inconsistent connected region, as shown in Figure 7(bi), which results in the detection of peaks in the sum of intensity profiles. Therefore, a candidate feature point with peaks on the sum of intensity profiles is discarded.

- STEP 5(b): Exclusion criterion 2

For the gradient interpolated patch associated with the retinal vessel as in Figure 8(ai), the cross-sectional lines with $L_{total} = 7$, $L_{distance} = 5$ and positioned perpendicular to the main orientation are fully intersected by the vessel. Therefore, the sum of the intensity profiles from these cross-sectional lines will consist of at least a valley, as depicted in Figure 8(aii). In opposite, no valley can be found on the sum of the intensity profiles extracted from the patch associated with noise, as shown in Figure 8(bii). Thus, this candidate feature point is discarded from further processing.

- STEP 5(c): Exclusion criterion 3

The valleys discovered in STEP 5(b) are further examined for their depth and positioned on the y -axis. For a candidate feature point located on a retinal vessel, the valley with the maximum depth is at the lowest position of the y -axis or global minimum, as shown in Figure 9(aii,bii). Therefore, a candidate feature point is discarded if the valley with the maximum depth is a local minimum, such as in Figure 9(cii).

- STEP 5(d): Exclusion criterion 4

The valley with maximum depth and global minimum from STEP 5(c) is examined for its position on the x -axis. The sum of the intensity profiles is divided into four sections of equal size. The valley with the maximum depth is expected to be at the second or third section on the x -axis if a candidate feature point on a retinal vessel is either normal or with central light reflex as shown in Figure 10a,b. Therefore, a candidate feature point is excluded if the valley with the maximum depth is located at the first and fourth sections, as in Figure 10c.

- STEP 5(e): Exclusion criterion 5

This criterion overlaid the sum of the intensity profiles from gradient and binary interpolated patches. The intersection can be found when a candidate feature point is located on a retinal vessel and vice versa, as depicted in Figure 11. Thus, the candidate feature point is discarded when the overlaid sums of the intensity profiles are apart from each other.

- (c) STEP 6: Selection process

The exclusion process removes the majority of the candidate feature points detected on noises. However, the remaining candidate feature points may include points detected on noises with a high structural similarity as the retinal vessels in the interpolated patches. Therefore, the selection process includes two main steps to select the final feature points, namely, distribution and selection weightage. The distribution will ensure the final feature points are selected throughout the image, while the selection weightage highlights the strength of the retinal vessel in the patch for each candidate feature point.

- STEP 6(a): Distribution

The distribution of the feature points all over the image is vital to ensure a high registration accuracy [42]. There are two procedures involved in distributing the feature points. First, the feature points are distributed across the hierarchical Gaussian scale space by computing the maximum number of feature points ($N_{p,q}$) for each Gaussian image $G_{p,q}$. $N_{p,q}$ is set proportionally inverse to the width of the Gaussian kernels used when building the scale space as described in [8,11,19]:

$$N_{p,q} = N_{total} \cdot F_{p,q} \tag{11}$$

The proportion of the feature points $F_{p,q}$ is given by:

$$F_{p,q} = \frac{f_0}{\mu^{(Q)p+q+1}} \tag{12}$$

The proportion in the initial image of the scale space f_0 and the constant factor μ can be expressed as:

$$f_0 = \frac{\mu^{P(Q)-1}}{\sum_{n=1}^{P(Q)} \mu^{n-1}} \tag{13}$$

$$\mu = 2^{\frac{1}{Q}} \tag{14}$$

where, P is the total octave with index $p \in [1, \dots, P - 1]$, Q is the total level with index $q \in [-1, \dots, Q - 2]$, n is the index of the images in the hierarchical Gaussian scale space and N_{total} is the total feature points in the hierarchical Gaussian scale space. In this study, N_{total} is set to 4500 points, which empirically shows to provide a reasonable amount of feature points to perform image registration. However, if the candidate feature points are detected at less than 4500 points, N_{total} is set to 90% of the total candidate feature points.

The second procedure distributes $N_{p,q}$ across partitioned grids in each Gaussian image $G_{p,q}$. This operation begins by partitioning $G_{p,q}$ into rectangle grids of 150×150 pixels. The maximum number of feature points N_u in a grid image of index u is computed as follows:

$$N_u = DC_u \cdot N_{p,q} \tag{15}$$

The distribution coefficient for a grid image (DC_u) represents a combination of three factors. These factors are entropy [43], peak deviation nonuniformity [44] and total candidate feature points detected in the grid image.

The first factor of the entropy (EG) [43] defines the texture of the grayscale grid image. The grid image with high contrast retinal vessels, regardless of the sizes, will yield a large entropy value and vice versa. However, the entropy value presents a minimal distinction between the grid image with low contrast retinal vessels and with only noises or retinal surface.

Therefore, peak deviation nonuniformity (UG) [44] is included as the second factor. This factor is sensitive to the changes in the grayscale level. Thus, it is beneficial in distinguishing between the grid image containing low contrast vessels and the grid image with only noises.

In the coarser grid image, particularly at the higher octave, fewer candidate feature points are detected compared to the finer grid image. However, the values of the entropy and peak deviation nonuniformity measured from the coarser and finer grid images only show a minimal difference. To compensate for these factors, the total candidate feature points detected in the grid image (TG) is considered as the third factor.

The distribution coefficient DC_u for a grid image u can be expressed as the combination of the three factors:

$$DC_u = W_{EG} \frac{EG}{\sum_u EG} + W_{UG} \frac{UG}{\sum_u UG} + W_{TG} \frac{TG}{\sum_u TG} \quad (16)$$

where, W_{EG} is the weight factor for the entropy, W_{UG} is the weight factor for the peak deviation nonuniformity, W_{TG} is the weight factor for the total candidate feature points, u is the index of the grid image with $u \in [1, \dots, U]$ and U is the total grid in a Gaussian image $G_{p,q}$. The weight factors are empirically set to $W_{EG} = 0.3$, $W_{UG} = 0.3$ and $W_{TG} = 0.4$ to give a distinctive representation in describing the grid image.

- STEP 6(b): Selection weightage

The selection process is continued by computing selection weightage for each candidate feature point. The selection weightage highlights the strength of the retinal vessels indicated by entropy, area of the intersected region and the mean histogram of gradient orientation at the vessel edges.

The entropy (EP) is computed as in [43] to describe the texture in the gradient interpolated patch. Next, the area of the intersected region (AP) is determined between the sums of intensity profiles from the gradient and binary interpolated patches in STEP 5(e), as depicted in Figure 12. The lowest intersection point on the y -axis is used as the reference level to approximate the area of the intersected region using the trapezoidal rule. The area of the intersected region expresses the strength of the retinal vessels in terms of size and contrast. For example, the intersected region has a larger area for a thicker and high contrast retinal vessel. The area decreases as the size and contrast of the retinal vessel decreases.

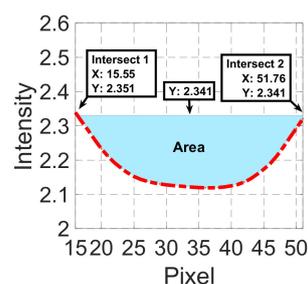


Figure 12. Area of intersected region between the sums of the intensity profiles from *exclusion criterion 5*.

The mean histogram of the gradient orientation at the vessel edges (HP) is estimated using both gradient and binary interpolated patches. Initially, the partial derivative is performed on the gradient interpolated patch to obtain gradient orientation for each pixel. The partial derivative is approximated using the central difference as it gives a more accurate approximation compared to other techniques, such as forward and backward approximations.

Then, the binary interpolated patch is used to obtain the vessel edges by performing binary dilation to increase the thickness of the edges. Once the pixels on the vessel edges are identified, the gradient orientation is extracted. The gradient orientation for these pixels is organized into a histogram of 36 bins, as shown in Figure 13. In this histogram, the non-zero frequencies are averaged to represent the mean histogram of the gradient orientation at the vessel edges. The mean histogram will yield a high value for a high contrast retinal vessel as the edges are thicker and the gradient orientation is more uniform. In contrast, the mean histogram will yield a low value for the low contrast retinal vessel as the edges are thinner and the gradient orientation is less uniform.

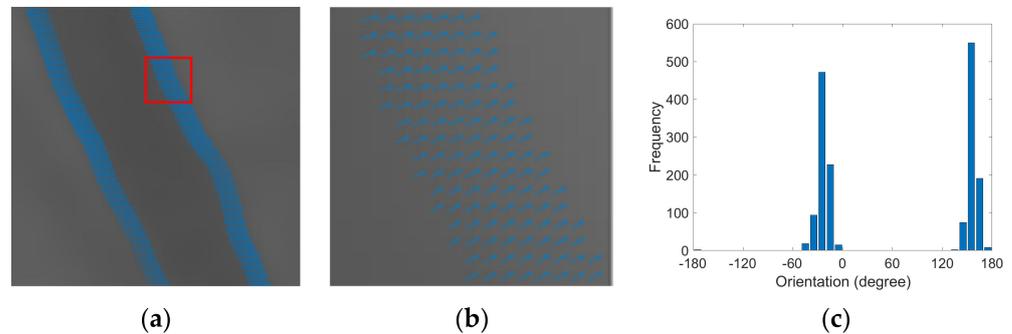


Figure 13. (a) Example of gradient orientation at the edges of the retinal vessel in a gradient interpolated patch. (b) Close-up from the red rectangle region. (c) Histogram of 36 bins generated for the gradient orientation in (a). The frequency in the histogram signifies the total occurrence of the gradient orientation within the respective bin.

