Anatomical Point-of-Interest Detection in Head MRI Using Multipoint Feature Descriptor

SAI LI1, SHUCHAO CHEN1, HAOJIANG LI2, GUANGYING RUAN2, SHUAI REN1, TIANQIAO ZHANG1, LIZHI LIU2, AND HONGBO CHEN1

1School of Life and Environmental Science, Guilin University of Electronic Technology, Guilin 541004, China
2Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

Corresponding author: Hongbo Chen (hongbochen@163.com)

This work was supported in part by project grants from the National Natural Science Foundation of China under Grant 81760322 and Grant 81460273.

ABSTRACT Automatic detection of specific anatomical points of interest (POIs) in head magnetic resonance imaging (MRI) is a technical bottleneck in medical image registration and big data analysis for head diseases. A technique for automatically retrieving POIs in head MRI scans is explored in this study. A Haar-like feature of the image is introduced to generate the feature description vector of the POI, and the most appropriate standard vector is selected from multiple marked images. A fixed five-point joint feature description is proposed for solid POI detection, and an adaptive three-point joint feature description is proposed for cavity POI detection. A total of 516 head MRI volumes were used for anatomical point detection. A solid point detection experiment was conducted using the POIs of the right/left internal acoustic pore (RIA/LIA). The POIs of the right/left ascending segment of the internal carotid artery in the posterior cavernous sinus (RAS/LAS) were used in a cavity point detection experiment. The experimental results show that the accuracies of the solid point detection for LIA and RIA are 81.8% and 84.7%, respectively. Those of cavity point detection for LAS and RAS are 66.7% and 76.2%. The performance of the proposed method is better than those of BRIEF and SIFT algorithm. The proposed method can facilitate the marking of anatomical points for doctors, thus providing technical support for head image automatic registration and big data analysis for head diseases.

INDEX TERMS Haar-like feature, point of interest (POI), head MRI, medical image.

I. INTRODUCTION

Head magnetic resonance imaging (MRI) plays an increasingly important role in auxiliary diagnosis and prognosis prediction of nasopharyngeal carcinoma (NPC). Dynamic contrast-enhanced MRI and readout-segmented diffusion-weighted imaging (DWI) are effective in differentiating NPC from nasopharyngeal lymphoma (NPL) because combining the parameters delivers maximal diagnostic efficiency [1]. For differential diagnosis between recurrent NPC and post-treatment sequelae, images of nasopharyngeal lesions obtained via turbo spin-echo DWI are of superior quality and have higher diagnostic capability than those obtained via echo-planar DWI [2].

Radiomics and big data image analysis technology have become the primary means of NPC prognosis research [3], [4]. Radiomics nomograms have been developed that demonstrate excellent prognostic estimation for NPC patients using a noninvasive MRI method [5]. Deep learning positron-emission tomography or computed-tomography (PET/CT) based radiomics can serve as a reliable and powerful tool for prognosis prediction and may act as a potential indicator for individual induction chemotherapy in advanced NPC [6]. MRI-based radiomics can be used as an aid tool for the evaluation of local recurrence based on individual local recurrence risk assessment in NPC patients before initial treatment [7].

Image registration is the key technology for big data image analysis of head MRI scans [8], [9]. Image registration is the process of aligning two or more images of the same scene captured from different viewpoints or via different modalities [10], [11]. Image registration has been applied extensively in many fields, such as biomedical image analysis, remote sensing, computer vision, and pattern recognition.
matching [12]. Biomedical imaging involves different modalities, with different physics [13]. Each modality provides different information that are complementary to each other, and registration of multiple images is often required to combine the complementary information [13]. In medical pattern recognition, it is often necessary to align all images in an atlas. Spatial alignment between atlases is an important preprocessing step for training machine learning models, with tremendous consequences for learning performance [14]. For example, comparative studies of biomedical image data often require the aligning of images of different subjects to a normative atlas [15]. The image registration process is broadly classified into intensity-based and feature-based methods [16]. The intensity-based methods use similarities between pixel intensities to determine the alignment between two images. The primary similarity measures used are cross correlation and mutual information [17]. Feature-based methods extract salient features and use the correlation between those features to determine the optimal alignment [17].

Image registration based on points of interest (POIs) is a common method in medical image registration technology that requires the selection of some feature points in the image before registration. In image-guided otological surgery, the target registration error of measurement points at various depths from the surface of the head is measured to achieve a higher degree of registration accuracy [18]. To reduce the time required, both preoperatively and intraoperatively, skin fiducials may offer an efficient alternative method of navigation registration to bone fiducials [19]. An automatic method has been used to quantitatively assess the registration of retinal images based on the extraction of similar vessel structures and a modified Hausdorff distance [20]. In the preoperative planning for liver surgical resection treatment, point-based registration is used to deform the mesh of the atlas to the vein branch [21], [22]. To evaluate fusion accuracy, Riva et al. divided anatomical landmark pairs into test and control structures based on distinct involvement with brain shift [23]. To analyze the positional error of MRI on the accuracy of neuronavigational localization for posterior fossa lesions when the operation is performed with the patient in the prone position, the accuracy of fiducial point localization in supine and prone MRI scans was investigated by Dho et al. while taking surface anatomy into consideration [24]. An iterative closest point algorithm based on anatomical landmarks has been used for image-guided neurosurgery [25] and maxillofacial surgery [26], while a geometry projection similarity algorithm was used for brain image registration based on the anatomical features of human head [27]. Optical tracking was used to register a plastic skull to its preoperative CT images with paired-point registration [28].

