IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 1 March 2024, accepted 17 May 2024, date of publication 22 May 2024, date of current version 20 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3404242

RESEARCH ARTICLE

Retinal Image Registration Using Partial Intensity Invariant Feature Descriptor and Redundant Keypoints Elimination Techniques

KITTIPOL WISAENG
Technology and Business Information System Unit, Mahasarakham Business School, Mahasarakham University, Maha Sarakham 44150, Thailand
e-mail: kittipol.w@acc.msu.ac.th

This work was supported by Mahasarakham University of Thailand.

ABSTRACT Retinal image registration is crucial for improving the accuracy of diagnosing and monitoring retinal diseases. This paper proposes a specific feature region technique for retinal image registration, involving five steps for accurate registration. The technique combines area and feature-based methods, including extracting the retinal vascular tree, detecting distinctive points, eliminating redundant keypoints, and matching features for image registration using affine transformation modes. First, the retina's vascular tree is extracted using a Top-Hat operation and optimal thresholding technique. Next, the Harris-PIIFD detector identifies key points in the binary image, and redundant keypoints are removed to reduce computational load. Finally, bilateral matching and best-bin-first algorithms compute the similarity matrix for registering the images. If the image pair is accepted, the points are controlled using the simplest affine transformation modes for the highest registration success rate. The simulation results on 134 pairs of FIRE datasets demonstrate the effectiveness and robustness of the proposed algorithm. The experimental results for the proposed method are satisfactory. The result obtained for the proposed approach for retinal image registration is precision 0.98240, recall 0.98312, RMSE 0.01280, ERR 0.01716, and matching score 0.98284, with the computational time taken 3.01s. This hybrid image registration approach is an efficient and reliable tool for retinal image registration, leading to more accurate diagnosis and monitoring of retinal diseases.

INDEX TERMS Retinal registration, retinal image, redundant keypoints elimination, bilateral matching.

I. INTRODUCTION

Medical image registration is an intricate process that plays a pivotal role in monitoring and analyzing the progression of certain medical conditions over time. It involves aligning and comparing images obtained from different sources, viewpoints, or under different conditions to accurately assess the development of diseases such as Glaucoma, Hypertension, and Diabetes that often lead to retinal lesions like exudates, microaneurysms, and hemorrhages. By precisely registering retinal images, medical professionals can gain crucial insights into the progression of these conditions, which can help them make informed decisions about the best course of treatment for their patients. This process is particularly crucial in assessing small vessels in the retina, as it can help promote accurate diagnosis and progression monitoring of retinal diseases, ultimately leading to better patient outcomes.

According to statistics from 2015, diabetes caused around 1.6 million deaths worldwide. Shockingly, there are about 12 million diabetic individuals around the world, and this number is predicted to double by the year 2025 [1]. Furthermore, diabetes is projected to emerge as the seventh most fatal disease worldwide by the year 2030 [2]. This alarming forecast underlines the pressing need for effective measures to curb the rising prevalence of diabetes across the globe. However, the advancements in retinal image processing and analysis have brought about significant improvements in the diagnosis and treatment of various retinal diseases, including but not limited to glaucoma, age-related macular degeneration, and others. Retinal images contain a wealth of crucial

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

local and temporal information about the human retina, making retinal image registration an increasingly important tool for ophthalmologists to diagnose and cure retinal diseases accurately. This technique enables medical professionals to align and compare different retinal images taken at different times, allowing them to detect even the slightest changes in the retina and take timely corrective measures. Retinal image registration is a crucial and fundamental technique used in ophthalmology. It involves aligning two retinal images that could have been captured at different times or via different devices to facilitate the analysis and monitoring of the progression of eye diseases. Combining a reference image with a target image makes identifying changes in the retina easier, which can help detect and treat various ocular conditions at an early stage. Accurate image registration using this technique can provide valuable insights into the structure and function of different body parts. This can assist medical professionals in making more informed decisions, ultimately improving patient outcomes [3]. However, image registration is a challenging problem, especially when the image pairs have differences in illumination, color, and contrast and only small overlapping areas. Therefore, there is a need to develop automated registration methods that can handle the various challenges of retinal image registration. Retinal image registration, which involves aligning two or more images taken from different imaging modalities, has been a crucial area of research for more than two decades. Despite progress in this area, achieving fully automatic and reliable multimodal image registration is still challenging. This is because imaging modalities frequently generate images with non-linear intensity differences, and images acquired from clinics and hospitals are commonly affected by pathologies and noise. Some examples of retinal images are shown in Fig. 1. For instance, optical coherence tomography, which uses coherent light to produce retina images, is susceptible to speckle noise, a form of cohesive noise that causes significant degradation in spatial resolution and quality. As a result, developing a reliable method for multimodal image registration that considers these challenges remains an active area of research.

Retinal image registration possesses a unique ability to facilitate non-invasive in vivo vascular and neuronal tissue observation. However, the registration of color retinal images presents complicated challenges, including variations in illumination, image distortion, and artifacts, which can lead to erroneous diagnoses if not dealt with properly. Despite these challenges, researchers are actively exploring various approaches to improve the accuracy of retinal image analysis through image registration. As a result, deep learning methods have gained immense popularity in medical image registration in recent years. These methods have shown significant advantages over classical methods, particularly in computing similarity metrics between images using convolutional neural networks (CNNs) [4]. By leveraging CNNs, relevant features can be learned to accurately estimate

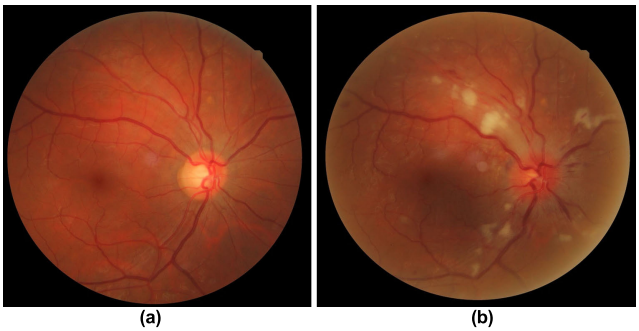

FIGURE 1. The provided figure showcases two comparative retinal images, (a) represents a normal retinal image, (b) depicts a retinal image of a patient diagnosed with both diabetes and hypertension. This image also demonstrates the presence of speckle noise, which adds complexity to the retinal image analysis.

the matching between images. However, retinal image registration poses a unique set of challenges due to this domain's complex features, making it a complicated problem to solve. Very few studies are available on retinal image registration, and only one recent deep-learning method [5] is known to exist. This is because traditional approaches used in medical image registration are typically based on intensity, making them less effective and unfavorable for retinal image registration. Although these methods have succeeded in other medical areas, they may only be somewhat suitable for retinal image registration and require modifications. The process of registering multi-source retinal images for analysis can be challenging and complex due to the significant dissimilarities between the images. To overcome this issue, this paper proposes five major steps. First, the vascular structure of the retina is extracted through an efficient Top-Hat operation using an optimal thresholding technique. Next, a new and innovative local feature extraction method based on the Harris-PIIFD algorithm has been proposed to detect the maximum keypoints information and distinctive points based on the binary image. However, this process leads to redundant image-matching keypoints, increasing the computation time. Prior work regarding existing retinal registration algorithms has been thoroughly researched and analyzed to identify the drawbacks and limitations of these methods. To overcome this, a redundant keypoints elimination (RKE) has been proposed to remove the redundant keypoints of the Harris-PIIFD algorithm, thereby reducing the overall computational load. Then, the features are matched with bilateral matching based on the best-bin-first algorithms for computing the similarity matrix for registering the images. The select transformation mode will be involved if the image pair is accepted. The transformation mode is a process of changing an image's position, orientation, and scale. In this case, the simplest modes of affine transformation have been employed, which involve altering the image by combining rotation, translation, and scaling operations. This paper presents a comprehensive approach to registering multi-source retinal images that considers this process's unique challenges and complexities. The novel local feature extraction method, combined with

the redundant keypoints elimination technique and hybrid framework, provides a powerful solution capable of producing accurate and reliable results even in poor-quality retinal image pairs.

II. RELATED WORK

The following section provides a comprehensive overview of retinal registration and the different models used for their abnormality diagnosis. It also highlights the limitations of existing retinal registration algorithms and the importance of fast and accurate registration techniques. Medical image analysis has always been an interesting field of research, and the registration of retinal images, particularly color retinal image registration, has been the focus of many studies. Traditional methods are the most effective in this field, providing the best performance compared to other techniques [6], [7]. However, these findings underscore the need for further research to develop more advanced and reliable diagnostic tools for retinal abnormalities. Matching retinal images for medical purposes is complex and challenging, requiring sophisticated techniques. One method to address retinal image registration is eye modeling and pose estimation, which involves matching images with keypoints and bifurcations in a two-step process. A multi-step process is typically employed to achieve an accurate registration. The first step often involves using random sample consensus to find an initial approximation. This initial approximation is refined in the second step using more advanced techniques such as particle swarm optimization and a complex transformation model. This refinement process is repeated multiple times to ensure the optimal solution is found from all the possible candidates. Bifurcations such as crossovers are detected using traditional methods that consider the point correspondence of the vascular tree. In the study conducted by Sureshjani et al. [8], filter bank orientations were employed as a technique to identify and detect the existence of landmarks. These landmarks could serve as important reference points for image registration. However, recent studies have proposed deep learning methods to enhance traditional approaches for feature extraction [9], [10]. Incorporating deep learning approaches for detecting point correspondence is also a promising development that could further enhance the method's accuracy. Therefore, the retinal image registration domain has yet to fully embrace its utilization and leverage the power of these advanced techniques to unlock new possibilities and enhance the accuracy and efficiency of retinal image registration. However, specific deep-learning pipelines have been developed for this purpose. A method that can be employed to align moving images involves directly predicting their transformation. This method uses a regression model to recover the transformation matrix parameters and align the images. This approach has been documented in literature [11]. Unfortunately, the availability of labeled data is often limited, making supervised learning challenging. Therefore, some methods rely on unsupervised learning to align the images. In contrast, deep learning maps the corresponding pair of images to the deformation field that aligns the images [12], [13]. However, deep learning methods in this area have been cautious due to retinal image registration's unique requirements and challenges. These challenges include preserving sparsely detailed structures, such as vascular, over relatively uniform backgrounds, the progression of diseases, and the expectedly large displacement transformations. Despite these challenges, deep learning methods have been successful in other medical areas and continue to hold promise for retinal image registration. Despite the impressive results of deep learning in image classification and segmentation, its use in retinal image registration still needs to be improved, with only a few studies exploring its potential. However, these algorithms can learn to map complex relationships between images and accurately align them, even on large datasets. This has the potential to improve the accuracy of retinal image registration and make the process more efficient, which is crucial for diagnosing and monitoring eye diseases. Mahapatra et al. [14] introduced leveraged generative adversarial networks to achieve end-to-end registration. Zou et al. [5] proposed an innovative retinal image registration approach involving an unsupervised structure-driven regression learning method. The method employs a parameterized deformation function that maps the images and calculates multi-scale similarity in conjunction with contextual structures. While these techniques offer great potential for improving the accuracy and efficiency of retinal image registration, they often require synthetic data or image augmentation techniques to augment the training dataset. However, these augmentation techniques may only sometimes be sufficient to account for the variability in multisite or multi-source images. Despite this limitation, these methods are up-and-coming and potentially significantly enhance retinal image registration performance. The study by Lee et al. [15] aimed to address the challenge of multi-modality retinal images using a feature-based learning approach. The proposed technique uses CNNs to learn a deep and complex representation of the image data. CNN's architecture is designed to process multi-modality images and convert them into meaningful feature spaces. The trained CNNs is then used to extract features from the multi-modality images, and a conventional approach is employed to complete the registration process. Wang et al. [16] introduced a content-adaptive weakly-supervised deep learning framework, which combines vascular segmentation, feature detection, and outlier rejection strategies. This framework has proven to be highly effective in clinical settings. However, despite the benefits of deep learning features, their interpretation and usability can still be improved. Therefore, further research is required to enhance the interpretability and usability of these features. One of the most critical tasks in medical imaging is image registration, which involves aligning two or more images. Traditional image registration methods are classified into area- and feature-based methods. The former uses images' intensity or color information to determine

