Affine registration of thermal images of plantar feet using convolutional neural networks.

Asma Aferhane[1,2], Doha Bouallal[1,2], Hassan Douzi[1], Rachid Harba[2]

[1]IRF-SIC Laboratory,Ibn Zohr University,Agadir,Morocco.
[2]PRISME Laboratory,Orléans University,Orléans, France.

Abstract

The use of a thermal camera to detect abnormal plantar foot temperature changes can be an effective way to identify the early signs of diabetic foot (DF) ulceration. In this work, we performed the affine registration of the plantar foot thermal images using four methods based on convolutional neural networks. These methods include two parts: an affine registration model for estimating transformation parameters and a spatial transformer for getting the registered image. The performances of the four models were evaluated using the Dice similarity coefficient (DSC), Mean Square Error (MSE), and peak signal-to-noise ratio (PSNR). In the first step, Methods were applied to register the left and right feet of the same subject, called "contralateral registration" and in the second step, the methods were evaluated on a pair of images of the same subject taken in two different times (T0 and T10) using a cold stress test protocol. Results showed that the used convolutional neural networks are more robust in both types of registration (contralateral and multitemporal), and they are suitable for our database of thermal images of DF, with the Dice score of 95 % for contralateral registration and a Dice score of 92 % for multitemporal registration.Futhermore, a transversal clinical study was perform on diabetic patients, that classified individuals into ischemic and non-ischemic groups. The objective was to analyze the coherence between the thermal results and medical data. The mean absolute point-to-point temperature difference $|\Delta T|$ between left and right foot is lower in non-ischemic patients than in those with ischemia, with a significance level of $p<0.05$. This result indicate that using thermal camera temperature assessment could help in the diagnosis of diabetic foot.

Keywords: Plantar foot thermal images, Deep learning, thermography, Medical images affine registration, convolutional neural network.

1 Introduction

People with diabetes are susceptible to a range of complications that affect multiple organs like the eyes (blindness), the nervous system (neuropathy), the kidneys (nephropathy) , and particularly the lower limbs known as the diabetic foot (DF) [1, 2] which represents the core of our work. Reduced blood flow to the feet can lead to untreatable ulceration that is preceded by inflammation or infection. In severe cases, it can lead to the lower limb amputations. According to diabetes experts, the occurrence of an ulcer is often associated with a temperature variation which can be identified using a thermal camera [3], that appears to be a useful predictor of foot ulceration. Numerous studies have shown a correlation between increased temperature in plantar foot of a individual and diabetic foot (DF) complications [4, 5]. Furthermore, the detection of abnormal plantar foot temperature of individuals with diabetes can serve as an early indicator of infection or ulceration in diabetic feet (DF). The most commonly used criterion is the point-to-point temperature difference $|\Delta T|$ of the corresponding area of the right and left foot, if it is greater than 2.2 [6] , this can be an indicative of abnormality. This is called hyperthermia as shown in Fig.1 with a high $|\Delta T|$ in the region of ulcers (yellow and red region) for the left foot of the person.

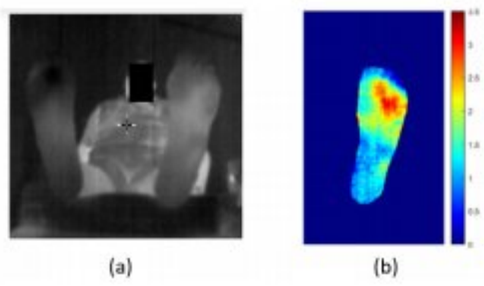


Figure 1: *(a) Original thermal image of the DF subject with ulcer on the left foot. (b) Thermal difference Map ΔT.*

Thermography is a non-invasive and contactless technology that allows visualization of a precise temperature distribution. In this context, the thermal images are taken using a friendly and mobile acquisition protocol based on a smartphone equipped with a dedicated thermal camera (FLIR ONE Pro)[7] .

The first step of the full automatic processing of the data is the segmentation [8, 9, 10] which consists in separating the plantar foot thermal image from the rest of the background. The second is the registration process for aligning the two feet and detecting the hyperthermia areas that are the core of our work. In [11], the authors used a classical rigid registration method which is the Iterative Closest Point (ICP) to register the two feet that have two transformations

Scanné avec CamScanner

translation and rotation. However, As a result of changes in the capture viewpoint and camera distance between the acquisitions at time T_0 and T_{10}, the rigid registration methods were found to have limitations, as reported by [12]. In this context, additional transformation parameters such as shearing and scaling are required. This type of transformation is known as affine registration, which aims to predict the matrix transformation T to align two feet with varying acquisition conditions.

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

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

In recent years, deep learning techniques have made a significant role in the development of models for medical image affine registration, particularly those based on convolutional neural networks (CNNs). Several CNN models have been developed for affine registration. The Self-Supervised Affine Registration (AIRNet) framework proposed by Chee et al. [14], which uses the use of two separate CNNs to predict the transformation matrix. One of the CNNs takes the fixed image as input, while the other takes the moving image as input. Another recent approach proposed by Tang et al. [15] is the unsupervised end-to-end Affine and Deformable Medical Image Registration (ADMIR) framework,which consists of two networks: Affine ConvNet for affine registration and a U-net architecture for deformable registration [16, 17]. The Affine ConvNet architecture uses a single network that recieves the concatenation of moving and fixed image to predict the transformation parameters.

Furthermore, the authors in this study used two well-known CNN architectures, namely the Visual Geometry Group with 16 layers (VGG-16) [18] and Densely Connected Convolutional Networks with 121 layers (DenseNet121) that was introduced by Huang et al [19, 20], for the task of affine registration. the concept is to incorporate these architectures with fully connected layers to enable the prediction of transformation parameters. Once this parameters have been predicted , the moving image is registered to the fixed image using a spatial transformer (resampler) adapted from VoxelMorph [21], which make the moving image and the fixed image fully aligned. Overall, The approach proposed in this study combines the pre-trained CNN architectures and spatial transformers to achieve precise and efficient affine registration of diabetic plantar foot.

In this paper, we performed plantar foot registration using the convolutional neural networks specifically AIRNet, Affine ConvNet, Vgg16 and DenseNet121 to obtain two fully registered feet images. The authors compared the performance of these CNN models for both types of registration. In order to calculate the temperature difference $|\Delta T|$ between contralateral and multitemporal feet (Section.3), which is important for thermal monitoring analysis. In addition to the affine registration of the plantar foot, we will present the results of a transversal clinical study on diabetic patients divided into two groups based on the presence or absence of ischemia. In order to analyze the coherence between the thermal results that we obtain and the medical data obtained by doctors.

The remaining part of this study is organized as follows. Section.2 describes The architectures of the four Convolutional neural Networks of affine registration. Section.3 presents the process of affine registration for plantar diabetic foot using convolutional neural networks (CNNs), including image acquisition and preprocessing. Section.4 presents registration results of the tested methods, for contralateral and multitemporel feet and for the last task, Section.5 shows a transversal clinical study on diabetic patients classified in ischemic and non-ischemic groups. Finally, in section. 6, we conclude and suggest some directions for future work.

2 Convolutional neural networks for affine registration

Convolutional neural networks are a type of neural network that has shown good results for image classification and registration. It has contributed to increase the performance of registration of medical thermal images. For Affine registration, the CNN models consists of obtaining a transformation matrix T, in order to align two images (fixed and moving images) with the application of spatial transformation.

A fundamental CNN architectures is a combination of convolution layers which extract different features and pooling layers to reduce the size of the image and the number of parameters which can reduce computational complexity. Then, a fully connected layers predict the transformation parameters using the extracted features. Affine registration networks are generally constructed into two ways, one approach is to use a two-pathway network and Another approach contains a single pathway encoder that they receive the concatenation of the images. Therefore, the registration methods chosen in this paper are AirNet, Affine ConvNet, and two other

Scanné avec CamScanner

models based on VGG16 and DenseNet121. In the following, we will present the details of these four affine registration architectures.

2.1 AIRNet architecture:

The encoder of AIRNet architecture (Fig.2) consists of two separate subnetworks that share the same parameters, the first one takes the fixed image F as input, while the second receives the moving image M, These subnetworks are adapted from the 2D DenseNet model [19].

