A Method for Detecting Femur Fracture Based on SK-DenseNet

Yu Miao
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

Peng-Fei Zhao
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

Xiong-Feng Tang
Department of Orthopedics
The Second Hospital of Jilin University
Changchun, China

Yu-Qin Li
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

Li-Yuan Zhang
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

Wei-Li Shi*
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China
shiweili@cust.edu.cn

Ke Zhang
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

Hua-Min Yang*
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China
yhm@cust.edu.cn

Jian-Hua Liu
School of Computer Science and Technology
Changchun University of Science and Technology
Changchun, China

ABSTRACT

In clinical diagnosis, automatic fracture detection can reduce misdiagnosis which caused by fatigue and inexperience of radiologists, and provide support for reducing patient suffering and preventing disease progression. This paper proposes a SK-DenseNet model to detect femur fracture based on DenseNet network. It uses SK module to adjust the size of receptive fields adaptively, adopts Focal loss function to update parameter, and uses Grad-CAM method to visualize the femoral detection results. It can improve the interpretability of the fracture detection model. This paper used the femoral dataset to verify the performance of the mode. Our model achieves an accuracy of 0.9117, with the kappa coefficient is 0.8228. The experimental results show that the performance of the proposed model is better than the traditional deep learning model.

AIAM 2019, October 17–19, 2019, Dublin, Ireland

https://doi.org/10.1145/3358331.3358402

CCS CONCEPTS

•Computing methodologies ~ Object detection; Computing methodologies ~ Object identification; Computing methodologies ~ Interest point and salient region detections

KEYWORDS

Femoral fracture; Deep learning; CAD; SK-DenseNet

ACM Reference format:

Yu Miao, Peng-Fei Zhao, Xiong-Feng Tang, Yu-Qin Li, Li-Yuan Zhang, Wei-Li Shi, Ke Zhang, Hua-Min Yang, Jian-Hua Liu. 2019. A method for detecting femur fracture based on SK-DenseNet. In *AIAM' 19: 2019 International Conference on Artificial Intelligence and Advanced Manufacturing, October 17–19, 2019, Dublin, Ireland. ACM, New York, NY, USA, 7 pages.* https://doi.org/10.1145/3358331.3358402

1 INTRODUCTION

Femur fracture is the most common disease in clinical orthopedics and traumatology. It is usually caused by direct external force or the patient's own osteoporosis, which breaks the integrity or continuity of bone or trabecular. X-ray imaging technology is the firstly and main diagnostic method to evaluate patients with fractures. Due to the influence of scattered light and various noises, X-ray images generally have shortcomings such as blurred edges and low contrast, which seriously affect the quality of femur images [1] and increase the difficulty of diagnosis for doctors. A study by Krupinski [2] shows that when radiologists diagnose a large number of X-ray images in a day, they may easily misdiagnose due to excessive workload. Misdiagnosis led to patients' pain and high risk of disease deterioration, and seriously affecting patients' physical health. The technology of computer-aided diagnosis (CAD) aims to help doctors make more effective disease decision effective, thus reducing the burden of doctors. This has become an indispensable auxiliary means [3].

CAD based on machine learning and image processing technology [4] is a research hotspot in radiological diagnostics [5], and it has been widely used in the diagnosis of various diseases. In the early CAD system based on the medical image, the key features needed to be extracted manually in advance, besides, the difference between the relevance of different features and problems directly affected the performance of machine learning. Due to the limitation of image analysis algorithm, it is usually to mark the pathological region of image at the cost of many non-pathological regions, which makes it difficult to use the deep information hidden by high-dimensional features of image [6]. As an emerging technology in the field of machine learning, deep learning does not rely on middle and low-level visual features of images but obtains task-specific visual features through training a large amount of data, which can provide an accurate diagnosis for clinical applications. Deep learning has shown great application prospects in prediction of diabetic retinopathy [7], risk assessment of cardiovascular diseases [8], dermatological classification [9] and histopathological analysis [10].