In medical images, some stable anatomical POIs are often used as feature points for registration. These stable anatomical points are not affected by the displacement caused by breathing and other factors in the process of image capturing. Thus, selecting these points can greatly improve the accuracy of registration. The landmarks of anatomical points need to be obtained before registration and these landmarks commonly rely on special equipment, invasive implants, and manual marking [29], [30]. Manual marking is a time-consuming operation for doctors. Thus, it makes perfect sense to design an algorithm that allows a computer to be programmed to automatically detect the target anatomical points.

The core task of the computer automatically detecting the POIs in the image involves extracting the features of the pixel points in the image. A feature in an image refers to a specific meaningful structure in the image [31]. The features can be structures in the image such as points and edges. These features are included in a descriptor, which specifies elementary properties of the object, such as shape, color, and texture [32]. In the feature description, the local geometric information around a keypoint is extracted and stored in a high-dimensional vector [33]. A Haar-like feature is a robust point feature description method [34]–[36] and can be used to extract fiducial facial feature points [37], [38], the lane line edge [39] and voxel-wise classification in medical image [40].

Fig. 1 shows the stable anatomical POIs in the head. They are the right/left internal acoustic pore (RIA/LIA) and the right/left ascending segment of the internal carotid artery in the posterior cavernous sinus (RAS/LAS). RIA and LIA are solid points that locate the intersection of two anatomical structures and have prominent surrounding features. The RAS and LAS are cavity points that locate the center of blood vessels and have fewer neighboring features. In this study, an automatic anatomical POI detection algorithm for medical image registration is proposed. The main contributions of this article are as follows:

![Anatomical points in head MRI](image)

**FIGURE 1.** Anatomical points in head MRI. (a) The location marked by red dots is the location of the anatomical site; (b) Special region display of right/left internal acoustic pore (RIA/LIA) and right/left ascending segment of the internal carotid artery in the posterior cavernous sinus (RAS/LAS).

1. We propose an automatic detection method of anatomical points in medical images based on pixel retrieval. Good results have been achieved in the detection of four anatomical points in head MR images.
algorithm in the registration method based on gray scale being kind of method can overcome the problem of the optimization and then calculates the image registration parameters. This keypoint matching based on the descriptor of the keypoint, optimization [50].

keypoint acquisition, keypoint matching, transformation Registration methods based on keypoints primarily include studies focus on these algorithms for improvement [47]–[49]. Generally, the keypoint is one of the most important features [26]. For feature-based algorithms, structures in different images must have the same physical features and structures in different images as correspondence extracted from the images [44].

The gray-based methods estimate the correspondence on gray information and registration based on features [43]. The registration involves finding the transformation matrix $T$ to maximize the similarity between the reference image and the image to be registered, which can be expressed as

$$M(x, y) = \arg \max_{T(x, y)} [\varphi(S(T(x, y)), I(x, y))]$$  (1)

Common registration methods include registration based on gray information and registration based on features [43]. The gray-based methods estimate the correspondence between images to be registered by analyzing the original intensity information directly, whereas the feature-based methods implement the estimation by matching features extracted from the images [44].

Feature-based approaches to registration use the same features and structures in different images as correspondence for alignment. This is based on the premise that anatomical structures in different images must have the same physical features and structure [11]. For feature-based algorithms, the keypoint is one of the most important features [26]. Keypoint-based registration is a popular research direction in feature-based registration methods [45], [46]. Generally, the Harris algorithm, scale invariant feature transform (SIFT) algorithm, and the speeded-up robust features (SURF) algorithm are used to extract keypoints, and most contemporary studies focus on these algorithms for improvement [47]–[49].

Registration methods based on keypoints primarily include keypoint acquisition, keypoint matching, transformation parameter solving, and transformation model optimization [50].

A registration method based on keypoints first establishes keypoint matching based on the descriptor of the keypoint, and then calculates the image registration parameters. This kind of method can overcome the problem of the optimization algorithm in the registration method based on gray scale being difficult to converge. Keypoint matching accuracy depends primarily on two factors:

1. The matching ability of descriptors, i.e., the ability to represent common image features among images.
2. The ability of the post-processing technology to identify wrong keypoint matching, i.e., the ability of the post-processing technology to identify wrong keypoint matching caused by insufficient characterization ability of the descriptor when the descriptor matching performance is certain.

In an image registration method based on keypoint, the keypoints of general images are usually obtained using some keypoint detection operators, such as SIFT, SURF, Harris, and other algorithms. The keypoints detected by these algorithms are typically the corner points with drastic gray changes. However, corner points in medical images are not necessarily the best points for registration, and the images of different patients sometimes vary greatly, which may cause the problem of mismatching keypoints. Therefore, many registration methods rely on placing human contact points in the image to help coordinate transformation [51]. In the registration of medical images, specific anatomical points are usually selected as keypoints, and these specific points need to be manually marked by doctors to be obtained.