The model comprises 6 of 2D convolutional

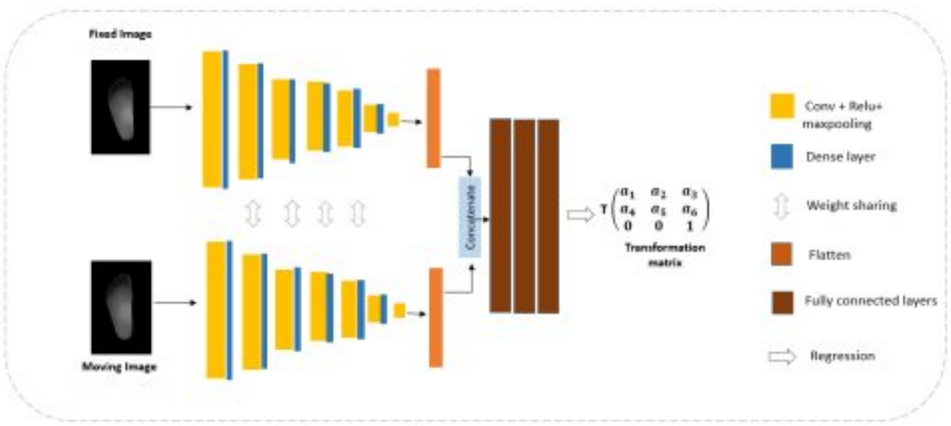


Figure 2: *AIRNet Architecture.*

layers using small $\times 1$ convolution filters and max-pooling layers which followed by a activation function (relu) and dense blocks as shown in Fig.3. A convolution and pooling layer is used before flattening the features that the outputs of the encoder are concatenated and passed them into several fully connected layers. Each layer of these fully connection layers is composed of linear layers and relu activation functions to produce the transformation matrix.

Figure 3: *Dense block with one Dense layer.*

2.2 Affine ConvNet architecture:

This architecture contains one subnetwork instead of two subnetworks to obtain the matrix of the transformation that the encoder received the concatenation of the moving (M) and fixed image (F) (Fig. 4).

The model consists of a series of 6 stridden convolution layers 3×3 convolution filters with a stride of 2, followed by a relu activation function. For the full connection part, each layer consists of linear layers and relu activation functions to estimate the six parameters of transformation to generate the matrix T.

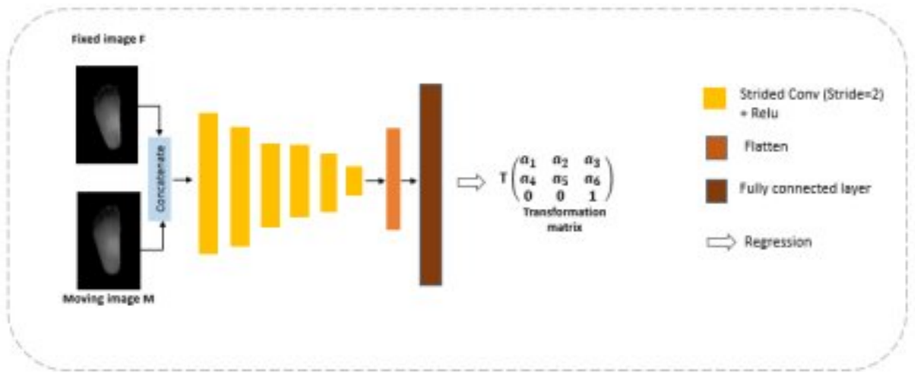


Figure 4: *Affine ConvNet Architecture.*

2.3 VGG16 architecture:

VGG16 is a convolutional neural network architecture that receives the moving and duplicate fixed image (3-channels). It consists of 5 blocks, where each block contains two or three convolutional layers followed by a max pooling. The convolutional layers use small 3×3 filters with a stride and a padding of 1.Then ,the max pooling layers have a 2×2 filter and a stride of 2 for decrease the spatial size of the feature as demonstrate in Fig.5.

To adapt VGG16 for predicting six transformation parameters for the matrix T, the fully connected layers at the top of the network are removed, and a new dense layer and ReLU activation is added. This layer is followed by a final dense layer with 6 outputs which producing the transformation parameters that formuled by the matrix T.

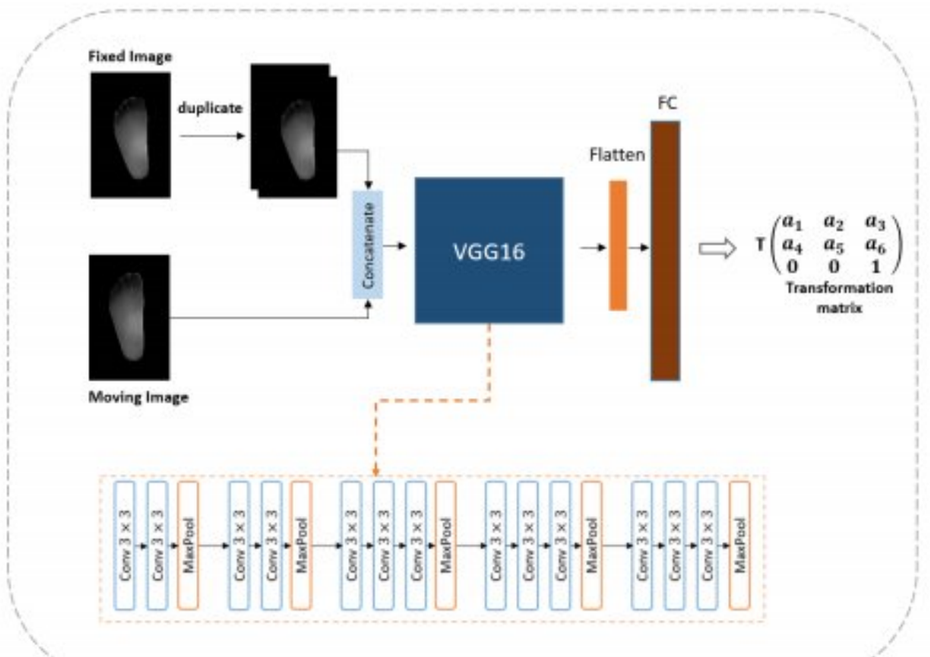


Figure 5: *VGG16 Architecture.*

2.4 DensNet121 architecture:

The DenseNet121 architecture is based on the concept of densely connected layers, where each layer is connected to every other layer in a feedforward way.

Scanné avec CamScanner

This architecture receives the moving and duplicate fixed image (3-channels) as input and is comprised of four dense blocks, with a transition layer between each pair of dense blocks. The architecture has a total of 121 layers, as depicted in the [19].

Each dense block has several convolution layers with

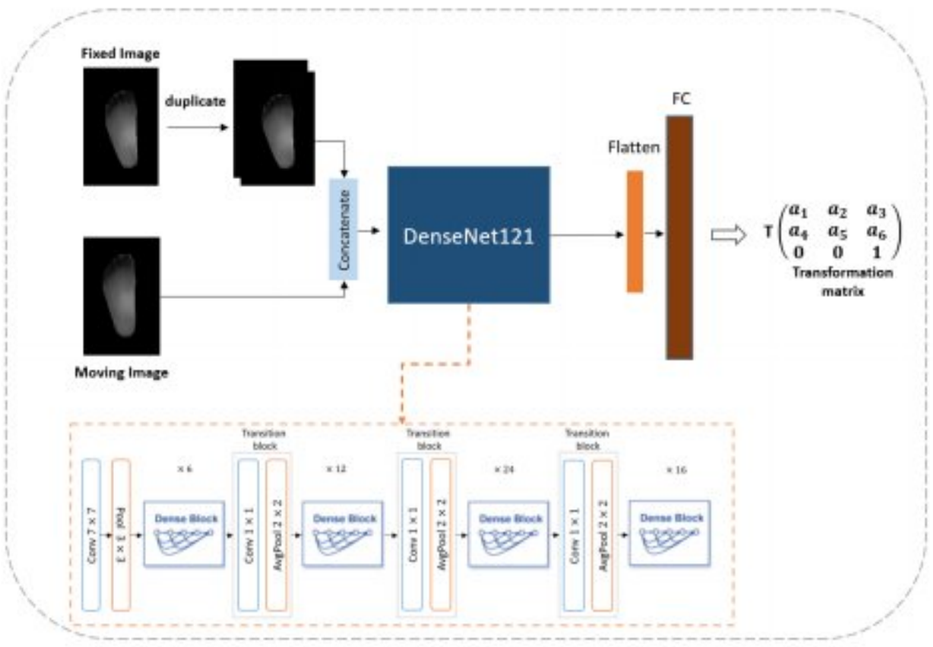


Figure 6: *DenseNet121 Architecture.*

fillers of 3×3 and 1×1 followed by a batch normalization layer and a ReLU activation function(Fig.3), and each transition layer has convolution layer with a small 1×1 filters and average pooling layer. Contrary the traditional CNNs, that only link output of one layer to the input of the subsequent layer, DenseNet connects all preceding layers to each subsequent layer, , allowing for the combination of feature maps from all preceding layers, which are then used as input for any successive layer as described Fig.6.