Quiet amount of research has done by scholars on diagnosing fracture based on the deep learning method. Lindsey [6] extended the U-Net network model to realize the detection of wrist bone fracture, and the proposed model can be used to assist clinicians to improve the ability of wrist bone fracture detection. Rajpurkar [11] used the DenseNet network with a 169-layer to detect upper limb fractures, and finally could comparable to the radiologist performance in detecting on finger and wrist fractures. DenseNet model deepened the layer number of traditional convolutional neural network model, solved the problem of gradient disappearance in the process of back propagation by adjusting the architecture, and significantly improved the performance. However, the model was only improved from the depth and architecture of the model, but it was not considered from the level of the receptive field, so the spatial information of different sizes could not be obtained, and it was difficult to get a more comprehensive feature set, making the model less effective. However, SKNet network [13] could obtain multi-scale information of images through adaptive adjustment of the size of convolution kernel, and fuse this information to recalibrate the features, which could effectively make up for the shortcomings of DenseNet network.

Therefore, in view of the problem of automatic detection of femoral fracture, this paper combines the design idea of SKNet network on the basis of DenseNet, and proposes SK-DenseNet network to optimize the network from the spatial structure. Then, combined with the impact of sample proportion imbalance and hard-to-identify conditions in the sample on the model, the Focal loss function is used to update the network weight and enhance the robustness of the model. Finally, gradient-weighted Class Activation Mapping (Grad-CAM) is used to determine the location of the classification technology to achieve the visualization of model recognition features and enhance the interpretability of the model. The proposed model can assist general radiologists or non-radiologists to interpret femoral fracture images for patients, reduce the occurrence of misdiagnosis, and improve the diagnostic efficiency. it plays an important role in promoting the quality of diagnosis and examination of femurs.

2 METHOD

2.1 SK-DenseNet Model

In this paper, the new network model adds SK module behind the transition layer of the traditional DenseNet model, and forms the network structure of SK-DenseNet, as shown in Figure 1. The network structure is mainly composed of four Denseblock, three transition layers and three SK modules. Through this design, Denseblock, a dense connection mode in DenseNet network, is retained. The features of the preceding convolution are directly used as the input of the subsequent convolution, which ensures the maximum information flow among the layers and avoids over-fitting. At the same time, the spatial dimension information of the transition layer features can be obtained by adjusting the size of the receptive field convolution kernel, and the feature recalibration can be realized by fusing information.

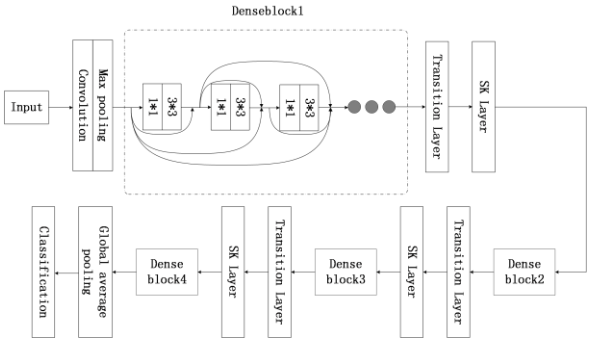

Figure 1: SK-DenseNet network model

Among them, SK module is the embodiment of the core idea of SKNet network. By adjusting the size of convolution kernel, the network can obtain the information of different receptive fields and realize the re-calibration of features, and its structure is shown in Figure 2.

