Affine Registration of Plantar Foot Thermal Images with Deep Learning: Application to Diabetic Foot Diagnosis

Asma Aferhane, Doha Bouallal, Hassan Douzi, and Rachid Harba

Abstract Early prevention of diabetic foot ulceration is possible by using the plantar foot temperature that can be measured with a thermal camera. In this work, we performed the plantar foot registration using three Deep Learning methods. These methods include two parts: an affine registration module for estimating transformation parameters and a spatial transformer for getting the registered image. All three models performances were evaluated using the Dice similarity coefficient (DSC), Mean Square Error (MSE), and peak signal-to-noise ratio (PSNR). Our aim was to find an accurate, fully convolutional neural network suitable for our database of thermal images of diabetic feet. Results showed that Affine ConvNet and DLIR (affine part) models produce the best plantar foot affine registration results with a Dice score of 95%.

Keywords Plantar foot thermal images · Thermography · Medical images affine registration · Convolutional neural network

1 Introduction

People with diabetes are exposed to many complications that affect the eyes, the nerves, the kidney, and mainly the feet. Having not enough blood flowing to the legs and feet can make it hard for an infection to heal which may lead to untreatable ulceration and in some cases may lead to lower limb amputations of the foot. In diabetic foot, the occurrence of an ulcer is often associated with a high temperature. In this context, various studies showed a strong correlation between plantar foot temperature elevation and foot complications in diabetes. Moreover, detection of abnormal plantar foot temperature in diabetic patients [19] can be a sign of diabetic

A. Aferhane · D. Bouallal · H. Douzi
IRF-SIC Laboratory, Ibn Zohr University, Agadir, Morocco

A. Aferhane (✉) · D. Bouallal · R. Harba
PRISME Laboratory, Orléans University, Orléans, France
e-mail: asma.aferhane@edu.uiz.ac.ma

N. Gherabi et al. (eds.), *Advances in Intelligent System and Smart Technologies*,
Lecture Notes in Networks and Systems 826,
https://doi.org/10.1007/978-3-031-47672-3_37

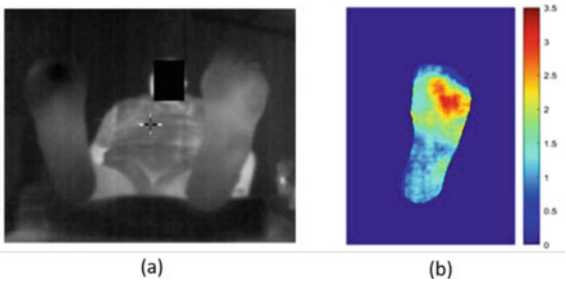

Fig. 1 **a** Originale thermal image **b** Thermal difference Map for a patient

foot (DF) ulcer in an early stage. In [1], the point-to-point temperature difference between the left and the right foot is greater than 2.2C and this is considered abnormal.

The increase in temperature may be present for a week before a foot ulcer occurs, if dedected may reduce the possibility of ulceration by 70%. According to [2], instead of using a complex and non-contact system acquisition protocol, we can use a user-friendly and mobile acquisition protocol for thermal images of the plantar foot Fig. 1 a smartphone equipped with a dedicated thermal camera (FLIR ONE Pro) [3].

After the first step of the full automatic processing of the data that is segmentation [4, 5]. The further step is the registration of the two feet (flipped left foot and right foot). Image registration refers to the process of aligning images, so that comparable characteristics can be easily related to one another.

During this process, pixels from one image are mapped to corresponding pixels in the other image. For the application of thermal image foot registration, we used in the first time the classical rigid registration method the iterative closest point (ICP) [6, 7, 15] to register the two images (flipped left and right foot image). For this method the image is translated and rotated to fit another image. With the use of acquisition protocols [5] and non-ideal camera angle, we need other transformation parameters [8] two fit two images. For that, we choose the affine registration methods which allows shearing and scaling in addition to rotation and translation to acquire better registration of two images. The transformation matrix of affine registration can be formulated by a transformation parameters by: translation parameter (t_x, t_y), rotation parameter θ, scaling (s_{cx}, s_{cy}) and shearing (sh_{cx}, sh_{cy}). The transformation matrix T expressed in 1.

$$T = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Where a_1, a_2, a_4 and a_5 represent relations with parameters of rotation, shearing, and scaling and a_3 and a_6 represent translation parameters according to x and y axes, respectively.

