



An Unsupervised Multispectral Image Registration Network for Skin Diseases

Songhui Diao^{1,2}, Wenxue Zhou^{2,3}, Chenchen Qin², Jun Liao²,
Junzhou Huang⁴, Wenming Yang³, and Jianhua Yao²(✉)

¹ Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences,
Shenzhen, China

² Tencent AI Lab, Shenzhen, China
jianhuayao@tencent.com

³ Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

⁴ University of Texas at Arlington, Arlington, USA

Abstract. Multispectral imaging has a broad, promising and advantageous application prospect in the diagnosis of skin diseases. However, there are inherent deviations such as rigid or non-rigid deformation among multispectral images (MSI), which makes accurate and robust registration algorithms desirable to extract reliable multispectral features. Existing registration algorithms are susceptible to significant and nonlinear amplitude differences and geometric distortions among MSI, resulting in an unsatisfactory estimation of the registration field (RF). In this study, we propose an end-to-end multispectral image registration (MSIR) network with unsupervised learning for human skin disease diagnosis. First, we propose a basic adjacent-band pair registration (ABPR) model to obtain the corresponding RFs through simultaneously modeling a series of image pairs from adjacent bands. Second, we introduce a multispectral attention module (MAM) for extraction and adaptive weight allocation of the high-level pathological features of multiple MSI pairs. Third, we design a registration field refinement module (RFRM) to rectify and reconstruct a general RF solution. Fourth, we propose an unsupervised center-toward registration loss function, combining a similarity loss for features in the frequency domain and a smoothness loss for RF. In addition, we built a MSI dataset of multi-type skin diseases and conducted extensive experiments. The results show that our method not only outperforms state-of-the-art methods on MSI registration task, but also contributes to the subsequent task of benign and malignant disease classification.

Keywords: Image registration · Multispectral image · Unsupervised registration · Registration field refinement

S. Diao and W. Zhou—These authors contributed equally to this work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43999-5_68.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14229, pp. 720–729, 2023.
https://doi.org/10.1007/978-3-031-43999-5_68

1 Introduction

Skin disease is common in clinic, which is characterized by the complexity of pathological morphology and etiology, as well as the diversity of disease types and locations. Multispectral imaging (MSI) has the characteristics of non-tissue contact puncture, no radiation and no need for exogenous contrast agents. The imaging mechanism characterizes specific, correlated and complementary tissue features, which makes it having a broad, promising and advantageous application prospect in the diagnosis of skin diseases [1]. Correspondingly, multispectral imaging also has some shortcomings. On one hand, wavelength and focal length vary with frequency, resulting in non-rigid deformation such as scaling deviation among MSI. On the other hand, the motion of imaging device or patient may introduce further deviation among images. Consequently, MSI registration, that is, the identification and mapping of the same or similar structure or content at the pixel level, is a fundamental and critical process for subsequent tasks such as image fusion, pathological analysis, disease identification and diagnosis.

The difficulties for MSI registration are twofold. Firstly, conventional registration method is to register two images [2–5], while group image registration (GIR) is the joint registration of a group of related images. Current GIR research focuses on time-series MRI. MSI contains multiple images with significant and nonlinear amplitude differences and geometric distortions, not only making the pair-wise image registration not applicable, but also bringing great challenges to GIR due to the inability to take advantage of image intensity or structural similarity. Secondly, the type and location of diseases both affect the light reflection coefficient of skin tissue, making it challenging to find a general registration field (RF) for GIR.

In the field of image registration, many inspiring methods based on traditional or deep learning techniques have been developed and applied to computer vision tasks in medical imaging, remote sensing, etc. Whereas, the traditional methods [6, 7] are not suitable for MSI dataset with significant non-rigid deformation, gray jump, noise and other factors, which will lead to low efficiency and poor accuracy. The supervised deep learning methods [8, 9] have the limitation of relying on the groundtruth of RF which is difficult to obtain in medical images. For the unsupervised deep learning method, G. Balakrishnan et al. [10] proposed a VoxelMorph framework for deformable and pairwise brain image registration based on image intensity. Y. Ye et al. [11] presented a MU-Net framework, which stacks several DNN models on multiple scales to generate a coarse-to-fine registration pipeline. L. Meng et al. [12] proposed an DSIM network for MSI registration, which utilized pyramid structure similarity loss to optimize the network and regress the homography parameters. Although the existing algorithms can achieve relatively accurate registration for images with weak or repeated texture, they are susceptible to significant and nonlinear amplitude differences and geometric distortions among MSI, and generally have the disadvantages of poor accuracy, low robustness and low efficiency in realizing group image registration, leading to unsatisfactory RF results.