To adopt DenseNet121 for image affine registration, we remove the fully connected layers at the top of the network. Then, the outputs from the final dense block are concatenated and passed through a new dense layer and ReLU activation, that is followed by a final dense layer with 6 outputs, which produce the transformation matrix.

3 Methodology

3.1 background

Several studies have shown that the most common techniques used for analyzing diabetic foot thermograms are asymmetric and external stress analysis. Asymmetric analysis [5, 22, 23, 24] consists in analyzing skin temperature distribution in the left and right foot of a patient to detect any asymmetry that may indicate hyperthermia and potential foot ulcers. Research has revealed that a temperature variation between the two feet(left and right) greater than 2.2 is a sign of an ulcer in diabetic foot (DF). On the other hand, external stress analysis [5, 25] consists in observing the response of the body's thermoregulation system after the application of specific stress. In [9] , researchers evaluated the relationship between plantar foot temperature variation and cold stress test applied on diabetic patients. They performed a multitemporal analysis by calculating the temperature difference of the same foot at two different times (time (T_0 and 10 minutes later called timeT_{10}).

Through these thermal analysis, several preprocessing steps were required for the foot thermograms. The first step consists in segmenting the two feet from the rest of the background (Fig.9). The second is registration, which consists aligning each pixel of the two segmented feet to calculate the absolute temperature difference point-to-point using either the asymmetric (contralateral) or the multitemporal analysis. In this work, we aim to perform plantar foot registration using CNN architectures to obtain registered foot thermal images, in order to calculate the temperature difference $|\Delta T|$ between the contralateral and multitemporal feet.

3.2 Image acquisition and preprocessing

The thermal images of plantar foot are taken using a smartphone (Galaxy S8) equipped with a dedicated thermal camera FliROne Pro. In this section, we will first discuss the rationale behind our choice of this thermal camera. then we will describe the acquisition protocol of our thermal images.

The chosen camera: The FliROne Pro is a high-resolution thermal camera that has two sensor: the first is a thermal sensor that has the ability to capture thermal images at a resolution of 160 x 120 pixels and a spectral range of 8-14 m and the second is an RGB sensor designed to acquire visible images in parallel with the thermal core. This thermal camera can detect relative temperature differences of 0.1 making it suitable for detecting abnormal temperature of 2.2 in DF. It can be easily attached to a smartphone which makes it convenient for clinical setting use.

Acquisition protocol. To create our dataset of thermal images of plantar feet, we followed two acquisition protocol:

a) Single acquisition: patients are first subjected to a medical examination. Following this examination, they are then requested to rest barefoot for a 15-minute interval to allow the feet to regain their normal temperature. Once the 15-minute resting period has elapsed, thermal and RGB images of the feet are acquired [18]. As detailled in Fig.7.

b) Cold stress test acquisition: After the patient signed the informed consent form, we provided them with a 15-minute rest period to allow their feet to return to their normal state. During this time, the

Scanné avec CamScanner

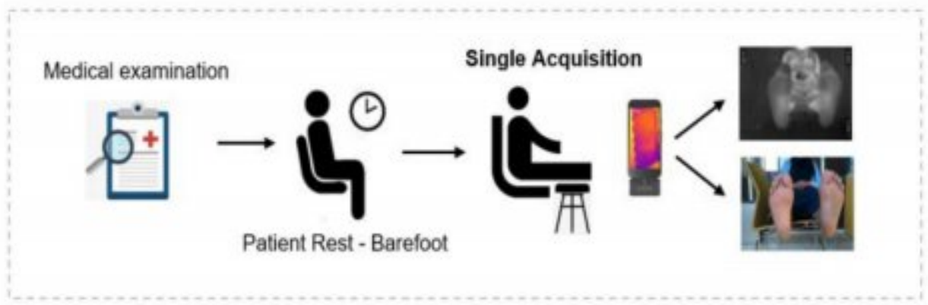


Figure 7: *Single acquisition (Contralareral Acquisition).*

patient was instructed to lie down on a stretcher with their feet placed at the end in a vertical position, with a distance of 10 cm between them. Baseline thermal and RGB images were captured using acquisition system. Following this, the patient was asked to immerse their feet, which were covered with thin plastic, in water at 15°C for 60 seconds. After 10 minutes(time T_{10}, a new plantar thermal and RGB image was taken. Then, two images of the same patient at two different times are taken [9]. As depicted in Fig.8.

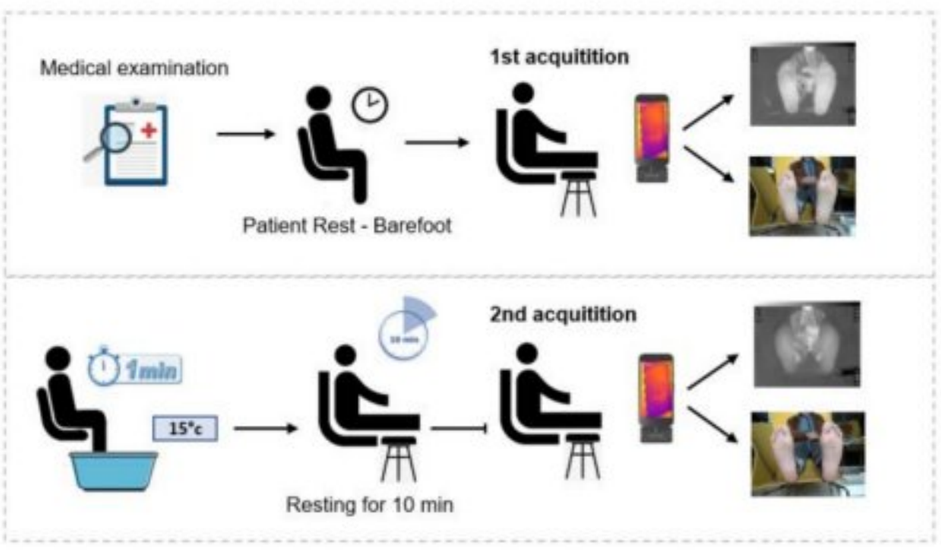


Figure 8: *Cold Stress Test Protocol (Multitemporal Acquisition).*

Data preprocessing Our dataset was generated from two groups: The first group, consisting of 102 healthy individuals who were not diabetic. The second group was comprised of 145 diabetic patients who agreed to participate in this study at the Diabetic Foot Service of the National Hospital Dos De Mayo de Lima in Peru [22]. The images of diabetic patients were captured under the supervision of specialists and diabetologists, using the cold stress test protocol, that two thermal images were taken at two different times for each patient, as shown in Fig.8.
After these acquisitions, The first step to build our the data consists in segmenting the plantar foot and separating them from the rest of the background by applying the ground truth to the thermal images (Fig.9). In this study, we carried out of two dataset: the first one is for a contralateral registration, which consists of aligning the right and left feet of the same patient image. And the second is for a mulitemporal

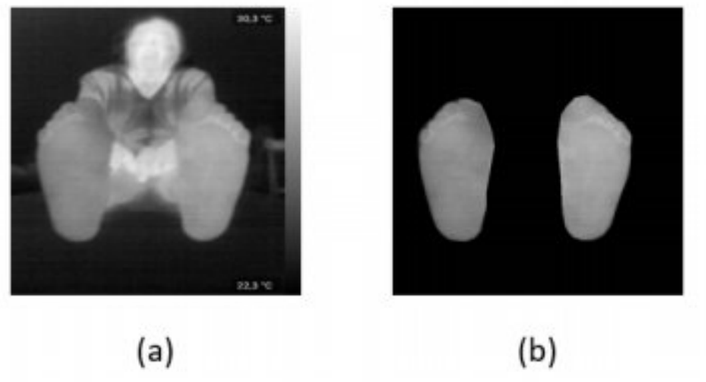


Figure 9: *(a) Thermal Image, (b) Segmented image.*

registration that applied on a pair of images of the same feet acquired in two different moments (T_0 and T_{10}) with the cold stress test.

3.3 Affine registration process of DF

This paper proposes the use of convolutional neural networks (CNNs) to perform affine registration of thermal images of the plantar foot of patients, for the two type of registration.

Types of registration .

-Contralateral registration: this type consists of aligning the image of right and left feet of the same patient by dividing the segmented images into two (right and left). The right foot image is used as the fixed image, while the left foot image is flipped vertically to match the orientation of the right foot and is considered the moving image., in order to calculate the temperature difference between the two feet after aligning them. The process is detailed in Fig.10

-Multitemporal registration: This second type consists of aligning a pair of thermal images of the same feet taken at two different times (Section.3.2) , before and after a cold stress test that lasted for 10 minutes. The image of the right foot captured at the later time called time T_{10} is treated as a moving image, while the image of the same foot taken earlier(time T_0 is considered the fixed image. The same approach is taken for the left foot. The objective is to align the two images of each foot for calculating the temperature differences acquired at time T_0 and T_{10}.