B. FEATURE DESCRIPTION

Local visual features are still one of the main tools for visual analysis tasks [52]. Distinctive and robust local feature description is crucial for image processing, such as image matching and image retrieval. Feature point description is used to match the same point in different images. A good feature descriptor will uniquely describe a feature point, allowing it to be correctly identified and matched in subsequent images [53]. In feature description, the local image descriptor of each feature point is calculated, and the descriptor has neighborhood information of feature points to identify the same feature points in various images. There are three commonly used feature point description operators in image processing technology: a feature point description operator based on Haar-like features in the SURF algorithm, a feature point description operator based on gradient features in the SIFT algorithm, and the binary robust independent elementary features (BRIEF) feature description operator.

BRIEF is a binary feature descriptor [54], [55]. In an $S \times S$ area around the feature points, $n$ pairs of pixel points are randomly selected. Then, the gray scale of each pair of pixel points are compared separately and assigned a 1 or 0. Therefore, each descriptor in BRIEF is really just an encoding of 1 or 0, and the number of pixel pairs $n$ is equal to the number of bits of binary value. Next, the Hamming distance is used to compare and match descriptors.

The Haar-like feature reflects the change in image gray scale [56]. The feature description algorithm based on Haar-like features takes a feature point as the center and constructs an $S \times S$ square window. The window is then divided into smaller $4 \times 4$ sub-regions, and the sum of Haar-like eigenvalues in the $x$ and $y$ directions in the sub-regions is calculated:
A feature point descriptor based on image gradient features is constructed using a mathematical model to describe the gradient features of several pixels in the neighborhood of the feature points. First, the region with a size of $S \times S$ pixels is taken as the center of the feature point and the gradient of each pixel is calculated. Then, the gradient values of eight directions ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$) of each square window are counted in each sub-region ($4 \times 4$) to form a 128-dimensional feature description vector. This descriptor is used in the SIFT algorithm.

The three different feature descriptors introduced above have their distinct advantages and disadvantages. In the SIFT, the feature point descriptor algorithm based on gradient has been proved to be time consuming by scholars. The algorithm based on Haar-like features is an accelerated version of the algorithm based on gradient, using integral images to speed up the computation and consequently consuming less time than the algorithm based on gradient [57], [58]. Compared with the algorithm based on Haar-like features and the algorithm based on gradient, the biggest advantage of the BRIEF algorithm is that its calculation speed is faster. Therefore, the BRIEF algorithm is often used in situations with high real-time requirements. The disadvantage of BRIEF is that it is very sensitive to noise.

### III. METHOD
The aim of this study is to extract the specific anatomical POIs in an unmarked head MRI scan. A three-step retrieving processing is illustrated in Fig. 2.

**Step 1:** Description of anatomical POIs. An anatomical POI, which is the point we want to retrieve in the head MRI scan, was marked by a doctor. The pixel feature is extracted by the point feature descriptor and a standard feature vector is defined for further retrieving processing.

**Step 2:** Description of test head MRI. An unmarked test MRI head image was prepared for retrieving anatomical POIs. To speed up the search processing, we limited the search scope to a $60 \times 60$ rectangular area near the target anatomical POI based on prior information. The features of all pixels in this rectangle are independently described with...
the feature description method, and each pixel has a specific feature vector.

Step 3: Retrieval of anatomical POIs in the head MRI scan. The feature distance between the description vectors of each pixel in the test head MRI and the standard feature vector are calculated using Euclidean distance, and a feature distance matrix is obtained. The smaller the feature distance, the closer the pixel is to the target anatomical POI. The result region is obtained via a threshold, which is set to the minimum value of the feature distance matrix plus 0.3. The center of the result region is the target anatomical POI.

A. FEATURE DESCRIPTION OF A POINT

The task of finding point correspondences between two images of the same scene or object is an element of many computer vision applications [59], and the central objective is to describe the feature of the pixels in the image and generate a feature description vector. To make the description vector possess good rotational invariability, it is necessary to determine a main direction for each pixel point. We calculate the Haar wavelet response in the x and y directions of each pixel in the circle region with the described pixel as the center of the circle, 12 pixels as the radius, and 8 pixels as the edge length of the Haar wavelet filter. We make a 60° sector in this circular region, count the sum of the Haar wavelet response in this sector, then rotate the sector and count the sum of the Haar wavelet response. The direction with the maximum sum of the Haar wavelet response is the main direction. To generate the description vector, a square region centered around the described point is constructed, and the size of this square region is selected as $S \times S$ mm. As shown in Fig. 3, the region is split up regularly into smaller $4 \times 4$ square sub-regions. For each sub-region, we compute Haar wavelet responses at $5 \times 5$ regularly spaced sample points, and the filter size for the Haar-like feature is $S/10$. The Haar wavelet responses in the x (horizontal) and y (vertical) directions in the sub-region are calculated as $dx$ and $dy$, respectively. The x and y directions are relative to the main direction. Finally, the feature values of $\sum dx$, $\sum |dx|$, $\sum dy$, $\sum |dy|$ in each sub-region are calculated and a feature vector with 64 dimensions is obtained for a specific point. In this study, S is given as 16 mm for POI detection.