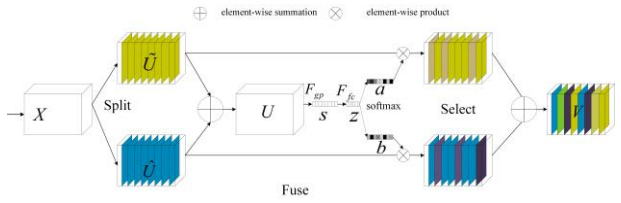

Figure 2: SK module structure

SK module is mainly composed of Split, Fuse and Select. In the Figure 2, X represents the input feature, and the Split operation convolute feature X using two different size convolution kernels to get two different output features $\tilde{U}$ and $\hat{U}$. In the Fuse operation, the output of different branches is first combined by adding elements by element, and the formula is expressed as follows:

$$U = \tilde{U} + \hat{U} \tag{1}$$

Secondly, the global pooling is used to obtain the global information of each layer of the feature channel after fusion, and the formula is expressed as follows:

$$s_c = F_{gp}(U_c) = \frac{1}{P}\sum_i \sum_j U_c(i,j) \tag{2}$$

where P represents the number of image pixels, C represents the number of layers, U_c represents the C layer features after the fusion of different branch output features, and s_c represents the C layer features obtained after global pooling. Then the global information is compressed by the full connection layer, and a compression feature is obtained, which is expressed as follows:

$$z = F_{fc}(s) = \delta\left(\beta(Ws)\right) \tag{3}$$

where δ is the ReLU function, β denotes batch Normalization, $W \in R^{d \times C}$, z represents compression feature obtained after full connection. In the compression process, a compression ratio r is used to control the dimension of the output, and the formula is expressed as follows:

$$d = \max(C/r, L) \tag{4}$$

where d represents the feature dimension after the full connection, and L denotes the minimal value of d, $L = 32$ is a typical setting in our experiments; The Select operation is to get the weights corresponding to different convolution branches through a Softmax activation function, and then get the spatial features through the weighted sum of corresponding convolution features. The formula is expressed as follows:

$$V_c = a_c \cdot \tilde{U}_c + b_c \cdot \hat{U}_c, a_c + b_c = 1 \tag{5}$$

where V_c denotes the C layer feature obtained after feature weight calibration, while a_c and b_c denote the C layer weights corresponding to $\tilde{U}_c$ and $\hat{U}_c$ after the softmax activation function.

In this paper, there is also a comparison network SE-DenseNet, which improves the network from the aspect of the feature dimension. The network draws lessons from the design idea of flight delay prediction [14], and applies this design idea to DenseNet169 network model, which the 3*3 convolution of Denseblock is followed by the SE unit of the core of the SENet network [15]. The specific structure is shown in Figure 3:

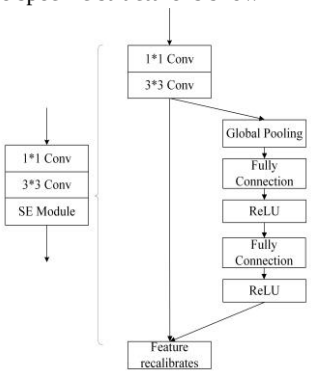

Figure 3: SE module added in Denseblock

Firstly, the feature information after 3*3 convolution is transferred to the global pooling layer, and the global information corresponding to the feature is obtained after the global pooling layer. Then, the relationship between the feature channels is obtained after the full connection layer, ReLU activation function, full connection layer and Sigmoid activation function. The process is the same as that of Fuse operation in SK module. Finally, the feature weights obtained are weighted with the feature information obtained after 3*3 convolution, and the feature re-calibration is realized. This operation corresponds to Select in SK module.

Compared with the SE-DenseNet network, SK-DenseNet can obtain multi-scale information by adjusting the size of convolution core. On the other hand, it also considers the relationship between feature channels after information fusion, which improves the generalization ability of the model.