Affine registration :
In this work, the affine registation is used to align thermal images of the plantar foot in order to detect hyperthermia areas and localize regions at risk of ulceration. the process consists in using the four convolutional neural network architectures (Section.2) to predict the optimal affine transformation parameters between the moving and fixed images, which can be formulated by the transformation matrix T. Once the matrix is obtained, a spatial transformation is applied to align the moving image with the fixed one.

Scanné avec CamScanner

Fig.10 demonstrated the process used to perform more effective affine registration on the plantar foot of the thermal images on the both plantar foot of the thermal images dataset (contralateral and multitemporal) by using the four CNNs models.

4 Experiments and results

In this section, we present the results of four convolutional neural network models that were applied to two types of acquisitions (contralateral and multitemporal).

4.1 Traning and parameters

For contraleteral registration, the dataset was split into two subsets. Th training set consisted of 315 pairs of feet, including the right foot (fixed) and the left foot flipped horizontally (moving), while the testing set included 79 pairs. Multitemporal registartion, the training set consisted of 234 pairs of the same feet acquired at two different times (T_0 and T_{10}) and the testing set included 58 pairs. The training sets was used to train the model and optimize the connection parameters for the different neurons. while the test set was only used to test registration accuracy and model reliability. The objective was to assess the effectiveness of these models in both registration (contralateral and multitemporal).

The original images were in PNG format with a size of 640x280 pixels. However, they were resized to 256x256. For the convenience, the image intensity was normalized to a range of [0, 1]. In the present work, all models were trained with Tensorflow and Keras using Adam with a learning rate of 104 as an optimizer and a batch size was set to 1 for all models. To register the moving image, the spatial transformer module adapted from the open-source code in VoxelMorph was used.

To enhance the performance of affine registration and improve the robustness of the network, we used data augmentation for the training datasets. In the both cases of registration, the networks were trained using thermal images from the training set and tested on images from the test set that the networks has not seen before.

4.2 Evalution metric

We evaluated the effectiveness of four CNN models using the three most commonly used metrics for affine registration: the Dice Similarity Coefficient (DSC), which was introduced for the image registration dissimilarity assessment. the Mean Squared Error (MSE), which measures the pixel-wise difference between two samples (the fixed and registered images), and Peak Signal to Noise Ratio (PSNR), which indicates the quality of similarity between the fixed image and registered image, a higher PSNR value signifies greater similarity between two images.

- The Dice similarity coefficient (DSC) is defined by (Equ.2) :

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

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

- The mean squared error is defined by (Equ.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 (Equ.4):

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

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

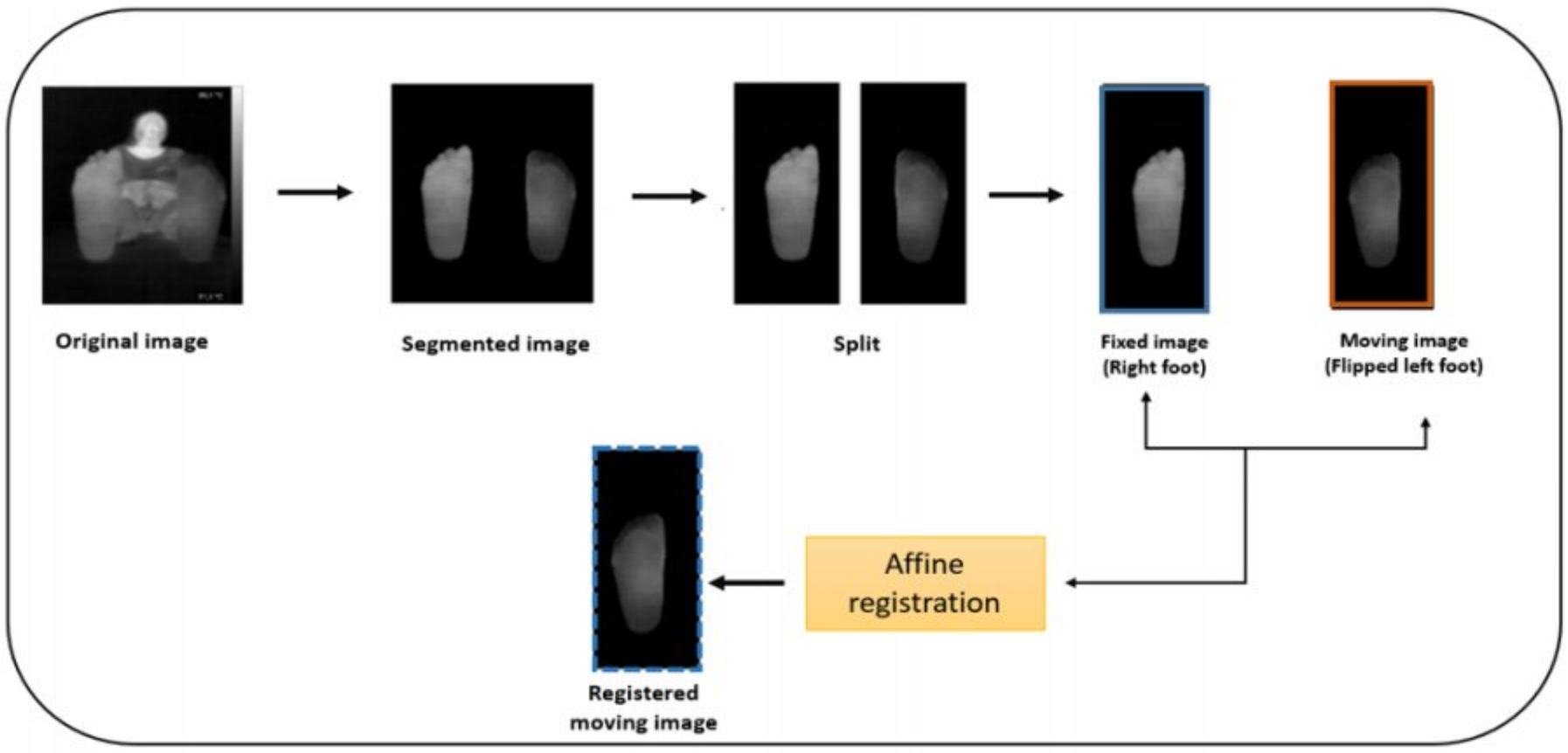
4.3 Quantitative Results

The performance of the four Convolutional Neural Networks (CNNs) for affine registration was evaluated on two thermal image datasets (contralateral and multitemporal).