B. MULTIPOINT JOINT DESCRIPTION SCHEME

The surrounding feature of an anatomical POI varies in different head MRI scans, which brings challenges for POI retrieval. If we describe only the surrounding feature of a single point, the retrieval accuracy for the POI may be reduced due to instability of the feature.

To enhance the robustness of the method, a multipoint joint feature description is introduced for detecting solid POIs in head MRI scans. As shown in Fig. 4A, four points (A1–A4) of the described point (P) are selected as auxiliary points for generating the feature description vectors, then a feature vector with $5 \times 64$ dimensions is generated to describe a specific solid point. These four auxiliary points are located above, below, and to the left and right of the described point in the image, and they are at equal intervals (m) from the described point.

As shown in Fig. 4B, the size of the tissue near the RIA is about 4–6 mm, and the tissue near the LIA is symmetrical to the RIA. The size of the cavity with the RAS is also about 4–6 mm, and the cavity with the LAS is symmetrical to the RAS. These results were obtained based on an investigation of the size of the tissue structure near the target points in 516 head MRI scans. To ensure that all auxiliary points for solid POI detection fall on the tissue when retrieving the target points of the RIA or LIA, we fixed the interval (m) between the auxiliary and described point at 2 mm.

The features around the RAS and LAS are not abundant because they are located in the vascular cavity. Furthermore, the shape of the vascular cavity in MRI scans is also different. The vascular cavity appears as a regular circle in some images (Fig. 4C), and as an irregular semicircle in some other images (Fig. 4D). Therefore, the multipoint joint feature description method does not adapt to detect the cavity point and an adaptive multipoint joint feature description scheme is proposed for cavity POI detection. As shown in Fig. 4E, two auxiliary points are used in the adaptive method and are located on the boundary of the cavity. For the RAS, the two auxiliary points are located at the left and lower boundaries of the cavity in the image while for the LAS, they are located at the right and lower boundaries. These selected auxiliary points have rich surrounding features, which is helpful for feature expression of cavity points. The interval between the auxiliary points and described point is not fixed, and is determined by the distance between the cavity boundary and the described point. Furthermore, a feature vector with $3 \times 64$ dimensions is generated to describe a specific cavity point.

C. SELECTING THE STANDARD VECTOR

The standard vector is selected from the feature description vector generated by the anatomical points manually marked by the doctor. There are some differences in the description vectors of marker points and their auxiliary points in different
marker images. Choosing the most suitable standard vector is the problem discussed in this section.

Due to the differences between images, there will be some distance between the description vectors of the same anatomical points in different images. However, because these vectors are from the same anatomical points in the image, the distance is very small. Therefore, when selecting the standard vector, we selected multiple marked images and calculated the distance between the feature description vectors of anatomical points in different images. In this study, the standard vector selected has the least distance from the other 49 vectors. This method is used to select the standard vectors of the anatomic points and auxiliary points.

**D. DETERMINING THE TARGET ANATOMICAL POI**

When determining the location of the anatomical point to be detected, the feature distance between he description vector of the most appropriate point and the standard vector should theoretically be minimum. However, the shape and brightness of the same anatomical points in different images are different, the feature distance between the description vector of the most appropriate point and the standard vector is small but not necessarily the minimum. Therefore, we select alternative points via a setting threshold, and the feature distance between the description vector of alternative points and the standard vector is less than this threshold. After observing the distance matrix between the standard vector and the description vector of the retrieved point, we set the threshold to the minimum value of the distance matrix plus 0.3. The selection of threshold has no effect on the result within a certain range. There are many points in the search range for which the feature distance is less than the set threshold. As shown in Fig. 5, these points may come in multiple clusters. The maximal cluster is extracted, and the target POI is considered to be in this area, located in the centroid of the result region.

**IV. RESULTS**

**A. MATERIALS**

The MRI data in the experimental test were obtained from the Cancer Center of Sun Yat-sen University. A total of 516 head
MRI volumes were used for anatomical POI detection. T2 transverse images were acquired using a fast recovery fast spin echo (FRFSE) sequence. The size of the image is 512 pixels $\times$ 512 pixels, and the field-of-view is 204.8 mm $\times$ 204.8 mm. In each image, four points in the head were labeled by a doctor possessing more than five years work experience and confirmed by a professor of medical imaging. The four anatomical POIs are the right/left internal acoustic pore (RIA/LIA) and the right/left ascending segment of the internal carotid artery in the posterior cavernous sinus (RAS/LAS).