The selection weightage denoted by SW_i is computed for each candidate feature point (i) to highlight the strength of the retinal vessels as expressed below:

$$SW_i = W_{EP} \frac{EP}{\sum_i^{TC} HP} + W_{AP} \frac{AP}{\sum_i^{TC} AP} + W_{HP} \frac{HP}{\sum_i^{TC} HP} \tag{17}$$

where, W_{EP} is the weight factor for the entropy, W_{AP} is the weight factor for the area of the intersected region, W_{HP} is the weight factor for the mean histogram of the gradient orientation at the vessel edges, i is the index of the candidate feature point with $i \in [1, \dots, TC]$ and TC is the total candidate feature point in a Gaussian image $G_{p,q}$. The weight factors are empirically set to $W_{EP} = 0.3$, $W_{AP} = 0.4$ and $W_{HP} = 0.3$ to distinctively highlight the strength of the retinal vessels.

Finally, a total of N_u candidate feature points with the highest value of the selection weightage SW are selected as feature points in each grid image. Then, the positions of the selected feature points are refined to sub-pixel accuracy at the respective $G_{p,q}$, as in [15,45]. The feature points with refined positions are converted from the position at the respective scale space to the coordinate system of the initial Gaussian image $G_{0,-1}$ follows:

$$K_m = 2^p \cdot K_{m,p,q} \tag{18}$$

where, K_m is the feature point of index m in the coordinate system of the initial Gaussian image $G_{0,-1}$ and $K_{m,p,q}$ is the feature point of index m in the coordinate system at the respective octave p and level q .

3.3. STAGE 3: Feature Descriptor

SIFT descriptor [15] is assigned to each feature point extracted from fixed and moving images. VLFeat toolbox [46] with default settings is used to compute the SIFT descriptor.

3.4. STAGE 4: Matching

The matches are obtained by establishing pairwise distances between SIFT descriptors. The distances are computed using the sum of squared differences (SSD). The outliers in the matches are eliminated using M-estimator SAmple Consensus (MSAC) algorithm [47]. MSAC eliminates the outliers when the distance between the matches in the fixed image and the projected matches from the moving image exceeds a specified threshold. The projection is performed according to the non-reflective similarity transformation and estimated from two randomly selected matches. The distance threshold is set between 1 and 100 with an increasing step of 0.1. The random trial is repeated 5000 times, and the desired confidence is set to 99%.

3.5. STAGE 5: Geometrical Transformation

Similarity and local weighted mean transformations [48] are estimated for each image pair from the established inliers. Only the transformation that gives the best registration accuracy is chosen for evaluation. The radius of influence for local weighted mean transformation is set in the range of 10 to the total inliers with an increasing step of two.

4. Experimental Setup

The CURVE-SIFT was implemented in MATLAB R2016b running on a virtual machine from Google Cloud Engine with specifications of Intel Xeon® E5 2.6GHz (24 vCPUs) and 40 GB of RAM. Toolboxes employed were Image Processing, Computer Vision, Signal Processing and VLFeat [46].

The evaluation was divided into two parts. First, CURVE was evaluated in extracting feature points on retinal vessels. The performance of CURVE was compared with five feature extraction methods from the existing feature-based RIR techniques, namely, Harris corner [26], SIFT [15], SURF [27,28], Ghassabi's [8] and D-Saddle [14]. Then, CURVE-SIFT was evaluated in registering image pairs from three retinal image registration applications and compared with five existing feature-based RIR techniques; GDB-ICP [13], Harris-PIIFD [10], Ghassabi's-SIFT [8], H-M 16 [16], H-M 17 [9] and D-Saddle-HOG [14]. In the experiment, these five existing feature-based RIR techniques were utilized exactly as they are. H-M 16, H-M 17 and D-Saddle-HOG were originally developed using FIRE dataset for super-resolution, image mosaicking and longitudinal study applications, while Ghassabi's-SIFT, GDB-ICP and Harris-PIIFD were developed for image mosaicking and low-quality image using other datasets.

4.1. Datasets

A total of five public datasets at the original image size were employed in the evaluation. The original image size was used in the experiment as decreasing the spatial resolution of the fundus image can degrade its quality and led to an inaccurate diagnosis and treatment of retinal diseases [49]. Four of the datasets evaluated the feature extraction performance, namely, CHASE_DB1 [33,34], DRIVE [35,36], HRF [37,38] and STARE [39,40]. These datasets contain fundus images affected by pathological cases. The provided ground truth images are in the form of the segmented vessels performed by experts. This enables the evaluation of the extracted feature points on the retinal vessels. The details of CHASE_DB1, DRIVE, HRF and STARE datasets are described in Table 2.

Table 2. Descriptions of CHASE_DB1, DRIVE, HRF and STARE datasets for evaluating feature extraction performance.

Descriptions	Datasets			
	CHASE_DB1	DRIVE	HRF	STARE
Total images	28	40	45	20
Image size (pixels)	999 × 960	564 × 584	3504 × 2336	605 × 700
Total patients	14	40	45	20
Age (Years)	9–10	25–90	N/A	N/A
Pathological cases	Vessel tortuosity	33 images without sign of diabetic retinopathy 7 images with mild early diabetic retinopathy	15 images of healthy patients 15 images of diabetic retinopathy 15 images of glaucomatous	Abnormalities that obscure the blood vessel appearance, such as hemorrhaging, etc.
Field of view	30°	45°	45°	35°
Year	2012	2004	2009	2000
Ground truth images	56	60	45	40
Intensity distribution ¹	22.6136	49.3307	34.9433	49.5126

¹ Described by peak deviation nonuniformity intensity. Values close to 0 indicates non-uniform intensity distribution in the image.

The registration performance of CURVE-SIFT is evaluated in the Fundus Image Registration dataset (FIRE) [41]. This dataset is the only public fundus image registration dataset with ground truth annotation. The FIRE dataset consists of 134 image pairs divided into super-resolution, image mosaicking and longitudinal study applications, as described in Table 3. All image pairs are affected by diabetic retinopathy where vessel tortuosity, microaneurysms and cotton-wool are visible on the images. Each image pair includes 10 corresponding ground truth annotations identified by experts.

Table 3. Descriptions of FIRE dataset for evaluating registration performance.

Descriptions	Retinal Image Registration Applications		
	Super-Resolution	Image Mosaicking	Longitudinal Study
Total images	71	49	14
Image size (pixels)		2912 × 2912	
Total patients		39	
Age (Years)		19–67	
Pathological cases		Diabetic retinopathy	
Field of view		45°	
Year		2006 to 2015	
Ground truth images		10 corresponding points for each image pair	
Anatomical differences ¹	No	No	Yes
Scale	≈1	≈1	≈1
Overlapping area (%)	86–100	17–89	95–100
Rotation (°)	0°–12°	6°–52°	1°–4°

¹ Anatomical differences observed between fixed and moving images.

Registering image pairs from super-resolution, image mosaicking and longitudinal study applications involve a combination of several challenges, namely, overlapping area and rotation. The overlapping area is an intersection region between fixed and moving images. A small overlapping area limits the amount of similar information between images, which can be insufficient to estimate an accurate geometrical transformation. The rotation in the fundus image is introduced to access part of the retina or due to involuntary movement by the patient. The rotation alters the orientation of similar information between images. This alteration can be challenging for the feature-based RIR technique to establish correspondences.

The super-resolution application combines multiple fundus images with a large overlapping area and small rotation. The super-resolution application is performed to increase

the density of the spatial sampling, which can resolve the blurred edges of the retinal vessels caused by patient movements or improper imaging setup.

The image mosaicking application aligns multiple fundus images to generate an image with a wider view of the retina. The wide view image of the retina can be used to view the full extent of the retinal disease in one big picture during diagnosis [50,51] and during the preparation of eye laser treatment for diabetic retinopathy [52]. However, registering image pairs from the image mosaicking application can be challenging as it involves a combination of small overlapping areas and large rotation.

The longitudinal study application combines multiple fundus images that are acquired at different screening sessions. Therefore, the anatomical changes due to progression or remission of retinopathy such as increased vessel tortuosity, microaneurysms and cotton-wool spots can be observed between fixed and moving images. The longitudinal study application is essential in monitoring the progression of retinal diseases, such as glaucoma and age-related macular degeneration, which usually undergoes a long degeneration process [53].