To address the aforementioned issues, we propose an end-to-end multispectral image registration (MSIR) network with unsupervised learning for multiple types of human skin diseases, which improves the capability of CNN architecture to

learn the cross-band transformation relationship among pathological features, so as to obtain an efficient and robust RF solution. First, we design a basic adjacent-band pair registration (ABPR) model, which simultaneously models a series of image pairs from adjacent bands based on CNN, and makes full use of the feature transformation relationship between images to obtain their corresponding RFs. Second, we introduce a multispectral attention module (MAM), which is used to achieve extraction and adaptive weight allocation for the high-level pathological features. Third, we design a registration field refinement module (RFRM) to obtain a general RF solution of MSI through rectifying and reconstructing the RFs learned from all adjacent-band MSI pairs. Fourth, we propose an unsupervised center-toward registration loss function, combining a similarity loss for features in the frequency domain and a smoothness loss for RF. We perform extensive experiments on a MSI dataset of multi-type skin diseases. The evaluation results demonstrate that our method not only outperforms prior arts on MSI registration task, but also contributes to the subsequent task of benign and malignant disease classification.

2 Methodology

As shown in Fig. 1, the proposed MSIR framework consists of four main components: (1) an ABPR model to extract pixel-wise representations of corresponding features from a pair of images synchronously. (2) a MAM to reassign the weight of the high-level features. (3) a RFRM to yield a general RF of MSI. (4) a center-toward registration loss function to optimize the effect of GIR.

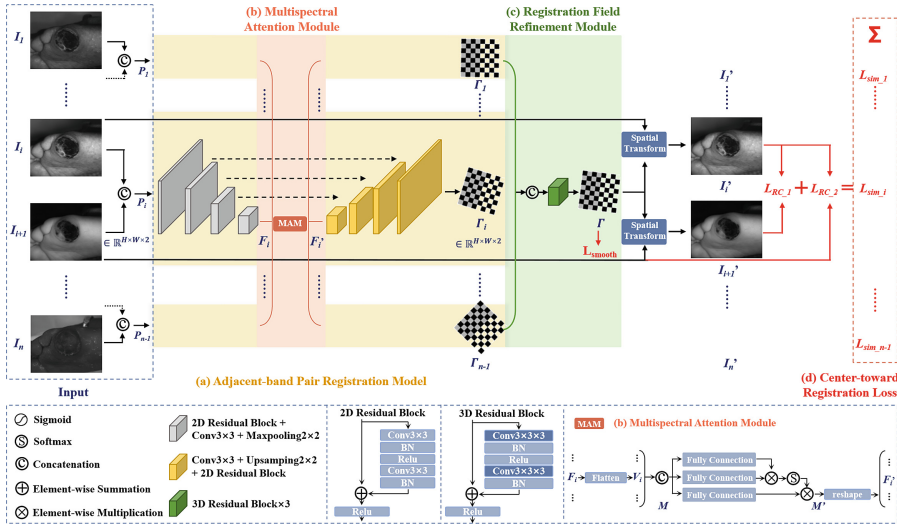


Fig. 1. The overall framework of our proposed MSIR network for multiple types of human skin diseases.

2.1 Adjacent-Band Pair Registration (ABPR) Model

The ABPR model is based on Unet [13] cascaded with 2D residual blocks [14]. In the feature encoder represented by the gray part, 4 convolution blocks with size of 3×3 and stride of 2 are used for down-sampling, and feature maps with channel number of 16, 32, 32 and 32 are obtained. Correspondingly, in the decoder represented by the yellow part, the convolution blocks contain interpolation operations for up-sampling.

The input to ABPR model is the concatenation of a pair of MSI images in adjacent bands $P_i \in R^{H \times W \times 2}$ (images I_i and I_{i+1} , $1 \leq i \leq n - 1$. n is the number of bands). The module constructs a function $\Gamma_i = F_\rho(P_i)$ to synchronously extracts the mutual transformation relationship. F represents the registration function fit by the designed Unet architecture. ρ refers to the weights and bias parameters of the kernels of the convolutional layers. $\Gamma_i \in R^{H \times W \times 2}$ is the RF between the given P_i , where 2 denotes two channels along the x-axis and y-axis.