### B. EVALUATION METRIC

The results of this experiment are the coordinates (x, y) of the anatomical points. We obtain the spatial distance by calculating the difference in value between the coordinates of the tests and the coordinates of the doctor’s markings. The performance of the algorithm is evaluated based on the spatial distance (mean value $\pm$ standard deviation). As shown in Fig. 4B, the anatomical point has a size. It can be considered as a correct detection if the detected result point locates within the organization area. Therefore, it can be considered as a correct result when the spatial distance between the result point and the labeled point is less than 4 mm. The metric accuracy is defined as $N/516 \times 100\%$, where $N$ is the number of POIs of the correct detections. The computational cost of the proposed POI detection is evaluated as well. The experiment was run on a personal computer with Intel Core i5-8250U CPU and 8 GB memory, using MATLAB 2016a on Microsoft Windows 10 home.

To verify the performance of the proposed algorithm, it was compared with the results of the same batch of data using the feature description method in BRIEF [55] and SIFT [60] algorithm. The BRIEF operator generates a 128-dimensional binary feature description vector. Thus, it calculates the Hamming distance of the vector instead of the Euclidean distance when calculating the distance between feature vectors. The feature description method in SIFT algorithm can obtain 128-dimensional feature description vector. We calculate Euclidean distance between vectors to measure the similarity between vectors. The other steps are exactly the same as those described above.

### C. EXPERIMENTAL RESULTS FOR SOLID ANATOMICAL POI

For the solid anatomical POI detection, the multipoint joint feature description with Haar-like features gets the minimum spatial distance and highest accuracy among all POI detection schemes. It can be found in Table 1 that the detection accuracy achieved for the two solid anatomical points (LIA and RIA) are 81.8% and 84.7%, respectively. The mean spatial distances between the result point and the standard point are 2.7 mm and 2.6 mm. The BRIEF algorithm has the shortest computation time. The feature description vector generated using Haar-like features is robust for anatomic POI detection. Although the BRIEF algorithm is very fast, its performance is poor due to its sensitivity to noise and lack of rotation invariance. Haar-like features are robust to noise and lighting changes since Haar-like features describe the ratio between dark and bright areas in a kernel, which can effectively express the details in the described region [56]. The feature description method in SIFT algorithm is superior to Haar-like method in the performance of rotation invariance, but inferior to Haar-like method in the brightness invariance [61]. The images we used are more different in brightness and do not rotate at a large angle, so the performance of Haar-like method is the best.

We can hardly use a one-point description vector to accurately describe the feature information. In addition,
the differences in anatomical points in different images and some accidental factors, the accuracy of the single-point description scheme is poor. The multipoint joint description scheme based on Haar-like features can effectively help for anatomic POI detection since Haar-like descriptor can extract local features of images and effectively shield the influence of brightness, rotation change and noise. However, the multi-point joint description scheme is ineffective for the BRIEF and SIFT algorithms because of sensitive to noise, rotation and brightness in the image respectively. Compared with single point description, multi-point joint description scheme increases the amount of information, but aggravates the sensitivity of the above factors. Therefore, if the multi-point joint description scheme is adopted, the description vector generated by the latter two methods will be too sensitive to the changes between images.

D. EXPERIMENTAL RESULTS FOR CAVITY ANATOMICAL POI

The cavity points are usually round holes or semi-enclosed round holes with different diameters and large spans. The size of the cavity point varies among different subjects. While the fixed multipoint joint feature description was not efficient for cavity point detection, adaptive multipoint joint feature description obtains satisfactory results. As shown in Table 2, the accuracy achieved for LAS and RAS detection was up to 66.7% and 76.2%, respectively. The mean spatial distances between the result point and the standard point are 3.9 mm and 3.4 mm. The BRIEF algorithm has the shortest computation time as well.

The adaptive method can effectively determine the edge position of the cavity point, and the edge point contains rich grayscale information, so the method of using the adaptive method to determine the position of the auxiliary point is
better than the method of fixedly selecting the auxiliary point.

E. SELECTION OF THE SIZE OF SQUARE WINDOW IN THE HAAR-LIKE FEATURE DESCRIPTOR

The parameter of length $S$ of the square window centered on the described point is important in the Haar-like feature descriptor. If the $S$ is too large, a lot of redundant features may be generated. While, if the $S$ is too small, the feature description for the POI may be incomplete, which may result in a decrease in POI detection accuracy. To verify the parameter settings for $S$, the results of POI detection were obtained for $S = 8$ mm, 16 mm, and 24 mm (Table 3). It can be seen that, for solid POI detection, the best result is obtained at $S = 16$ mm, and the detection accuracy for the cavity point is lower than that for the solid point because of the complicated surrounding features. The best result for RAS detection is obtained at $S = 16$ mm, while the best result for LAS detection is obtained at $S = 8$ mm.

| Anatomical POI | Accuracy (%) |
|----------------|--------------|
|                | $S = 8$ mm   | $S = 16$ mm | $S = 24$ mm |
| RIA            | 58.5         | 84.7        | 76.4        |
| LIA            | 55.6         | 81.8        | 78.5        |
| RAS            | 74.8         | 76.2        | 57.0        |
| LAS            | 69.2         | 66.7        | 48.4        |

TABLE 3. POI detection accuracy for different sizes of the square window in the Haar-like Feature Descriptor.

