


RESEARCH

Open Access



Palpation localization of radial artery based on 3-dimensional convolutional neural networks

Qiliang Chen^{1,2}, Yulin Huang², Xing Zhu², Hong Lu⁴, Zhongzhi Ji², Jiacheng Yang² and Jingjing Luo^{2,3*} 

*Correspondence:
luojingjing@fudan.edu.cn

¹ School of Basic Medicine,
Guangzhou University of Chinese
Medicine, Guangzhou 510006,
China

² Jihua Laboratory,
Foshan 528200, China

³ Academy for Engineering
and Technology, Fudan
University, Shanghai 210043,
China

⁴ School of Computer
Science, Fudan University,
Shanghai 210043, China

Abstract

Palpation localization is essential for detecting physiological parameters of the radial artery for pulse diagnosis of Traditional Chinese Medicine (TCM). Detecting signal or applying pressure at the wrong location can seriously affect the measurement of pulse waves and result in misdiagnosis. In this paper, we propose an effective and high accuracy regression model using 3-dimensional convolution neural networks (CNN) processing near-infrared picture sequences to locate radial artery upon radius at the wrist. Comparing with early studies using 2-dimensional models, 3Dcnn introduces temporal features with the third dimension to leverage pulsation rhythms, and had achieved superior performance accuracy as 0.87 within 50 pixels at testing resolution of 1024×544 . Model visualization shows that the additional dimension of the temporal convolution highlights dynamic changes within image sequences. This study presents the great potential of our constructed model to be applied in real wrist palpation location scenarios to bring the key convenience for pulse diagnosis.

Keywords: 3Dcnn, Palpation localization, Temporal convolution, Video analysis, Traditional Chinese Medicine

1 Introduction

Radial artery pulse diagnosis is an indispensable part of the principle 4-methods of diagnosis in Traditional Chinese Medicine (TCM) [1, 2]. It provides with rich physiological information for health evaluation of patients [3], and is regarded as an important tool in non-invasive diagnostic practice [4]. However, the localization of radial artery pulse, divided into 'Cun' (Inch), 'Guan/Gwan' (Bar), and 'Chi' (Cubit) in TCM [5], relies heavily on the doctor's personal experience currently. This has disadvantages of low efficiency or poor reproducibility [6]. Techniques for accurately, automatically and efficiently locating these three important radial pulse positions can make great contributions for modernization of TCM diagnosis.

Several advanced methods have been used to objectively locate radial artery. One example is to use tactile sense or pressure sensor array [7, 8], but it suffers from low positioning accuracy [9] which relies on the sizes and conformity of sensors [10, 11]. Another study performs non-contact image detection, locating radial artery using

thermal imagery [3]. This approach prevents direct contact; however, it depends on the sensitivity of an infrared thermal imaging equipment. Individual variations in the shape of the wrist can also cause irreducible deviations. This paper proposes video location method based on deep learning models, hoping to reduce installation cost while improves location accuracy and repeatability.

Video analysis for localization has been increasingly used to detect the location of monitored objects. Related studies include finding object's boundaries [12] [13], detecting stellar position [14], as well as locating human position in complex real-life scene [15]. For medical applications, videos are also used for non-contact monitoring of vital signs [16], monitoring of blood perfusion in free flaps [17], and detection of muscle tension dystonia [18] etc. Contraction and relaxation of the heart's ventricles produce rhythmic circulation changes, which are reflected in such blood volume waveform. In fact, "Guan" of wrist pulse positions has the most obvious periodic beating signal, and studies have shown prominent periodic pulsation signals detected around "Guan" in videos [19, 20]. We propose for the first time to use video analysis combined with vital signal to directly locate "Guan" TCM pulse positions of the radial artery.

In this study, we evaluate deep learning models for video analysis in locating the position of the radial artery. Convolution neural networks (CNN) are usually applied to automatically and adaptively detect spatial hierarchies of features [21] and have already been widely used in the medical field in recent years [22]. Our earlier study has achieved advanced accuracy using 2-dimensional CNN model with image resolution of 1024×544 [20]. However, this resolution is relatively low, and the 2Dcnn method extracts information from a single picture with only spatial rather than temporal patterns. We introduced 3-dimensional CNN model [23], which has 3D convolution kernel and learns information not only from spatial but also temporal features by analysing the relationship between image sequences. In fact, 3Dcnn has been used both in the classification task [24] and the regression task [25] in medical image detection with convincing performance achieved. We hope to make use of information from rhythmic pulse beating process as the typical temporal information at high resolution.

The main contributions of this paper consist of three parts: (1) we constructed wrist image dataset of our own, which contains 500 labelled videos of TCM pulse localization. (2) We proposed an advanced way to construct model of 3Dcnn by adding temporal rhythms, and to improve object localization accuracy. (3) We optimized the structure of the traditional CNN model by ablation experiments, and explained the effectiveness of this model from the perspective of model's visualization.

The rest of this paper is organized as follows: Sect. 2 shows the method of model's construction; Sect. 3 reports the results based on the proposed 3Dcnn model; Sect. 4 shows the discussion, and Sect. 5 shows the conclusion finally.

2 Methods

To improve the accuracy of palpation localization, we propose an advanced localization model based on 3D convolutional neural network and applied to infrared video of radial artery. The 3D convolution kernel is introduced to enrich model's ability to extract spatial and temporal features from circulatory pulsations. This can more accurately predict the location of the radial artery in the wrist.

2.1 Data acquisition

In our research, to erase the interference from environmental light inference, the near infrared camera from HIKVISIONH MV-CA050-20GN was used for collecting high quality wrist video of volunteers' forearm (Fig. 1). In this study, we recruited a total of 50 people to participate in data collection.