Since P_i in multi-band imaging has unequal contribution to the solution of the final RF, we introduce a multispectral attention module (MAM) into ABPR model, which can not only facilitate the extraction of mutual complementary non-local features between MSI, but also guide the model to pay more attention to features from specific spectra with high impact.

Specifically, we carry out global average pooling operation on the high-level features F_i of the encoder for image pair P_i , and flatten F_i into a feature vector V_i . Then we obtain a feature map $M \in R^{(n-1) \times H' \times W' \times C'}$ by concatenating the accumulated $n - 1$ feature vectors in column. The formula for reallocating attention weights is defined as [15]:

$$M' = softmax(\frac{MW_Q MW_K^T}{\sqrt{d_k}}) MW_v \quad (1)$$

where W_Q , W_K and W_V are weights of the fully connection layers. d_k represents the dimension of V_i . Next, we reshape M' to obtain the updated high-level features F'_i and continue the subsequent decoding calculation in ABPR model.

2.2 Registration Field Refinement Module (RFRM)

In order to make full use of the potential complementarity of corresponding features between cross-band images, we design a RFRM to obtain a general RF of MSI through rectifying and reconstructing a series of RFs learned from multiple adjacent-band image pairs, which is more conducive to further improving the accuracy and generalization of the MSI registration network.

First, RFRM concatenates all $\Gamma_i \in R^{H \times W \times 2}$ learned from $n - 1$ adjacent-band image pairs. The obtained RF $\Gamma \in R^{(n-1) \times H \times W \times 2}$ is then refined through three 3D residual blocks, and the reliable $\Gamma \in R^{H \times W \times 2}$ is finally generated. T_Γ means a coordinate mapping that is parameterized by Γ and its spatial transformation. That is, for each pixel $p \in I_i$, there is a Γ such that $I_i(p)$ and $T_\Gamma(I_i)(p)$ are two corresponding points in adjacent bands.

2.3 Center-Toward Registration Loss Function

The ultimate goal of image registration is to obtain a RF from MSI with significant amplitude difference and geometric distortion, so that the perceived images corrected by the RF have the best similarity with each other. The pixel similarity in spatial domain among MSI is prone to large difference due to spectral intensity, while the feature similarity in the frequency domain is more stable. In order to optimize the adaptive center-toward registration of a group of images simultaneously, we propose an unsupervised loss function, including a similarity loss for features in the frequency domain and a smoothness loss for RF. The specific scheme is described in detail as follows:

(1) The first is a similarity loss function based on the residual complexity of features in the frequency domain, which is used to penalize differences in appearance and optimize the registration effect under different lighting conditions. The residual complexity loss function is defined as [16]:

$$L_{RC}(I, I') = \frac{1}{m} \sum_{j=1}^m \log[(DCT(I - I'))^2 / \alpha + 1] \quad (2)$$

where m is the number of pixels of the images I and I' . DCT denotes discrete cosine transform whose weight is regulated by a hyperparameter α with an empirical value of 0.05. It is worth noting that the similarity loss consists of two components. For the image with band i , we first use the fused Γ and its spatial transformation to obtain the warped image $I'_i = T_\Gamma(I_i)$. Then we evaluate its similarity to the reference image with the adjacent band I_{i+1} and the warped image $I'_{i+1} = T_\Gamma(I_{i+1})$, which not only ensures that the transformed images do not deviate from the original spatial distribution, but also realizes center-toward registration of a group of images synchronously and uniformly. Then the total similarity loss can be obtained through adding the residual complexity results from $n - 1$ image pairs. The formula is defined as:

$$L_{sim} = \frac{1}{n-1} \sum_{i=1}^{n-1} [L_{RC}(I'_i, I_{i+1}) + L_{RC}(I'_i, I'_{i+1})] \quad (3)$$

(2) The second is an auxiliary loss function that constrains the smoothness of RF and penalizes the local spatial variation. L_{smooth} is constructed based on Γ_i through differentiable operation on spatial gradient, and the formula is as follows [10]:

$$L_{smooth} = \frac{1}{n-1} \sum_{i=1}^{n-1} \|\nabla(\Gamma_i)\|^2 \quad (4)$$

where ∇ is the gradient operator calculated along the x-axis and y-axis. Then the loss function of our module L_{total} is the weighted sum of L_{sim} and L_{smooth} , which is defined as:

$$L_{total} = L_{sim} + \lambda L_{smooth} \quad (5)$$

where λ is a hyperparameter used to balance the similarity and smoothness of RF, and the empirical value 4 is taken, which is consistent with the Voxelmorph [10] method. All these components are unified in the MSIR network and trained end-to-end.