4.2. Evaluation Metrics

4.2.1. Feature Extraction Performance

(a) Extraction accuracy

The extraction accuracy expresses the ability of a feature extraction method to extract feature points on retinal vessels. The extraction accuracy for an image can be computed by:

$$ExAc = \frac{\text{total feature points extracted on vessels}}{\text{total feature points}} \times 100\% \quad (19)$$

where, *ExAc* is the extraction accuracy in percentage.

The extraction accuracy for an image is set to 0% when the feature points extracted are below the minimum requirement of three points to perform a transformation. One-way Analysis of Variance (ANOVA) with Tukey's post hoc was performed to compare the extraction accuracy between methods.

(b) Factors influencing the extraction accuracy

Two factors influencing the feature extraction accuracy were investigated. These factors are changes in image size and intensity distribution throughout the image. The relations were investigated using Spearman's rank-order correlation. The image size and the intensity distribution of the fundus images in CHASE, DRIVE, HRF and STARE datasets are summarized in Table 2. The intensity distribution is described by peak deviation nonuniformity [44].

4.2.2. Registration Performance

(a) Success rate

Success rate measures the ability of a feature-based RIR technique to register image pairs and meet the specified requirement of target registration error (TRE). TRE is the mean distance in pixel between the ground truth annotations in a fixed image to the transformed ground truth annotations from the moving image. A perfect registration for an image pair is represented by TRE values equal to 0.

However, achieving a perfect registration can be challenging in a real-world application. Thus, the registration for an image pair is considered successful if the obtained TRE is below one pixel for super-resolution applications and five pixels for image mosaicking and longitudinal study applications [54]. The success rate can be computed as given below:

$$\text{Success rate} = \frac{\text{total image pairs with successful registration}}{\text{total image pairs}} \times 100\% \quad (20)$$

The one-way ANOVA with Tukey's post hoc was performed to compare the success rate between the feature-based RIR techniques.

(b) Factors influencing the success rate

Factors of overlapping area and rotation were investigated for their influence on the success rate using Spearman's rank-order correlation. It should be noted that for this evaluation, the successful registration was set below five pixels for all image pairs despite its registration application. As the details of the overlapping area and rotation are not initially provided by the FIRE dataset, this information is measured as follows.

The overlapping area in percentage is obtained from the overlap area between the fixed image and transformed moving image. The moving image is transformed to the orientation of the fixed image using affine transformation inferred from the corresponding ground truth annotations. The rotation for an image pair is measured from the average angle between corresponding ground truth annotations without considering the effect of translation, as in [14]. Thus, results were in the larger angle of rotation. The range of overlapping area and rotation in the FIRE dataset is provided in Table 3.

5. Results

5.1. Feature Extraction Performance

CURVE extracts an average of 2482 feature points from the CHASE, DRIVE, HRF and STARE datasets, where 2149 of them are accurately associated with retinal vessels. This constitutes an average feature extraction accuracy of 86.021% with a variation of 9.199% between images, as outlined in Table 4. Furthermore, the one-way ANOVA analysis shows that the feature extraction accuracy of CURVE was significantly outperformed by all existing feature extraction methods ($p < 0.001^*$). CURVE obtained the biggest accuracy difference with SIFT detector (69.857%) and the smallest difference with Harris corner (44.408%). Examples of CURVE feature points extracted from four datasets are depicted in Figure 14.

Table 4. Overall feature extraction accuracy (%) in CHASE_DB1, DRIVE, HRF and STARE datasets.

Feature Extraction Method	Total Images	Mean	Standard Deviation	Min	Max
Harris corner	133	41.613	21.317	0.000	92.857
SIFT detector	133	16.164	5.411	5.241	30.299
SURF	133	18.929	4.206	9.502	30.412
Ghassabi's	133	28.280	5.975	17.055	44.197
D-Saddle	133	20.509	4.791	12.221	31.273
CURVE	133	86.021	9.199	59.677	97.842

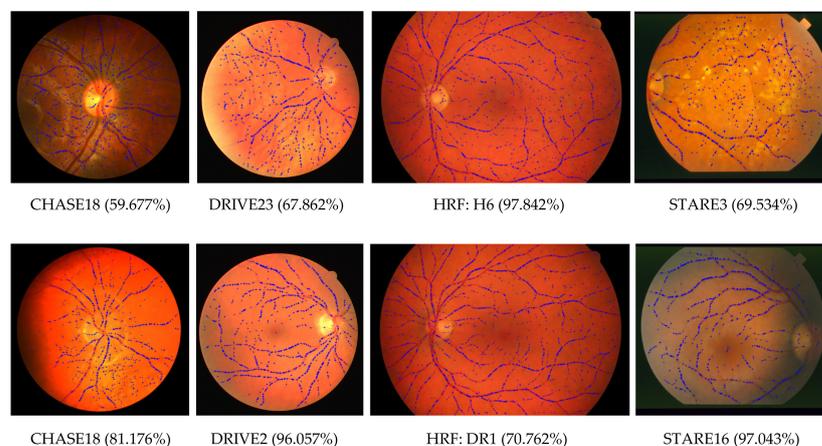


Figure 14. Examples of feature points extracted by CURVE. Top row: Images with the lowest extraction accuracy in each dataset. Bottom row: Images with the highest extraction accuracy in each dataset.

The high feature extraction accuracy of CURVE is contributed to by the utilization of both the retinal vessels and noise characteristics in the feature detection and selection modules. Thus, enabling accurate discrimination between the retinal vessels and noises. Contrarily, Ghassabi's and D-Saddle enhanced the fundus image to increase the visibility of the retinal vessels. However, this enhancement also increases the visibility of the noise, which led both methods to yield low feature extraction accuracy. The extracted feature points located on the noises for these methods were observed to be on the edge of the optic disc, retinal nerve fiber layer, underlying choroidal vessels and macula. The other feature extraction methods such as Harris corner, SIFT detector and SURF are without a specific feature selection module to extract feature points on retinal vessels. These feature extraction methods were used in the existing feature-based RIR techniques [10,13,16], where the authors focused on the development of the feature descriptor and transformation model.

Other than that, the minimal usage of rigid thresholds or variables allows CURVE to accurately extract feature points from fundus images with varying sizes. This is shown by the smallest Spearman's rho among all methods and insignificant correlation between the changes in image size and the extraction accuracy of CURVE ($r_s = -0.032$, $p = 0.712$) as presented in Table 5. Furthermore, the extraction accuracy of D-Saddle ($r_s = -0.138$, $p = 0.114$) and Ghassabi's ($r_s = -0.142$, $p = 0.104$) exhibit insignificant correlation with the changes in image size but their correlations are stronger than CURVE. In contrast, SIFT detector is very sensitive to the changes in image size among all methods where its feature extraction accuracy decreases in larger images ($r_s = -0.649$, $p < 0.001^*$).

Table 5. Correlation between extraction accuracy and factors.

Feature Extraction Method	Image Size		Intensity Distribution	
	r_s	p -Value	r_s	p -Value
Harris corner	-0.178	0.041 *	0.360	<0.001 **
SIFT detector	-0.649	<0.001 **	0.138	0.113
SURF	0.590	<0.001 **	-0.398	<0.001 **
Ghassabi's	-0.142	0.104	0.314	<0.001 **
D-Saddle	-0.138	0.114	0.386	<0.001 **
CURVE	-0.032	0.712	0.342	<0.001 **

r_s : Spearman's rho. **: Correlation is significant at the 0.01 level (2-tailed). *: Correlation is significant at the 0.05 level (2-tailed).

However, CURVE performance is significantly affected in the presence of non-uniform intensity distribution in the image ($r_s = 0.342$, $p < 0.001^*$). CURVE is sensitive towards the non-uniform intensity distribution because it highly depends on the intensity changes to locate the curvature of the retinal vessels in the feature detection module. Furthermore, CURVE does not incorporate any feature enhancement algorithm. The feature enhancement algorithms, such as DoG and *ii*DoG operators, can suppress the non-uniform intensity distribution and increase the visibility of the retinal vessels but at the cost of increasing the visibility of the noises. Thus, it is avoided in the proposed CURVE. Contrarily, the correlation between SIFT detector and the intensity distribution is not significant and the weakest among all feature extraction methods ($r_s = 0.138$, $p = 0.113$).

5.2. Registration Performance

The evaluation continues by accessing the registration performance of CURVE-SIFT and six existing feature-based RIR techniques [8–10,13,14,16]. From the experimental results outlined in Table 6, CURVE-SIFT successfully registered a total of 59 image pairs in the FIRE dataset with a success rate of 44.030%. The one-way ANOVA analysis shows that the success rate of CURVE-SIFT significantly outperformed GDB-ICP at $p = 0.007^*$, whereas Harris-PIIFD, Ghassabi's-SIFT, H-M 16, H-M 17 and D-Saddle-HOG at $p < 0.001^*$. The biggest success rate difference was observed between CURVE-SIFT and Harris-PIIFD (40.299%), while the smallest difference was with GDB-ICP (16.418%).

Table 6. Success rate (%) in the FIRE dataset.