During experiment, the distance from camera to wrist was set to approximately the same, and the resolution of each video is 2048×1088 . For each volunteer, any obvious bracelet in the wrist was asked for removal, and there was no scar on the wrist. 10 different inclination degree of forearm position were obtained, and each position was recorded for 8 s at 30 frames per second. Moreover, for each recording video, volunteer's radial artery location was checked by the experimenter and presenting pixel was recorded. All procedures performed in studies involving human participants were in accordance with the ethical standards of the Ethics Committee of Fudan University School of Life Science and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

2.2 Data preprocessing

The frequency of the artery pulsation is 60–100Hz, which means that it can contain at least 2 periods for each 2 s. To reduce the computation complex, we selected 2 s (60 frames) from each recording as the video for model training. Each pixel range from 0 to 255 at recording, and we use zero-mean normalization (z -score) to normalize pixels at each video frame. In this part, both frame reducing and normalizing can help to reducing over-fitting during the training period and reduce computational redundancy. Here shows the function of z -score (1), \hat{x} is the mean value of input data, σ is the standard deviation of input data:

$$z - \text{score}(x) = \frac{x - \hat{x}}{\sigma} \quad (1)$$

2.3 Design and setting of the localization networks

To increase the location accuracy by adding temporal information, 3D convolution kernels are presented at the convolutional feature map to extract features from local

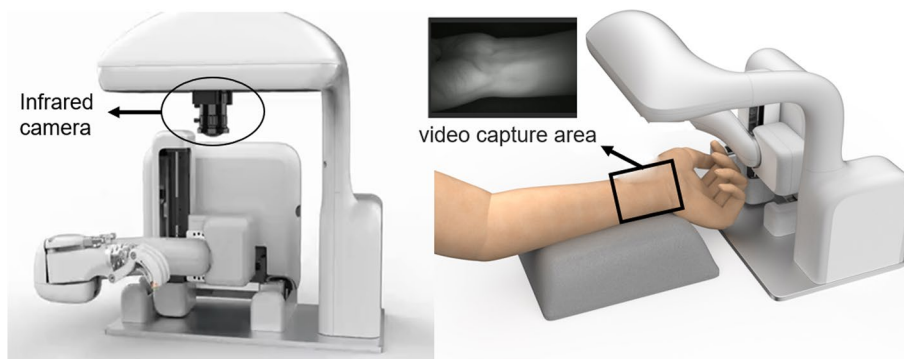


Fig. 1 The infrared camera with the video type it collects. Within the range of black elliptical wire-frame is the near-infrared camera used in this experimental acquisition. The video acquisition content is shown in the black rectangular wire-frame. During the experimental acquisition, the wrist is placed directly under the camera and kept as stable as possible

neighborhood. Therefore, the value of unit at each 2Dcnn position (x, y) in the j th feature map in the i th layer [26], denoted as the function of (2), is added with one more dimension, and the value of each 3Dcnn position (x, y, z) on the j th feature map in the i th layer is given by the function of (3). We employ the Mean Squared Error (MSE) loss for measurement of estimation differences (4):

$$v_{ij}^{xy} = \tan h \left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} \right) \tag{2}$$

$$v_{ij}^{xyz} = \tan h \left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)} \right) \tag{3}$$

$$\text{Loss(MSE)} = \sum_i^N (y_i - \hat{y}_i)^2 \tag{4}$$

$$f(x) = \max(0, x) \tag{5}$$

We construct the model with the main structure as follows: input layer, convolutional layers (3D-conv), pooling layer, fully connected layer, and output layer (Fig. 2). The proposed network contains 3 blocks and 2 fully connected layers, each block contains a 3D convolution layer and a pooling layer. As a kind of nonlinear down sampling, pooling layer can help model to lower dimension, removing redundant information, compressing

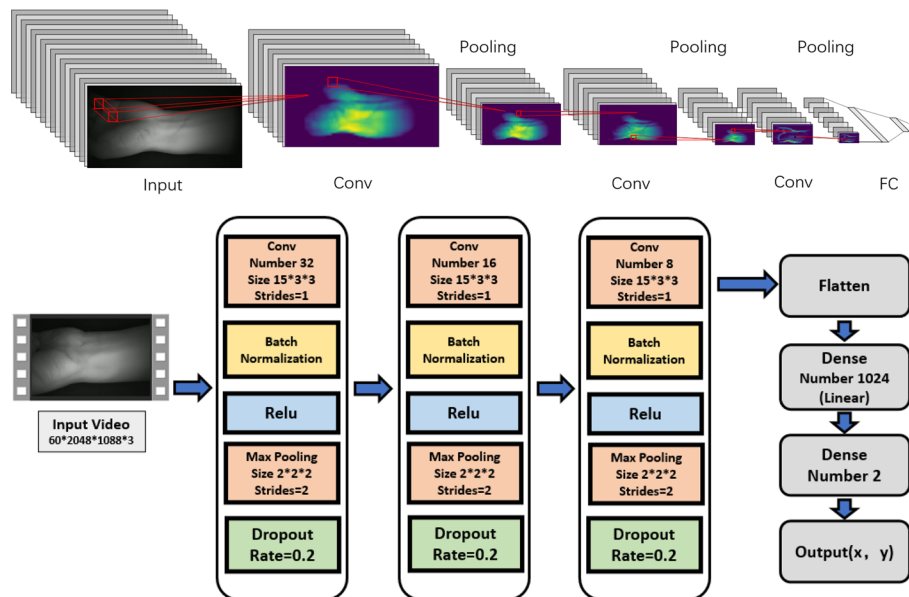


Fig. 2 The whole structure of this 3Dcnn network. The input of this method is video (60 × 2048 × 1088 × 3). Using 3D convolution kernel, a convolution neural network for video processing is constructed. The total number of layers of the network is three, and each layer contains the corresponding convolution layer and pooling layer. Finally, the full connection layer neural network is constructed as the output, and the final output is a two-dimensional array (coordinates)

features, simplifying network complexity. It greatly diminishes the number of parameters that need to be optimized.