their similarity. In contrast, the latter is computationally efficient and less affected by intensity and rotation. Feature-based methods are commonly used in image registration tasks as they consist of three essential steps. The first step in feature-based registration methods is extracting sufficient feature points from the reference and target images. These points are then used to establish correspondences between the two images. To extract feature points in retinal image registration, several popular image descriptors are used, including the Scale-Invariant Feature Transform (SIFT) [17], the Edge-Oriented Histogram-Scale Invariant Feature Transform (EOH-SIFT) [18], and the Speed-Up Robust Features (SURF) [19]. These descriptors are designed to be consistent with changes in scale, rotation, and illumination, making them reliable and effective in image registration tasks [20]. The second step in feature-based registration methods aims to align feature point sets from different sources, times, and viewpoints, referred to as point set registration. This process involves estimating correspondence between the points in the sets and updating the transformation to align them [21]. Several point-set registration methods have been proposed, each with its unique strengths. One such method is the Coherent Point Drift (CPD) algorithm, which evaluates the correspondence between point sets using the Euclidean distance and applies a uniform distribution for outlier modeling [22]. Another method is the Global and Local Mixture Distance Transformed Point Sets (GLMDTPS), which improves feature description for point set registration by utilizing global and local mixture distance features [23]. A multi-feature-based finite mixture model has been proposed, combining SIFT with different geometrical features for point set registration [24]. These methods have all shown great promise in aligning feature point sets from various sources, times, and viewpoints. They have been successfully applied to retinal image registration, where they have demonstrated significant improvements in accuracy and robustness. For example, Ma et al. proposed a non-rigid point set registration method called PR-GLS based on the CPD algorithm. This method employs shape context to estimate the correspondence between feature point sets. It manually assigns an a priori probability based on the correspondence to guide the posterior probability function of the Gaussian mixture model. Finally, the authors applied PR-GLS to retinal image registration to align multi-modal retinal images and correct eye motion artifacts in images with optical coherence tomography. This approach is highly effective in improving the accuracy and robustness of retinal image registration [23], [25], [26]. To summarize, the registration of retinal images is a complex task requiring advanced techniques and methods. Table 1 provides an overview of the different retinal registration methods reviewed.

As previously stated, retinal image registration is a challenging task that requires careful consideration of the strengths and limitations of different computational methods. While robust, DNNs have high computational requirements for training and inference, mainly when dealing with large-scale data prone to overfitting. Additionally, DNNs need more transparency and interpretability in the learned representations, which can hinder understanding the registration process. SDRN, another popular method, requires careful design and tuning of the network architecture and loss functions to effectively capture retinal structures, which can be time-consuming and resource-intensive. However, it has limited generalization to unseen retinal image variations, particularly in pathological changes or uncommon retinal conditions not well-represented in the training data. While effective, CNNs have fixed input sizes and spatial transformations that may limit the ability to handle geometric distortions or variances in retinal image sizes or orientations.

Additionally, CNNs may have difficulty capturing long-range dependencies and contextual information across large retinal images due to local receptive fields in convolutional layers. GANs have training instability and mode collapse issues, especially when balancing the generator and discriminator networks for retinal image synthesis or transformation tasks. Moreover, GANs need more control over the generated output and potential challenges in ensuring the fidelity and consistency of synthesized retinal images for registration purposes. The problem arises mainly when the global transformations fail to capture local variations accurately. This issue is particularly significant in retinal images, where even subtle changes can have critical diagnostic implications. These methods also have constraints on the deformable patterns or transformations they can model, which may restrict their applicability to diverse retinal image registration scenarios. CPD algorithms are sensitive to initialization parameters and hyperparameters, which can impact the convergence and accuracy of the registration. Additionally, CPD has limited scalability to large-scale point clouds or high-dimensional feature spaces, potentially affecting the efficiency and performance of CPD for retinal image alignment. SIFT has computational complexity in feature extraction and matching processes, which may be prohibitive for real-time or high-throughput retinal image registration applications. Furthermore, SIFT has limited adaptability to changes in scale, rotation, and illumination across retinal images, potentially leading to suboptimal feature correspondences and transformations. Finally, GMM is another popular method for retinal image registration, but they assume Gaussian distributions, which may only sometimes hold for complex and multimodal retinal image features. This leads to limitations in capturing the variability and diversity of retinal structures and sensitivity to initialization and model selection. This can impact the robustness and accuracy of GMM-based registration methods, particularly in the data's presence of noise or outliers. Therefore, it is necessary to continuously develop and improve registration techniques to meet the ever-evolving challenges of retinal image registration.

This study proposes a hybrid image registration framework for retinal images, combining area and feature-based methods. The proposed method applies five major steps for retinal registration: vascular tree extraction, detection of the

TABLE 1. Classification of image registration according to eight different methods. These include DNNs, SDRN, CNNs, GANs, Global and local geometric structures, CPD, SIFT, and GMM. Each method has advantages and limitations, making them suitable for different image registration applications.

Authors	Methods	Advantage	Disadvantages
X. Cheng (2018) [4], B.D. de Vos (2019) [12], J. Lee (2019) [15], Y. Wang J. (2021) [16]	Deep Neural Networks (DNNs)	DNNs have proven to be particularly advantageous for retinal image registration. These networks excel at learning complex patterns and relationships within images, which is required for the accurate and reliable registration of retinal images. By leveraging deep learning techniques, retinal image registration algorithms can achieve higher precision and robustness, even when dealing with low-quality or noisy images.	One of the main disadvantages of DNNs for retinal image registration is the requirement for large amounts of training data. Since DNNs have many parameters, they require significant data to learn the patterns and features of retinal images. In addition, the training process can be computationally expensive and time-consuming, especially when dealing with large datasets.
B. Zou (2020) [5], D. Motta (2019) [6], B.	Structure-Driven Regression Network (SDRN)	SDRN exploits the anatomical structures in retinal images, such as blood vessels, optic discs, and fovea, to guide the registration process. This ensures that the registered images are aligned based on relevant anatomical features, improving alignment accuracy.	SDRN offers a targeted and anatomically guided approach for retinal image registration. However, it is important to consider potential disadvantages in terms of limitations in generalization, sensitivity to noise, the complexity of training, lack of flexibility, and interpretability challenges.
S. Miao (2016) [11], G. Balakrishnan (2019) [13]	Convolutional Neural Network (CNNs)	CNNs can automatically learn hierarchical features from retinal images, capturing relevant patterns and structures for registration without requiring manual feature extraction. This allows the network to adapt to the specific characteristics of retinal images and improve registration performance.	Training CNNs for retinal image registration can be computationally intensive, especially for deeper architectures and large datasets. This complexity may require access to high-performance computing resources, which can be a barrier for clinical settings with limited computational capabilities.
D. Mahapatra (2018) [14]	Generative Adversarial Networks (GANs)	GANs can facilitate domain adaptation by learning the underlying distribution of retinal images from different sources or imaging devices. Adapting the network to different domains through GAN-based techniques can improve the registration accuracy and transferability of the model across diverse datasets.	GANs may amplify existing biases in the training data, leading to biased or unrealistic retinal images that do not accurately represent the underlying distribution of real-world images. Inflexible data generation can introduce systematic errors and undermine the registration model's generalization performance on diverse datasets.
D. Bi (2019) [20]	Global and Local Geometric Structure	Global geometric structure provides valuable contextual information about the spatial relationships between retinal structures, while local geometric structure offers detailed information about specific regions or structures within the image. Integrating contextual information from global and local features allows the registration algorithm to leverage spatial constraints and anatomical priors for more accurate and context-aware alignment of retinal images.	Integrating global and local geometric structures often involves tuning multiple parameters and balancing the weighting of different features to achieve optimal registration performance. Parameter optimization can be challenging and time-consuming, requiring careful parameter selection, validation, and fine-tuning to ensure effective utilization of both global and local features for accurate retinal image alignment.
J. Ma (2016) [23]	Coherent Point Drift (CPD)	CPD combines global and local alignment strategies by considering the overall spatial coherence of retinal structures (global context) along with the fine-grained correspondence of key points or landmarks (local features). This global-to-local alignment approach enables CPD to capture both the macroscopic layout of retinal images and the detailed anatomical structures, facilitating accurate and comprehensive registration outcomes.	CPD's flexibility and adaptability to noise and variability in retinal images can also pose a risk of overfitting the registration model to specific features or patterns in the training dataset, potentially reducing its generalization performance to unseen or diverse retinal image samples. Overfitting can lead to biased registration results, limited transferability to new data, and reduced robustness in settings where the training data must fully represent the variability in real retinal images.
K. Yang (2017) [24]	Scale-Invariant Feature Transform (SIFT)	SIFT features are designed to be robust to variations in illumination, contrast, and noise in retinal images, making the algorithm resilient to imaging artifacts and environmental factors affecting image quality. The robustness of SIFT descriptors to noise and lighting conditions enhances the algorithm's ability to reliably detect and match features across different retinal images, ensuring consistent and accurate registration results in clinical applications.	SIFT generates high-dimensional feature vectors representing local image structures and characteristics, resulting in large feature spaces that can challenge storage, processing, and matching in retinal image registration. High-dimensional feature descriptors may increase memory usage, computational complexity, and computational time for feature extraction, matching, and correspondence estimation, potentially impacting the algorithm's scalability and applicability to large-scale retinal imaging studies.

TABLE 1. *(Continued.)* **Classification of image registration according to eight different methods. These include DNNs, SDRN, CNNs, GANs, Global and local geometric structures, CPD, SIFT, and GMM. Each method has advantages and limitations, making them suitable for different image registration applications.**

Authors	Methods	Advantage	Disadvantages
C. Liu (2016) [26]	Gaussian Mixture Model (GMM)	GMM is inherently robust to noise and outliers in the data, as it models the underlying distribution as a mixture of smooth Gaussian components. This robustness to noise can improve the accuracy and reliability of retinal image registration by reducing the impact of artifacts or imperfections in the image data on the registration process.	GMM requires specifying the number of Gaussian components in the mixture model a priori, which can be challenging in practice, mainly when the number of clusters or modes in the data is still being determined. Choosing an inappropriate number of components can result in either underfitting or overfitting the data, impacting the quality of the registration process.

maximum keypoints information, redundant keypoints elimination, and keypoints matching for computing the similarity matrix for registering the images and selecting transformation mode. The proposed method is focused on enhancing the conventional method of retinal image registration by gradually replacing each algorithm involved with a more advanced hybrid technique. The proposed method is known for achieving high-quality results through a combination of area-based and feature-based strategies.

The paper is divided into several sections that discuss various aspects of the proposed methods. Section III introduces the framework in detail, along with the preprocessing of retinal images, registration of retinal images, keypoints detection, redundant keypoints elimination, keypoints matching, and selection transformation mode that makes it stand out. The section comprehensively explains how the framework works and its potential applications. Section IV focuses on the datasets and evaluation criteria for testing the proposed framework. Finally, the paper concludes with a discussion of the results and a summary of the keypoints findings in Section V. The conclusion also highlights the potential implications of the framework and how it can be used to solve retinal image registration problems.