2.2 Focal loss

The Focal loss function [16] is used to update the parameters in back propagation, and compared with the traditional cross-entropy loss function in experiments. In the process of model training, Focal loss function can dynamically assign different weights to samples to make them pay more attention to those samples which are difficult to identify. The form of Focal loss function is as follows:

$$\begin{aligned} L_{fl} = &-\alpha y(1-y')^{\gamma} \log y' \\ &-(1-\alpha)(1-y)\, y'^{\gamma} \log(1-y') \end{aligned} \tag{6}$$

where y and y' represent the label and prediction label of the real sample, respectively. γ is weight index, which is used to reduce the loss of easy to classify samples and make the model focus more on difficult to classify samples in training. α is the

equilibrium coefficient, which is used to balance the uneven proportion of positive and negative samples. Cross entropy, as a common loss function in image classification, is expressed as follows:

$$L_{ce} = -y \log y' - (1-y)\log(1-y') \tag{7}$$

2.3 Grad-CAM

The Grad-CAM technology [17] is used to realize the visualization of model identification features. This technology can help researchers to understand the characteristics of femoral fractures identified by the deep network model and plays a guiding role in improving the interpretability of the model. Grad-CAM is similar to the basic idea of Class Activation Mapping (CAM) [18]. It calculates the weight corresponding to each feature map and obtains the final heat map by weighted summation. But Grad-CAM uses the global average gradient to calculate the weight, which solves the problem that the CAM technology needs to change the model framework or retrain in the process of implementation. In order to obtain the gradient weight corresponding to each feature map, the weight of the corresponding category c of the $k\text{-}th$ feature map in Grad-CAM is defined as follows:

$$\alpha_k^c = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{8}$$

where Z denotes the number of the pixels in the feature map, y^c denotes the gradient of the score in category c, and $A_{i,j}^k$ denotes the pixel value at the position (i, j) in the $k\text{-}th$ feature map. Then use the obtained feature map to perform weighted summation, and the final heat map is:

$$L_{Grad-CAM}^c = \mathrm{Re}\,LU\left(\sum_k \alpha_k^c A^k\right) \tag{9}$$

Among them, the application of ReLU activation function to the linear combination of the final weighted sum is to reflect the features that only focus on the positive impact on the target category in the visualization process.

3 Experiments and analysis

3.1 Dataset

The experimental data set in this paper is from the radiology department of the second hospital of Jilin university, with a total of 1283 X-ray images of the femoral, including 593 normal femur images, 690 femoral fractures (including 110 femoral neck fracture images, 213 intertrochanteric fracture images, and 367 femoral shaft fracture images).

The X-ray of femur has the characteristics of high noise, high overlap. In order to improve the image quality and highlight the image details, and effectively suppress the noise, the combination of histogram equalization in the spatial domain and frequency domain high-frequency enhancement filtering is used [19]. Figure 4(a) is the original femur image, and Figure 4(b) is the enhanced image. It can be observed that the influence of the overlap on the skeletal structure in the enhanced image is reduced, and the position of the fracture line is clearer. Meanwhile, in order to ensure the aspect ratio of the original image, the background color is used to expand the image to both sides to reach the same size, and then it is normalized to the uniform size (512*512). Figure 4(c) is the normalized result image.

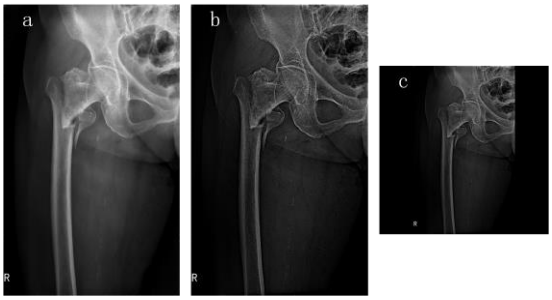

Figure 4: Image preprocessing, (a) is the original femur image, Figure (b) is the enhanced image, and (c) is the normalized size image

Considering that the limited data samples, in order to obtain as much effective information as possible from the limited learning data, and make the accuracy of the model more convincing, this paper uses 10 times 10-fold cross-validation method to evaluate the performance of the model. In each cross-validation, the data are divided into 10 parts, 9 of which are selected as training data, and the remaining one is used as test data. Finally, the average value is used as the final evaluation criterion of the model.