For each layer, activation function is selected as Rectified Linear Units (ReLU) (5) to introduce nonlinear calculating and increase the expressive ability of model.

The rest part of 3Dcnn networks is Fully Connected layer (FC layer), whose activation function can be Relu normally. However, in our regression task, the information feature extraction is mainly concentrated in the convolution layer, so we hope to reduce the network redundancy and accelerate the network training by simplified fully connected layer. Compared with Relu, linear function can be faster and higher efficiency.

3D-Unet based on 3D convolution kernel is also introduced to compare to the proposed model. Through its unique U-shaped structure, the convolution results of each layer are sampled down, sampled up and connected. This comparative 3D-Unet model we built only borrowed from the U-shaped structure, so the convolution Kernels setting as same type as our proposed 3Dcnn model. The output is a 2-dimensional graph, mapped to a two-dimensional matrix with the same resolution as the input image. Then, we extract the corresponding positioning coordinates from the output image as the final resulted coordinates. This model does not include a fully connected layer.

Finally, the pixel distance is used as the evaluation index, and the proportion of video whose predicted distance label is less than the threshold is calculated as the evaluation index. At the same time, for fair comparison with other methods, when calculating the model accuracy, the spatial pixel resolution of all the tests is 1024×544 (7):

$$N_i = \{V_i, V_i[(y_i - \hat{y}_i)^2 \leq T^2]\} \quad (6)$$

$$\text{Accuracy} = \frac{|N_i|}{|N|} \times 100\% \quad (7)$$

y_i is the base truth of video V_i , $(\hat{y})_i$ is the prediction of model, where we set threshold T , if $(y_i - \hat{y}_i)^2 \leq T^2$, V_i will belong to N_i which is the set of the videos whose prediction is considered as accurate. N is the set of valid videos. Threshold T will be set to 50, 40, 30, 20.

3 Experiments

We have collected 500 videos from different people by infrared camera, each video was labeled by professional physicians with location of radial artery. Then 100 videos are chosen as valid set, the rest 400 video for training.

3.1 Ablation study

Performance of the 3Dcnn model is tested by adjusting different hyperparameters. First of all, we have tested the influence of the number of convolutional layers (Fig. 3). The result shows that best performance is shown when the layer number is selected as 3 (Table 1). We keep the layer number as 3, and adjust the number of filters by using the controlled variable method. Best performance and the accuracy are shown when filter is equal to 32 (Table 2). Then, we also control the layer number and the filter number, to test the influence of the size of convolution kernel filters. Results show that $15 \times 3 \times 3$

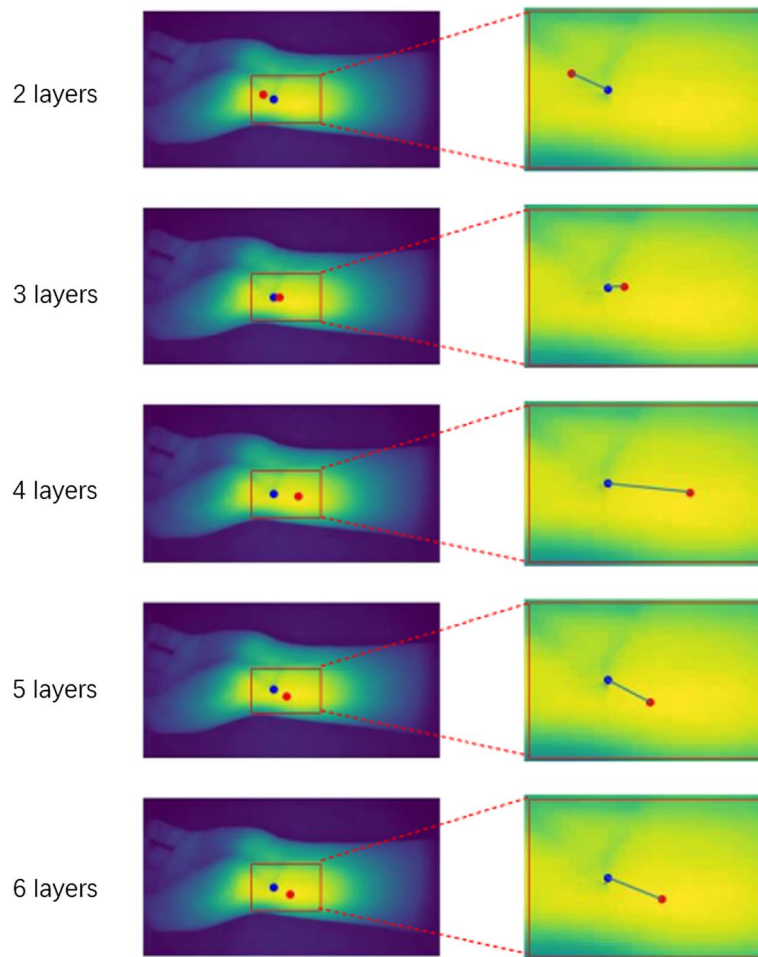


Fig. 3 The prediction performance from 3Dcnn model with different hidden layers. The figure on the left is the single frame schematic diagram of the collected wrist video, and the figure on the right is the enlarged picture of the red rectangular box on the left, blue point is the label pixel marked by experimenter, red point is prediction result

Table 1 Accuracy of 3Dcnn in different numbers of layers

Layers	Filters	50 pixel	40 pixel	30pixel	20 pixel
2	32	0.72	0.46	0.37	0.19
3	32	0.79	0.52	0.37	0.25
4	32	0.65	0.53	0.38	0.21
5	32	0.76	0.49	0.37	0.13
6	32	0.71	0.53	0.38	0.19

and with dropout method are best according our experiment with accuracy from 0.27 to 0.87 when pixel range from 20 pixel to 50 pixel (Table 3).