III. PROPOSED METHODOLOGY

The retinal registration framework proposed for image processing involves five crucial steps that work together to create a practical and effective approach. Firstly, the images undergo preprocessing using Top-Hat operation and optimal thresholding. In the second step, corner points are detected using a combination of Harris-PIIFD detectors. The third step includes extracting the PIIFD on Harris surrounding each corner point to identify the unique characteristics of the image. The fourth step involves eliminating redundant point correspondence based on the distance between the image point correspondence, which helps reduce the overall computational load. Finally, the Harris-PIIFD is matched with bilateral matching, a technique used to compare the similarity of two images. This step involves selecting the transformation mode that aligns with the reference image to ensure the accuracy of the image registration. Overall, these five steps create a robust framework for retinal image registration, as shown in Fig. 2.

A. PREPROCESSING OF RETINAL IMAGES

Retinal image registration is a highly intricate and challenging process that involves aligning a series of retina images captured at different points in time. The primary objective of this process is to accurately align the retinal vascular structures, such as arteries and veins, to facilitate effective diagnosis and treatment of various eye conditions. Since these vascular structures can change over time, achieving precise alignment is crucial for accurately tracking and monitoring changes in the retina and detecting any abnormalities or potential diseases. Image preprocessing methods are essential to complete accurate registration. This involves four main processes that prepare the image for registration. One of these processes is segmentation, which separates the vascular from the background structures. However, this task is complex, as retinal vascular and background structures have similar intensities. Therefore, a sophisticated preprocessing approach is required to achieve optimal results. The preprocessing involves four main processes: 1) the input image format is converted to grayscale, which helps reduce the computational load and eliminate the color variations that may hinder the detection process; 2) median filtering is employed to reduce the artifacts or sparkle noise present in the images. This helps produce smoother images that are free from any unnecessary noise that may lead to false detections; 3) mathematical morphology based on Top-Hat operation is applied to enhance the contrast of bright structures against a dark background [27]. The retinal is filtered using an opening operation with a structuring element of diameter larger than the maximum width of the retinal vascular. The opened retinal image is defined as Eq. (1).

$$I_{enhanced} = I_{original} - \left(I_{original} \circ B\right) \tag{1}$$

where $I_{enhanced}$ is obtained by performing the Top-Hat operation on the original retinal image, $I_{original}$ is the original retinal image, B represents a disk-shaped structuring element, and the symbol $\circ$ represents the grayscale opening process. The disk structuring element with a diameter of 7 pixels is selected for this stage, and 4) the retinal image enhancement process involves transforming the enhanced image into a binary format. This is done by applying an optimal threshold based on the desired outcome [28], [29]. This step aims to obtain a clear and precise visualization of the vascular

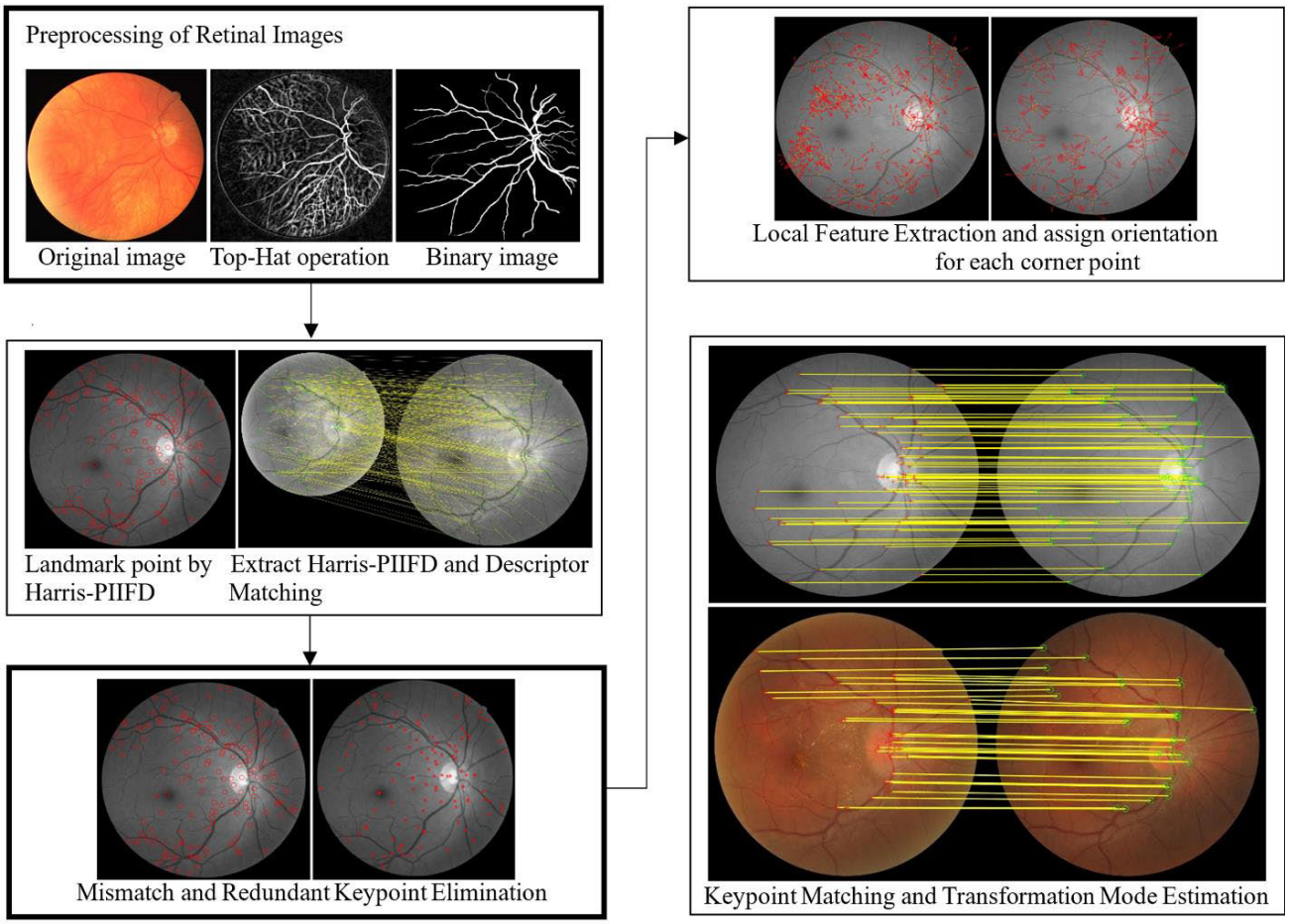

FIGURE 2. Our retinal image registration framework has been represented through a flowchart, where the most significant contribution of this work is emphasized in bold.

structure in the retina. The best threshold can be calculated as Eq. (2).

$$T_O = \frac{1}{2}\left(P_{\max} + P_{\min}\right) \tag{2}$$

To calculate the threshold, we need to consider the original threshold value, denoted by T_O, along with the maximum and minimum values, represented by P_{max} and P_{min}, respectively. Segmenting a retinal image involves separating it into two parts: vascular and background. Once this segmentation is done, we can compute the mean of these two parts, which are $P_{vascular}$ and $P_{background}$, respectively. This can be achieved by following Eq. (3) and Eq. (4).

$$P_{vascular} = \frac{\sum_{f(i,j)>T_k} f(i,j)\cdot w(i,j)}{\sum_{f(i,j)>T_k} w(i,j)}; \tag{3}$$

$$P_{background} = \frac{\sum_{f(i,j)\le T_k} f(i,j)\cdot w(i,j)}{\sum_{f(i,j)\le T_k} w(i,j)}; \tag{4}$$

The function *f(i, j)* represents the pixel's gray value at the coordinates *(i, j)*. On the other hand, the coefficient of the point *(i, j)* is denoted by *w(i, j)* and determines the contribution of that pixel to the overall image processing. For the function in Eq. (4), the coefficient of each pixel is set to a constant value of 1. The algorithm utilizes a formula expressed as Eq. (5) to determine the optimal threshold.

$$T_{k+1} = \left(P_{vascular} + P_{background}\right)/2 \tag{5}$$

This equation considers various factors to compute the ideal threshold value to yield the desired outcome. If $T_k = T_{k+1}$, exit; otherwise, k ← k + 1, and return to Eq. (3) and (4). The retina image has been processed to be smoother using a mathematical morphology operator known as the Top-Hat operation. Additionally, the morphology opening operator has also been applied. The resulting image can be observed in Fig. 3.

B. KEYPOINTS DETECTION USING HARRIS DETECTOR

The Harris detector is a popular method for detecting image keypoints [30], [31], [32]. This technique is widely used due to its computational efficiency and ease of implementation. In particular, the Harris detector is often applied to retinal images due to its invariance to rotation, making it a valuable tool for medical imaging and analysis. The Harris detector is

FIGURE 3. An example of the retinal image smoothed using a mathematical morphology operator with Top-Hat transform and optimal thresholding, (a) original image, (b) the result of applying the operator based on the Top-Hat operation, and (c) the thresholding image.

an algorithm used to detect corners in images. It works by analyzing the changes in image intensity that occur when a Gaussian window is coevolved in all directions. To achieve this, the algorithm first calculates the gradient of the image, which is then used to compute the second-order derivatives of the image in both the x and y directions. These derivatives are then combined to form a matrix known as the structure tensor. In the case of retinal image I, the traditional image gradients are assumed to be given as described in Eq. (6). This implies that the image is being evaluated regarding its gradient values, which measure the rate of change of the image's brightness and color values.

$$\begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \partial I/\partial x \\ \partial I/\partial y \end{bmatrix} \tag{6}$$

The Harris detector utilizes the structure tensor to detect and localize corners in an image. This detector calculates the eigenvalues of the structure tensor using mathematical equations, specifically Eq. (7) and (8), to identify the points in an image with a high-intensity variation in multiple directions.

$$M = \begin{bmatrix} G_x^2 & G_x G_y \\ G_y G_x & G_y^2 \end{bmatrix} * h \tag{7}$$

$$R = \det(M) - k \cdot tra^2(M) \tag{8}$$

The formula for detecting corner points in an image involves several variables and calculations. Specifically, the gradient values of the original image, denoted by G_x and G_y, are used in conjunction with a Gaussian window (h) and a constant value ($k = 0.05$ in our study). The calculations also involve the determinants and traces of the matrix (*det* and *tra*, respectively), which ultimately result in a value R. When the value of R exceeds zero, it indicates that the point in the original image is a corner point. This means that the particular point represents the intersection of two or more edges in the image and can be used to extract valuable information about the object or scene in the image. The proposed retinal image registration framework involves several stages, with the extraction of Harris-PIIFD being one of the most time-consuming. The runtime to complete this stage is proportional to the number of corners and potential control points. In our experiments, we utilized an automated tuning process to detect approximately 169 Harris corner points by adjusting the sigma value of the Gaussian window.

C. REDUNDANT KEYPOINTS ELIMINATION

The Harris-PIIFD corner detector is a popular technique for identifying and extracting keypoints in an image. However, in some cases, the detector may remove redundant keypoints, which can negatively impact the accuracy and efficiency of subsequent image-processing tasks. To address this issue, the redundant keypoints elimination method has been proposed [33], [34]. The process involves two main steps.

1) Keypoints are identified and extracted using the Harris-PIIFD corner detector in the reference image. The number of extracted keypoints is determined based on the complexity of image locations. In complex structures, there tends to be a larger number of keypoints.