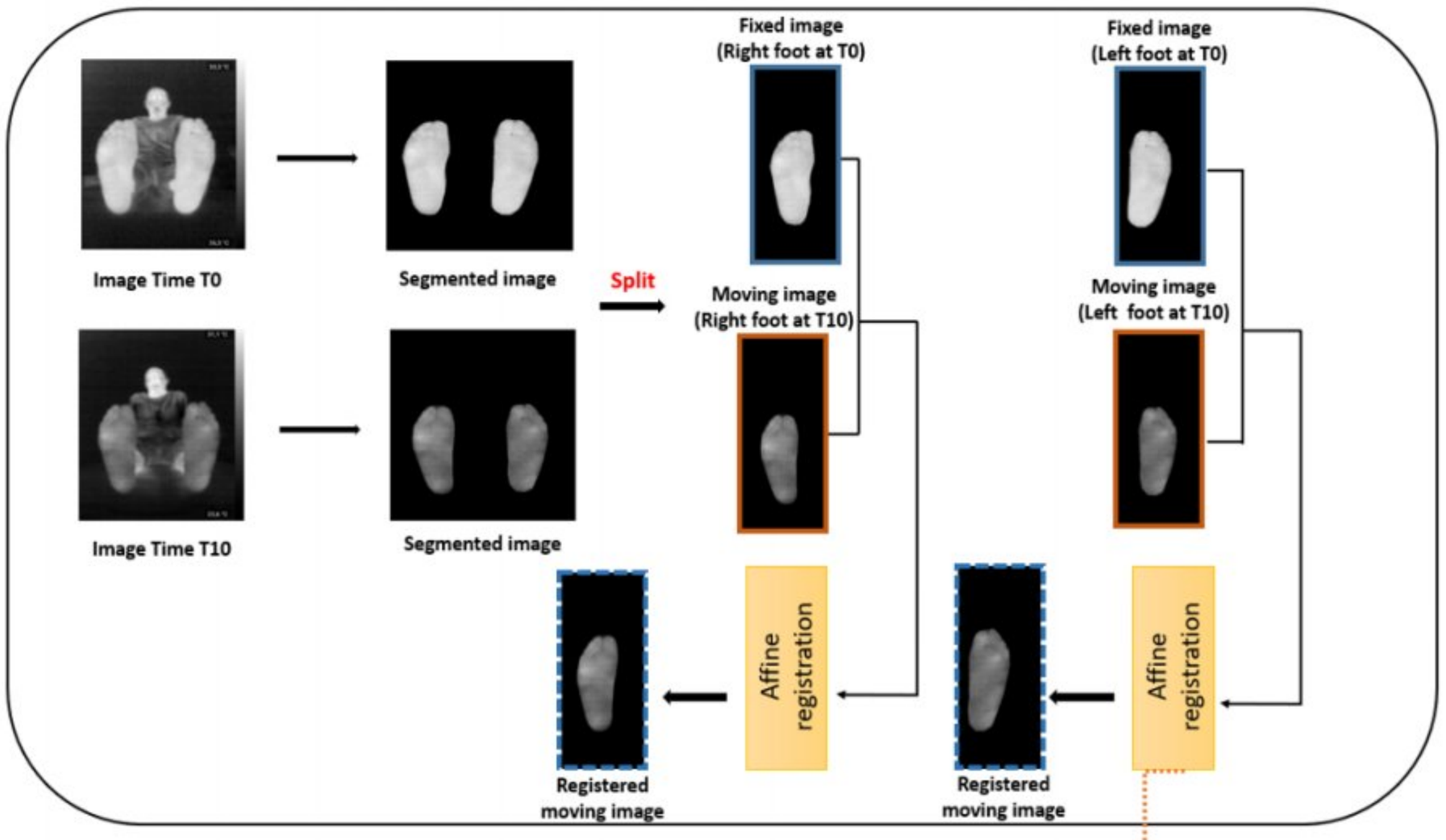
As shown in table.1, the results of contralateral registration indicate that the Affine ConvNet and VGG16 models achieved a DSC^2 value of approximately 95%, followed by AIRNet model with DSC of 94%, and the DenseNet121 model achieved a DSC of 90%. These results represent a significant improvement over the initial registration, which had a DSC of only 76%, which demonstrate the effectiveness of the CNNs models aligning thermal images accurately.

For multitemporal registration, the camera distance and position can change between time T_0 and T_{10}, as well as the freehand manner in which the cold stress test images are captured using smartphones can lead to changes in scale between the two feet and making the registration task more complicated. The results presented in table.[1] clearly demonstrate the complexity of this task compared to the the contralateral registration, as the dice scores achieved by the

Scanné avec CamScanner



(a) Contralateral affine Registration Process.



(b) Multitemporal affine Registration Process.

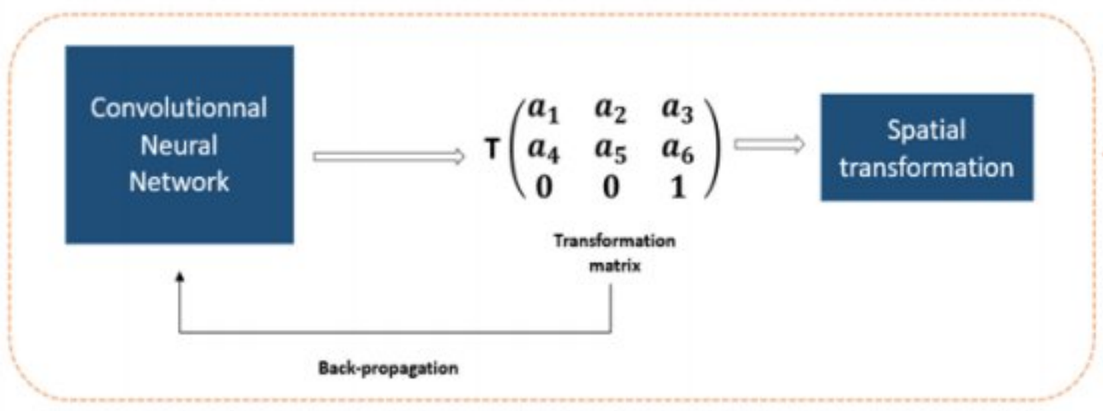


Figure 10: *Overiew of affine registration:(a) contralateral affine registration process between left and right foot and (b) Mutitemporel affine registration process. For affine registration CNNs archituctures used for predict the transformation matrix and a spatial transformation is applied to registred moving image by the matrix the transformation obtained with CNNs models.*

Scanné avec CamScanner

Table 1: *Quantitative Result of Contralateral and Multitemporal plantar Foot Registration.*

	Metrics \ *Models*	Initial	AIRNet	Affine ConvNet	VGG16	DenseNet121
Contralateral	Dice	0.7661	0.9463	0.9512	0.953	0.9007
	MSE	0.0098	0.0023	0.0015	0.0017	0.0047
	PSNR	21.63	27.26	28.62	28.88	23.67
Multitemporal	*Dice*	0.7567	0.9112	0.9232	0.9225	0.8808
	MSE	0.0106	0.0029	0.0025	0.0026	0.0044
	PSNR	20.86	25.41	26.95	26.24	24.05

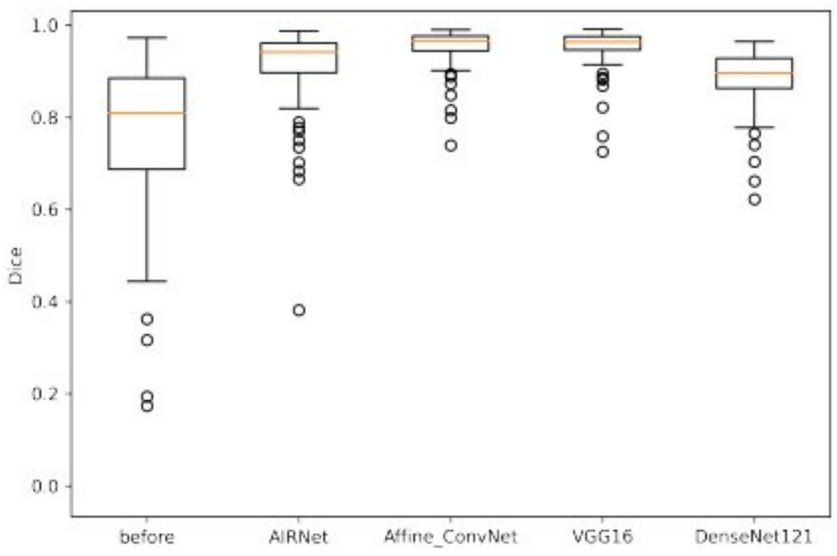


Figure 11: *Box-plot of Dice similarity coefficient (Contralateral).*

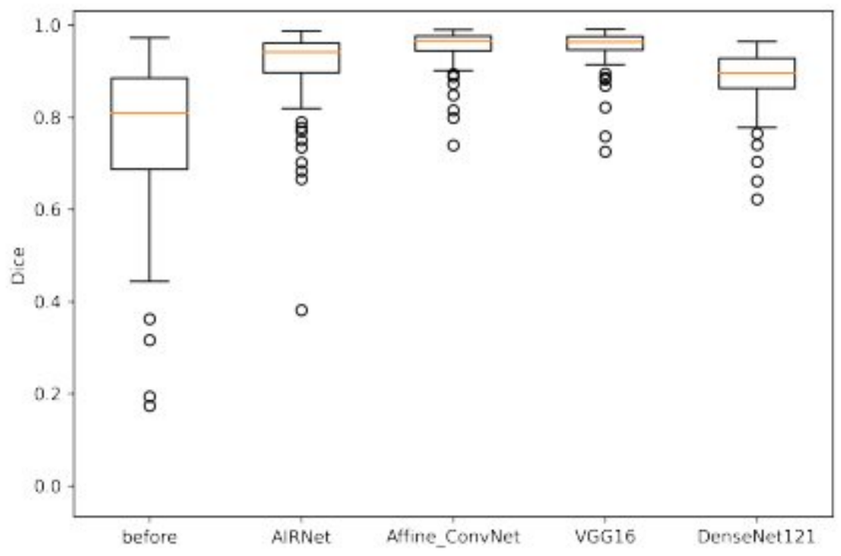


Figure 12: *Box-plot of Mean Squared Error (Contralateral).*

four CNN models were slightly lower than those of contralateral registration.

The Affine ConvNet, VGG16, and AIRNet models achieved DSC values of around 92%, followed by the DenseNet121 model with a DSC of 88%. These results indicate a significant improvement over the initial registration, which achieved a score of only 76%.