3 Experiments

3.1 Dataset and Implementation Details

We validated the proposed MSIR network on an in-house MSI dataset containing 85 cases with multi-type skin diseases collected from our partner hospital from November 2021 to March 2022, including 36 cases with benign diseases (keloid, fibrosarcoma, cyst, lipoma, hemangioma and nevus) and 49 cases with malignant diseases (eczematoid paget disease, squamous cell carcinoma, malignant melanoma and basal cell carcinoma). Specifically, each case consists of 22 scans with wavelengths ranging from 405 nm to 1650 nm and a uniform size of 640×512 . For each case, a clinician manually annotated four corresponding landmarks for registration on each scan, and their positions were further reviewed by an experienced medical expert. These landmarks are used to evaluate the registration accuracy. The detailed information for the training set and test set is shown in Table 1. We conducted a 4-fold cross-validation in training set to select models and hyperparameters.

Table 1. Characteristics for in-house dataset.

Characteristic	Entry	Training Set	Test Set
Patient Demographics	No. of cases	59	26
	No. of scans	1298	572
Prevalence No. of scans (percentage) with	Benign	528(28.24%)	264(14.12%)
	Malignant	770(41.18%)	308(16.47%)

3.2 Quantitative and Qualitative Evaluation

In this paper, we adopt three evaluation indexes to assess the registration performance of different methods, including normalized mutual information (NMI), registration feature error (RFE) and target registration point error (TRE) [17].

Table 2 shows the quantitative comparisons of the registration performance among our MSIR network and three state-of-the-art competitive methods, including Voxelmorph [10], DSIM [12] framework and MU-Net [11]. The mean value of the initial TRE is 8.77 pixels. It can be observed that our method achieves the superior performance. Specifically, it improves the best NMI to 0.547, RFE to 1.166 and TRE to 4.984 pixels, which significantly outperforms the other competitors.

Table 2. Quantitative comparisons of different methods.

Method	Input	NMI	RFE	TRE
Voxelmorph [10]	Multidimensional Image	0.498(± 0.057)	1.909(± 0.246)	5.734(± 3.265)
DSIM [12]	Multispectral Image	0.479(± 0.064)	2.122(± 0.393)	5.763(± 3.560)
MU-Net [11]	Multispectral Image	0.513(± 0.027)	1.851(± 0.891)	5.692(± 3.206)
MSIR (Ours)	Multispectral Image	0.547(± 0.029)	1.166(± 0.642)	4.984(± 3.471)

Table 3. Quantitative results of ablation study on components and augmented dataset.

NCC Loss	Our Loss	MAM	Augmented Dataset	NMI	RFE	TRE
✓	×	×	×	0.523(± 0.051)	1.331(± 0.819)	5.227(± 4.441)
×	✓	×	×	0.539(± 0.046)	1.168(± 0.674)	5.085(± 3.905)
✓	×	✓	×	0.526(± 0.078)	1.245(± 0.795)	4.935(± 6.422)
×	✓	✓	×	0.547(± 0.029)	1.166(± 0.642)	4.984(± 3.471)
×	✓	✓	✓	0.576(± 0.024)	1.109(± 0.547)	4.736(± 3.353)

We further conduct ablation study to verify the contributions of individual components, as shown in Table 3. NCC refers to the normalization cross correlation, which describes the relevance and similarity of targets. Compared with NCC loss, the proposed loss function significantly improves the registration performance (0.142 reduction in TRE, 0.163 reduction in RFE and 0.016 increase in NMI). The introduction of MAM further improves the registration performance. On this basis, we introduce an augmented dataset to expand the training set to 50 times of its original size by performing affine transformation operations on MSI, such as translation, rotation, scaling, clipping, oblique cutting, etc. The adoption of the augmented dataset makes the indexes continuously optimized. Quantitative results validate the effectiveness of MSIR network and augmented dataset in improving registration performance.

Next, we conduct ablation experiments to explore the registration performance based on images with different subsets of bands, as illustrated in Table 4. VIS and NIR represent the visible bands (405 nm–780 nm) and the near infrared bands (780 nm–1650 nm), respectively. ALL represents the whole bands (405 nm–1650 nm). The results demonstrate that our method can achieve remarkable registration improvement for images of different bands.