In recent years, convolutional neural networks have shown state-of-the-art performance on the task of medical image registration. Serval approaches have been proposed for affine image registration with the convolutional neural network (CNN).

A Self-Supervised Affine Registration (AIRNet) framework for affine registration was introduced by Chee et al. [9], that the transformation matrix (Matrix. 1) in AIRNet is regressed by two separate CNNs: the first take fixed image like input and the input of other cnn is the moving image. In addition, in the recent work by Tang et al. [10], an unsupervised end-to-end Affine and Deformable Medical Image Registration (ADMIR) was proposed. There are two sub-networks in this model : Affine ConvNet for affine registration and a U-net-like architecture for deformable registration [11, 19]. The affine transformation framework in the ADMIR called Affine ConvNet uses one CNN take the concatenation of two images (fixed and moving image), in order to regress the affine transformation matrix. Similar to the Affine ConvNet, the pooling layers used instead of stridden convolutional layers in the affine part in an unsupervised Deep Learning Image Registration DLIR [11]. And then the moving image is warped through a spatial transformer adapted VoxelMorph [12], to yield the final registration result.

In this communication, we performed the plantar foot registration using three deep learning methods. In order to find an accurate fully convolutional neural network suitable for our database.

The remaining part of this paper is organiwed as follow. Section 2 describes the dataset of thermal image of plantar foot. Section 3 details the three methods of affine registration using deep learning, including the proposed process followed for this task Sect. 3.1 and the Sect. 4 presents the experimental results on our thermal images dataset.

2 Dataset and Materials

The first step was to build our database with the thermal images of plantar feet from healthy (86 images) and diabetic patients (292 images) by using a thermal camera the FliROne Pro and a smartphone Samsung Galaxy S8. This camera has a thermal image resolution of 160×120 pixels and a spectral range of $8 - 14m$, that can detect temperature differences of 0.1C, which is sufficient to detect possible hyperthermia variations.

The two protocol followed for the image acquisition: the first one allows to obtain a thermal image of plantar feet after letting the patient rest for 15 min. The second one called Cold Stress test protocol that consists in taking two acquisitions of the same patient in two different moments as detailed in [5, 17]. After the acquisition, the original images were segmented manually to get the groud truth the image. Then the thermal images, after being segmented, are divided to have two images, one for each foot as shown in Fig. 2.

The right foot image is considered as the fixed foot and the left one is vertically flipped in order to have the same orientation (moving image) (Fig. 2).

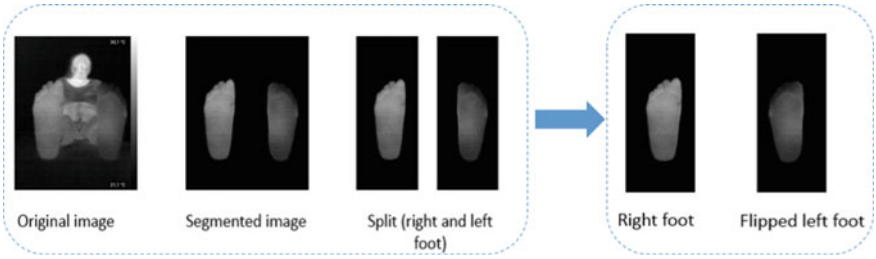

Fig. 2 Segmentation and split: Right foot (fixed image) and flipped left foot (moving image)

3 Image Registration by Deep Learning

3.1 Process

Plantar foot registration of thermal images represents the last step in the process of automated diagnosis of diabetic foot, the aim of it is to calculate the absolute point-to-point temperature difference between the right and the left foot for each pixel. Figure 3 presents the process adopted to perform more effective affine registration on plantar foot of thermal images dataset using Deep learning.

After segmenting the thermal image(1-channel) that is taken with FLiROne camera pro (Sect. 2), it is divided into two images: right and left foot. We Denote the fixed image (right foot) as F or the moving image(Flipped left foot) as M.

Affine registration image applied throughout this contribution aims to establish the six affine transformation parameters between F and M using neural networks. Each of these transformation parameters can be formulated by a matrix T. (Matrix. 1).