To conduct further analysis , box plots in Fig.11 and Fig.12 were used to display the distribution of the Dice similarity coefficient (DSC) and mean squared error (MSE) for contralateral registration (Section.4.2. These values were obtained by analyzing a visually correct overlap between the fixed image (right foot) to the registered moving image for each tested method. The results illustrated that the DSC values for Affine ConvNet,VGG16 and AIRNet models were approximately 5% higher than the DenseNet121 model, and is improved by about 18% than the DSC obtained before any CNN models were applied. Additionally, the MSE values of Affine ConvNet and VGG16 were approximately 0.003 smaller than that of DenseNet121, with a difference of 0.008 compared to the initial registration.

In the context of multitemporal registration, Fig.13

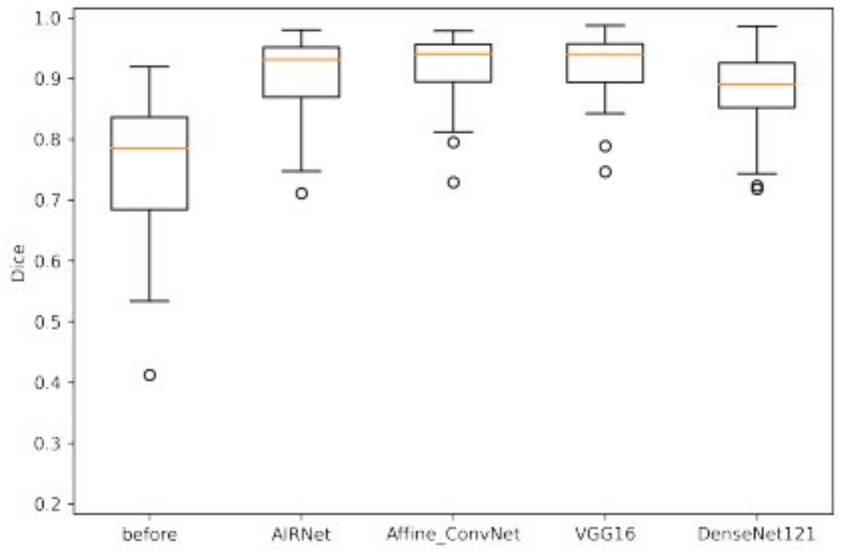


Figure 13: *Box-plot Dice similarity coefficient (Multitemporal).*

and Fig.14 demonstrate also the box-plots of The Dice similarity coefficient (DSC) and mean squared error (MSE) respectively, for each tested method. Our results indicate that these Four CNNs models can improve further the registration accuracy by about 13% compared with the initial registration.

4.4 Qualitative results

The qualitative results are in total agreement with the results presented in table.1, indicating that all three CNN models effectively achieve a precise visual overlap between the fixed thermal image and the registered foot. Fig.15 illustrates that the Affine ConNet, AIRNet, and VGG16 models demonstrate

Scanné avec CamScanner

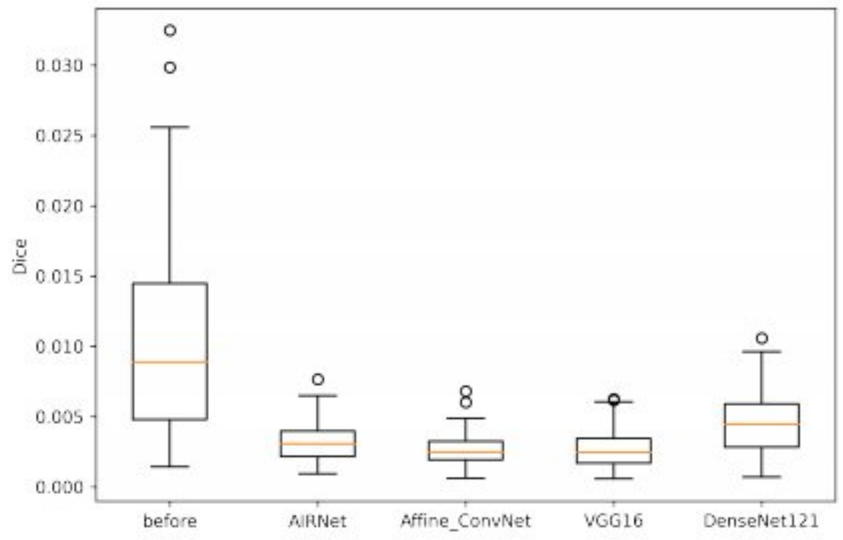


Figure 14: *Box-plot Mean Squared Error (Multitemporal).*

the most effective foot registration, followed by the DenseNet model in case of contralateral registration images and multitemporal registration, while the challenges due to changes viewpoint and distance from the camera between acquisitions at different times ($T_0 and T_{10}$)

Despite these challenges, the four CNN models evaluated in this study have demonstrated their robustness, as illustrated in Fig.16, that the effectiveness of these models is evidenced by their ability to accurately align the foot images acquired at the same and different times. Overall, the CNNs models have the potential to registred the feet with all the complexity.

5 Clinical analysis

5.1 Data collection

Medical data include age, time of diagnosis (TOD), body mass index (BMI), the neurological assessment and the vascular assessment. The neurological assessment included the evaluation of total symptom score (TSS) and The vascular assessment consists of collecting the ankle brachial pressure index (ABI) for the right and left foot, respectively ABI_R and ABI_L. [26, 27]. The risk classification based on the clinical examination results is presented below [7]:

- R0 (Low risk): no ischemia, no neuropathy.
- R1 (Medium risk): no ischemia, neuropathy.
- R2 (High risk): ischemia, neuropathy.

Based on the classification by the risk. The authors in this study classified The diabetic patients into two groups:

- Non-Ischemic patients: low and medium risk were combined $(R_0 + R_1)$.
- Ischemic patients: High risk (R_2).

On the other hand, the thermal images used in this study were acquired using a Samsung Galaxy S8 and

Table 2: *Mean and standard deviation of age, time of diagnosis (TOD), total symptom score (TSS), ABI_L and ABI_R of each group.*

		Groups	
Clinical data	**General**	**Non-ischemic ($R_0 + R_1$)**	**Ischemic (R_2)**
Age	62.81 (±10.4)	61.7 (±10.7	65 (±9.56)
TOD	10.17 (±8.85)	8.94 (±8.25	12.7 (±9.60)
TSS	3.38 (±2.75)	3.35 (±2.93)	3.44 (±2.39)
ABI_R	1.01 (±0.2)	1.09 (±0.107)	0.83 (±0.21)
ABI_L	0.982 (±0.2)	1.07 (±0.106)	0.79 (±0.2)
Total	128	86	42

Flir One Pro camera. After being segmented, they are split into two images , one for each foot. Accurate alignment the two contralateral feet (right and left feet) is necessary to calculate the point-to-point temperature difference between the left and right feet, denoted as $|\Delta|$ using CNNs model of affine registration. A set of thermal information was extracted from the acquired images and temperature maps(Fig.17). The mean temperature of the plantar foot surface at time 0 and 10 minutes for both the left and right and the mean temperature of $|\Delta T|$ was calculated for two time T_0 and T_{10}. These thermal values were calculated for the left and the right foot. As results are almost identical for both feet, we have opted to present the results for the left foot only. Table.[3] contains the mean values of the thermal data for each group.

Table 3: *Mean values and standard deviations of thermal information for each group*

		Groups	
Thermal data (°C)	**General**	**Non-ischemic ($R_0 + R_1$)**	**Ischemic (R_2)**
T_0	27.87 (± 2.18)	27.90 (± 2.15)	27.79 (±2.26)
$\|\Delta T\|_0$	0.49 (±0.28)	0.455 (±0.259)	0.56 (±0,31)
T_{10}	25.84 (±2.55)	25.8 (±2.74)	25.82 (±2,14)
$\|\Delta T\|_{10}$	0.58 (±0.30)	0.544 (±0.289)	0.66 (±0.31)