Feature-Based RIR Technique	Total Image Pairs ¹	Mean	Standard Deviation	TRE (Pixels)	
				Min	Max
Overall					
GDB-ICP	37	27.612	44.875	2.354	10.416
Harris-PIIFD	5	3.731	19.024	3.319	1486.255
Ghassabi's-SIFT	17	12.687	33.407	3.082	322.616
H-M 16	22	16.418	37.183	2.857	410.087
H-M 17	26	19.403	39.694	2.920	60.875
D-Saddle-HOG	16	11.940	32.548	4.583	27.266
CURVE-SIFT	59	44.030	49.829	1.928	1016.330
Super-resolution					
GDB-ICP	17	23.944	42.978	0.486	4.575
Harris-PIIFD	2	2.817	16.663	0.785	12.850
Ghassabi's-SIFT	13	18.310	38.950	0.665	15.798
H-M 16	18	25.352	43.812	0.554	13.903
H-M 17	20	28.169	45.302	0.489	5.696
D-Saddle-HOG	10	14.085	35.034	0.748	9.327
CURVE-SIFT	28	39.437	49.219	0.613	9.696
Image Mosaicking					
GDB-ICP	16	32.653	47.380	1.946	6.323
Harris-PIIFD	0	0.000	0.000	10.041	3870.632
Ghassabi's-SIFT	0	0.000	0.000	7.358	578.494
H-M 16	0	0.000	0.000	7.976	129.658
H-M 17	1	2.041	14.286	3.327	41.192
D-Saddle-HOG	2	4.082	19.991	3.082	366.401
CURVE-SIFT	26	53.061	50.423	1.787	19.799
Longitudinal Study					
GDB-ICP	4	28.571	46.881	2.354	10.416
Harris-PIIFD	3	21.429	42.582	3.319	1486.255
Ghassabi's-SIFT	4	28.571	46.881	3.082	322.616
H-M 16	4	28.571	46.881	2.857	410.087
H-M 17	5	35.714	49.725	2.920	60.875
D-Saddle-HOG	4	28.571	46.881	4.583	27.266
CURVE-SIFT	5	35.714	49.725	1.928	1016.330

¹ Total image pairs with successful registration.

Moreover, the overall success rate of D-Saddle-HOG (14.085%) reported in this study is much lower than in [14] because this study evaluates D-Saddle performance on the FIRE dataset at the original image size of 2912×2912 pixels. Contrarily, the work presented in [14] evaluates D-Saddle-HOG performance on the FIRE dataset at the smaller image size of 583×583 pixels. The extraction accuracy of D-Saddle is insignificantly correlated to the changes in image size, as shown in Table 5. However, D-Saddle-HOG employed a Histogram of Oriented Gradients (HOG) descriptor [55] in its framework where a larger image can decrease the number of correct matches or inliers established between the computed HOG descriptor [56]. Insufficient amounts of the established inliers can lead to the estimation of inaccurate geometrical transformation.

The most noticeable performance of CURVE-SIFT is observed in the image mosaicking application. The image pairs from the image mosaicking application involved the combination of smaller overlapping areas (17–89%) and larger rotation (6° – 52°) in the dataset. Despite these challenges, the success rate of CURVE-SIFT (53.061%) is significantly outperformed for all existing feature-based RIR techniques ($p < 0.001^*$). This performance is contributed to by CURVE's ability to accurately extract feature points on retinal vessels and distribute them throughout the image to increase the chances of the inliers being established within the overlapping area. Furthermore, the employed SIFT descriptor has the ability to establish over 60% of inliers when the rotation is below 90° [57]. These abilities are also expressed in the established Spearman's rank-order correlations in Table 7, where CURVE-SIFT yields smaller Spearman's rho values indicating weaker correlations with the

overlapping area and rotation compared to Harris-PIIFD, Ghassabi's, H-M 16, H-17 and D-Saddle. In contrast, the existing feature-based RIR techniques recorded a much lower success rate with less than 32.653%, whereas Harris-PIIFD, Ghassabi's and H-M 16 were unable to register any of the image pairs in the image mosaicking application.

Table 7. Correlation between success rate and factors.

Feature-Based RIR Technique	Overlapping Area		Rotation	
	r_s	p -Value	r_s	p -Value
GDB-ICP	0.443	<0.001 **	−0.380	<0.001 **
Harris-PIIFD	0.732	<0.001 **	−0.723	<0.001 **
Ghassabi's-SIFT	0.795	<0.001 **	−0.766	<0.001 **
H-M 16	0.785	<0.001 **	−0.763	<0.001 **
H-M 17	0.773	<0.001 **	−0.765	<0.001 **
D-Saddle-HOG	0.769	<0.001 **	−0.745	<0.001 **
CURVE-SIFT	0.415	<0.001 **	−0.382	<0.001 **

r_s : Spearman's rho. **: Correlation is significant at the 0.01 level (2-tailed).

The image pairs from the super-resolution application are the least challenging in the FIRE dataset as they involve a large overlapping area (86–100%) and small rotation (0° – 12°). However, the super-resolution application requires a very accurate registration with a TRE of less than one pixel. For this reason, CURVE-SIFT only recorded a success rate of 39.437% in this application, where the TRE of the failed registration ranged between 1.003 pixels to 9.696 pixels. The success rate of CURVE-SIFT outperformed all existing feature-based RIR techniques evaluated in this study but only significant with Harris-PIIFD ($p < 0.001^*$), Ghassabi's-SIFT ($p = 0.030^*$) and D-Saddle-HOG ($p = 0.004^*$).

The image pairs from the longitudinal study application are the most challenging for CURVE-SIFT to register, where it obtained the lowest success rate (35.714%) among the applications in the FIRE dataset. Furthermore, no significant difference can be noted between the success rate of CURVE-SIFT and existing feature-based RIR techniques. This shows that the registration performance of CURVE-SIFT is affected when the anatomical appearance is varied between images in the pair. Particularly, CURVE-SIFT failed to register image pairs when the prominent differences of vessel thickness and tortuosity were observed between images. The difference in vessel thickness between fixed and moving images leads to different descriptors being computed for local features at the same part of the vessels. As a result, these local features were unable to establish a correspondence, resulting in low registration accuracy. In the event of increased tortuosity, the corresponding local features were appropriately established. However, the tortuosity causes the vessels to bend and alters the actual physical position of the vessels on the eyeball. Consequently, the registration was performed between local features on the same part of the vessels but at different physical positions, which resulted in high TRE. For existing feature-based RIR techniques, the invariant features utilized in their works were extracted throughout the image. Thus, minimize the impact of vessel thinning and tortuosity compared to our work. Examples of registered image pairs for CURVE-SIFT in each application are depicted in Figure 15.

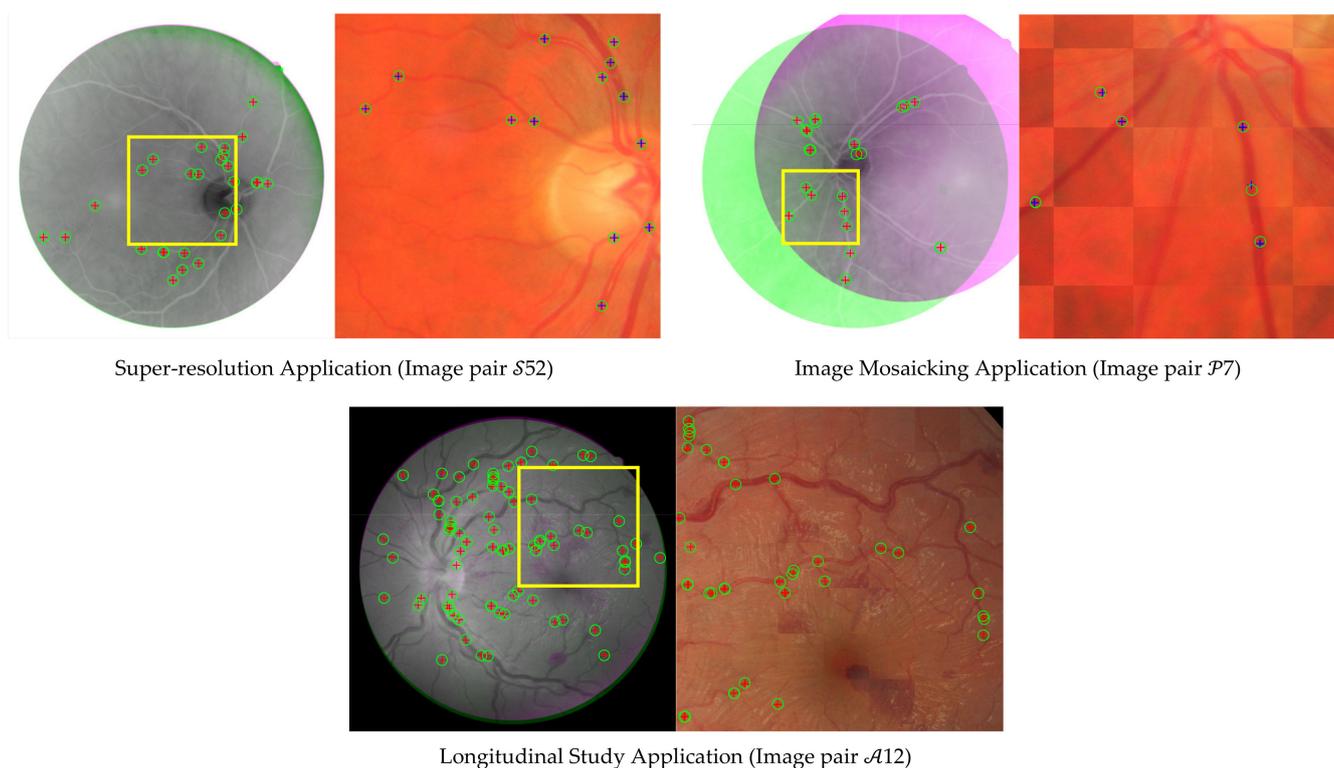


Figure 15. Examples of the successfully registered image pairs for CURVE-SIFT in the FIRE dataset. The green markers are inliers on the fixed image, while red/blue markers are inliers on moving image. Right images: Close-up for yellow square area as checkerboard image containing alternating rectangular regions from fixed image and moving image.