The input of these neural networks are 2-channel (fixed and moving image) to produce the transformation matrix (Matrix. 1) by using the back-propagated to learn

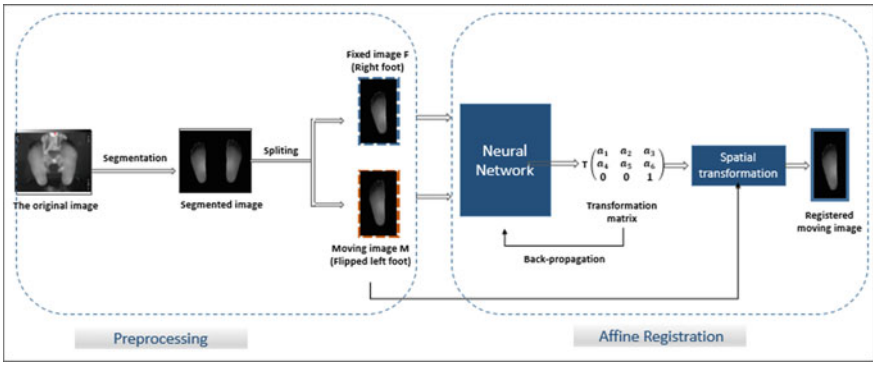

Fig. 3 Affine registration process

the optimal transformation parameters. Following that, The spatial transformation [12] is applied when warping the M image by the matrix of the transformation. (Matrix. 1).

3.2 *Training Convolutional Neural Networks*

With the development of various deep learning techniques, convolution neural networks based on affine registration have greatly contributed to the increase of affine registration performance of medical thermal images. This affine registration consists of aligning each pixel of two images (left foot F and flipped right foot M) into the same correspondence pixel. Hence, CNNs are a combination of convolution layers and pooling layers that extract different characteristics from the images and the fully connected layers to wrap up the CNN architecture. The goal of this registration is to determine optimal transformation parameters in a way that the registered moving image becomes similar to the fixed image.

Thus, all affine registration networks proposed for medical image affine registration are built in two ways: the first network in which the encoder correspondents to one pathway such as Affine ConvNet and DLIR(affine part) to predict the transformation parameters. Another example of networks is the AIRNet which takes two pathways to predict the affine transformation parameters.

In this section, we detail the architecture of these three networks. The first is the affine image registration network (AIRNet) which has been proposed by Chee et al. [9]. This architecture uses two separate subpathways (subnetworks) that share the same parameters, the first pathway takes the fixed image F as input and the other takes the moving image M, each of these subnetworks is adapted from 2D DenseNet [13] as in Fig. 4 to extract features from the input images.