2) To detect and match features between two images, the distance between each feature point must be calculated. The distance is calculated between each keypoint and all other keypoints in the reference image, based on a mathematical formula known as Eq. (9) and Eq. (10). This helps to identify and eliminate redundant keypoints, improving the accuracy and efficiency of subsequent image processing tasks.

$$d_1(p_m, p_n) = \sum_{i=1}^{l} |p_m(i) - p_n(i)| \tag{9}$$

$$d_2(p_m, p_n) = \sqrt{\sum_{i=1}^{l} (p_m(i) - p_n(i))^2} \tag{10}$$

where $p_m(i)$ and $p_n(i)$ are the i^{th} coordinates of the keypoints, specific points of interest in the compared images. The keypoints detection is a fundamental task that involves identifying significant points in an image. These points can describe the structure of objects or scenes, typically represented as coordinates in a two-dimensional space. A key metric for evaluating the quality of a keypoints detection algorithm is the sum of the distances between each keypoints and every other keypoints in the image. This metric, commonly referred to as Eq. (11), measures the distance between all pairs of keypoints in the image.

$$SD(p_m) = \sum_{j=1}^{N} d(p_m, p_j) \tag{11}$$

where N represents the total number of keypoints present in the retinal image, $d(p_m, p_j)$ denotes the distance between keypoints p_m, and keypoints p_j is determined based on Eq. (9) or Eq. (10). The redundancy index (*RI*) of a keypoints pm is defined based on Eq. (12), which considers the possibility of redundancy in a place with a complex structure. In practical terms, the more complex the structure of the image, the smaller the distance between keypoints, leading to a greater possibility of redundancy. Therefore, calculating the redundancy index of each keypoints in an image is an important step in understanding its structural complexity. This can be useful in image recognition, object tracking, and registration.

$$RI(p_m) = \frac{1}{SD(p_m)} \tag{12}$$

where $SD_{(Pm)}$ is a widely used mathematical formula to measure the similarity between two images. This formula calculates the sum of distances between each key point in one image and all others in the second image. The keypoints are identified by detecting significant points in the image, such as corners or edges. The distance between each keypoints in the two images is calculated and summed up to obtain the final similarity score.

3) To enhance the performance and accuracy of keypoints matching algorithms, selecting a threshold value that can effectively remove redundant keypoints is necessary. The threshold value determines the minimum distance between keypoints that should be considered for matching. Increasing the threshold value can lead to faster processing time and negatively impact matching precision. To remove redundant keypoints, the algorithm compares the distances between them (excluding the distance of each one from itself) and removes those within the threshold value. Among these keypoints, the one with the lowest distance value is removed. A keypoints with a smaller distance value sum is considered more redundant since it is closer to multiple keypoints. By extracting the most redundant keypoints, the quality and accuracy of the matching method are improved. It should be noted that all of these steps are carried out independently on the sensed image without any external influence or bias. Overall, selecting an appropriate threshold value and removing redundant keypoints can significantly enhance the performance and accuracy of keypoints matching algorithms.

D. LOCAL FEATURE EXTRACTION

Assigning a main orientation to each corner point relative to the local gradients is essential to extracting local features from an image [35]. This ensures that the local feature can be represented in this orientation and achieves invariance with image rotation. Our study introduced a continuous method that utilizes the averaging squared gradients technique to assign the orientation to each corner point or keypoints. This method calculates the averaged perpendicular direction to the gradient, which is then used to represent the keypoints orientation. It is important to note that the orientation has been limited from 0 to π. This limitation is crucial in determining the orientation of the keypoints. For each image sample, $I(x, y)$, the gradient vector $[G_x(x, y)\ G_y(x, y)]^T$ is defined as Eq. (13). This gradient vector plays a crucial role in determining the orientation of the keypoints. Using the averaging squared gradients technique, we can assign a reliable and accurate orientation to each keypoints in an image [31].

$$\begin{bmatrix} G_x(x, y) \\ G_y(x, y) \end{bmatrix} = \text{sgn}\left(\frac{\partial I(x, y)}{\partial y}\right) \begin{bmatrix} \partial I(x, y)/\partial x \\ \partial I(x, y)/\partial y \end{bmatrix} \tag{13}$$

A gradient vector is used to compute the orientation of an image. The gradient vector consists of two elements, one indicating the magnitude of change in the x direction (G_{sx}) and the other indicating the change in the y direction (G_{sy}). However, to avoid any confusion, it is essential to note that the second element of the gradient vector is always chosen to be positive. This consistency is necessary because opposite gradient directions indicate equivalent orientations in symmetric descriptors. However, there is an issue with gradient vectors when averaging them directly. This is because opposite gradient vectors will cancel each other out, even though they indicate the same orientation. To overcome this problem, a proposed solution involves squaring the gradient vector, which is considered a complex number, before averaging. In this approach, the squared gradient vector $[G_{sx}(x, y)\ G_{sy}(x, y)]^T$ is defined by Eq. (14) to ensure that the averaging process yields accurate and reliable results. By squaring the gradient vector, the magnitude of the gradient is preserved, while the direction is made unambiguous. This allows for averaging gradient vectors and the computation of image orientation.

$$\begin{bmatrix} G_{sx}(x, y) \\ G_{sy}(x, y) \end{bmatrix} = \begin{bmatrix} G_x^2(x, y) - G_y^2(x, y) \\ 2G_x(x, y)\, G_y(x, y) \end{bmatrix} \tag{14}$$

Then, the Gaussian weighted average squared gradient can be a useful metric in retinal image processing. This metric is calculated by taking the Gaussian weighted horizontal gradient (G_{sx}) and the Gaussian weighted vertical gradient (G_{sy}) at each pixel location. Once this is done, the result is averaged over a neighborhood area. A Gaussian weighted circular window determines the size of a neighborhood area. A standard deviation determines this window (σ) value, calculated using Eq. (15). The standard deviation value measures the spread of values in a dataset. In this context, it controls the window size for processing the image. By adjusting the value of σ, the window size can be increased or decreased, allowing for more or less neighboring pixels to be included in the calculation. In other words, a larger σ value will result in a larger neighborhood area being used to calculate the average squared gradient.

$$\begin{bmatrix} \overline{G}_{sx} \\ \overline{G}_{sy} \end{bmatrix} = \begin{bmatrix} G_{sx} * h_\sigma \\ G_{sy} * h_\sigma \end{bmatrix} \tag{15}$$

The formula presented in Eq. (15) involves a Gaussian-weighted kernel along with the convolution operator represented by the symbol $*$. The outcome of applying this formula can determine the most significant direction of each locality ϕ, where the value of ϕ will be between 0 and π. This dominant direction is determined through the use of Eq. (16).

$$\varphi = \frac{1}{2}\begin{cases} \tan^{-1}\left(\overline{G}_{sy}/\overline{G}_{sx}\right) + \pi & \overline{G}_{sx} \geq 0 \\ \tan^{-1}\left(\overline{G}_{sy}/\overline{G}_{sx}\right) + 2\pi & for\ \overline{G}_{sx} < 0 \cap \overline{G}_{sy} \geq 0 \\ \tan^{-1}\left(\overline{G}_{sy}/\overline{G}_{sx}\right) & \overline{G}_{sx} < 0 \cap \overline{G}_{sy} < 0 \end{cases} \tag{16}$$

Computing the symmetric keypoints descriptor involves assigning an orientation to each key coordinate, denoted by (x, y). This orientation is determined by $\phi(x, y)$. To extract the sub-descriptor, we compile a vector consisting of the

values of all the orientation histogram entries that correspond to the lengths of the arrows. This sub-descriptor provides a comprehensive representation of the orientation features present in the image.

E. KEYPOINTS MATCHING

In our retinal image registration process, we use a highly efficient algorithm called the best-bin-first algorithm to ensure that the correspondences between two images are accurately matched [31], [35]. This algorithm is specifically designed to identify the approximate closest neighbors of points in high-dimensional spaces, significantly reducing the computational cost and time required for processing. However, it's important to note that the results obtained using this algorithm are only sometimes 100% accurate, as it only returns the nearest neighbor with a high probability. Nonetheless, the algorithm's ability to efficiently identify the closest neighbors of points in high-dimensional spaces makes it a highly valuable tool in retinal image registration and other similar applications. To understand how this algorithm works in our system, let's consider two images, I_1 and I_2, and the sets of all symmetric descriptors of these images, DES_1 and DES_2, respectively.

For a given descriptor, $des \in DES_1$, a set of distances is defined as Eq. (17). The best-bin-first method then uses this set of distances to identify the approximate closest neighbor in DES_2 for each descriptor in DES_1. This matching process is crucial for our image processing system, allowing us to compare and analyze images accurately. We can efficiently and effectively match correspondences between two images by utilizing the BBF algorithm and defining sets of distances for each descriptor.

$$Dis_{-des} = \{des \cdot des_s \,|des_s \in DES_2\} \tag{17}$$

When working with vector data, we often need to compare the distances between vectors to identify patterns or similarities. One common way of doing this is by using the dot product of vectors, denoted by a dot symbol ($\bullet$). By applying this method to a set of vectors, we can obtain a list of all the distances between each vector and all the other vectors in the set. The set of des contains multiple values; among them, the largest value represents the nearest neighbor of Dis_{-des}. In other words, des is closest to the value in Dis_{-des}, which is the largest in magnitude. If the closest neighbor to a given descriptor is significantly closer than the second-closest neighbor, we can consider it a unilateral match or correspondence from DES_1 to DES_2. In other words, if the closest neighbor is much closer than any other potential match, we can be confident that it is the best match, and we can use it to establish a correspondence between the two sets of descriptors. On the other hand, if the distance between the closest and second-closest neighbors is relatively small, we cannot be sure which one is the best match, and the descriptor should be discarded to avoid introducing errors in the matching process. To ensure that we only consider high-quality matches when using the nearest-neighbor criterion, we can set a threshold value (T). This threshold value determines the maximum acceptable distance between the feature vectors of two potential matches. Any potential match that exceeds this threshold value will be discarded, reducing the likelihood of false matches and improving the overall accuracy of the matching process. By setting an appropriate threshold value, we can ensure that only reliable and accurate matches are considered when using the nearest-neighbor criterion. In our study, we set this value to 0.8 empirically. Any matches with a distance ratio above this threshold are discarded. However, the unilateral matching process needs to be more foolproof and can lead to mismatches. One common issue is when two descriptors in I_1 are matched to the same descriptor in I_2. To avoid this, we can use a bilateral matching process. This involves applying the same matching process to both sets of vectors ($M(I_1, I_2)$ and $M(I_2, I_1)$) and then identifying the matches that occur in both sets (i.e., the bilateral matches). Overall, identifying matches or correspondences between sets of vectors involves careful consideration and calibration of various parameters and techniques. Using a combination of dot product calculations, nearest-neighbor criteria, and bilateral matching processes, we can obtain high-quality matches useful for retinal image registration. Bilateral matching is where two sets of objects are to be paired based on specific criteria. The objects on the left set are matched with the ones on the right set, and each pair is assigned a score based on their compatibility. Bilateral matching aims to find the optimal pairs that maximize the overall score while ensuring that each object is matched with exactly one other object.