3.2 Compare with 2Dcnn

The extracted features from kernel convolution reflect information of the wrist radial artery. We visualised each layer processing results to recognize model effects. From left

Table 2 Accuracy of 3Dcnn in different numbers of filters

Layers	Filters	50 pixel	40 pixel	30pixel	20 pixel
3	256	0.76	0.51	0.35	0.20
3	128	0.75	0.49	0.35	0.19
3	64	0.63	0.47	0.37	0.18
3	32	0.79	0.52	0.37	0.25
3	16	0.71	0.53	0.38	0.21

Table 3 Accuracy of 3Dcnn in different kernel sizes

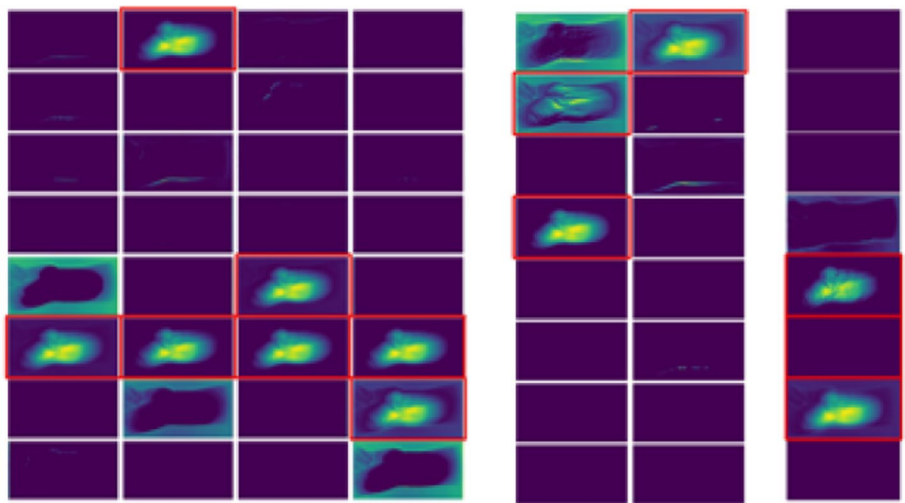
Kernel	50 pixel	40 pixel	30 pixel	20 pixel	Dropout
$60 \times 5 \times 5$	0.63	0.39	0.28	0.15	Yes
$60 \times 3 \times 3$	0.79	0.52	0.37	0.25	Yes
$60 \times 1 \times 1$	0.74	0.50	0.34	0.18	Yes
$30 \times 3 \times 3$	0.75	0.54	0.38	0.17	Yes
$15 \times 3 \times 3$	0.87	0.61	0.45	0.27	Yes
$15 \times 3 \times 3$	0.81	0.56	0.34	0.22	No
$8 \times 3 \times 3$	0.78	0.59	0.31	0.20	Yes

to right, they represent the output of first layer, second layer and third layer, 2Dcnn has already reflected much information learned from the edge, contour, texture of the wrist. However, results are blurry, with the edge overlapped with background and the contour indistinct. Overall, the information of skin texture can be well reflected, but doesn't make the wrist artery and other parts obviously differentiated. We also tagged the relevant filtering result using red rectangles, shown as Fig. 4a. With limited number of red rectangles, result shows that 2Dcnn can get limited useful information. For 3Dcnn with added temporal information, result shows that the rhythmic pulsation presents at radial artery at wrist in a large portion of the filtered results, shown as (Fig. 4b). Information with low correlation with frequency was discarded reasonably, and this result helps to validate that 3Dcnn can extract targeted rhythmic features better than 2Dcnn.

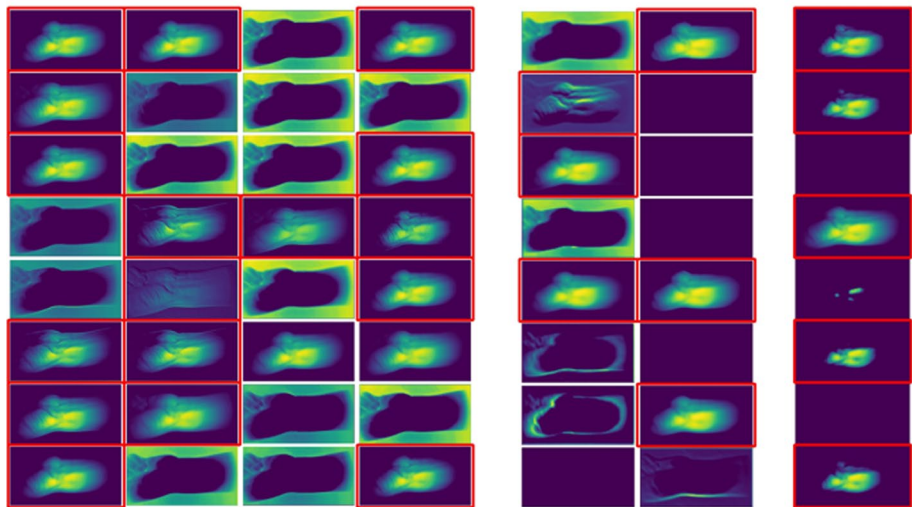
After we apply both 2Dcnn and 3Dcnn models to fit for radial artery location at the wrist. Here is one example is as shown (Fig. 5), it shows that the distance from the 3Dcnn prediction point to the target point is closer than the distance from the 2Dcnn prediction point to the target point. Then, we had compared the euclidean distance between models results and target labeled point. It can be clearly seen from the figure that as the accuracy required for positioning gradually increases, 3Dcnn model has obviously higher accuracy than the 2Dcnn model (Fig. 6).