To verify the generalization of the proposed method, we test the model using a synthetic dataset (with [5] degrees of rotation, [0.02] of scaling, [6] and [8] pixels of translation along the x-axis and y-axis). The visualization results are shown in Fig. 2. It can be seen that MSIR method can achieve accurate registration not only for the MSI of the raw dataset, but also for that of the synthetic dataset with more complex transformation that are challenging for registration.

In addition, in order to verify the effect of this registration method to subsequent tasks, we further conduct a classification task for benign and malignant diseases based on the established MSI dataset. Table 5 shows the quantitative

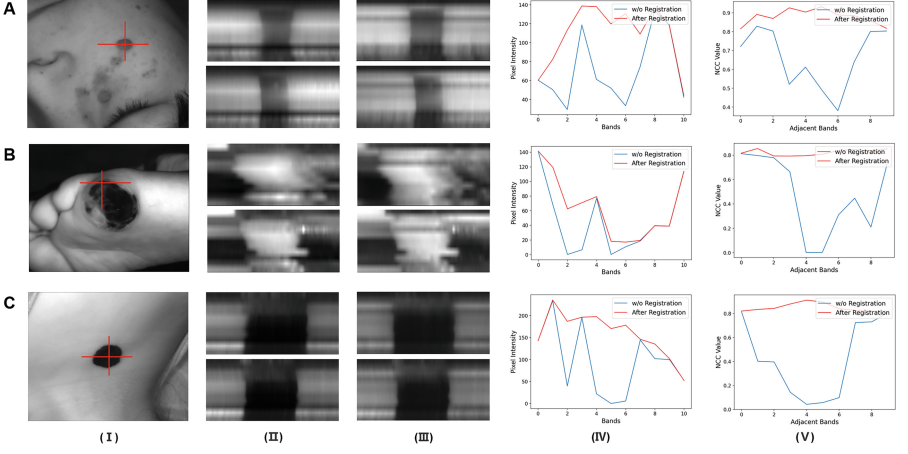


Fig. 2. Visual comparisons of test results on raw dataset and synthetic dataset. A, B and C represent three patient cases respectively. Column (I) shows the example of the original image at $650nm$, in which the boxes marked with red crosses indicate the position of zooming for column (II) to column (III). Column (II) and column (III) are the qualitative results of the raw dataset processed without and with our method, respectively. The first row represents the sectioning of the red cross on the x-axis, and the second row represents that on the y-axis. Column (IV) and column (V) are the quantitative results of two indicators in the red cross region of the synthetic dataset, namely, the pixel intensity among bands and the normalized correlation coefficient (NCC) between adjacent bands. The blue and red lines represent the results without and with registration, respectively.

comparisons of different classifiers using single-band and all-bands images. Four classifiers are compared, namely KNN [18], SVM [19], Resnet18 [20] and Inception V3 [21]. Columns 2 and 3 represent the worst and best classification accuracy based on single-band images, which come from the wavelengths of $450nm$ and $525nm$, respectively. The fourth column is the classification results based on the original MSI dataset without registration, and the last column shows that based on the images processed by MSIR network. It can be seen that the MSI contains more abundant information than the single-band image, which is more conducive to the subsequent analysis. More importantly, due to the contribution of MSIR network for image registration, the classification accuracy on

Table 4. Quantitative mean(\pm std) results of ablation experiments on bands.

Bands	w/o Registration			MSIR (Ours)		
	NMI	RFE	TRE	NMI	RFE	TRE
VIS	0.295(\pm 0.103)	1.505(\pm 1.197)	8.625(\pm 7.055)	0.517(\pm 0.089)	1.287(\pm 0.788)	4.632(\pm 3.947)
NIR	0.227(\pm 0.101)	1.324(\pm 1.241)	12.448(\pm 15.093)	0.576(\pm 0.048)	1.013(\pm 0.718)	5.595(\pm 5.406)
ALL	0.271(\pm 0.097)	1.323(\pm 0.904)	10.070(\pm 9.544)	0.547(\pm 0.029)	1.166(\pm 0.642)	4.984(\pm 3.471)

Table 5. Quantitative comparisons of different classifiers on band and registration operation.