3.2 Evaluation of the Model

In machine learning, the final accuracy (ACC) of the model on the test dataset is used to verify the fitting of the model. However, in clinical diagnosis, it is inevitable that misdiagnosis will occur. In order to accurately grasp the condition of the model in the diagnosis process, the sensitivity and specificity of the model in the test group will also be recorded. Sensitivity, also known as true positive rate, it is a probability which sick persons are diagnosed correctly; specificity is also called true negative rate, it is a probability which normal person are diagnosed correctly. In addition, the ability of the training model to detect the presence of fractures in the femur X-ray was evaluated on the test data by calculating the Receiver Operating Characteristic Curve (ROC) and calculating the area under the curve (AUC). where the AUC is 1.0, it means that the training model can perfectly predict the reference standard, and when the AUC is 0.5, it means that the model is no better than chance. Since there may be no significant difference between the two models in AUC, in order to prevent this kind of situation, it is necessary to further consider the kappa coefficient [20]. Kappa coefficient is an ideal index to describe the consistency of diagnosis, so it has been widely used in clinical trials. The expression of the Kappa coefficient is as follows:

$$k = \frac{p_0 - p_c}{1 - p_c} \tag{10}$$

where p_0 denotes the proportion of the number of correctly classified samples to the total number of samples, p_c denotes the theoretical consistency rate. Assuming that the number of real samples for each class is $t_1, t_2, \cdots, t_e$, and the number of predicted samples for each class is $f_1, f_2, \cdots, f_e$, and the total number of samples is n , then:

$$p_c = \frac{t_1 \times f_1 + t_2 \times f_2 + \cdots + t_e \times f_e}{n \times n} \tag{11}$$

3.3 Training of the Model

The femoral part can usually be divided into four categories according to the fracture area: femoral neck, intertrochanteric, femoral shaft and normal. However, the difficulty of diagnosis is also different in different femoral regions, so no one has unified the diagnosis in the femoral region. Therefore, in this experiment, the deep learning method is used to make a unified determination of the femur region. Since it is only to determine whether the fracture is caused, the task of the model is defined as a simultaneous binary classification and conditional semantic segmentation problem. In this way, a single probability value will be obtained, indicating the possibility of fracture. When the model makes sure that there is a fracture in the input X-ray image, the heat map will be displayed to the clinician. In the heat map, each pixel location represents the confidence that the corresponding pixel location in the input radiograph is part of a given fracture.

In the training process, the data are enhanced by horizontal flip and random rotation of 30%. In order to improve the generalization ability of the model and reduce the training time of the network, the Adam [21] algorithm is used to optimize the network parameters. The number of iteration steps is determined to be 80, the batch size is 32, the dropout is set to 0.2, and the learning rate is initialized to 0.001. The learning rate does not change in the first 40 steps and decreases exponentially in the last 40 steps, as shown in Figure 5.

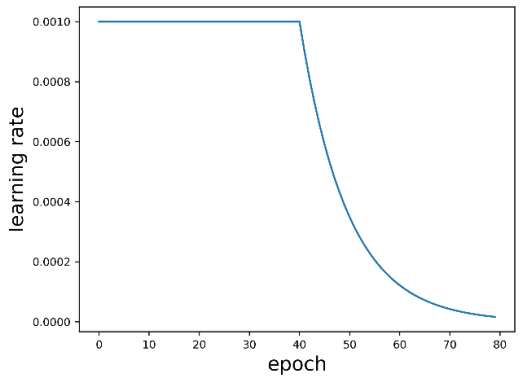

Figure 5: Learning rate curve

3.4 Result

The server used in the experiment was based on the 64-bit Ubuntu16.04 operating system, the GPU accelerated graphics card was NVIDIA Tesla V100, and the python3.6 environment built the deep learning framework Pytorch. By learning and training the model proposed in this paper (SK-DenseNet, SK-DenseNet + Focal), different evaluation standard values are calculated in the test data set, as shown in table 1:

Table 1: The performance of DenseNet and its improved model on the different evaluation standard on the test dataset

	Densenet169	SE-DenseNet
Acc (95%, CI)	0.9006(0.8953,0.9059)	0.8856(0.8799,0.8913)
Sensitivity (95%, CI)	0.8910(0.8838,0.8981)	0.8823(0.8744,0.8900)
Specificity (95%, CI)	0.9116(0.9030,0.9201)	0.8895(0.8799,0.8991)
AUC (95%, CI)	0.9624(0.9593,0.9654)	0.9556(0.9522,0.9590)
Kappa (95%, CI)	0.8005(0.7898,0.8111)	0.7704(0.7589,0.7817)
	SK-DenseNet	SK-DenseNet + Focal
Acc (95%, CI)	0.8983(0.8932,0.9034)	0.9117(0.9072,0.9163)
Sensitivity (95%, CI)	0.8891(0.8820,0.8962)	0.9063(0.8988,0.9138)
Specificity (95%, CI)	0.9090(0.8997,0.9183)	0.9180(0.9106,0.9254)
AUC (95%, CI)	0.9619(0.9587,0.9650)	0.9697(0.9673,0.9720)
Kappa (95%, CI)	0.7959(0.7857,0.8062)	0.8228(0.8136,0.8320)

Table 1 shows the performance of DenseNet and its improved model in Acc、Sensitivity、Specificity、AUC and Kappa five evaluation standards. The proposed SK-DenseNet model is superior to SE-DenseNet in overall performance, and its effect is comparable to that of DenseNet model. After replacing the loss function on the basis of the SK-DenseNet model, the overall effect is significantly improved. The ACC of the final model was 0.9117, and the kappa coefficient was 0.8228. The experimental results show that the model is very consistent with the orthopedic surgeon who created the dataset. Figure 6. shows the comparison of AUC on the test dataset for the four network models at the first ten-fold cross-validation.

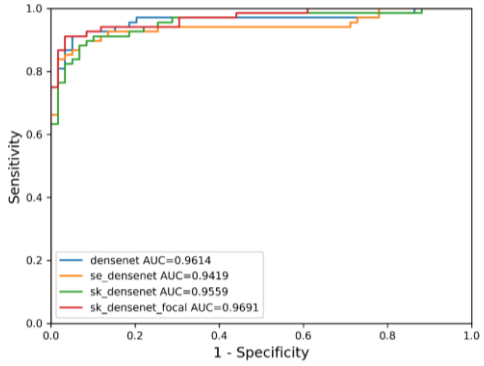

Figure 6: Roc crave on diffidence model

Figure 7 shows some femur images and their corresponding heat maps. It can be seen from the figure that for different femoral fracture images in Figure 7(a-d), the proposed model can accurately locate the fracture area. For normal femoral images 7(e), the model can also be judged as normal. However, for the normal femoral image 7(f), the model is incorrectly interpreted, and some

information that was not the femoral area was incorrectly used as a standard for fracture.

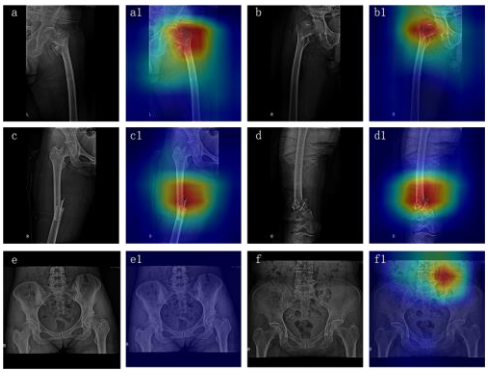

Figure 7: Femur image (a-f) and its corresponding heat map (a1-f1), where (a) is a femoral neck fracture, (b) is an intertrochanteric fracture, (c) and (d) are femoral shaft fractures, (e) is a normal femur image, and (f) is a normal femur image that is judged to be abnormal.