Scanné avec CamScanner

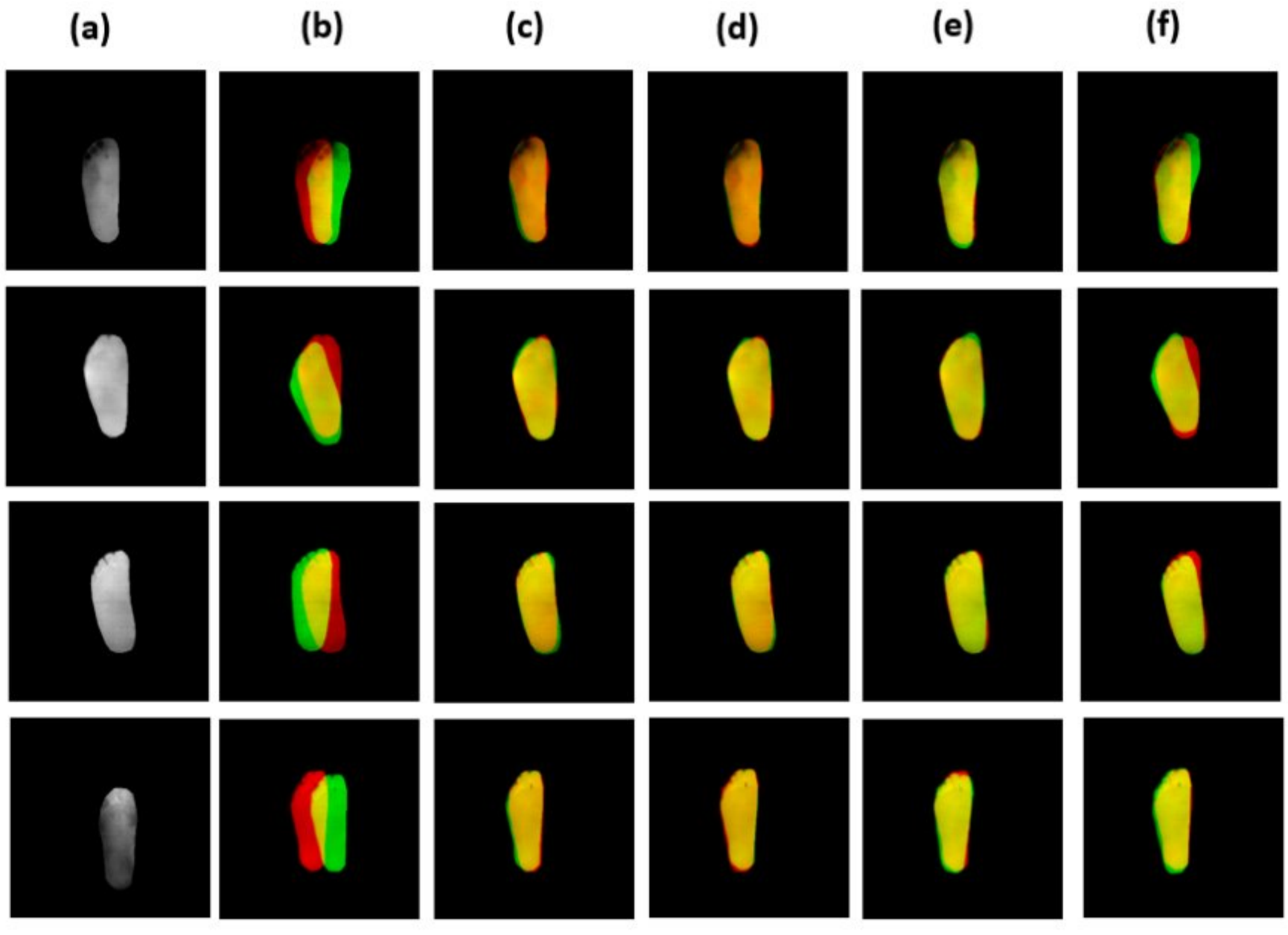


Figure 15: *Contralateral Registration . The foot in green is the fixed foot (Right foot) and the foot in red is the moving foot (Flipped left foot). (a) fixed image (b)before registration (c) AIRNet (d) Affine ConvNet (e) VGG16 (f) DenseNet121.*

5.2 Results

A Student t-test was performed on the age, TOD, TSS to assess if there were significant differences in these two groups(Ischemic and no Ischemic) based on table.2. ABI (ABI_R and ABI_L) was not included in this test since the risk group was mainly based on this variable.

From (Table.4), there are no significant differences in age or total symptom score (TSS) were observed between the two groups. However, the time of diagnosis (TOD) was found to be significantly different among the groups($p < 0.05$) indicating that TOD is an important factor in the classification of patients by risk.

The t-test was also performed on the thermal information (Table.3) on both groups to assess the statistically significant difference between them based on the extracted information from the images and temperature maps.

Table 4: *Related student t-test of clinical data*

	t-test	
Clinical data	**t-value**	**p-value**
Age	−1.76	0.0804
TOD	- 2.175	0.0348
TSS	- 1.1854	0.8531

Table 5: *Related student t-test of thermal data*

	t-test	
Thermal data	**t-value**	**p-value**
T_0	0.2426	0.8087
$\|\Delta T\|_0$	-2.1271	0.0357
T_{10}	-0.0456	0.9640
$\|\Delta T\|_{10}$	-2.1572	0.0329

Ischemic patients have poor blood circulation and low blood flow to their legs due to artery narrowing. Clinically, these patients typically have cold feet. From the table.5, we observed that the ischemic group (R_2) has a higher temperature difference between contralateral feet at time 0 and 10 minute ($|\Delta T|_0 = 0.56$ and $|\Delta T|_{10} = 0.66$) compared to the non-ischemic group ($R_0 + R_1$).

Furthermore, the results of the Student t-test in table.5 demonstrated that there is a significant difference between the two groups based on $|\Delta T|_0$ and $|\Delta T|_{10}$ ($p < 0.05$). These results suggest that the temperature difference $|\Delta T|$ can be used as a complementary tool to medical examination. This will allow for distinguishing between ischemic and non-ischemic groups.

Scanné avec CamScanner

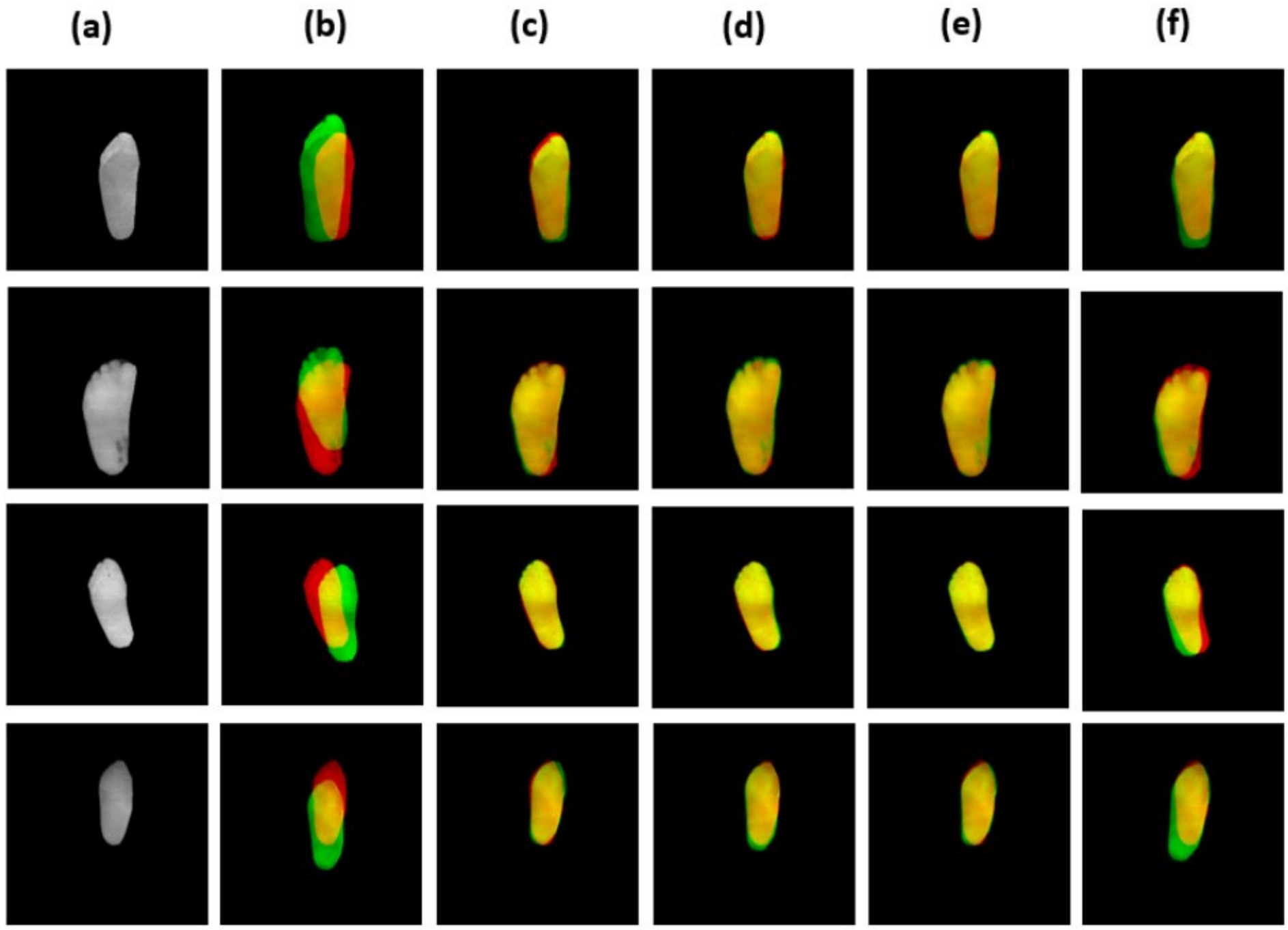


Figure 16: *Multitemporal Registration. The foot in red is the fixed foot (right at T_0) and the foot in green is the moving foot((right at T_{10})). (a) fixed image (b)before registration (c) AIRNet (d) Affine ConvNet (e) VGG16 (f) DenseNet121.*