V. CONCLUSION AND DISCUSSIONS

Automatic extraction of specific anatomical POIs is a technical bottleneck in big data analysis of medical image. For example, anatomical points need to be manually marked for medical image registration based on a structure with obvious characteristics, which is a waste of time and energy. In this study, the Haar-like feature of the image is introduced for generating the feature description vector of the POI, and the most appropriate standard vector is selected from multiple marked images. A fixed five-point joint feature description is proposed for solid POI detection and an adaptive three-point joint feature description for cavity POI detection. Experimental results show that the method proposed in this study is effective for POI detection in head MRI images.

Although our method has demonstrated its robust and accurate detection capability upon the large-scale dataset, we are aware that the proposed method could be challenged in certain situations.

- Our method has only been verified at four anatomical points. Although it is proven that the proposed method is effective for detection of solid and cavity POIs, its detection performance for the POI with other surrounding characteristics has not been studied.
- The results in this study were obtained based on the head MRI with the same resolution, all of which are $512 \times 512$ in size and $204.8$ mm $\times 204.8$ mm in field-of-view. The proposed method is to detect the POI based on the point feature description which may be sensitive to image resolution. The performance of the proposed method upon the datasets with other image resolution is not investigated yet.

- Our images were T2 transverse images. The performance of the proposed method upon multimodal MR imaging (i.e., T1, PWI, DTI, DWI) is not investigated yet.

In general, our method can assist the physician to complete the task of anatomical points marking on the data set we used. There are two interesting research directions in our future work. First, we will study the method for POI detection using the new scheme, i.e., coordinate regression with convolutional neural networks [62], which was used to locate the facial landmark [63] and may improve the performance of the POI detection. Second, we will study the method for POI detection in the 3D head MRI.

REFERENCES

[1] C. Song, P. Cheng, J. Cheng, Y. Zhang, M. Sun, S. Xie, and X. Zhang, “Differential diagnosis of nasopharyngeal carcinoma and nasopharyngeal lymphoma based on DCE-MRI and RESOLVE-DWL,” Eur. Radiol., vol. 30, no. 1, pp. 110–118, Jan. 2020.

[2] W. Huang, J. Liu, B. Zhang, L. Liang, X. Luo, Y. Mei, and S. Zhang, “Potential value of non-echo-planar diffusion-weighted imaging of the nasopharynx: A primary study for differential diagnosis between recurrent nasopharyngeal carcinoma and post-chemoradiation fibrosis,” Acta Radiol., vol. 60, no. 10, pp. 1265–1272, Oct. 2019.

[3] C. Resteghini, A. Trama, E. Borgonovi, H. Hosni, G. Corrao, E. Orlandi, G. Calareso, L. De Cocco, C. Piazza, L. Mainardi, and L. Licitra, “Big data in head and neck cancer,” Current Treatment Options Oncol., vol. 19, no. 12, p. 62, 2018.

[4] K. Scheckenbach, “Radiomics: Big data instead of biopsies in the future?”

[5] X. Ming, R. W. Ooi, R. Zhai, F. Kong, C. Du, C. Hu, W. Hu, Z. Zhang, H. Ying, and J. Wang, “MRI-based signatures are a quantitative prognostic biomarker for nasopharyngeal carcinoma,” Sci. Rep., vol. 9, no. 1, p. 10412, Dec. 2019.

[6] H. Peng, D. Dong, M.-J. Fang, L. Li, L.-T. Tang, L. Chen, W.-F. Li, Y.-P. Mao, W. Fan, L.-Z. Liu, L. Tian, A.-H. Lin, Y. Sun, J. Tian, and J. Ma, “Prognostic value of deep learning PET/CT-based radiomics: Potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma,” Clin. Cancer Res., vol. 25, no. 14, pp. 4271–4279, Jul. 2019.

[7] L. Zhang, H. Zhou, D. Gu, J. Tian, B. Zhang, D. Dong, X. Mo, J. Liu, X. Luo, S. Pei, Y. Dong, W. Huang, Q. Chen, C. Liang, Z. Lian, and S. Zhang, “Radiomic nomogram: Pretreatment evaluation of local recurrence in nasopharyngeal carcinoma based on MR imaging,” J. Cancer, vol. 10, no. 18, pp. 4217–4225, 2019.

[8] M. Toews, C. Wachinger, R. S. J. Estepar, and W. M. Wells, “A feature-based approach to big data analysis of medical images,” Inf. Process. Med. Imag., vol. 9123, pp. 339–350, Jun. 2015.

[9] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, and J. Wang, “Applications of deep learning to MRI images: A survey,” Big Data Mining Anal., vol. 1, no. 1, pp. 1–18, Mar. 2018.

[10] H.-H. Chang, G.-L. Wu, and M.-H. Chiang, “Remote sensing image registration based on modified SIFT and feature slope grouping,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 9, pp. 1363–1367, Sep. 2019.

[11] M. Chen, A. Carass, A. Jog, J. Lee, S. Roy, and J. L. Prince, “Cross contrast multi-channel image registration using image synthesis for MR brain images,” Med. Image Anal., vol. 36, pp. 2–14, Feb. 2017.