6. Conclusions

This paper introduces a new feature extraction method known as CURVE for the feature-based RIR technique. The proposed CURVE aims to extract feature points on retinal vessels and throughout the fundus image, which is important to ensure accurate registration of fundus images. However, in the local patches, the noises, such as retinal nerve fiber layer, underlying choroidal vessels, microaneurysm and exudates can also appear similar to retinal vessels. Therefore, CURVE incorporates both characteristics of the retinal vessels and noises in its modules to enable accurate discrimination between them.

The ability of CURVE to extract feature points on retinal vessels was demonstrated on the CHASE_DB1 [33,34], DRIVE [35,36], HRF [37,38] and STARE [39,40] datasets. Then, the CURVE performance was compared with five feature extraction methods from the existing feature-based RIR techniques, namely, Harris corner [26], SIFT detector [15], SURF [27,28], Ghassabi's [8] and D-Saddle [14]. From the experiment, CURVE accurately extracts an average of 86.021% of the feature points on retinal vessels and significantly outperformed the existing feature extraction methods ($p < 0.001^*$). Further analysis shows that the impact of image size on CURVE performance is minimal ($r_s = -0.032$, $p = 0.712$) but significantly affected in the presence of non-uniform intensity distribution in the image ($r_s = 0.342$, $p < 0.001^*$).

The registration performance when utilizing CURVE feature points in the feature-based RIR technique was demonstrated on the FIRE dataset. CURVE was paired with the SIFT descriptor [41], and the registration performance of CURVE-SIFT was compared with five existing feature-based RIR techniques; GDB-ICP [13], Harris-PIIFD [10], Ghassabi's-SIFT [8], H-M 16 [16], H-M 17 [9] and D-Saddle-HOG [14]. Overall, CURVE-SIFT successfully registered 44.030% of the image pairs in the FIRE dataset, while the success rate of the existing feature-based RIR techniques is less than 27.612%. The one-way ANOVA analysis showed that CURVE-SIFT is significantly outperformed GDB-ICP at $p = 0.007^*$ whereas Harris-PIIFD, Ghassabi's-SIFT, H-M 16, H-M 17 and D-Saddle-HOG at $p < 0.001^*$. CURVE-

SIFT obtained the highest success rate (53.061%) in the image mosaicking application, while the success rates of the existing feature-based RIR techniques were only between 0% to 32.653%. The image mosaicking application consists of image pairs with smaller overlapping areas compared to other applications in the FIRE dataset. Thus demonstrating the ability of CURVE to extract feature points on retinal vessels throughout the image. This is crucial to increase the chances of the inliers being established within the overlapping area to estimate an accurate geometrical transformation. In the future, we will focus our efforts to improve CURVE in extracting feature points from fundus images with non-uniform intensity distribution. Moreover, we will explore the possibility of a fusion strategy to combine deep convolutional neural network (CNN) with local feature point for feature extraction [58]. However, at the time of this study, the size of the public RIR dataset was small, which may result in model overfitting or underfitting [59]. The study will begin once a larger dataset or suitable pre-trained model for RIR is available publicly.

Author Contributions: Conceptualization, R.R. and N.K.A.K.; methodology, R.R. and M.Y.I.I.; software, R.R. and A.W.A.W.; validation, R.R. and K.H.; investigation, R.R., M.Y.I.I., K.H., A.W.A.W. and N.K.A.K.; writing—original draft preparation, R.R. and M.Y.I.I.; writing—review and editing, R.R. and K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundamental Research Grant Scheme (FRGS) (FP003-2021).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study namely, CHASE_DB1 [31,32], DRIVE [33,34], HRF [35,36], STARE [37,38] and (FIRE) [39] are openly available at <https://blogs.kingston.ac.uk/retinal/chasedb1/> (10 February 2019), <https://drive.grand-challenge.org/> (10 February 2019), <https://www5.cs.fau.de/research/data/fundus-images/> (10 February 2019), <https://cecas.clemson.edu/~ahoover/stare/> (10 February 2019) and <https://projects.ics.forth.gr/cvrl/fire/> (10 February 2019).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Mathematical symbols and notation.

No.	Symbol	Description	No.	Symbol	Description
1	$\check{\sigma}_{p,q}$	Relative Gaussian kernel at octave p and level q .	28	$L_{distance}$	Distance between the parallel cross-sectional lines. Exclusion criterion 1 : $L_{distance} = 3$ pixels. Exclusion criterion 2, 5 : $L_{distance} = 5$ pixels.
2	β	Central intensity value.	29	L_{length}	Length of the cross-sectional lines.
3	ε	Offset, $\varepsilon = 0.0010$.	30	L_{total}	Total of the cross-sectional lines, an odd number. Exclusion criterion 1 : $L_{total} = 5$ pixels. Exclusion criterion 2, 5 : $L_{total} = 7$ pixels.
4	μ	Constant factor.	31	m	Index of the feature point.
5	σ_0	Base width of Gaussian kernel, $\sigma_0 = 1.6$.	32	n	Index of the images in the hierarchical Gaussian scale space.
6	σ_s	Sampling Gaussian kernel, $\sigma_s = 0.5$.	33	N_{total}	Total feature points in the hierarchical Gaussian scale space, $N_{total} = 4500$ points.
7	$\sigma_{p,q}$	Absolute Gaussian kernel at octave p and.	34	$N_{p,q}$	The maximum number of feature points in $G_{p,q}$.
8	a_j	Pixel for inner ring test.	35	N_u	The
9	AP	Area of the intersected region between the sums of intensity profiles from the gradient and binary interpolated patches.	36	p	Octave index, $p \in [0, \dots, P - 1]$.
10	b_l	Pixel for outer ring test.	37	P	Total octave in the scale space, $P = 3$.

Table A1. Cont.

No.	Symbol	Description	No.	Symbol	Description
11	B	The blue channel.	38	q	Level index within an octave, $q \in [-1, \dots, Q - 2]$.
12	DC_u	Distribution coefficient for a grid of index u .	39	Q	Total level in each octave, $Q = 6$.
13	e	Extremum.	40	\mathcal{R}	The red channel.
14	EG	Entropy of a grid image.	41	s_{bin}	Side length of the binary interpolated patch. Initial side length of the patch in pixels. $s_{initial}$ is set according to the image size of the initial Gaussian image $G_{0,-1}$.
15	EP	Entropy of a gradient interpolated patch.	42	$s_{initial}$	$s_{initial} \begin{cases} 35 \text{ pixels} & \text{if } G_{0,-1} > 1000 \times 1000 \text{ pixels} \\ 25 \text{ pixels} & \text{if } G_{0,-1} \leq 1000 \times 1000 \text{ pixels} \\ & > 600 \times 600 \text{ pixels} \\ 21 \text{ pixels} & \text{if } G_{0,-1} \leq 600 \times 600 \text{ pixels} \end{cases}$
16	f_0	Proportion of the feature points at the initial Gaussian image $G_{0,-1}$.	43	s_p	Side length of the patch at octave p .
17	$F_{p,q}$	Proportion of the feature points at $G_{p,q}$.	44	SW_i	Selection weightage for a candidate feature point of index i
18	\mathcal{G}	The green channel.	45	TG	Total candidate feature points detected in a grid image.
19	G	Hierarchical Gaussian scale space.	46	u	Index of the grids in $G_{p,q}$.
20	$G_{p,q}$	Gaussian image at octave p and.	47	U	Total grids in a Gaussian image $G_{p,q}$.
21	i	Index of the candidate feature point in a Gaussian image.	48	UG	Peak deviation nonuniformity of a grid image.
22	I	Input image in grayscale.	49	W_{AP}	Weight factor for the area of the intersected region, $W_{AP} = 0.4$.
23	I_{a_j}	Intensity of inner ring pixel a_j in grayscale, $I_{a_j} \in [0, 1]$.	50	W_{EG}	Weight factor for the entropy, $W_{EG} = 0.3$.
24	I_{b_l}	Intensity of outer ring pixel b_l in grayscale $I_{b_l} \in [0, 1]$.	51	W_{EP}	Weight factor for the entropy, $W_{EP} = 0.3$.
25	j	Index of inner ring pixels, $j \in [1, \dots, 8]$.	52	W_{HP}	Weight factor for the mean histogram of the gradient orientation at the vessel edges, $W_{HP} = 0.3$.
26	K_m	Feature point of index m .	53	W_{TG}	Weight factor for the total candidate feature points, $W_{TG} = 0.4$.
27	$K_{m,p,q}$	Feature point of index m in the coordinate system at the respective octave p and level q .	54	W_{UG}	Weight factor for the peak deviation nonuniformity, $W_{UG} = 0.3$.