F. SELECTION TRANSACTIONS MODE

In our registration framework, this particular subsection holds immense importance. While our study has multiple focus areas, this subsection is crucial to the overall process. Our framework employs a range of transformation modes, including linear transformations [36], [37], affine transformations [38], [39], and second-order polynomial transformations [31] to choose the most suitable option for the task. The choice of transformation mode depends on the number of matches found during the process. The linear conformal transformation, being the simplest, requires a minimum of two pairs of control points to apply. Hence, no transformation is applied to the floating image in cases where only one match exists. However, we use the linear conformal transformation mode if there are at least two matches. Similarly, we use the affine transformation mode if there are three to six matches. Finally, we opt for the second-order polynomial transformation mode if there are six or more matches. After selecting the appropriate transformation mode, we apply it to the floating image and superimpose the transformed image on the fixed image to create a mosaic image. This approach ensures that we achieve the best possible results while maintaining the integrity of the registration framework. The image pair contains 69 corresponding points, carefully selected to cover the overlapping

surface area between the images comprehensively. The correspondence points have been placed across the image to ensure broad coverage, enabling us to calculate the accuracy across the whole image. The majority of the points of correspondence are located in vessels and crossings, as these areas have a well-defined image structure and can be manually selected with accuracy. The annotator provided precise initial markings for these points, which is quite challenging in other image areas lacking such clear structure. The number of correspondences was carefully chosen by balancing the trade-off between the annotator's time availability, the annotations' accuracy, and the number of marked images.

IV. EXPERIMENTAL RESULTS

The primary objective of the experiments was to explore the most effective configuration of all the components involved in the Harris-PIIFD and RKE frameworks for developing a precise and reliable retinal image registration technique. These experiments involved multiple steps, such as detecting points using the Harris detector, eliminating redundant keypoints, extracting local features, matching keypoints, and selecting the most appropriate transformation mode. Additionally, the experiments aimed to conduct a comprehensive and quantitative evaluation of the proposed method compared to other state-of-the-art retinal image registration methods.

A. CONFIGURATION AND SIMULATION RESULTS

The experiments for the proposed retinal image registration framework were conducted on Intel(R) Core(TM) i7-6700K CPU @ 4.00GHz Desktop with 16 GB RAM running on Windows 10 Professional operating system. The framework was implemented using the powerful and widely used MATLAB R2024 applications, ensuring accuracy and reliability in the results.

Three preprocessing operations are performed in our suggested Harris-PIIFD with the RKE framework to detect control point candidates. First, the input image format is converted to grayscale. Second, the intensities of the input image are scaled to the total intensity range [0, 255]. Third, the image is resized to a fixed size of approximately 756×756 pixels. The framework comprises five steps, starting with corner points as control point candidates instead of bifurcations. This is because corner points are sufficiently and uniformly distributed across the image domain. Two subsets of control point candidates are assumed, which could be identically matched across two images. In our experiments, a constant value of $k = 0.05$ and approximately 169 Harris corner points are used, with the sigma of the Gaussian window being automatically tuned (step 1). However, in some cases, the detector may extract redundant keypoints, which can negatively impact the accuracy and efficiency of subsequent image-processing tasks. Therefore, redundant keypoints elimination criteria are used to remove them. The assumption is that there are more correct matches than incorrect matches. In our framework, this assumption is satisfied when the threshold for the nearest neighbor criterion in bilateral matching is set to equal to or less than 0.7 (step 2). To extract local features from an image, assigning a main orientation to each corner point relative to the local gradients is essential. The Gaussian window size is used to control the processing of the image, and in this work, the Gaussian window size is set to 6 pixels empirically (step 3). Next, the best-bin-first algorithm matches the correspondences between two images. This algorithm identifies the approximate closest neighbors of points in high-dimensional spaces. In this study, the threshold for the nearest-neighbor criterion is set to 0.8 empirically (step 4). Finally, an adaptive transformation mode is applied to register the image pairs based on these matched control point candidates (step 5). These steps are all part of our suggested Harris-PIIFD with RKE framework, designed to accurately and efficiently detect and match control points in images.

B. DATASETS

The proposed method involves a two-step approach for detecting and extracting vascular trees from retinal images, which is crucial for accurate retinal image registration. To achieve this, we first train our model on the DRIVE dataset [40], a publicly available dataset containing retinal images with ground truth segmentations. We use this dataset to learn how to correctly detect and extract keypoints and bifurcations in the vascular trees. This step is critical as it allows us to improve the accuracy of the proposed method. Once our model is trained on the DRIVE, we validate its performance on the FIRE dataset, another publicly available dataset containing retinal images with appropriate registration labeling. By testing our model on the FIRE dataset, we can evaluate its effectiveness in detecting and extracting vascular trees in a different dataset, ensuring the proposed method is robust and reliable. For instance, the DRIVE dataset is a collection of retinal images of size 584×565, relatively smaller than the FIRE dataset. The FIRE dataset, on the other hand, has a resolution of $2,912 \times 2.912$, which is significantly higher and more detailed. It's worth mentioning that the DRIVE dataset comprises images of individuals who have been diagnosed with diabetes. However, only a small subset of these individuals, precisely seven out of the forty images, exhibit signs of diabetic retinopathy, a background retinopathy affecting the eye. In contrast, the FIRE dataset consists of 134 image pairs, each of which can be classified (S, P, and A) based on their unique features [41]. This highlights that although both datasets originate from the same underlying disease, the symptoms and characteristics can vary significantly between datasets, adding a layer of complexity when analyzing and interpreting the data. Therefore, it is essential to consider the dataset characteristics while working on retinal images carefully.

C. EVALUATION CRITERIA

Various methods are developed and evaluated in retinal image registration based on parameters crucial in determining their

accuracy and efficiency in aligning retinal images. These parameters include root mean square error, precision, recall, matching accuracy, and registration error. These metrics assess the method's performance in terms of its ability to align retinal images with high precision and accuracy. These metrics are given as follows.

The Root Mean Square Error (RMSE) is a widely used statistical measure for assessing the accuracy of a given registration method [42]. The RMSE value is determined by quantifying the difference between the experimental matching points obtained by the registration method and the reference matching points obtained through manual calibration. The RMSE value is calculated using the formula in Eq. (18), which involves computing the squared differences between the registered and reference matching points, averaging them, and then taking the square root of the result.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_C} \left\| (x_i, y_i) - \left(x_i^{ref}, y_i^{ref}\right) \right\|_2}{Number\ of\ matching\ points}} \tag{18}$$

where (x_i, y_i) and (x_i^{ref}, y_i^{ref}) represent the coordinates of the i^{th} experimental and referenced matching points, respectively. This value is calculated by measuring the distance between the reference and target points, and the lower the RMSE value, the more accurate the matching points are. On the other hand, a higher RMSE value indicates that the matching points are inaccurate, and the registration process has resulted in a larger error. Therefore, it is essential to strive for lower RMSE values to ensure the accuracy of the registration process during experiments involving image registration.

Precision is a metric commonly used to evaluate the accuracy of a registration method [43]. It is the ratio of correctly matched points to the total number of matches obtained. The calculation of precision is typically done using Eq. (19), which takes into account the number of true positives (i.e., correctly matched points), false positives (i.e., incorrectly matched points), and false negatives (i.e., missed matches). Precision is essential in assessing the effectiveness of registration techniques, as it indicates the method's ability to align two or more data sets accurately. A high precision score means the registration method is reliable and produces fewer false positives and negatives.

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

The term *TP* (True Positive) refers to the situation in which a keypoints is correctly matched to its corresponding point in an image. At the same time, *FP* (False Positive) denotes the situation in which a keypoints is incorrectly matched.

Recall is a metric used in feature-matching methods to determine their effectiveness [43]. It measures the number of correct matches obtained through feature-matching by comparing them to the number of putative accurate matches determined manually after detecting feature points. In technical terms, recall is the ratio of correct matches to the number of putative accurate matches, as stated in Eq. (20). In other words, recall measures how many correct matches were identified out of all the precise potential matches detected.

$$Recall = \frac{\#\ Correct\ Matches}{\#\ Correctsponding\ Features} \tag{20}$$

where # Corresopnding Feature refers to the count of feature points correctly matched between two images. The matching accuracy metric is then calculated by determining the ratio of correctly matched keypoints pairs to the total number of pairs [44]. The formula to compute the matching accuracy can be found in Eq. (21).

$$Accuracy = \frac{Total\ number\ of\ correctly\ matched\ pairs}{Total\ number\ of\ pairs} \tag{21}$$

Finally, the success of the registration process hinges on calculating a parameter called "registration error." This error is computed by measuring the average distance, in pixels, between 64 reference images and their corresponding target images [44]. The registration error is a quantitative measure of the degree of similarity between the two sets of images. A lower value of the registration error indicates that the images are more closely aligned, and hence, the registration process is more accurate. A mathematical formula in Eq. (22) calculates the registration error.

$$Registration\ error = \frac{1}{N} \sum_{i=1}^{N} \left\| c_i - c_j \right\|_2 \tag{22}$$

D. EXPERIMENTAL ANALYSIS

The performance of the Harris-PIID with the RKE algorithm and the original Harris-PIID algorithms have been evaluated through 134 image pairs of tests. The test set includes various retinal images, and both the basic algorithms (Harris-PIID) and proposed methods (Harris-PIID with RKE) have been compared based on the extracted keypoints. The precision and recall have been evaluated to determine the accuracy of the process. Moreover, the last experiment assessed the performance of the keypoints matching in the registration process. The threshold value in both images (the reference and the target images) has been considered the same for simplicity. The effect of the proposed method will be discussed in the next section in more detail.

1) EXPERIMENT OF REGISTRATION PROCESS

Retinal image registration is vital to aligning two images of the same scene captured under different conditions. The mechanism ensures that the reference and target images are geometrically aligned, which is essential for detecting and analyzing retinal diseases. Therefore, it is crucial to improve the retinal image registration process. In this regard, a new method called Harris-PIIFD was proposed, and its effectiveness was tested through a series of experiments on simulated retinal images captured at different times. The results showed

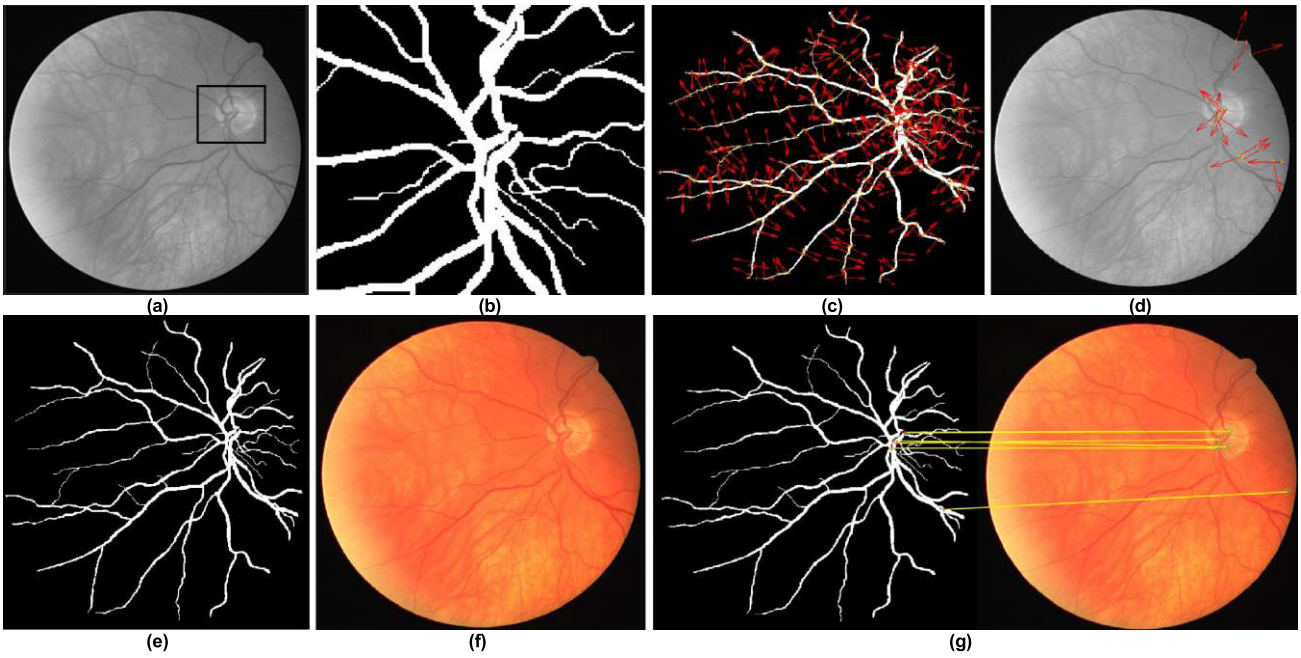

FIGURE 4. Visual overview of the step-by-step process for registering a pair of retinal images. (a) the reference images are displayed in greyscale, and a black square indicates a closer look at the vascular detail, (b) vascular extraction, (c) shows the orientation of the corresponding keypoints using the Harris-PIIFD, (d) the orientation of the corresponding keypoints using Harris-PIIFD and RKE is illustrated, (e) the source image, (f) the target image, (g) the process involves projective registration by keypoints matching. This process is repeated for all adjacent frame pairs, with the initial frame serving as the anchor frame.