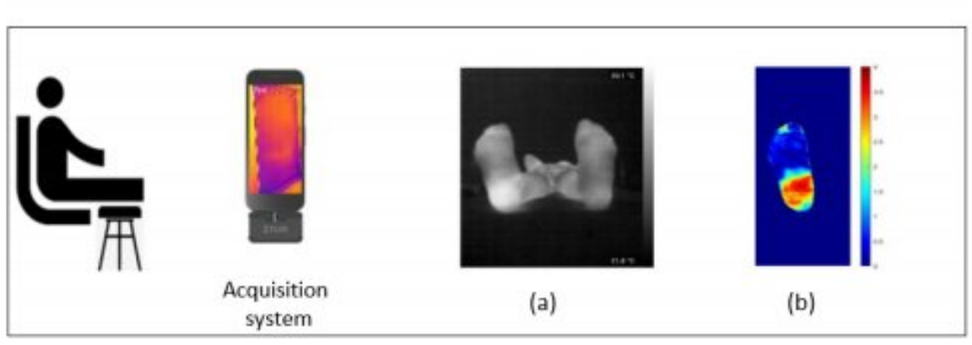


Figure 17: *Extraction of Thermal informations. (a) Mean of T_0 and T_{10} and (b) Temperature difference $|\Delta T|$.*

6 Conclusion

In this paper, we attempted to use affine registration methods based on CNNs to align thermal images for diabetic foot application. For the registration of contralateral feet, the foor models were evaluated which the results in table.1 and Fig.15 show that the Affine ConNet, VGG16 and AIRNet models, followed by DenseNet performed better, in comparison with the initial registration. However, we observed slightly lower performance in multitemporal registration compared to contralateral registration. This could be due to changes in capturing viewpoint and distance between the two moments $(T_0 and T_{10})$, which increased the complexity of registration. The Dice coefficients obtained showed a DSC value of 0.92 for affine ConvNet and VGG16 in the case of multitemporal registration, compared to a DSC value of 0.95 for contralateral registration. Despite this difference, our tested methods proved to be effective, as demonstrated by the results presented in the table.1 and Fig.16, when compared to inital registration.

his study focuses on using CNN models to register feet with all complexities associated with the adopted acquisition protocols (single acquisition and cold stress test), which are based on a smartphone and a dedicated thermal camera. The purpose of this registration task is to identify abnormalities in the plantar foot of patients with diabetic foot (DF) disorder.

In the future, we intend to expand our dataset to to enhance the accuracy of the CNNs models. Additionally, we plan to improve all the tested models and combine them with the segmentation task to create a complete system of diagnostic that could be used by clinicians during diabetic

Scanné avec CamScanner

foot examination. The ultimate goal is to identify hyperthermia areas and detect DF disorders in their early stages.

Acknowledgments 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

References

[1] van Netten, Jaap J., et al. "Definitions and criteria for diabetic foot disease." Diabetes/metabolism research and reviews 36 (2020): e3268.

[2] Evangeline, N. Christy, S. Srinivasan, and E. Suresh. "Application of non-contact thermography as a screening modality for Diabetic Foot Syndrome–A real time cross sectional research outcome." Biomedical Signal Processing and Control 79 (2023): 104054.

[3] Kumar, Prasoon, et al. "Applications of thermal imaging with infrared thermography in Orthopaedics." Journal of Clinical Orthopaedics and Trauma 24 (2022): 101722.

[4] David G Armstrong, Katherine Holtz-Neiderer, Christopher Wendel, M Jane Mohler, Heather R Kimbriel, and Lawrence A Lavery. Skin temperature monitoring reduces the risk for diabetic foot ulceration in high-risk patients. The American journal of medicine, 120(12):1042– 1046, 2007.

[5] Hernandez-Contreras, D., et al. "Narrative review: Diabetic foot and infrared thermography." Infrared Physics Technology 78 (2016): 105-117.

[6] Alshayeji, Mohammad H., and Silpa ChandraBhasi Sindhu. "Early detection of diabetic foot ulcers from thermal images using the bag of features technique." Biomedical Signal Processing and Control 79 (2023): 104143.

[7] 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, 2015, Toronto, Canada (pp. 228-231). Springer, Cham.

[8] Bougrine, Asma, et al. On the segmentation of plantar foot thermal images with Deep Learning. 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019.

[9] Bouallal, D., Bougrine, A., Douzi, H., Harba, R., Canals, R., Vilcahuaman, L. and Arbanil, H., 2020, July. *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.

[10] Bougrine, Asma, et al. "Segmentation of Plantar Foot Thermal Images Using Prior Information." Sensors 22.10 (2022): 3835.

[11] L. Vilcahuaman et al., « Automatic Analysis of Plantar Foot ThermalImages in at-Risk Type II Diabetes by Using an Infrared Camera », in World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada, vol. 51, D. A. Jaffray, Ed. Cham: Springer International Publishing, 2015, p. 228-231.

[12] Bouallal, Doha, Hassan Douzi, and Rachid Harba. Registration Methods for Thermal Images of Diabetic Foot Monitoring: A Comparative Study.

[13] Chen, Xu, et al. "Learning unsupervised parameter-specific affine transformation for medical images registration." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021.

[14] Chee, E., Wu, J.*Airnet: self-supervised affine registration for 3D medical images using neural networks*. Preprint (2018). arXiv:1810.02583

[15] K. Tang, Z. Li, L. Tian, L. Wang and Y. Zhu, ADMIR–Affine and Deformable Medical Image Registration for Drug-Addicted Brain Images, in IEEE Access, vol. 8, pp. 70960-70968, 2020, doi: 10.1109/ACCESS.2020.2986829.

[16] 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)

[17] Shi, Huabang, et al. "Joint few-shot registration and segmentation self-training of 3D medical images." Biomedical Signal Processing and Control 80 (2023): 104294.

Scanné avec CamScanner

[18] K. a. Z. A. Simonyan, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[19] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger. *Densely connected convolutional networks*. In IEEE Conference on Computer Vision and Pattern Recognition, pages 2261–2269, 2017.

[20] Kong, Lingzhi, and Jinyong Cheng. "Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion." Biomedical Signal Processing and Control 77 (2022): 103772.

[21] Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., Dalca, A. V. (2019). *VoxelMorph: a learning framework for deformable medical image registration*. IEEE transactions on medical imaging, 38(8), 1788-1800.

[22] Bouallal, Doha, Hassan Douzi, and Rachid Harba. Registration Methods for Thermal Images of Diabetic Foot Monitoring: A Comparative Study.

[23] Bouallal, Doha, Asma Bougrine, Rachid Harba, Raphael Canals, Hassan Douzi, Luis Vilcahuaman, and Hugo Arbanil. *STANDUP database of plantar foot thermal and RGB images for early ulcer detection*. Open Research Europe 2, no. 77 (2022): 77.

[24] Kaabouch, N.; Chen, Y.; Anderson, J.; Ames, F.; Paulson, R. Asymmetry analysis based on genetic algorithms for the prediction of foot ulcers. Vis. Data Anal. 2009, 7243, 724304.

[25] Fushimi, Hisako, et al. "Abnormal vasoreaction of peripheral arteries to cold stimulus of both hands in diabetics." Diabetes research and clinical practice 32.1-2 (1996): 55-59.

[26] E.. J. Bastyr, K. L. Price, V. Bril, et MBBQ Study Group, « Development and validity testing of the neuropathy total symptom score-6: questionnaire for the study of sensory symptoms of diabetic peripheral neuropathy », Clin Ther, vol. 27, no 8, p. 12781294, august 2005, doi: 10.1016/j.clinthera.2005.08.002.

[27] Tang, Yi, et al. "The role of carotid stenosis ultrasound scale in the prediction of ischemic stroke." Neurological Sciences 41 (2020): 1193-1199.

Scanné avec CamScanner