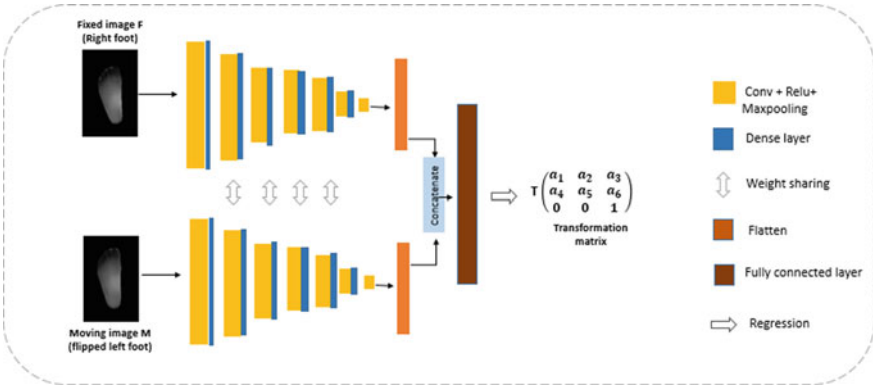

Fig. 4 Architecture of the AIRNet for affine image registration

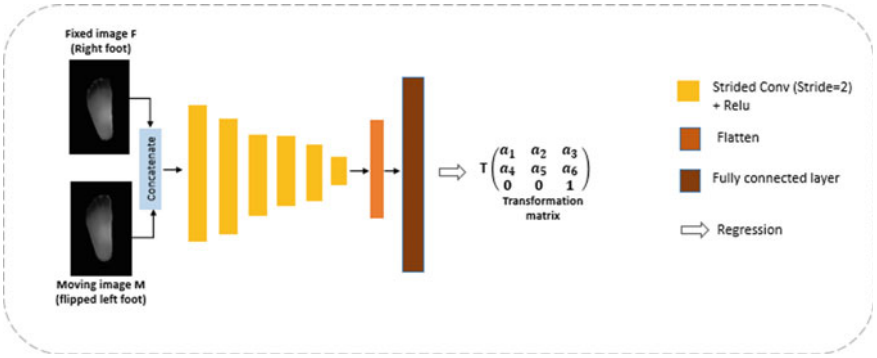

Fig. 5 Architecture of the Affine ConvNet for affine image registration

It comprises 2D convolutional blocks with a kernel size of 1×1 and max-pooling layers with a stride of 2 and this reduces the spatial dimension of the fixed and moving images followed by dense blocks. Then they concatenated these features and used them as an input to several fully connected layers. For the fully connection part, each layer consists of linear layers and relu activation functions. Affine ConvNet was developed by Tang et al. [10] for medical image registration, and the architecture contains one subnetwork instead of two subnetworks to obtain the matrix of the transformation. In Fig. 5, we concatenated moving and fixed images together and fed them into the network.

This architecture is made up of a series of 6 stridden convolution blocks with a kernel size of 3×3 and a stride of 2, followed by a relu activation function. For the full connection part, each layer is followed by a Dropout layer to estimate the six parameters of transformation.

The last architecture affine part network in DLIR which this model has the same architecture as the architecture affine ConvNet. Instead of using stridden convolutional layers to reduce the spatial dimension of the fixed and moving images, the pooling layer is used.

Finally, the spatial transformation is applied when warping the moving image(flipped left foot) by the affine transformation matrix (Matrix. 1).

4 Experiments and Results

To choose the best approach for registering the moving and fixed images in our dataset among of the three deep learning registration methods (Sect. 3). In this section, we present the quantitative and qualitative results of these methods and then we evaluate the robustness and the performance of registration on plantar foot thermal images with a widely used metric of image registration.

4.1 Dataset and Training

The generated dataset contained 392 pairs of thermal images of plantar feet, 292 of these pairs were obtained by using the cold stress test acquisition protocol with the FlirOne Pro camera [17]. These thermal images are in PNG format and their original size is 640×280 pixels. Furthermore, we resized these thermal images into 640×640 and then we normalized them into the range of [0, 1]. Our dataset was divided into 315 pairs of plantar foot images (left and flipped right foot) for training and 79 pairs for testing. We considered the right foot as a fixed image F and the other foot as a moving image M (Sect. 2).

We trained the networks using Adam with a learning rate of 10^{-4} as the optimizer in the network. In the Affine ConvNet and DLIR(Affine part) methods, every thermal image of flipped left foot in the training set is sequentially selected as the moving image and is concatenated with the fixed image (right foot) as a whole that is fed into the CNN to predict the affine parameter of transformation. In the AIRNet network, instead of concatenating images, before jumping into the convolutional networks, they used separate networks to achieve affine registration.

4.2 Data Augmentation

The number of thermal images in our database is not sufficient to feed a deep learning neural network. For that, we used image augmentation techniques to increase samples in our dataset, so we can improve the performance of the affine registration models and avoid overfitting.

Our image augmentation parameters were rotation, translation, and zooming [14]. After data augmentation, the database size has been increased by nine to contain 2800 pairs of thermal images of the plantar foot instead of 315 pairs of thermal images in the training set.

4.3 Evalution Metric

The networks were evaluated using three widely used metrics: Dice similarity coefficient (DSC) which was introduced for the image registration dissimilarity assessment. Mean squared error (MSE) is used to measure the difference per pixel between two samples (fixed and registered image) and Peak Signal to Noise Ratio (PSNR) is determined as a quality measurement between the fixed image and registered image, a higher PSNR value indicates a good similarity between two images [16].

– The Dice similarity coefficient (DSC) is defined by (Eq. 2) :

$$DSC = 2\frac{F \cap W}{F \cup W} \tag{2}$$

where F is the fixed image and W is the registered image.

– The mean squared error is defined by (Eq. 3):

$$MSE = \frac{1}{N}\frac{1}{N}\sum\sum(F(i,j) - W(i,j))^2 \tag{3}$$

That i and j denotes the pixel location in the fixed image F and registred image W based on x-y coordinates and N represents the total number of pixels in the image.

– Peak Signal to Noise Ratio (PSNR) is defined by (Eq. 4):

$$PSNR = 10.log_{10}(\frac{MAX_I^2}{MSE}) \tag{4}$$

where MAX_I is the maximum possible pixel value of the image.

4.4 Quantitative Results

We evaluated these three deep learning models in our dataset of thermal images. Table 1 shows that the DLIR (Affine part) and Affine ConvNet models give the best results. The quantitative results are presented in (Table 1).