3.3 Compare with other classical networks

Expect our earlier developed 2Dcnn model, we also compare the proposed 3Dcnn method with classical networks for object detection such as AlexNet, VGG16, VGG19 and 3D-Unet, here is the accuracy of them (shown as Fig. 7). The final results show that compared with these classical models, 3dcnn improves the accuracy from 7 to 19% at 50 pixels, and improves the accuracy from 7 to 20% at 20 pixels.



(a) The outputs of each convolutional layer from 2Dcnn



(b) The outputs of each convolutional layer from 3Dcnn

Fig. 4 From left to right, they represent the output of first layer, second layer and third layer. The light area demonstrates localised features that the model learned, and the red rectangles emphasize desired information detected at each convolutional layer. Through comparison, we can see that 3Dcnn can extract more key feature information than 2Dcnn

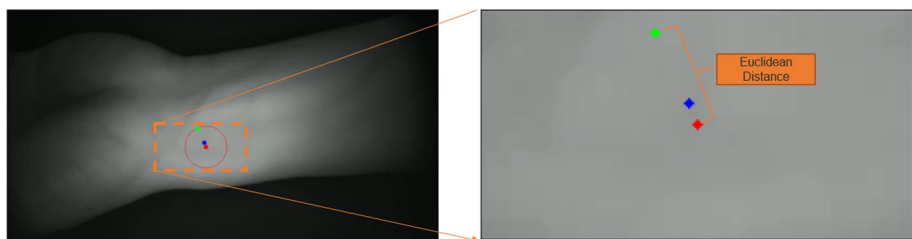


Fig. 5 The prediction performance between normal 2Dcnn model and our 3Dcnn model. The figure on the left is the single frame schematic diagram of the collected wrist video, and the figure on the right is the enlarged picture of the orange rectangular box on the left. Among them, Red point is the label pixel marked by experimenter, green point is prediction results from 2Dcnn, blue point is prediction results from 3Dcnn

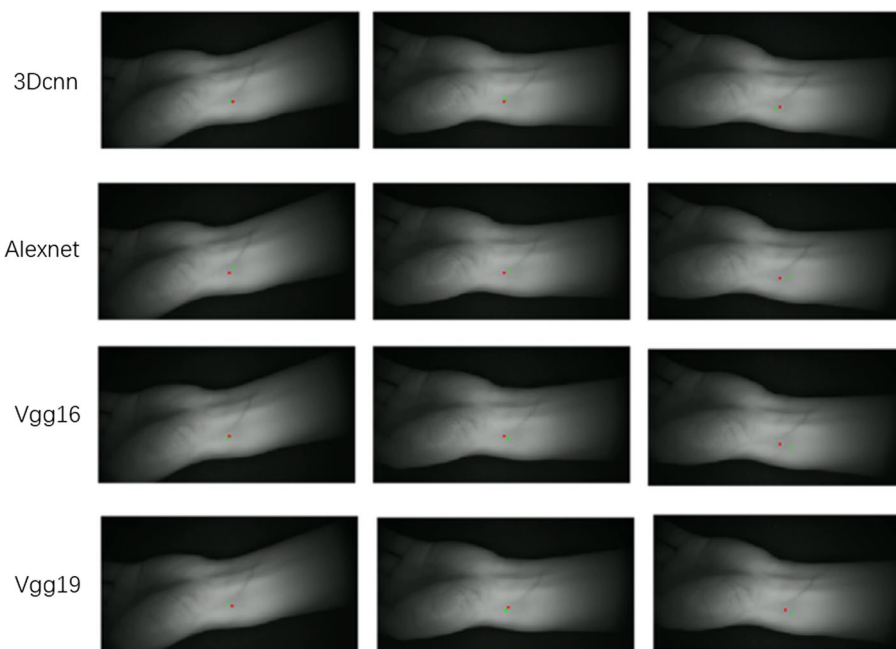


Fig. 6 The prediction performance of each 2D Classical Networks. In this figure, we compare the prediction effects of 3Dcnn with classical 2Dcnn models which mainly including Alexnet, Vgg16, Vgg19. Red point is the label pixel marked by experimenter, green point is prediction results from each model