[12] H. Yang, X. Li, L. Zhao, and S. Chen, “A novel coarse-to-fine scheme for remote sensing image registration based on SIFT and phase correlation,” Remote Sens., vol. 16, no. 5, p. 1833, 2019.

VOLUME 8, 2020

175247
[3] S. Nandish, G. Prabhu, and K. V. Rajagopal, “Multiresolution image registration for multimodal brain images and fusion for better neurosurgical planning,” *Biomed. J.*, vol. 40, no. 6, pp. 329–338, Dec. 2017.

[4] Z. Wang, X. Liu, and W. Chen, “Head and neck CT atlases alignment based on anatomical priors constraint,” *J. Med. Imag. Health Informat.*, vol. 9, no. 9, pp. 2004–2011, Dec. 2019.

[5] M. Toews and W. M. Wells, “Efficient and robust model-to-image alignment using 3D scale-invariant features,” *Med. Image Anal.*, vol. 17, no. 3, pp. 271–282, Apr. 2013.

[6] S. Paul and U. C. Pai, “Remote sensing optical image registration using modified uniform robust SIFT,” *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1300–1304, Sep. 2016.

[7] W. Ma, Z. Wen, Y. Wu, L. Iao, M. Gong, Y. Zheng, and L. Liu, “Remote sensing image registration with modified SIFT and enhanced feature matching,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 3–7, Jan. 2017.

[8] N. Matsumoto, M. Yamashita, B. Cho, N. Komune, and M. Hashizume, “Asymmetrical surface scanning registration for image-guided otorhinolaryngology: A phantom study,” *Auris Nasus Larynx*, vol. 47, no. 4, pp. 574–579, Aug. 2020, doi: 10.1016/j.anl.2020.01.007.

[9] R. Zamora, M. P. Punt, C. Christian-Skilleker, C. Yildirim, J. C. Shapton, and E. U. Conrad, “Are skin fiducials comparable to bone fiducials for registration when planning navigation-assisted musculoskeletal tumor resection in a cadaveric simulated tumor model?” *Clin. Orthopaedics Rel. Res.*, vol. 477, no. 12, pp. 2692–2701, Dec. 2019.

[10] Y. Shu, Y. Feng, G. Wu, J. Kang, and H. Li, “An automatic evaluation method for retinal image registration based on similar vessel structure matching,” *Med. Biol. Eng. Comput.*, vol. 58, no. 1, pp. 117–129, Jan. 2020.

[11] O. I. Alir and A. A. Abd. Rahmi, “Automatic atlas-based liver segmental anatomy identification for hepatic surgical planning,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 2, pp. 239–248, Feb. 2020.

[12] J. K. Rinne, A. Ehrlich, J. Ward, V. Väyrynen, M. Laine, I. H. Kellokumpu, Y. Shu, Y. Feng, G. Wu, J. Kang, and H. Li, “An automatic evaluation method for retinal image registration based on similar vessel structure matching,” *Med. Biol. Eng. Comput.*, vol. 58, no. 1, pp. 117–129, Jan. 2020.

[13] J. L. Y. Zhang, J. G. G. G. Van der Mark, A. C. M. Viergever, and M. A. Viergever, “Real-time and multimodal brain slice-to-volume registration using 3D scale-invariant features,” *IEEE Med. Imaging Conf.*, vol. 34, no. 8, pp. 1694–1704, Aug. 2015.

[14] X. Wang, J. Li, J. Li, and H. Yan, “Multilevel similarity model for high-resolution remote sensing image registration,” *Inf. Sci.*, vol. 505, pp. 294–305, Dec. 2019.

[15] C. Wachinger, M. Toews, G. Langs, W. Wells, and P. Golland, “Keypoint transfer for fast whole-body segmentation,” *IEEE Trans. Med. Imaging.*, vol. 39, no. 2, pp. 273–282, Feb. 2020.

[16] J. Z. Wang, X. Yao, T. Hu, and X. Xu, “An automatic spatial registration method for image-guided neurosurgery system,” *J. Craniofacial Surg.*, vol. 30, no. 4, pp. 1225–1235, Jul. 2018.

[17] J. Wang, Y. Shen, and S. Yang, “A practical marker-less image registration method for augmented reality oral and maxillofacial surgery,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 14, no. 5, pp. 763–773, May 2019.

[18] J. Li, Y. Peng, H. Chen, T. Potter, and Y. Zhang, “An image registration approach based on 3D geometric projection similarity of the human head,” *J. Med. Biol. Eng.*, vol. 39, no. 1, pp. 126–138, Feb. 2019.

[19] J. L. K. Wang, Y. Shen, and S. Yang, “Real-time and multimodal brain slice-to-volume registration using CNN,” *Expert Syst. Appl.*, vol. 133, pp. 86–96, Nov. 2019.

[20] Z. Chen, Z. Xu, W. Yi, X. Yang, W. Hsu, M. Ding, and O. Graniuch, “Real-time and multimodal brain slice-to-volume registration using CNN,” *Expert Syst. Appl.*, vol. 133, pp. 86–96, Nov. 2019.