Classifier	Accuracy			
	Single Band		All Bands	
	Worst (450 nm)	Best (525 nm)	w/o Registration	MSIR (Ours)
KNN [18]	0.524(± 0.017)	0.567(± 0.079)	0.607(± 0.063)	0.644(± 0.058)
SVM [19]	0.539(± 0.064)	0.583(± 0.072)	0.615(± 0.084)	0.672(± 0.049)
Resnet18 [20]	0.540(± 0.072)	0.602(± 0.051)	0.647(± 0.051)	0.754(± 0.037)
Inception V3 [21]	0.551(± 0.061)	0.606(± 0.046)	0.659(± 0.022)	0.762(± 0.018)

MSI dataset has been significantly improved, which verifies the necessity and effectiveness of MSIR network.

4 Conclusion

In this study, an efficient and robust framework for multispectral image registration is proposed and validated on a self-established dataset of multiple types of skin diseases, which holds great potentials for the further analysis, such as the classification of benign and malignant diseases. We intend to release the MSI dataset in future. The quantitative results of experiments demonstrate the superiority of our method over the current state-of-the-art methods.

References

1. Spreinat, A., Selvaggio, G., Erpenbeck, L., Kruss, S.: Multispectral near infrared absorption imaging for histology of skin cancer. *J. Biophotonics* **13**(1), e201960080 (2020)
2. Mok, T.C.W., Chung, A.C.S.: Large deformation diffeomorphic image registration with Laplacian pyramid networks. In: Martel, A.L., et al. (eds.) *MICCAI 2020*. LNCS, vol. 12263, pp. 211–221. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59716-0_21
3. Shu, Y., Wang, H., Xiao, B., Bi, X., Li, W.: Medical image registration based on uncoupled learning and accumulative enhancement. In: de Bruijne, M., et al. (eds.) *MICCAI 2021*. LNCS, vol. 12904, pp. 3–13. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87202-1_1
4. Meng, M., Bi, L., Feng, D., Kim, J.: Non-iterative coarse-to-fine registration based on single-pass deep cumulative learning. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) *MICCAI 2022*. LNCS, vol. 13436, pp. 88–97. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16446-0_9
5. Chen, J., Frey, E.C., He, Y., Segars, W.P., Li, Y., Du, Y.: Transmorph: transformer for unsupervised medical image registration. *Med. Image Anal.* **82**, 102615 (2022)
6. Ma, W., et al.: Remote sensing image registration with modified sift and enhanced feature matching. *IEEE Geosci. Remote Sens. Lett.* **14**(1), 3–7 (2016)

7. Cao, X., et al.: Deformable image registration based on similarity-steered CNN regression. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10433, pp. 300–308. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66182-7_35
8. Blendowski, M., Hansen, L., Heinrich, M.P.: Weakly-supervised learning of multi-modal features for regularised iterative descent in 3D image registration. *Med. Image Anal.* **67**, 101822 (2021)
9. Guo, H., Kruger, M., Xu, S., Wood, B.J., Yan, P.: Deep adaptive registration of multi-modal prostate images. *Comput. Med. Imaging Graph.* **84**, 101769 (2020)
10. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Voxelmorph: a learning framework for deformable medical image registration. *IEEE Trans. Med. Imaging* **38**(8), 1788–1800 (2019)
11. Ye, Y., Tang, T., Zhu, B., Yang, C., Li, B., Hao, S.: A multiscale framework with unsupervised learning for remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **60**, 1–15 (2022)
12. Meng, L., et al.: Investigation and evaluation of algorithms for unmanned aerial vehicle multispectral image registration. *Int. J. Appl. Earth Obs. Geoinf.* **102**, 102403 (2021)
13. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
15. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
16. Myronenko, A., Song, X.: Intensity-based image registration by minimizing residual complexity. *IEEE Trans. Med. Imaging* **29**(11), 1882–1891 (2010)
17. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 694–711. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_43
18. Uddin, S., Haque, I., Lu, H., Moni, M.A., Gide, E.: Comparative performance analysis of k-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Sci. Rep.* **12**(1), 1–11 (2022)
19. Therese, M.J., Devi, A., Kavya, G.: Melanoma detection on skin lesion images using k-means algorithm and SVM classifier. In: Handbook of Deep Learning in Biomedical Engineering and Health Informatics, pp. 227–251. Apple Academic Press (2021)
20. Vasu, K., et al.: Effective classification of colon cancer using Resnet-18 in comparison with squeezeNet. *J. Pharm. Negat. Results* 1413–1421 (2022)
21. Khamparia, A., Singh, P.K., Rani, P., Samanta, D., Khanna, A., Bhushan, B.: An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. *Trans. Emerg. Telecommun. Technol.* **32**(7), e3963 (2021)