In recent years, researchers have made corresponding contributions to the detection of fractures in different parts of the femur [22-23], and each study has achieved good results. Considering that the dataset used for detection in different studies are different, in this paper, we apply those models which has used for femoral fracture detection into the supplied dataset for unified detection. It can be observed from table 2 that the model they choose does not perform as well as the model in this paper on the femoral dataset, indicating that the network in this paper is more effective in the unified detection of the femur.

Table 2: Performance comparison results of different depth network models on femur data sets

Dataset	Model	Accuracy	AUC
Femoral dataset	VGG16 [22]	0.895875	0.956983
	GoogleNet [23]	0.894254	0.959028
	SK-DenseNet + Focal	0.9117	0.9697

3.5 Discussion

In this paper, the SK module is added on the basis of DenseNet network, which can optimize the network from the aspect of adjusting the receptive field size adaptively. SK module reflects the idea of SE in practical operation, SK module consider the weight of multi-channel convolution besides channel's weight. It can also be seen from the experimental results that the effect of SK-DenseNet network is significantly improved compared to the SE-DenseNet network, but it is not greatly improved compared with DenseNet network. the core idea of the two modules is that learning the feature weight through the loss function, but SK adds dynamic convolution in the design, so the performance of the network adding SK is better than the network that added SE. In general, the traditional cross entropy loss function does not guarantee that the network after adding module can achieve best result. In order to solve this problem and ensure that the model can train difficult samples better, the original cross entropy loss function is changed to the Focal loss function, and the evaluation index of the final model is higher than other representative methods.

Previous researchers [22,23] needed to cut the interested region before detecting the skeletal X-ray images, and then put it into the network for training, so that the model can be avoided from being affected by other noise areas, and focus attention on the target visual features to be identified, and so as to improve the accuracy of the network model. General, this method is not suitable for the doctor's clinical diagnosis. In clinical, doctors need take all the information about full image into consideration. In this paper uses the entire X-ray image. By observing the heat map, it can be found that the input of the whole X-ray image will increase the recognition area of the model to a certain extent, thereby reducing the accuracy of the model.

The method in this paper also has some limitations. First, this study is a retrospective study based on femoral fractures. Although it has been confirmed in the experiment that deep learning has a consistent effect on the detection of the entire femoral region and has achieved good results in the detection process, it has not been applied to clinical practice. Second, the diagnosis of the model is based on the visible contents of X-ray images, while a more effective clinical study should be based on the overall situation of patients. Last, datasets play a crucial role in deep learning, and only a large amount of accurate data can train better models. Due to the lack of sample size, it is unclear whether this model can achieve similar accuracy in more pathological analysis of femur fractures, which requires larger data samples for verification. Future research will focus on more extensive clinical data validation, abnormal detection of fractures and classification of different fracture types.

4 CONCLUSION

This paper proposes SK-DenseNet based on SKNet and DenseNet network, has applied to increasing the detection accuracy of femoral fractures. At the same step size, the experimental results showed that the strengthened network is more effective in the detection of femoral fracture. At the same time, in order to determine the credibility of the model, this paper uses Grad-CAM classification discrimination and positioning technology to generate heat map, and carries out visual operation on the model identification features. Due to the small number of data sets, although the data is enhanced through data enhancement, the advantages of deep learning cannot be fully played, it has not been applied in clinical applications. However, with the improvement of technology, CAD will play an important role in fracture diagnosis.

ACKNOWLEDGMENTS

We would like to acknowledge the data support from the Second Hospital of Jilin University. The experimental data obscured the patient's personal privacy. After the review and approval of the ethics committee of the second hospital of Jilin university, intelligent assisted diagnosis of acute bone and joint loss based on convolutional neural network was conducted. This work is supported by the Science & Technology Development Program of

Jilin Province, China (Nos. 20170204031GX, 20180201037SF, 20190201196JC and 20190302112GX), the National Key Research and Development Program of China under Grant No. 2017YFC0108303 and the Science & Technology Development Program of Jilin Province, China (No. 2018C039-1).