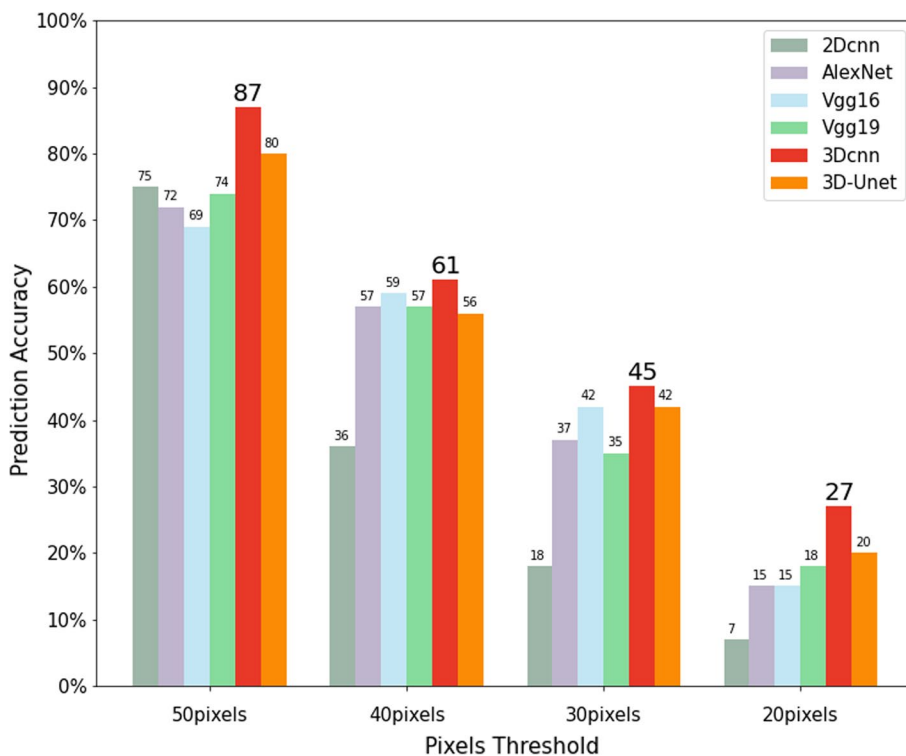


Fig. 7 Model accuracy comparison between 3Dcnn and others. In this figure, we compare the prediction accuracy of 3Dcnn with classical 2Dcnn models which mainly including normal 2Dcnn, Alexnet, Vgg16, Vgg19, and other 3D models mainly including 3D-Unet. The red column indicates 3Dcnn which shows the best performance than others, and the orange column indicates 3D-Unet which has same 3D convolution Kernels as 3Dcnn

AlexNet [27], Vgg16 [28], Vgg19 [29], which have 2 convolution kernel structure, are the classical 2D type networks. Due to the input of these networks model requires small size (224×224), we resize the images before feeding to networks for pulse localization. After the training, we expand the size proportionally to find detecting pixel locations. Results show that 3Dcnn model outperformed other models in all distance thresholds, as Table 4 showing the localization examples and (Fig. 6) showing the pulse localization results comparison. To valid our model has unique performance, we introduced 3D-Unet [30] that also has 3d convolution kernel structure to be compared. Results shows that 3D type model represented by both 3Dcnn and 3D-Unet have better performance than 2D type model. In the case of same 3D convolutional, 3Dcnn contains fully connected layers has higher accuracy in radial artery locating than 3D-Unet.

4 Discussion

4.1 The effect of location

Palpation localization of radial artery is the foundation of pulse diagnosis which is an indispensable part of the principle 4-methods of diagnosis in Traditional Chinese Medicine. In this paper, We assess the effectiveness of 3Dcnn on locating radial artery upon radius at the wrist by medical video data from infrared camera. The traditional 2Dcnn locating model focuses on analyzing a single picture [31], by extracting spatial feature from convolutional layers, 2Dcnn can learn relevant spatial information features, such as wrist skin texture [32]; therefore, for most standard groups, 2Dcnn can achieve nice localization results. However, from an anatomical point of view, the interlaced meridians of the body's wrists are complicated [33]. The relative position between the radial artery and the radius does not have a clear distance range in space, such as Oblique flying pulse [34]. In this special situation, the 2Dcnn, which can only learn spatial information, shows its original limitations [19]. To solve the problem of inaccurate positioning caused by insufficient information extraction in 2Dcnn, we increase the temporal and spatial information by introducing a 3D convolution kernel. The results show the comparison between the positioning prediction effect of 3Dcnn and 2Dcnn. To increase the credibility of the results, we also added four additional sets of comparative experiments, including three classic models, AlexNet, Vgg16, Vgg19, which have 2D convolution kernel structure and 3D-Unet that also has 3d convolution kernel structure.

The result shows that our proposed 3Dcnn model on task of locating with improvements of 20%, 12%, 12%, 9%, 7% on 20pixels than 2Dcnn, AlexNet, Vgg16, Vgg19, 3D-Unet, respectively. An improvement of 20% and 7% is achieved better than 2Dcnn and 3D-Unet is significant. Because, in addition to 3D convolution kernel, 2Dcnn and 3Dcnn are very similar in networks structure. Both 3D-Unet and 3Dcnn have 3D

Table 4 Comparing with other classical networks

Threshold	2Dcnn (%)	AlexNet (%)	Vgg16 (%)	Vgg19 (%)	3Dcnn (%)	3D-Unet (%)
50 pixels	75	72	69	74	87	80
40 pixels	36	57	59	57	61	56
30 pixels	18	37	42	35	45	42
20 pixels	7	15	15	18	27	20

convolution kernel, but 3D-Unet has more complex U-shaped network structure than 3Dcnn, thereby it becomes a challenging task to adjust the complexity of our model with in 3D convolution kernel. Even in this rare case, we find that 3Dcnn can provide more visually accurate results. Therefore, we believe that our proposed architecture can serve as a viable location tool for palpation localization of radial artery.

4.2 The limitation of this study and further work

There is limitation in our study in this work. First of all, there was only 50 young people with 500 videos are used to train and verify the accuracy of the system, all of them in the same age range at middle twenties. With the age increasing, the human collagen will be losing, the skin becomes loose and the pulsation of the wrist becomes emerge, so the radial artery position of the wrist of the elderly is more easily detected [35], so it is not considered. In addition, all this test sample has standard BMI, and with high BMI subject, the skin will be thicker [36], pulse beats information is getting difficult to be captured by camera, incomplete training samples will greatly reduce the universality of the model. Meanwhile, the detection accuracy of the system is affected by the accuracy of the infrared camera [37] itself which including photosensitive components, lens quality, color depth,. However, the higher the accuracy, the higher the price of the infrared camera itself.