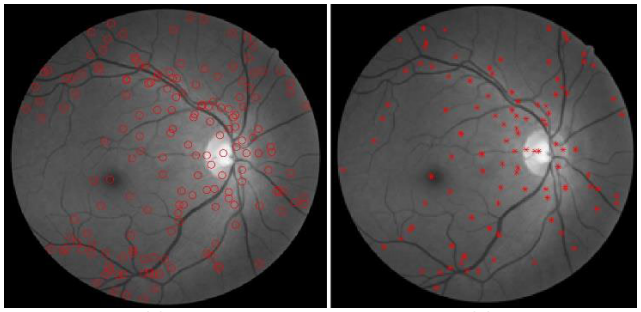

FIGURE 5. Retinal images were taken at different times and compared using two methods, (a) the traditional keypoints correspondence based on Harris-PIIFD. However, this method may encounter difficulty detecting specific landmark features, which can lead to the failure of registering the two images, (b) Extracting key points using Harris-PIIFD and eliminating redundant keypoints correspondence. This method offers higher accuracy and reliability in detecting keypoints and, thus, is more suitable for analyzing retinal images taken at different times.

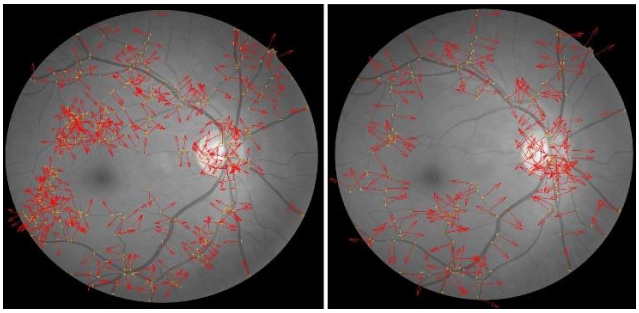

FIGURE 6. The symmetric descriptor identifies specific points of interest, the keypoints, to extract meaningful information from an image, (a) and (b) the gradient magnitude and the orientation at each sample point in the region around the keypoints location were calculated using Harris-PIIFD and Harris-PIIFD with eliminating redundant keypoints correspondence, respectively.

that the performance of the registration process improved significantly with the redundancy keypoints elimination (RKE) approach. The image parts not correctly registered by the Harris-PIIFD algorithm were well registered using the RKE method, as depicted in Fig. 4. Hence, the RKE method effectively improves the retinal image registration process.

2) COMPARISON OF KEYPOINTS EXTRACTION BY HARRIS-PIIFD AND HARRIS-PIIFD WITH RKE

This experiment's primary objective is to extract keypoints from retinal images. The process of extracting these results of extracting keypoints in retinal images can be seen in Fig. 5.

It is worth mentioning that the extracted keypoints are represented as points shown with a circle in the classic Harris-PIIFD algorithms. These points serve as a reference to the areas around which the overlapped keypoints are located. The original Harris-PIIFD algorithm produces a large number of keypoints that tend to overlap with each other. To overcome this issue, a technique called RKE is applied, removing the overlapped keypoints from the image, as seen in Fig. 5(b). This process ensures that the subsequent processes, including the symmetric descriptor, keypoints extraction (as shown in Fig. 6), and the matching process (as shown in Fig. 7), are not affected by the overlapping keypoints. The RKE technique proves to be highly effective in preventing interference and extracting meaningful information from the image.

3) EXPERIMENT ON IMAGES BY THE EVALUATION CRITERIA

This section has conducted a comprehensive experiment to determine an optimal threshold value for the proposed

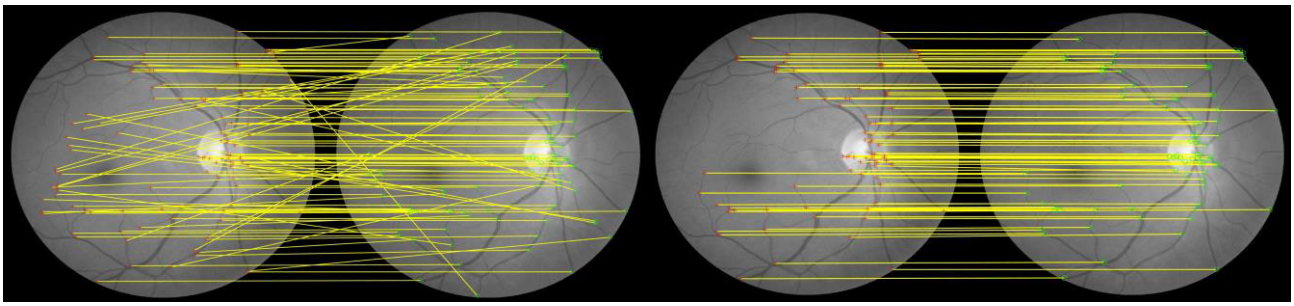

FIGURE 7. Matches that are identified by the proposed method between the two images in Fig. 5, (a) initial matching results using Harris-PIIFD, (b) results after removing the incorrect matches and eliminating redundant keypoints correspondence for improved accuracy and reliability.

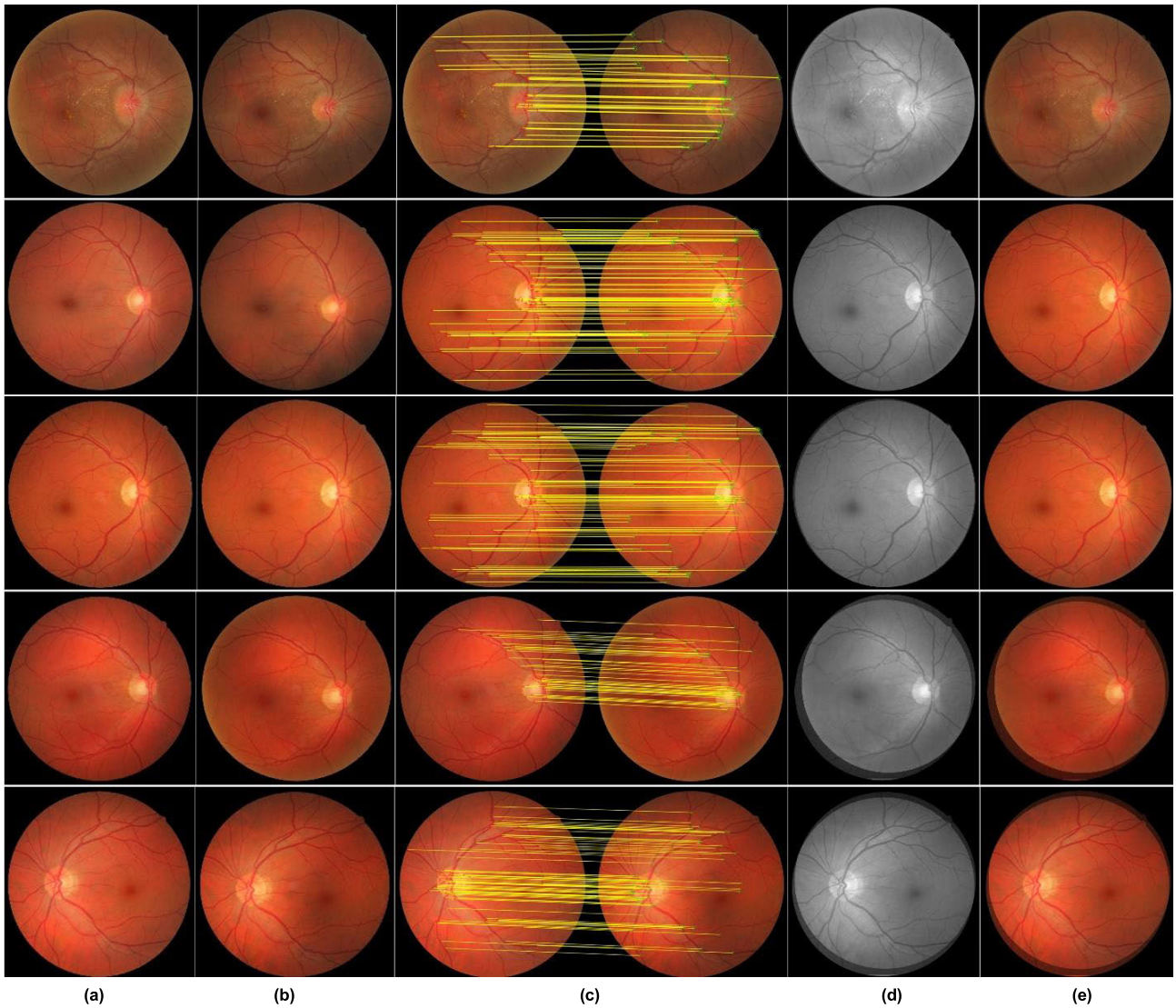

FIGURE 8. Qualitative comparison between source and target images on FIRE. We aim to create a foundational dataset that achieves super-matching accuracy in Harris-PIIFD with the RKE technique to compare source and target images qualitatively, (a) source image, (b) target image, (c) matched image, (d) registered image on grayscale, (e) registered image on RGB color space.

Harris-PIIFD and RKE methods. The experiment involved analyzing 134 pairs of images with varying imaging conditions, such as different categories, categories, and intensities.

In this experiment, a pair of images at different times on FIRE data is used, and the performance of registrations on the matching process is evaluated, which can be seen in Fig. 8. Fig. 9(a) and 9(b) show the iteration-wise plot of the registration error for traditional Harris-PIIFD and Harris-PIIFD with RKE, respectively. Interestingly, we notice that the error of Harris-PIIFD with RKE is significantly lower

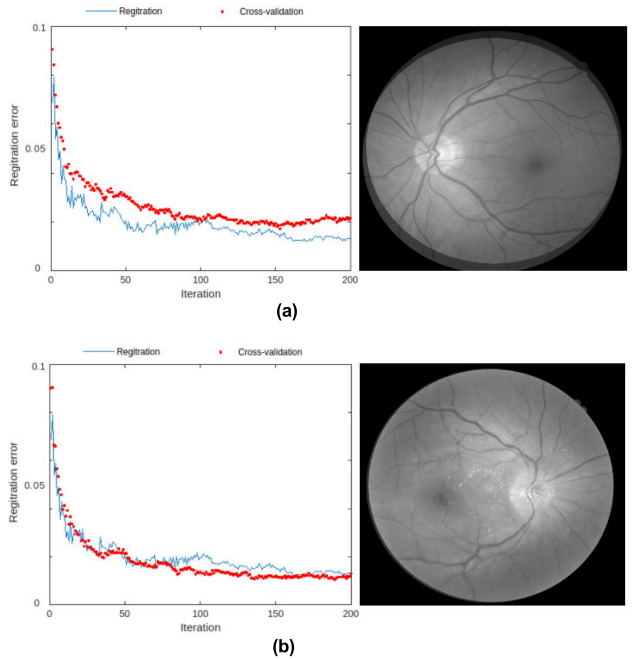

FIGURE 9. The iteration-wise plot of the registration error, (a) traditional Harris-PIIFD, (b) Harris-PIIFD and RKE (proposed). The blue lines represent the segmentation using different methods, and the red lines represent the performance comparison.