The DLIR(affine part) achieved better performance in terms of the Dice coefficient ($DSC = 0.953$) than the AIRNet network ($DSC = 0.946$), followed by the Affine ConvNet model ($DSC = 0.9512$). We clearly notice that the best affine registration result is given by the two architecture DLIR (Affine part) and Affine ConvNet. More specifically, we can deduce that the architectures based on one subnetwork are more powerful and efficient than other models (AIRNet) on our dataset (Sect. 4.1).

Furthermore, We used the Peak Signal to Noise Ratio (PSNR) to measure the quality and the performance of the affine registration between the fixed image F and the moving image M. Before applying registration, the PSNR value is 21.63. Table 1 shows that the DLIR (Affine part) achieved the best value ($PSNR = 28.85$), followed by the Affine ConvNet ($PSNR = 28.62$) and then the AIRNet ($PSNR = 27.35$).

Table 1 Quantitative plantar foot registration results of three CNNs models. Dice similarity coefficient 2 (DSC), Mean squared error 3 (MSE) and Peak Signal to Noise Ratio 4 (PSNR)

Metrics models	Initial	AIRNet	Affine ConvNet	DlIR (affine part)
$Dice$	0.7661	0.9463	0.9512	**0.9534**
MSE	0.0098	0.0023	**0.0015**	0.0016
$PSNR$	21.63	27.26	28.62	**28.85**

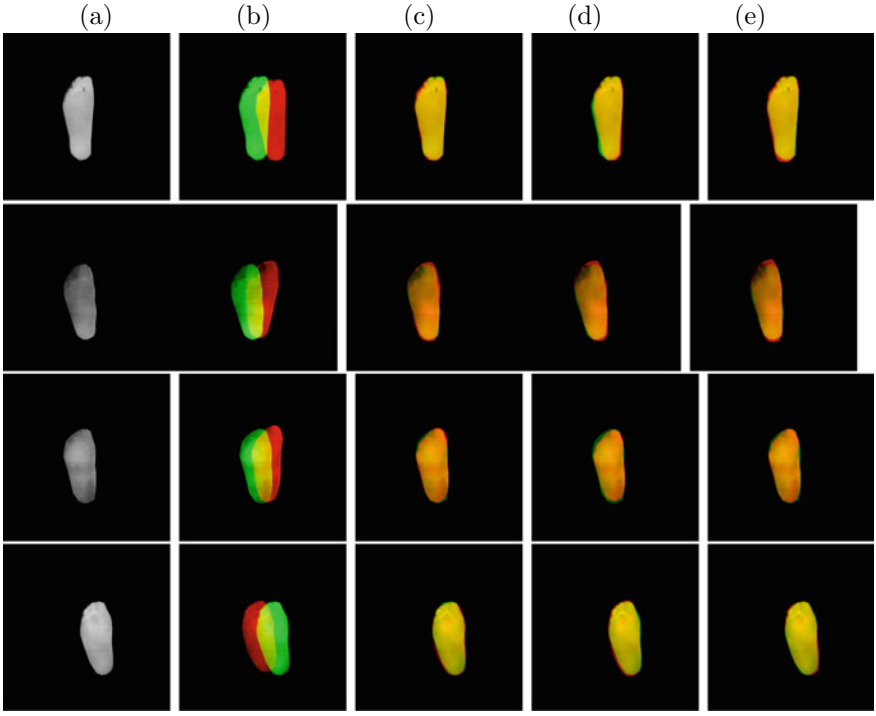

Fig. 6 Qualitative plantar foot registration results of three CNNs models with four examples. The foot in green is the fixed foot (Right foot) and the foot in red is the moving foot (Flipped left foot). **a** fixed image **b** before registration results **c** Affine ConvNet results **d** AIRNet results **e** DLIR(Affine part) results

4.5 Qualitative Results

We qualitatively evaluate the performance of these three tested deep learning models on the 79 thermal images of our dataset (Sect. 2). The results in Fig. 6 show that the DLIR (Affine part) and the AffinConvNet models give qualitatively good results, followed by the AIRNet method. We can conclude that deep learning affine registration methods give a visually correct overlap between the fixed image (right foot) and the registered foot (Flipped left foot).