In addition, on the algorithm we have just discussed the difference between 3Dcnn and 3D-Unet, the comparison model which contains 3D convolution kernel seems inadequate. At the same time, the 3D-Unet model we use is only a preliminary reference to the U-shaped structure, and we have not carried out in-depth exploration of feature optimization and algorithm details. Moreover, limited by the memory of the original computer, when we compared the effects of different neurons and neural network layers of 3Dcnn on the final result, we only performed a partial comparison. If you increase the size of the data set, perhaps a deeper network structure can build a better prediction model.

In the future work, We will expand the dimension of the data set, so that the training data of the model can cover more widely in terms of age, BMI and gender. At the same time, we will increase budget in order to adopt more precise and better quality near-infrared cameras to collect data. The last but not least, we will further optimize the positioning algorithm, on the basis of 3D structure, deep exploring the network structure of this model. In addition, other algorithms with the same 3D structure beside 3d-Unet Will be included in the comparative experiment, and the results will be reported promptly in the near future.

5 Conclusions

In this paper, a new approach of using 3Dcnn networks was explored to automatically and accurately find radial artery from wrist video. The core idea is to convolve the input image sequence by 3D convolution kernel. This proposed model makes full use of the abundant temporal information contained in radial artery pulsation. The ablation study shows that best performance at network structure of layer number as 3, the size of kernel as $15 \times 3 \times 3$. For different distance threshold setup, our model performs with accuracy as 0.87 at 50 pixels, accuracy as 0.61 at 40 pixels, accuracy as 0.45 at 30 pixels, accuracy

as 0.27 at 20 pixels. Meanwhile, this paper shows the advantages of 3Dcnn not only from the results, but also in the principle of filtering process. In addition, we showed that 3Dcnn can be easily scaled for the improved performance. In the future, we will continue this research work by collecting further labelled videos to improve and increase the model robustness and trying to fuse models to improve the accuracy and precision. By extending this research, the authors wish to achieve a valuable impact on development of Traditional Chinese Medicine.

Abbreviations

TCM	Traditional Chinese Medicine
2Dcnn	2-Dimensional convolutional neural networks
3Dcnn	3-Dimensional convolutional neural networks
CNN	Convolutional neural networks
FC layer	Fully connected layer
ReLU	Rectified linear units
MSE	Mean squared error
VGG16	Visual geometry group network 16
VGG19	Visual geometry group network 19
AlexNet	Alex ImageNet

Acknowledgements

Authors acknowledge technical support from our institutional research teams.

Author contributions

Conceptualization, QC and JL; methodology, YH and XZ; software, YH and JY; validation, QC, YH and ZJ; formal analysis, QC and YH; investigation, QC and JL; resources, QC and JL; data curation, ZJ, JY and XZ writing—original draft preparation, QC and YH; writing—review and editing, QC and JL; visualization, XZ; supervision, JL; project administration, XZ; funding acquisition, JL. All authors read and approved the final manuscript.

Funding

We thank the Key Area Support Plan of Guangdong Province for Jihua Laboratory (X190051TB190), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01), Special projects in key fields of Guangdong Universities (next generation of information technology) (2020ZDZX3003) for supporting this work.

Availability of data and materials

Some data can be shared on request.

Declarations

Competing interests

The authors declare no conflict of interest.

Received: 13 October 2021 Accepted: 10 June 2022

Published online: 18 July 2022

References

1. V. Rosemarie, An objective review of the technological developments for radial pulse diagnosis in traditional Chinese medicine. *Eur. J. Integr. Med.* **7**(4), 321–331 (2015)
2. M.M. Ko, M.S. Lee, S. Birch, J.A. Lee, The reliability and validity of instruments measuring pattern identification in Korean medicine: a systematic review. *Eur. J. Integr. Med.* **15**, 47–63 (2017)
3. X. Geng, S. Liu, Y. Zhang, J. Hou, H. Zhang, A noncontact method for locating radial artery above radial styloid process in thermal image. *Evid.-Based Complement. Altern. Med.* **2020**(2), 1–9 (2020)
4. H. Wang, Y. Cheng, A quantitative system for pulse diagnosis in traditional Chinese medicine. In: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (2006)
5. W. Nanyue, Y. Youhua, C. Yan, S. Zengyu, L. Tongda, C. Yanping, Research of length of cun, guan and chi in pulse taking. *China J. Tradit. Chin. Med. Pharm.* **24**, 1633–5 (2009)
6. L. Yan-Ping, H. Cen-Han, Traditional and modern research of 'taking pulse at cunkou alone' method. *China J. Tradit. Chin. Med. Pharm.* (2011)
7. P. Wang, W. Zuo, D. Zhang, A compound pressure signal acquisition system for multichannel wrist pulse signal analysis. *IEEE Trans. Instrum. Meas.* **63**(6), 1556–1565 (2014)
8. D. Wang, D. Zhang, G. Lu, A novel multichannel wrist pulse system with different sensor arrays. *IEEE Trans. Instrum. Meas.* **64**(7), 2020–2034 (2015)
9. Y.F. Chung, C.S. Hu, C.H. Luo, C.C. Yeh, X.C. Si, D.H. Feng, S.M. Yeh, C.H. Liang, Possibility of quantifying tcm finger-reading sensations: li. An example of health standardization. *Eur. J. Integr. Med.* **4**(3), e263-70 (2012)