TABLE 2. Our proposed registration method has been extensively tested on a wide range of detected image pairs of source and target images to guarantee improved accuracy on the FIRE dataset.

Methods	Performance				
	Precision	Recall	Accuracy	RMSE	Error
Harris-PIIFD	0.91325	0.91350	0.91322	0.21338	0.08687
Harris-PIIFD + RKE	0.98240	0.98312	0.98284	0.01280	0.01716

than that of Harris-PIIFD in the 200 iterations, although both have the same learning rate. This observation is consistent with the keypoints matching results depicted in Fig. 7 and 8 and is further validated by the test results presented in Table 2. The incorporation of RKE into Harris-PIIFD leads to a significant improvement in the accuracy of the registration process, as reflected in the lower error rates.

Table 2 presents an analysis of the effectiveness of the proposed method using various assessment criteria, such as precision, recall, matching, RMSE, and Error scores. The assessment was conducted first using the Harris-PIIFD algorithm and then using Harris-PIIFD with RKE. The results indicate that the proposed Harris-PIIFD with RKE method is more effective in assessing the above criteria. The analysis reveals that the original Harris-PIIFD algorithms have some unnecessary points that cause significant interference in the matching process. However, this study's proposed Harris-PIIFD with the RKE method has significantly improved the matching process. The effectiveness of the proposed method can be attributed to the fact that it reduces the irrelevant points that can cause interference and noise in the matching process, leading to better precision, recall, matching, RMSE, and error values.

TABLE 3. Our proposed registration method has been extensively tested on a wide range of detected image pairs of source and target images to guarantee improved accuracy on the FIRE dataset.

Methods	Harris-PIIFD	Harris-PIIFD+RKE
Total Time (s)	1,585	412

4) SPECIAL EXAMINATION

This experiment's main objective is to compare the execution times for Harris-PIIFD and RKE and the original Harris-PIIFD framework. The proposed algorithm undergoes an in-depth study of the execution times of its various components to pinpoint any potential bottlenecks that may arise. This analysis serves to identify areas for improvement and optimization, ensuring that the algorithm's performance is as efficient and effective as possible. The data presented in Table 3 provides the total registration time of 137 pairs, measured in seconds, for both the proposed Harris-PIIFD with the RKE method and the original Harris-PIIFD (average 3.0119 s per image) method (average 11.870s per image). The results show that the Harris-PIIFD with RKE method outperforms the original Harris-PIIFD method in speed, making it the fastest-performing method compared to other original Harris-PIIFD methods (see Table 3). However, it is worth noting that the original Harris-PIIFD method still falls within acceptable time limits despite taking longer than other competing needs to catch up on performance registration. This means that using the original Harris-PIIFD method for registering image pairs within an examination session is still feasible. Overall, the presented data provides valuable insights into the performance of these two methods and can help inform decisions regarding the most appropriate method to use for image registration.

The proposed method for detecting keypoints and matching image pairs is an outstanding solution with high efficiency and processing speed. It is capable of processing an image pair with a resolution of 758 × 758 pixels in just 3.0119 seconds. The process involves approximately 1.5847 seconds for keypoints detection using the Harris-PIIFD and RKE methods on both images. The remaining 1.4272 seconds are dedicated to the matching process. The fastest stage of the proposed methods takes execution times less than the original Harris-PIIFD method, which is 1,173 seconds. One of the key advantages of the proposed Harris-PIIFD method is its utilization of RKE, which offers several benefits over classical methods in computation time. The RKE method can efficiently extract descriptors from keypoints regions, which is crucial for fast and accurate matching. By combining the Harris-PIIFD and RKE methods, the proposed solution achieves superior performance in speed and accuracy. Then, the extraction process was implemented by utilizing acceleration techniques and redundancy keypoints eliminations, which are the primary issues that often arise during the keypoints extraction and matching stages and in the image formation stages. However, parallelization of these stages has yet to be implemented.

TABLE 4. Average keypoints extraction of the proposed Harris-PIIFD with RKE, compared with the original Harris-PIIFD method on FIRE (137 pairs).

A Total number of Keypoints Extraction (points)	Methods	
	Harris-PIIFD	Harris-PIIFD + RKE
	23,153	9,453

The extraction of keypoints is a time-consuming process, especially when it comes to bifurcation extraction which requires vascular segmentation. Table 4 provides an overview of the average keypoints extraction of the proposed Harris-PIIFD with RKE, compared with the original Harris-PIIFD method, for completing the registration process of an image pair.

Although the proposed method may use fewer keypoints than competing methods, it compensates for it by delivering highly accurate registration results. Overall, the Harris-PIIFD and RKE approach offers significant improvements over classical methods, making it a highly efficient and accurate solution for detecting keypoints and matching image pairs. The proposed methods reduce key point matching and computation time while providing accurate results.

5) STATE OF THE ART COMPARISON

We compare our approach with the popular registration methods using DNNs [7], global and local geometric structures [20], and GMM [26] with manually tuned parameters. DNNs is a machine learning algorithm that has proven advantageous for retinal image registration. DNNs can accurately align and register retinal images by utilizing multiple layers to process and extract features from the images, which is crucial for diagnosing and monitoring various eye diseases. Its ability to handle complex and large datasets and its adaptability to different image variations make it a powerful tool in ophthalmology. However, DNNs have a disadvantage when it comes to retinal image registration. The complexity of DNNs makes it challenging to understand the learned features, which is crucial in medical applications where interpretability is essential. Additionally, DNNs require a large amount of labeled data to train. This can be a limiting factor in the medical domain, where obtaining labeled data can be difficult and expensive. Having global and local geometric structures in a retinal registration process provides numerous advantages. The global geometric structure refers to the overall shape and orientation of the retina, while the local geometric structure pertains to the finer details and features within the retina. Additionally, using both global and local geometric structures helps reduce errors and increase the efficiency of the registration process. However, global and local geometric structures in the images can pose a significant disadvantage during the registration process. The global geometric structures, such as the optic disc and blood vessels, can cause misalignment. In contrast, the local geometric structures, such as retinal folds and wrinkles, can cause deformations and distortions. These factors can lead to inaccurate registration results, affecting the diagnosis and treatment of retinal diseases. GMM is particularly advantageous for retinal registration because they separate multiple image components, such as background noise and signal, and accurately model their distribution. By accurately modeling these components, the model can estimate the parameters of each feature and then use this information to align retinal images. Additionally, GMM can identify and remove outliers in the image, further improving the registration accuracy. However, GMM has been shown to have limitations in accurately registering retinal images due to its assumption of normality and the presence of noise in the images. Our algorithm has achieved impressive RMSE scores, indicating that it can accurately identify and classify data points. These results suggest that our algorithm has the potential to offer a more accurate and efficient solution for data analysis tasks, which are summarized in Table 5.

TABLE 5. Average of RMSE, and running times for DNNs (7), Global and local geometric structure [20], GMM [26], and Our method for retinal image registration.

Authors	Methods	RMSE	Times (s)
D. Bi (2019) [20]	Global and Local geometric structure	1.063	0.331
C. Liu (2016) [26]	Gaussian mixture model (GMM)	1.735	1.026
Our Methods	Harris-PIID with RKE	0.012	3.011

6) EXPERIMENTAL LIMITATIONS

The proposed method for image registration is a strong competitor to the best available registration methods. Despite being based on a simple pipeline, our method requires less parameter tuning than state-of-the-art methods. However, it is essential to note that our method may only sometimes produce the best results, as some state-of-the-art methods may outperform it. Therefore, analyzing results for each category is imperative to determine where the method performs well and falls short. We have observed that the results fluctuate depending on the dataset. For instance, our method can place second best in the overall state-of-the-art ranking in the FIRE dataset, which has high degrees of overlapping. The proposed method combines Harris-PIIFD and RKE techniques to register images. This technique employs a relatively simple transformation compared to all other available methods. Despite its simplicity, this technique offers several advantages, such as the minimal requirement for keypoints. This feature enables the registration of images with severe pathology progression. However, the method has a significant drawback: its limited capability to align images with low overlapping and high expected transformations. Therefore, while the proposed technique offers some distinct advantages, there may be better fits for some scenarios.

V. DISCUSSION AND CONCLUSION

The paper presents a novel framework for detecting domain-specific keypoints correspondences in retinal images. The proposed approach combines the Harris-PIIFD with RKE detectors with the best-bin-first algorithms for point matching, providing accurate results based on

state-of-the-art methods. This approach is the first to use RKE in a Harris-PIIFD method, enabling the detection of sparse, domain-specific keypoints correspondences and using them to register image pairs. The keypoints correspondence detector has been trained using the DRIVE dataset, which provides ground truth information for both crossovers and bifurcations. The proposed method was tested on the FIRE dataset containing registration labeling to evaluate its effectiveness. This step is crucial as it helps assess the performance of our proposed method even without an appropriate dataset. One of the significant advantages of the proposed method is that it does not require the expensive computation of advanced descriptors for each feature and keypoints. Instead, it uses domain-specific features, resulting in a lower number of feature detections when compared to classical keypoints detectors. This makes it computationally viable to test every possible landmark match combination. The required information for this method includes the area-based and local feature-based, differentiating between crossovers and bifurcations. This allows for less computation than state-of-the-art methods, making our proposal the fastest in the state of the art by orders of magnitude. While competing state-of-the-art methods take minutes, our approach takes less than a second to register each image pair. Features are matched with bilateral matching to register the images based on the best-bin-first algorithms for computing the similarity matrix. If the image pair is accepted, the simplest affine transformation mode is used for just two-point matching. Thus, our method is resistant to image pairs with severe pathology progress, which may impede retinal diseases. The proposed approach was validated by comparing it with the previous work based on classical methods that detect the same features. Our method outperformed the previous one, and as both share the same matching mechanism, we can affirm that detecting crossovers and bifurcations using Harris-PIIFD with RKE outperforms classical approaches. Moreover, unlike the current feature-based method, the proposed method can register the images in the FIRE dataset despite their category. The experimental results for the proposed method are satisfactory. The result obtained for the proposed approach for retinal image registration of precision 0.98240, recall 0.98312, RMSE 0.01280, ERR 0.01716, and matching score for the proposed method is 0.98284 with the computational time taken 3.01s.

In the upcoming work, the authors plan to expand the scope of their investigation by testing the transformation mode with more degrees of freedom than the similarity transformation used in the current proposal. The higher transformation mode is expected to provide greater flexibility in accommodating the larger displacements of the features, which in turn will help improve the overall results. Furthermore, the authors aim to extend their approach to detect and describe the crossings and bifurcations, which are key features in many clinical applications. Doing so can enhance the approach's diagnostic accuracy and effectiveness. Finally, the authors are exploring the possibility of developing novel feature extraction and matching to optimization methods that are more efficient and less computationally expensive. This would make their approach more accessible for widespread clinical use, enabling them to diagnose and treat patients more effectively and efficiently.