5 Conclusion

In this work, our purpose was to find an accurate deep learning architecture of affine registration for thermal images of plantar feet. However, these different methods: Affine ConvNet, DLIR(Affine part), and AIRNet were tested and evaluated using

79 thermal images of plantar foot. The results in Table 1 and Fig. 6 show that the models based on one subnetwork Affine ConvNet and DLIR(Affine part) networks to predict the affine transformation parameters produced the best plantar foot affine registration results with a Dice score of 95%.

The next task will be calculating the difference in the temperature point to point between the left and right foot to detect hyperthermia areas. In future work, we aim to develop thermal analysis approaches for the detection of hyperthermia of diabetic foot and to merge the segmentation and affine registration in one processus using convolution neural networks.

Acknowledgements This research work was supported by the European Union's project Standup Horizon 2020 777661. A research and innovation program under the Marie Sklodowska-Curie. Aiming to develop smartphone applications for prevention and supervision of diabetic foot ulcers.

References

1. Armstrong, D.G., Holtz-Neiderer, K., Wendel, C., Mohler, M.J., Kimbriel, H.R., Lavery, L.A.: Skin temperature monitoring reduces the risk for diabetic foot ulceration in high-risk patients. Am. J. Med. **120**(12), 1042–1046 (2007)
2. Fraiwan, Luay, AlKhodari, Mohanad, Ninan, Jolu, Mustafa, Basil, Saleh, Adel, Ghazal, Mohammed: Diabetic foot ulcer mobile detection system using smart phone thermal camera: a feasibility study. Biomed. Eng. Online **16**(1), 117 (2017)
3. Vilcahuaman, L., Harba, R., Canals, R., Zequera, M., Wilches, C., Arista, M.T., Arbanil, H. (2015). Automatic analysis of plantar foot thermal images in at-risk type II diabetes by using an infrared camera. In: World Congress on Medical Physics and Biomedical Engineering, June 7–12, Toronto, Canada, pp. 228–231. Springer, Cham (2015)
4. Bougrine, A., et al.: On the segmentation of plantar foot thermal images with Deep Learning. In: 2019 27th European Signal Processing Conference (EUSIPCO). IEEE (2019)
5. Bouallal, D., Bougrine, A., Douzi, H., Harba, R., Canals, R., Vilcahuaman, L., Arbanil, H.: Segmentation of plantar foot thermal images: application to diabetic foot diagnosis. In: 2020 International Conference on Systems, Signals and Image Processing (IWSSIP) (pp. 116-121). IEEE (2020)
6. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992). https://doi.org/10.1109/34.121791.
7. Vilcahuaman, L., et al.: Automatic analysis of plantar foot thermal images in at-risk type II diabetes by using an infrared camera. In: Jaffray, D.A. (ed.) World Congress on Medical Physics and Biomedical Engineering, June 7–12, 2015, Toronto, Canada, vol. 51, pp. 228–231. Springer International Publishing, Cham (2015)
8. Chen, X., et al.: Learning unsupervised parameter-specific affine transformation for medical images registration. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham (2021)
9. Chee, E., Wu, J.: Airnet: self-supervised affine registration for 3D medical images using neural networks (2018). arXiv:1810.02583
10. Tang, K., Li, Z., Tian, L., Wang, L., Zhu, Y.: ADMIR-affine and deformable medical image registration for drug-addicted brain images. IEEE Access **8**, 70960–70968 (2020). https://doi.org/10.1109/ACCESS.2020.2986829
11. De Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Išgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Med. Image Anal. **52**, 128–143 (2019)

12. 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)
13. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269 (2017)
14. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning (2017). arXiv:1712.04621
15. Bouallal, D., Douzi, H., Harba, R.: Registration methods for thermal images of diabetic foot monitoring: a comparative study
16. Tanabe, Y., Ishida, T.: Quantification of the accuracy limits of image registration using peak signal-to-noise ratio. Radiol. Phys. Technol. **10**(1), 91–94 (2017). https://doi.org/10.1007/s12194-016-0372-3.
17. Bouallal, Doha, Bougrine, Asma, Harba, Rachid, Canals, Raphael, Douzi, Hassan, Vilcahuaman, Luis, Arbanil, Hugo: STANDUP database of plantar foot thermal and RGB images for early ulcer detection. Open Res. Eur. **2**(77), 77 (2022)
18. Liu, Y., Polo, A., Zequera, M., Harba, R., Canals, R., Vilcahuaman, L., Bello, Y.: Detection of diabetic foot hyperthermia by using a regionalization method, based on the plantar angiosomes, on infrared images. In: 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), pp. 1389–1392 (2016)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241. Springer (2015)