Appendix B

Table A2. Characteristics of retinal vessels and noises in blue squares. Red lines in (ii) and (iv) are cross-sectional lines to extract intensity profiles in (iii) and (v).

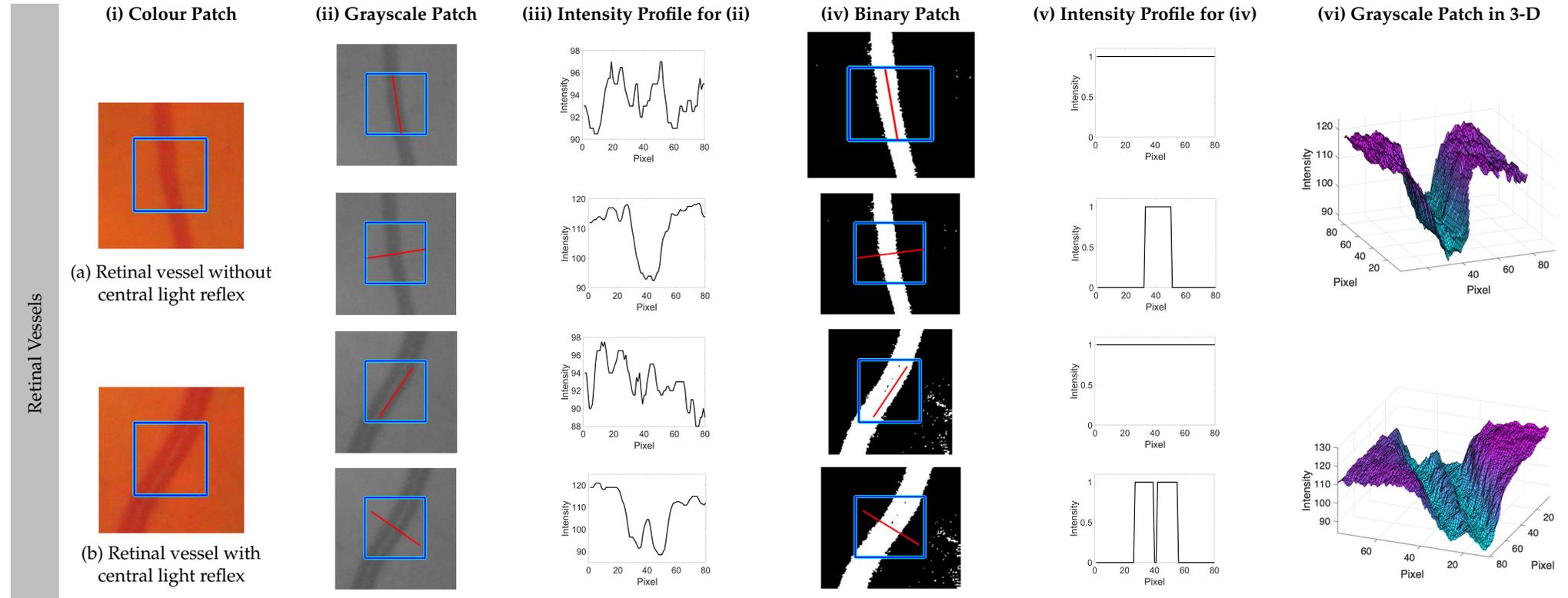


Table A2. Cont.

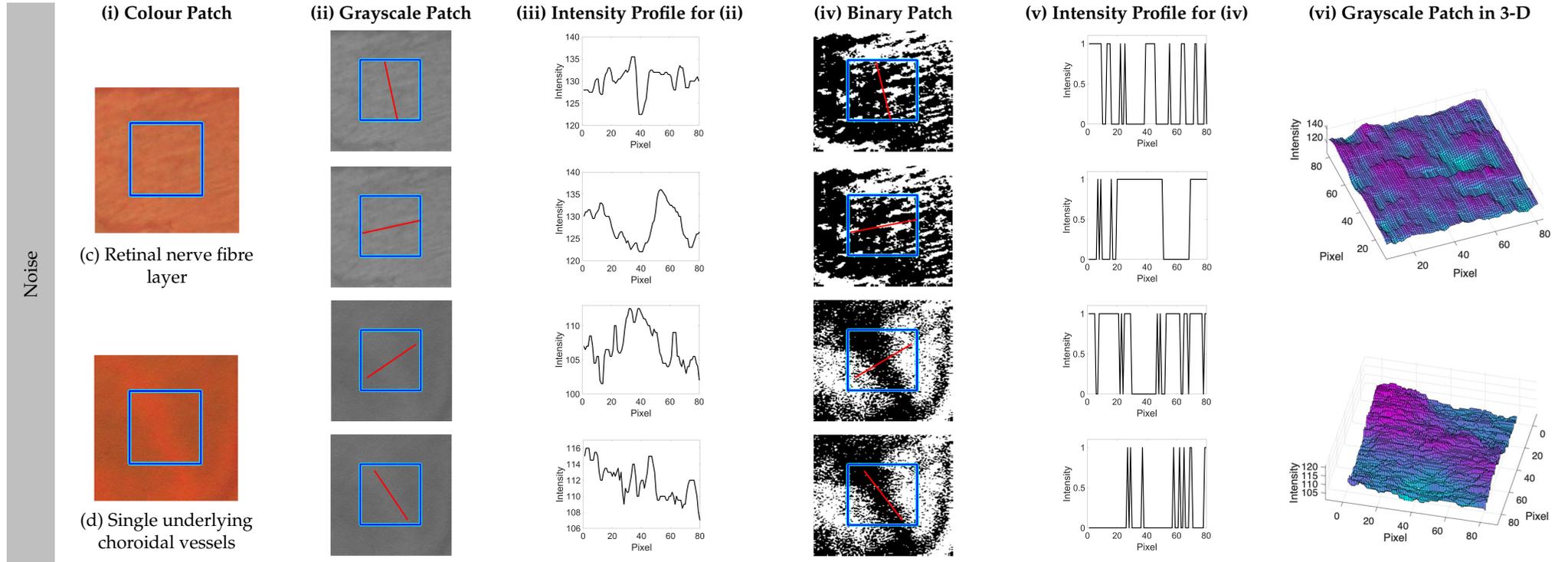


Table A2. Cont.