DECLARATION OF COMPUTING INTEREST

The authors declare no conflict of interest associated with this study.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and editors for their insightful comments and helpful suggestions, which improved the quality of our manuscript.

REFERENCES

[1] Diabetes, *Archived From the Original on 29*, World Health Organization, Geneva, Switzerland, Jan. 2023.

[2] K. Wisaeng and W. Sa-Ngiamvibool, "Exudates detection using morphology mean shift algorithm in retinal images," *IEEE Access*, vol. 7, pp. 11946–11958, 2019, doi: 10.1109/ACCESS.2018.2890426.

[3] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, "A survey of medical image registration—Under review," *Med. Image Anal.*, vol. 33, pp. 140–144, Oct. 2016, doi: 10.1016/j.media.2016.06.030.

[4] X. Cheng, L. Zhang, and Y. Zheng, "Deep similarity learning for multimodal medical images," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 3, pp. 248–252, May 2018, doi: 10.1080/21681163.2015.1135299.

[5] B. Zou, Z. He, R. Zhao, C. Zhu, W. Liao, and S. Li, "Non-rigid retinal image registration using an unsupervised structure-driven regression network," *Neurocomputing*, vol. 404, pp. 14–25, Sep. 2020, doi: 10.1016/j.neucom.2020.04.122.

[6] D. Motta, W. Casaca, and A. Paiva, "Vessel optimal transport for automated alignment of retinal fundus images," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 6154–6168, Dec. 2019, doi: 10.1109/TIP.2019.2925287.

[7] S. K. Saha, D. Xiao, A. Bhuiyan, T. Y. Wong, and Y. Kanagasingam, "Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: A review," *Biomed. Signal Process. Control*, vol. 47, pp. 288–302, Jan. 2019, doi: 10.1016/j.bspc.2018.08.034.

[8] S. Abbasi-Sureshjani, I. Smit-Ockeloen, E. Bekkers, B. Dashtbozorg, and B. T. H. Romeny, "Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Prague, Czech Republic, Apr. 2016, pp. 189–192, doi: 10.1109/ISBI.2016.7493241.

[9] Á. S. Hervella, J. Rouco, J. Novo, M. G. Penedo, and M. Ortega, "Deep multi-instance heatmap regression for the detection of retinal vessel crossings and bifurcations in eye fundus images," *Comput. Methods Programs Biomed.*, vol. 186, Apr. 2020, Art. no. 105201, doi: 10.1016/j.cmpb.2019.105201.

[10] H. Pratt, B. Williams, J. Ku, C. Vas, E. McCann, B. Al-Bander, Y. Zhao, F. Coenen, and Y. Zheng, "Automatic detection and distinction of retinal vessel bifurcations and crossings in colour fundus photography," *J. Imag.*, vol. 4, no. 1, pp. 1–14, Dec. 2017, doi: 10.3390/jimaging4010004.

[11] S. Miao, Z. J. Wang, and R. Liao, "A CNN regression approach for real-time 2D/3D registration," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1352–1363, May 2016, doi: 10.1109/TMI.2016.2521800.

[12] B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Isgum, "A deep learning framework for unsupervised affine and deformable image registration," *Med. Image Anal.*, vol. 52, pp. 128–143, Feb. 2019, doi: 10.1016/j.media.2018.11.010.

[13] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A learning framework for deformable medical image registration," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019, doi: 10.1109/TMI.2019.2897538.

[14] D. Mahapatra, B. Antony, S. Sedai, and R. Garnavi, "Deformable medical image registration using generative adversarial networks," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Washington, DC, USA, Apr. 2018, pp. 1449–1453, doi: 10.1109/ISBI.2018.8363845.
[15] J. Lee, P. Liu, J. Cheng, and H. Fu, "A deep step pattern representation for multimodal retinal image registration," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Seoul, (South) Korea, Oct. 2019, pp. 5076–5085, doi: 10.1109/ICCV.2019.00518.
[16] Y. Wang, J. Zhang, M. Cavichini, D. G. Bartsch, W. R. Freeman, T. Q. Nguyen, and C. An, "Robust content-adaptive global registration for multimodal retinal images using weakly supervised deep-learning framework," *IEEE Trans. Image Process.*, vol. 30, pp. 3167–3178, 2021, doi: 10.1109/TIP.2021.3058570.
[17] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Kerkyra, Greece, Sep. 1999, pp. 1150–1157, doi: 10.1109/ICCV.1999.790410.
[18] C. Aguilera, F. Barrera, F. Lumbreras, A. D. Sappa, and R. Toledo, "Multispectral image feature points," *Sensors*, vol. 12, no. 9, pp. 12661–12672, Sep. 2012, doi: 10.3390/s120912661.
[19] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008, doi: 10.1016/j.cviu.2007.09.014.
[20] D. Bi, R. Yu, M. Li, Y. Yang, K. Yang, and S. H. Ong, "Multiple image features-based retinal image registration using global and local geometric structure constraints," *IEEE Access*, vol. 7, pp. 133017–133029, 2019, doi: 10.1109/ACCESS.2019.2941256.
[21] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, no. 1, pp. 23–79, Jan. 2021, doi: 10.1007/s11263-020-01359-2.
[22] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010, doi: 10.1109/TPAMI.2010.46.
[23] J. Ma, J. Jiang, J. Chen, C. Liu, and C. Li, "Multimodal retinal image registration using edge map and feature guided Gaussian mixture model," in *Proc. Vis. Commun. Image Process. (VCIP)*, 2016, pp. 1–4, doi: 10.1109/VCIP.2016.7805491.
[24] K. Yang, A. Pan, Y. Yang, S. Zhang, S. Ong, and H. Tang, "Remote sensing image registration using multiple image features," *Remote Sens.*, vol. 9, no. 6, pp. 581–601, Jun. 2017, doi: 10.3390/rs9060581.
[25] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002, doi: 10.1109/34.993558.
[26] C. Liu, J. Ma, Y. Ma, and J. Huang, "Retinal image registration via feature-guided Gaussian mixture model," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 33, no. 7, p. 1267, 2016, doi: 10.1364/josaa.33.001267.
[27] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Trans. Inf. Technol. Biomed.*, vol. 3, no. 1, pp. 47–60, Mar. 1999, doi: 10.1109/4233.748975.
[28] E.-H. Zhang, Y. Zhang, and T.-X. Zhang, "Automatic retinal image registration based on blood vessels feature point," in *Proc. Int. Conf. Mach. Learn. Cybern.*, Beijing, China, 2002, pp. 2010–2015, doi: 10.1109/icmlc.2002.1175389.
[29] K. Wisaeng, "U-Net++DSM: Improved U-Net++ for brain tumor segmentation with deep supervision mechanism," *IEEE Access*, vol. 11, pp. 132268–132285, 2023, doi: 10.1109/ACCESS.2023.3331025.
[30] C. Harris and aM. J. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, 1998, pp. 147–152.
[31] J. Chen, R. T. Smith, J. Tian, and A. F. Laine, "A novel registration method for retinal images based on local features," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Vancouver, BC, Canada, Aug. 2008, pp. 2242–2245, doi: 10.1109/IEMBS.2008.4649642.
[32] J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine, "A partial intensity invariant feature descriptor for multimodal retinal image registration," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1707–1718, Jul. 2010, doi: 10.1109/TBME.2010.2042169.
[33] Z. Hossein-Nejad and M. Nasri, "RKEM: Redundant keypoint elimination method in image registration," *IET Image Process.*, vol. 11, no. 5, pp. 273–284, Jan. 2017, doi: 10.1049/iet-ipr.2016.0440.
[34] Z. Hossein-Nejad and M. Nasri, "Retinal image registration based on auto-adaptive SIFT and redundant keypoint elimination method," in *Proc. 27th Iranian Conf. Electr. Eng. (ICEE)*, Yazd, Iran, Apr. 2019, pp. 1294–1297, doi: 10.1109/IranianCEE.2019.8786443.
[35] L. Wei, L. Pan, L. Lin, and L. Yu, "The retinal image registration based on scale invariant feature," in *Proc. 3rd Int. Conf. Biomed. Eng. Informat.*, vol. 2, Yantai, China, Oct. 2010, pp. 639–643, doi: 10.1109/BMEI.2010.5640037.
[36] G. Wang, Z. Wang, Y. Chen, and W. Zhao, "Robust point matching method for multimodal retinal image registration," *Biomed. Signal Process. Control*, vol. 19, pp. 68–76, May 2015, doi: 10.1016/j.bspc.2015.03.004.
[37] Z. Ghassabi, J. Shanbehzadeh, and A. Mohammadzadeh, "A structure-based region detector for high-resolution retinal fundus image registration," *Biomed. Signal Process. Control*, vol. 23, pp. 52–61, Jan. 2016, doi: 10.1016/j.bspc.2015.08.005.
[38] J. A. Lee, J. Cheng, B. H. Lee, E. P. Ong, G. Xu, D. W. K. Wong, J. Liu, and A. T. H. Laude, "A low-dimensional step pattern analysis algorithm with application to multimodal retinal image registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1046–1053, doi: 10.1109/CVPR.2015.7298707.
[39] C. Liu, J. Ma, Y. Ma, and J. Huang, "Retinal image registration via feature guided Gaussian mixture model," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 33, no. 7, pp. 1267–1276, 2016.
[40] *DRIVE: Digital Retinal Images for Vessel Extraction*. Accessed: July 6, 2023. [Online]. Available: http://www.isi.uu.nl/Research/Databases/DRIVE
[41] *FIRE: Fundus Image Registration Dataset*. Accessed: Jul. 6, 2023. [Online]. Available: http://www.ics.forth.gr/cvrl/fire
[42] Y. Wang, J. Zhang, M. Cavichini, D. G. Bartsch, W. R. Freeman, T. Q. Nguyen, and C. An, "Study on correlation between subjective and objective metrics for multimodal retinal image registration," *IEEE Access*, vol. 8, pp. 190897–190905, 2020, doi: 10.1109/ACCESS.2020.3032348.
[43] Q. Jiang, Y. Liu, Y. Yan, J. Deng, J. Fang, Z. Li, and X. Jiang, "A contour angle orientation for power equipment infrared and visible image registration," *IEEE Trans. Power Del.*, vol. 36, no. 4, pp. 2559–2569, Aug. 2021, doi: 10.1109/TPWRD.2020.3011962.
[44] S. Saha, G. M. A. Rahaman, T. Islam, M. Akter, S. Frost, and Y. Kanagasingam, "Retinal image registration using log-polar transform and robust description of bifurcation points," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102424, doi: 10.1016/j.bspc.2021.102424.

KITTIPOL WISAENG received the Ph.D. degree in electrical and computer engineering in Thailand, in 2016. He is currently holds the position of an Associate Professor of computer science. He has conducted extensive research programs on deep learning and machine learning, addressing complex issues, such as optimization techniques, retinal segmentation, breast cancer detection, blood vessel extraction, exudates identification, optic disc localization, and brain tumor classification in challenging environments. With more than 26 technical articles published in this area and a H-index of seven. His primary research focuses on advancing artificial intelligence methods and integrating systems in various applications, particularly medical image segmentation. He has received numerous awards and recognitions for his outstanding contributions, including the senior researcher award, outstanding publication award, and recognition for the highest publication of research in the Scopus database, highest H-index, and highest citation per article. He has also received the Newton Fund Fellowship in Thailand, a program delivered by the British Council in collaboration with national government partners to support collaborative research projects and build the skills of researchers between Thailand and U.K. In addition, he has served as a guest reviewer for various esteemed journals.