REFERENCES

[1] Li Ce, Zhou Yannan, Ouyang Chengsu (2014). X-ray medical image enhancement algorithm based on hierarchical fuzzy membership. Computer Applications and Software, 31(8), 209-212.

[2] Krupinski E A, Berbaum K S, Caldwell R T, et al (2010). Long radiology workdays reduce detection and accommodation accuracy. Journal of the American College of Radiology, 7(9), 698-704.

[3] Zheng Guangyuan, Liu Xiabi, Han Guanghui (2018). Survey on medical image computer-aided detection and diagnosis systems. Journal of Software, 29(5), 1471-1514.

[4] Chen Shihui, Liu Weixiang, Qin Jing, et al (2017). Research progress of computer-aided diagnosis in cancer based on deep learning and medical imaging. Journal of Biomedical Engineering, 34(2), 314-319.

[5] Doi K (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Computerized Medical Imaging and Graphics, 31(4), 198-211.

[6] Lindsey R, Daluiski A, Chopra S, et al (2018). Deep neural network improves fracture detection by clinicians. Proceedings of the National Academy of Sciences, 115(45), 11591-11596.

[7] Gulshan V, Peng L, Coram M, et al (2016). Development and validation of a deep Learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402-2410.

[8] Poplin, R. et al (2018). Webster. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nature Biomedical Engineering, 2, 158–164.

[9] Esteva A, Kuprel B, Novoa R A, et al (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[10] Sirinukunwattana K, Raza S, Tsang Y W, et al (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Transactions on Medical Imaging, 35(5), 1196-1206.

[11] Pranav Rajpurkar, Jeremy Irvin, Aarti Bagul, et al (2018). MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs. The Conference on Medical Imaging with Deep Learning 2018, arXiv:1712.06957v4.

[12] Huang G, Liu Z, Laurens V D M, et al (2017). Densely Connected Convolutional Networks. The IEEE Conference on Computer Vision and Pattern Recognition, 4700-4708.

[13] Li X, Wang W, Hu X, et al (2019). Selective Kernel Networks. The IEEE Conference on Computer Vision and Pattern Recognition 2019, arXiv:1903.06586v2.

[14] Renbiao WU, Ting ZHAO, Jingyi QU (2019). Flight delay prediction model based on deep SE-DenseNet. Journal of Electronics and Information Technology, 40(6), 1510-1517.

[15] Hu J, Shen L, Albanie S, et al (2018). Squeeze-and-Excitation Networks. The IEEE Conference on Computer Vision and Pattern Recognition, 7132-7141.

[16] Lin T Y, Goyal P, Girshick R, et al (2017). Focal Loss for Dense Object Detection. The IEEE International Conference on Computer Vision, 2999-3007.

[17] Selvaraju R R, Cogswell M, Das A, et al (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. The IEEE International Conference on Computer Vision, 618-626.

[18] Zhou B, Khosla A, Lapedriza A, et al (2016). Learning Deep Features for Discriminative Localization. The IEEE Conference on Computer Vision and Pattern Recognition, 2921-2929.

[19] Xu Weijun, Liu Guozhong (2009). Space domain and frequency domain combination of image enhancement technology and its realization. China Measurement and Test, 35(4), 52-54.

[20] Sim J, Wright C C (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy, 85(3), 257-268.

[21] Kingma D P, Ba J (2015). Adam: A Method for Stochastic Optimization. Proceedings of the International Conference for Learning Representations 2015. arXiv:1412.6980v8.

[22] Takaaki U, Yuki T, Shinichi G, et al (2018). Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiology, 48(2), 239-244.

[23] Adams M, Chen W, Holcdorf D, et al (2018). Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures. Journal of Medical Imaging and Radiation Oncology, 63(1), 27-32.