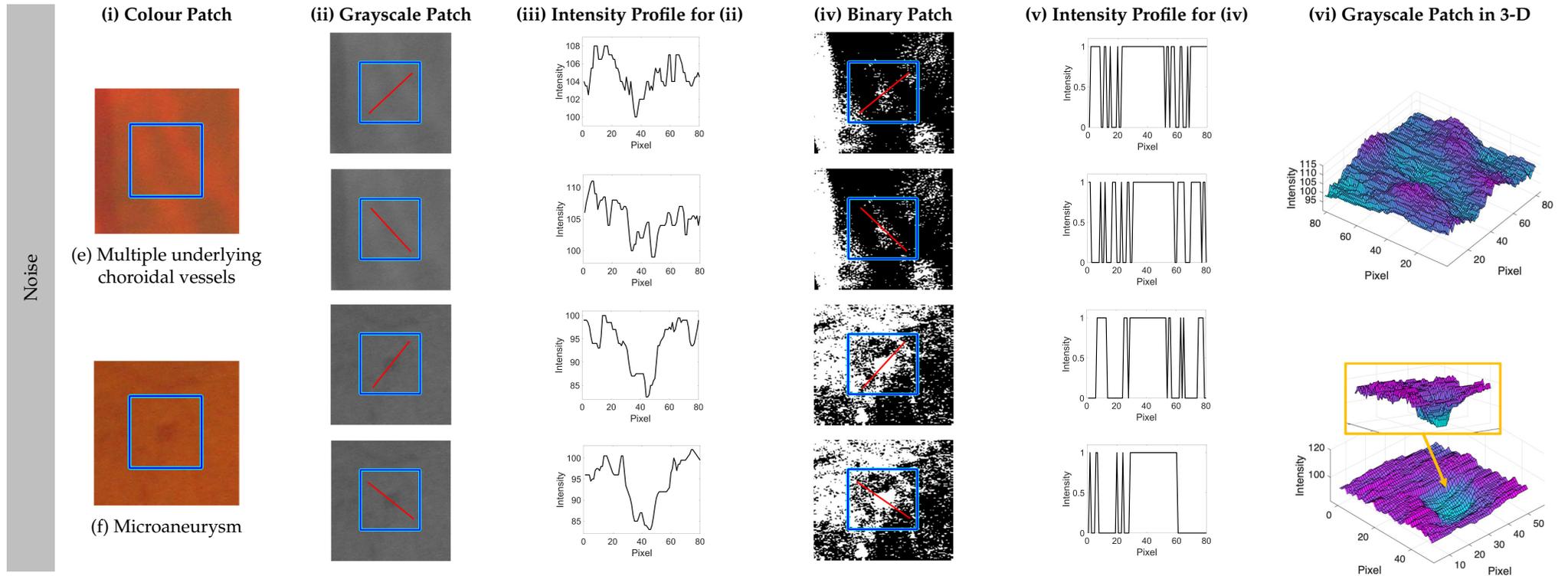
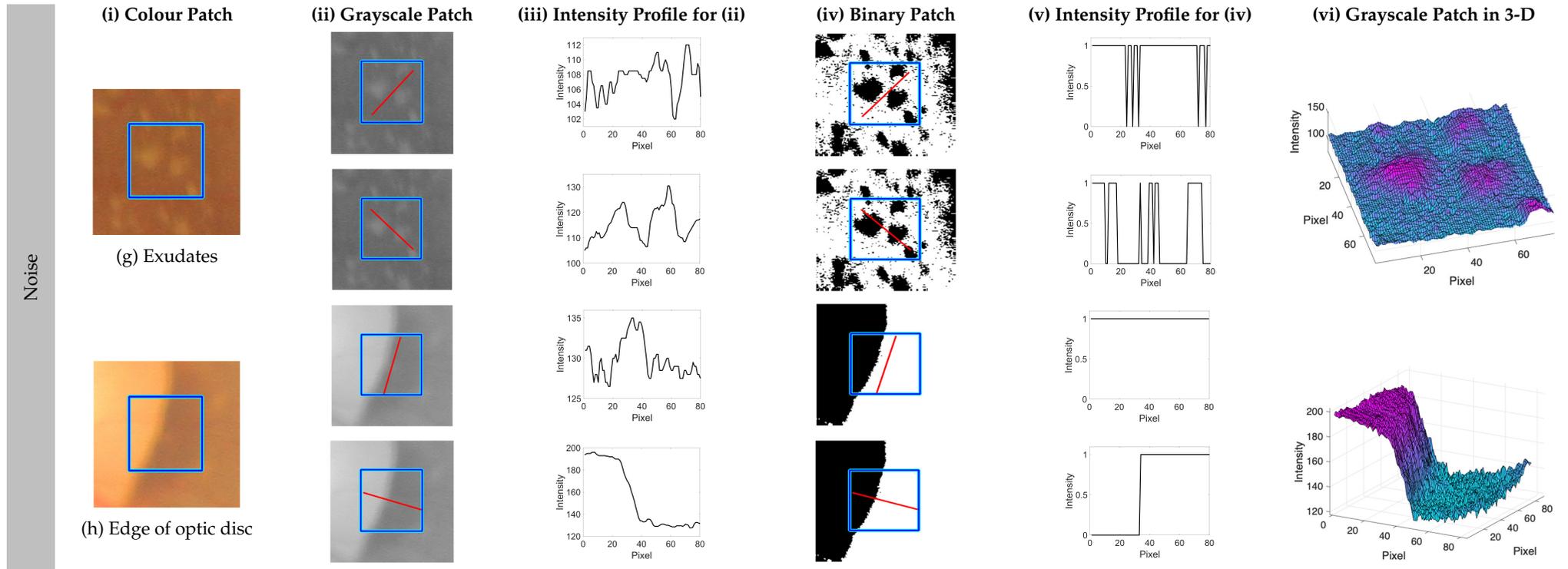


Table A2. Cont.



References

1. Hernandez-Matas, C.; Zabulis, X.; Argyros, A.A. Retinal image registration as a tool for supporting clinical applications. *Comp. Meth. Prog. Biomed.* **2021**, *199*, 105900. [[CrossRef](#)] [[PubMed](#)]
2. Legg, P.A.; Rosin, P.L.; Marshall, D.; Morgan, J.E. Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation. *Comput. Med. Imaging. Graph.* **2013**, *37*, 597–606. [[CrossRef](#)] [[PubMed](#)]
3. Nakagawa, T.; Suzuki, T.; Hayashi, Y.; Mizukusa, Y.; Hatanaka, Y.; Ishida, K.; Hara, T.; Fujita, H.; Yamamoto, T. Quantitative depth analysis of optic nerve head using stereo retinal fundus image pair. *J. Biomed. Opt.* **2008**, *13*, 064026. [[CrossRef](#)] [[PubMed](#)]
4. Kolar, R.; Sikula, V.; Base, M. Retinal Image Registration using Phase Correlation. *Anal. Biomed. Signals Images* **2010**, *20*, 244–252. [[CrossRef](#)]
5. Kolar, R.; Harabis, V.; Odstrcilik, J. Hybrid retinal image registration using phase correlation. *Imaging Sci. J.* **2013**, *61*, 369–384. [[CrossRef](#)]
6. Chanwimaluang, T.; Fan, G.L.; Fransen, S.R. Hybrid Retinal Image Registration. *IEEE Trans. Inf. Technol. Biomed.* **2006**, *10*, 129–142. [[CrossRef](#)]
7. Zitova, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [[CrossRef](#)]
8. Ghassabi, Z.; Shanbehzadeh, J.; Mohammadzadeh, A.; Ostadzadeh, S.S. Colour retinal fundus image registration by selecting stable extremum points in the scale-Invariant feature transform detector. *IET Image Process.* **2015**, *9*, 889–900. [[CrossRef](#)]
9. Hernandez-Matas, C.; Zabulis, X.; Argyros, A.A. An Experimental Evaluation of the Accuracy of Keypoints-Based Retinal Image Registration. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, Korea, 11–15 July 2017; pp. 377–381.
10. Chen, J.; Chen, J.; Tian, J.; Lee, N.; Zheng, J.; Smith, R.T.; Laine, A.F. A Partial Intensity Invariant Feature Descriptor for Multimodal Retinal Image Registration. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 1707–1718. [[CrossRef](#)]
11. Ghassabi, Z.; Shanbehzadeh, J.; Sedaghat, A.; Fatemizadeh, E. An efficient approach for robust multimodal retinal image registration based on UR-SIFT features and PIIFD descriptors. *Eurasip. J. Image Video Process.* **2013**, *2013*, 25. [[CrossRef](#)]
12. Ramli, R.; Idris, M.Y.I.; Hasikin, K.; Karim, N.K.A. Histogram-Based Threshold Selection of Retinal Feature for Image Registration. In Proceedings of the 3rd International Conference on Information Technology & Society (IC-ITS), Penang, Malaysia, 31 July–1 August 2017; pp. 105–114.
13. Yang, G.; Stewart, C.V.; Sofka, M.; Tsai, C.-L. Registration of challenging image pairs: Initialization, estimation, and decision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 1973–1989. [[CrossRef](#)] [[PubMed](#)]
14. Ramli, R.; Idris, M.Y.I.; Hasikin, K.; Karim, N.K.A.; Abdul Wahab, A.W.; Ahmedy, I.; Ahmedy, F.; Kadri, N.A.; Arof, H. Feature-Based Retinal Image Registration Using D-Saddle Feature. *J. Healthc. Eng.* **2017**, *2017*, 1489524. [[CrossRef](#)] [[PubMed](#)]
15. Lowe, D.G. Distinctive image features from scale-Invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
16. Hernandez-Matas, C.; Zabulis, X.; Argyros, A.A. Retinal image registration through simultaneous camera pose and eye shape estimation. In Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3247–3251.
17. Hernandez-Matas, C.; Zabulis, X.; Triantafyllou, A.; Anyfanti, P.; Argyros, A.A. Retinal image registration under the assumption of a spherical eye. *Comput. Med. Imaging Graph.* **2017**, *55*, 95–105. [[CrossRef](#)] [[PubMed](#)]
18. Tsai, C.; Li, C.; Yang, G.; Lin, K. The Edge-Driven Dual-Bootstrap Iterative Closest Point Algorithm for Registration of Multimodal Fluorescein Angiogram Sequence. *IEEE Trans. Med. Imaging.* **2010**, *29*, 636–649.
19. Sedaghat, A.; Mokhtarzade, M.; Ebadi, H. Uniform robust scale-Invariant feature matching for optical remote sensing images. *IEEE Trans. Geosci. Remote. Sens.* **2011**, *49*, 4516–4527. [[CrossRef](#)]
20. Frangi, A.F.; Niessen, W.J.; Vincken, K.L.; Viergever, M.A. Multiscale vessel enhancement filtering. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), Cambridge, MA, USA, 11–13 October 1998.
21. Vonikakis, V.; Chrysostomou, D.; Kouskouridas, R.; Gasteratos, A. A biologically inspired scale-Space for illumination invariant feature detection. *Meas. Sci. Technol.* **2013**, *24*, 074024. [[CrossRef](#)]
22. Lee, J.A.; Cheng, J.; Hai Lee, B.; Ping Ong, E.; Xu, G.; Wing Kee Wong, D.; Liu, J.; Laude, A.; Han Lim, T. A low-Dimensional step pattern analysis algorithm with application to multimodal retinal image registration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1046–1053.
23. Wang, G.; Wang, Z.C.; Chen, Y.F.; Zhao, W.D. Robust point matching method for multimodal retinal image registration. *Biomed. Signal Process. Control.* **2015**, *19*, 68–76. [[CrossRef](#)]
24. Hernandez-Matas, C.; Zabulis, X.; Argyros, A.A. Retinal image registration based on keypoint correspondences, spherical eye modeling and camera pose estimation. In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 5650–5654.
25. Lee, J.A.; Lee, B.H.; Xu, G.; Ong, E.P.; Wong, D.W.K.; Liu, J.; Lim, T.H. Geometric corner extraction in retinal fundus images. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 26–30 August 2014; pp. 158–161.

26. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the 4th Alvey Vision Conference, Manchester, UK, 31 August–2 September 1988; pp. 147–151.
27. Bay, H.; Tuytelaars, T.; van Gool, L. SURF: Speeded up Robust Features. In Proceedings of the 9th European Conference on Computer Vision (ECCV), Graz, Austria, 7–13 May 2006; pp. 404–417.
28. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [[CrossRef](#)]
29. International Telecommunication Union. Studio encoding parameters of digital television for standard 4:3 and wide-Screen 16:9 aspect ratios. In *Recommendation ITU-R BT.601–7*; ITU: Geneva, Switzerland, 2017; pp. 1–8.
30. Kanan, C.; Cottrell, G.W. Color-to-Grayscale: Does the Method Matter in Image Recognition? *PLoS ONE* **2012**, *7*, e29740. [[CrossRef](#)]
31. Burger, W.; Burge, M.J. SIFT—Scale-Invariant Local Features. In *Principles of Digital Image Processing: Advanced Methods*; Springer: London, UK, 2013; pp. 229–295.
32. Aldana-Iuit, J.; Mishkin, D.; Chum, O.; Matas, J. In the Saddle: Chasing Fast and Repeatable Features. In Proceedings of the 23rd International Conference on Pattern Recognition, Cancun, Mexico, 4–8 December 2016.
33. CHASE_DB1 Retinal Image Database. Available online: <https://blogs.kingston.ac.uk/retinal/chasedb1/> (accessed on 10 December 2017).
34. Fraz, M.M.; Remagnino, P.; Hoppe, A.; Uyyanonvara, B.; Rudnicka, A.R.; Owen, C.G.; Barman, S.A. An ensemble classification-Based approach applied to retinal blood vessel segmentation. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 2538–2548. [[CrossRef](#)] [[PubMed](#)]
35. Staal, J.; Abràmoff, M.D.; Niemeijer, M.; Viergever, M.A.; Van Ginneken, B. Ridge-Based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **2004**, *23*, 501–509. [[CrossRef](#)]
36. DRIVE: Digital Retinal Images for Vessel Extraction. Available online: <http://www.isi.uu.nl/Research/Databases/DRIVE/> (accessed on 10 December 2017).
37. HRF: High-Resolution Fundus Image Database. Available online: <https://www5.cs.fau.de/research/data/fundus-images/> (accessed on 10 December 2017).
38. Budai, A.; Bock, R.; Maier, A.; Hornegger, J.; Michelson, G. Robust vessel segmentation in fundus images. *Int. J. Biomed. Imaging* **2013**, *2013*, 154860. [[CrossRef](#)]
39. STARE: Structured Analysis of the Retina. Available online: <http://cecas.clemson.edu/~jahoover/stare/> (accessed on 10 December 2017).
40. Hoover, A.; Kouznetsova, V.; Goldbaum, M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imaging* **2000**, *19*, 203–210. [[CrossRef](#)] [[PubMed](#)]
41. Hernandez-Matas, C.; Zabulis, X.; Triantafyllou, A.; Anyfanti, P.; Douma, S.; Argyros, A.A. FIRE: Fundus Image Registration dataset. *J. Modeling Ophthalmol.* **2017**, *1*, 16–28. [[CrossRef](#)]
42. Saha, S.K.; Xiao, D.; Frost, S.; Kanagasingham, Y. A Two-Step Approach for Longitudinal Registration of Retinal Images. *J. Med. Syst.* **2016**, *40*, 277. [[CrossRef](#)]
43. Gonzalez, R.C.; Woods, R.E.; Eddins, S.L. Representation and Description. In *Digital Image Processing Using MATLAB*; Prentice Hall: Hoboken, NJ, USA, 2009.
44. Goerner, F.L.; Duong, T.; Stafford, R.J.; Clarke, G.D. A comparison of five standard methods for evaluating image intensity uniformity in partially parallel imaging MRI. *Med. Phys.* **2013**, *40*, 082302-1–082302-10. [[CrossRef](#)]
45. Brown, M.; Lowe, D.G. Invariant Features from Interest Point Groups. In Proceedings of the British Machine Vision Conference (BMVC), Cardiff, UK, 2–5 September 2002.
46. Vedaldi, A.; Fulkerson, B. VLFeat: An open and portable library of computer vision algorithms. In Proceedings of the 18th ACM international conference on Multimedia, Firenze, Italy, 25–29 October 2010; pp. 1469–1472.
47. Torr, P.H.; Zisserman, A. MLESAC: A new robust estimator with application to estimating image geometry. *Comput. Vis. Image Underst.* **2000**, *78*, 138–156. [[CrossRef](#)]
48. Goshtasby, A. Image registration by local approximation methods. *Image Vis. Comput.* **1988**, *6*, 255–261. [[CrossRef](#)]
49. Pauli, T.W.; Gangaputra, S.; Hubbard, L.D.; Thayer, D.W.; Chandler, C.S.; Peng, Q.; Narkar, A.; Ferrier, N.J.; Danis, R.P. Effect of Image Compression and Resolution on Retinal Vascular Caliber. *Investig. Ophthalmol. Vis. Sci.* **2012**, *53*, 5117–5123. [[CrossRef](#)]
50. Brown, D.M.; Ciardella, A. Mosaic Fundus Imaging in the Diagnosis of Retinal Diseases. *Investig. Ophthalmol. Vis. Sci.* **2005**, *46*, 2581.
51. Bontala, A.; Sivaswamy, J.; Pappuru, R.R. Image mosaicing of low quality neonatal retinal images. In Proceedings of the 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 720–723.
52. Lee, B.H.; Xu, G.; Gopalakrishnan, K.; Ong, E.P.; Li, R.; Wong, D.W.K.; Lim, T.H. AEGIS-Augmented Eye Laser Treatment with Region Guidance for Intelligent Surgery. In Proceedings of the 11th Asian Conference on Computer Aided Surgery (ACCAS 2015), Singapore, 9–11 July 2015.
53. Adal, K.M.; van Etten, P.G.; Martinez, J.P.; van Vliet, L.J.; Vermeer, K.A. Accuracy Assessment of Intra-and Intervisit Fundus Image Registration for Diabetic Retinopathy Screening Accuracy Assessment of Fundus Image Registration. *Investig. Ophthalmol. Vis. Sci.* **2015**, *56*, 1805–1812. [[CrossRef](#)] [[PubMed](#)]

54. Matsopoulos, G.K.; Asvestas, P.A.; Mouravliansky, N.A.; Delibasis, K.K. Multimodal registration of retinal images using self organizing maps. *IEEE Trans. Med. Imaging.* **2004**, *23*, 1557–1563. [[CrossRef](#)]
55. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–25 June 2005; pp. 886–893.
56. Patel, M.I.; Thakar, V.K.; Shah, S.K. Image Registration of Satellite Images with Varying Illumination Level Using HOG Descriptor Based SURF. *Procedia Comput. Sci.* **2016**, *93*, 382–388. [[CrossRef](#)]
57. Grabner, M.; Grabner, H.; Bischof, H. Fast approximated SIFT. In Proceedings of the Asian Conference on Computer Vision, Hyderabad, India, 13–16 January 2006.
58. Rashid, M.; Khan, M.A.; Sharif, M.; Raza, M.; Sarfraz, M.M.; Afza, F. Object detection and classification: A joint selection and fusion strategy of deep convolutional neural network and SIFT point features. *Multimed. Tools Appl.* **2019**, *78*, 15751–15777. [[CrossRef](#)]
59. Andrei Dmitri, G.; Alex, J.; Maya, V.; Jack, D. Preventing Model Overfitting and Underfitting in Convolutional Neural Networks. *Int. J. Softw. Sci. Comput. Intell. (IJSSCI)* **2018**, *10*, 19–28.