[21] Z. Chen, Z. Xu, W. Yi, X. Yang, W. Hsu, M. Ding, and O. Graniuch, “Real-time and multimodal brain slice-to-volume registration using CNN,” *Expert Syst. Appl.*, vol. 133, pp. 86–96, Nov. 2019.

[22] G. P. Penney, D. C. Barratt, C. S. K. Chan, M. Slomczykowski, T. J. Carter, P. J. Edwards, and D. J. Hawkes, “Cadaver validation of intensity-based ultrasound to CT registration,” *Med. Image Anal.*, vol. 10, no. 3, pp. 385–395, Jun. 2006.

[23] D. Mukherjee, Q. M. Jonathan Wu, and G. Wang, “A comparative experimental study of image feature detectors and descriptors,” *Mach. Vis. Appl.*, vol. 26, no. 4, pp. 443–466, May 2015.

[24] C. Barajas-Garcia, S. Solorza-Calderón, and E. Gutiérrez-López, “Scale, translation and rotation invariant wavelet local feature descriptor,” *Appl. Math. Comput.*, vol. 363, Dec. 2019, Art. no. 124954.
S. Li et al.: Anatomical POI Detection in Head MRI Using Multipoint Feature Descriptor

[55] S. Urban, M. Weinmann, and S. Hinz, “mdbrief—a fast online-adaptable, distorted binary descriptor for real-time applications using calibrated wide-angle or fisheye cameras,” Comput. Vis. Image Understand., vol. 162, pp. 71–86, Sep. 2017.

[56] S. Nigam, R. Singh, and A. K. Misra, “A review of computational approaches for human behavior detection,” Arch. Comput. Methods Eng., vol. 26, no. 4, pp. 831–863, 2019.

[57] C. Tao, Y. Duan, and X. Hong, “Discrimination of fabric frictional sounds based on Haar features,” Textile Res. J., vol. 89, no. 11, pp. 2067–2074, Jun. 2019.

[58] M. M. Goyani and N. M. Patel, “Evaluation of various classifiers for expression recognition using multi level Haar features,” Int. J. Next-Gener. Comput., vol. 9, no. 2, pp. 131–150, 2018.

[59] 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.

[60] P. Chiara, P. Marta, B. Guido, and R. Marco, “Validation of deformable registration in adaptive radiation therapy with scale invariant feature transform,” presented at the 9th IEEE Int. Symp. Biomed. Imag., 2012.

[61] L. Juan and O. Gwun, “A comparison of SIFT, PCA-SIFT and SURF,” Int. J. Image Process., vol. 3, no. 4, pp. 143–152, 2009.

[62] A. Nibali, Z. He, S. Morgan, and L. Prendergast, “Numerical coordinate regression with convolutional neural networks,” 2018, arXiv:1801.07372. [Online]. Available: http://arxiv.org/abs/1801.07372

[63] H. Zhang, Q. Li, and Z. Sun, “Joint voxel and coordinate regression for accurate 3D facial landmark localization,” 2018, arXiv:1801.09242. [Online]. Available: http://arxiv.org/abs/1801.09242

SAI LI is currently pursuing the M.Sc. degree in biomedical engineering with the School of Life and Environmental Sciences, Guilin University of Electronic Technology. His research interests include computer vision and medical image analysis.

SHUCHAO CHEN is currently pursuing the B.Sc. degree in biomedical engineering with the School of Life and Environmental Sciences, Guilin University of Electronic Technology. His research interests include computer vision and medical image analysis.

HAOJIANG LI is currently a Radiologist with the State Key Laboratory of Oncology in South China, Department of Radiology, and the Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangdong, China. His research interest includes data mining in medical imaging and statistics.

GUANGYING RUAN received the B.M. degree from Sun Yat-sen University, in 2018. She is currently a Research Assistant Fellow with the Sun Yat-sen University Cancer Center. Her research interests include medical image analysis and clinical medicine.

SHUAI REN is currently pursuing the M.Sc. degree in biomedical engineering with the School of Life and Environmental Sciences, Guilin University of Electronic Technology. His research interests include medical image analysis and machine learning.

TIANQIAO ZHANG received the Ph.D. degree from the University of Chinese Academy of Sciences, China, in 2017. He is currently a Lecturer with the School of Life and Environmental Sciences, Guilin University of Electronic Technology. His research interests include medical image analysis, computer vision, and machine learning.

LIZHI LIU received the B.Sc. degree from North Sichuan Medical University, in 1996, and the M.Sc. and D.Sc. degrees from Sun Yat-sen University, in 2008 and 2016, respectively. He was a Visiting Associate Professor with the Department of Radiology and Imaging Sciences, Emory University. He is currently a Professor and a Radiologist with the Sun Yat-sen University Cancer Center. His research interests include oncologic imaging and medical image analysis.

HONGBO CHEN received the B.Sc. degree in physics from Xiangtan Normal University, in 1996, and the M.Sc. degree in biomedical engineering with the School of Life and Environmental Sciences, Guilin University of Electronic Technology. His research interests include medical image analysis and medical pattern recognition.