10. C.S. Hu, Y.F. Chung, C.C. Yeh, C.H. Luo, Temporal and spatial properties of arterial pulsation measurement using pressure sensor array. *Evid.-Based Complement. Altern. Med.* **2012**(1741–427), 745127–745127 (2012)
11. J.H. Bae, Y.J. Jeon, J.Y. Kim, J.U. Kim, New assessment model of pulse depth based on sensor displacement in pulse diagnostic devices. *Evid.-Based Complement. Altern. Med.* **2013**(1), 938641 (2013)
12. R. Nevatia, Locating object boundaries in textured environments. *IEEE Transactions on Computers* C-25(11), 1170–1175 (1976)
13. C. Yan, Z. Li, Y. Zhang, Y. Liu, X. Ji, Y. Zhang, Depth image denoising using nuclear norm and learning graph model (2020)
14. H. Jiang, X.Y. Fan, Centroid locating for star image object by fpga. *Adv. Mater. Res.* **403–408**, 1379–1383 (2011)
15. S. Denman, M. Halstead, C. Fookes, S. Sridharan, Locating people in surveillance video using soft biometric traits (2017)
16. G. Ennio, A. Angela, B. Alberto, B. Laura, C. Enea, F. Sandro, P. Paola, R. Manola, S. Agnese, S. Susanna, Heart rate detection using microsoft kinect: validation and comparison to wearable devices. *Sensors* **17**(8), 1776 (2017)
17. L.Y. Fang, V. Christopher, W.P. Charles, P.N. Ray, I.J. Christian, F.P.A. Andrade, L.S. Choon-Sung, Noninvasive free flap monitoring using eulerian video magnification. *Case Rep. Otolaryngol.* **2016**, 1–4 (2016)
18. J. Adleberg, A.P. Ferster, D.A. Benito, R.T. Sataloff, Detection of muscle tension dysphonia using Eulerian video magnification: a pilot study. *J. Voice* (2019)
19. J. Luo, O. Chun, X. Nie, W. Yin, H. Lu, Y. Guo, Accurate targeting in robot-assisted tcm pulse diagnosis using adaptive sensor fusion. *Period. Eng. Nat. Sci.* **7**(1), 381–387 (2019)
20. B. Yang, K. Meng, H. Lu, X. Nie, G. Huang, J. Luo, X. Zhu, Pulse localization networks with infrared camera. In: *ACM International Conference on Multimedia in Asia* (2020)
21. A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks. In: *NIPS* (2012)
22. S.G. Armato, N.A. Petrick, R.L. Thomaz, P.C. Carneiro, A.C. Patrocínio, Feature extraction using convolutional neural network for classifying breast density in mammographic images **10134**, 101342 (2017)
23. K. Hara, H. Kataoka, Y. Satoh, Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet? In: *CVPR2018* (2018)
24. Y. Li, Y. Tian, B. Ge, Lung cancer classification using 3d-cnn with a scheduled learning strategy. In: *2018 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018)* (2018)
25. L. Chen, Z. Zhou, D. Sher, Z. Qiongwen, S. Jennifer, Combining many-objective radiomics and 3d convolutional neural network through evidential reasoning to predict lymph node metastasis in head and neck cancer. *Phys. Med. Biol.* (2019)
26. S. Ji, W. Xu, M. Yang, K. Yu, 3d convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intel.* **35**(1), 221–231 (2013)
27. Z.W. Yuan, J. Zhang, Feature extraction and image retrieval based on alexnet. In: *Eighth International Conference on Digital Image Processing (ICDIP 2016)* (2016)
28. B. Liu, X. Zhang, Z. Gao, C. Li, *Weld defect images classification with vgg16-based neural network* (Springer, Singapore, 2017)
29. B. Bermeitering, S. Donig, M. Christoforaki, A. Freitas, S. Handschuh, Vgg19 (2017)
30. Z. Xiao, B. Liu, L. Geng, F. Zhang, Y. Liu, Segmentation of lung nodules using improved 3d-unet neural network. *Symmetry* **12**(11), 1787 (2020)
31. J. Konstantinova, A. Jiang, K. Althoefer, P. Dasgupta, T. Nanayakkara, Implementation of tactile sensing for palpation in robot-assisted minimally invasive surgery: a review. *Sens. J. IEEE* **14**(8), 2490–2501 (2014)
32. J. Yu, B. Yang, J. Wang, J. Leader, J. Pu, 2d cnn versus 3d cnn for false-positive reduction in lung cancer screening. *J. Med. Imaging* **7**(5) (2020)
33. L. Terslev, S. Torp-Pedersen, N. Bang, M.J. Koenig, H. Bliddal, Doppler ultrasound findings in healthy wrists and finger joints before and after use of two different contrast agents. *Ann. Rheum. Dis.* **64**(6), 824–827 (2005)
34. M. Saint-Cyr, M. Mujadzic, C. Wong, D. Hatef, A.S. Lajoie, R.J. Rohrich, The radial artery pedicle perforator flap: vascular analysis and clinical implications. *Plastic Reconstr. Surg.* **125**(5), 1469–78 (2010)
35. A.J. Bailey, T.J. Sims, E.N. Ebbesen, J.P. Mansell, J.S. Thomsen, L. Mosekilde, Age-related changes in the biochemical properties of human cancellous bone collagen: relationship to bone strength. *Calcif. Tissue Int.* **65**(3), 203–210 (1999)
36. M. Falcone, M. Preto, M. Timpano, D. Ciclamini, P. Gontero, The surgical outcomes of radial artery forearm free-flap phalloplasty in transgender men: single-centre experience and systematic review of the current literature. *Int. J. Impotence Res.* 1–9 (2021)
37. E.K. Murphy, Use of an infrared camera to improve the outcome of facial nerve monitoring. *Am. J. Electroneurodiagnostic Technol.* **48**(1), 38–47 (2008)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.