Computer Methods and Programs in Biomedicine 171 (2019) 27–37

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images

Yoga Dwi Pranata[a], Kuan-Chung Wang[a], Jia-Ching Wang[a,*], Irwansyah Idram[b], Jiing-Yih Lai[b], Jia-Wei Liu[c], I-Hui Hsieh[c,*]

[a] Department of Computer Science and Information Engineering, National Central University, Jhongli County, Taoyuan City, Taiwan
[b] Department of Mechanical Engineering, National Central University, Jhongli County, Taoyuan City, Taiwan
[c] Institute of Cognitive Neuroscience, National Central University, Jhongli County, Taoyuan City, Taiwan

ARTICLE INFO

Article history:
Received 11 October 2018
Revised 29 January 2019
Accepted 11 February 2019

Keywords:
Convolutional neural networks
Calcaneus fracture
Computed tomography image
Residual network
Visual geometry group

ABSTRACT

Background and objectives: The calcaneus is the most fracture-prone tarsal bone and injuries to the surrounding tissue are some of the most difficult to treat. Currently there is a lack of consensus on treatment or interpretation of computed tomography (CT) images for calcaneus fractures. This study proposes a novel computer-assisted method for automated classification and detection of fracture locations in calcaneus CT images using a deep learning algorithm.

Methods: Two types of Convolutional Neural Network (CNN) architectures with different network depths, a Residual network (ResNet) and a Visual geometry group (VGG), were evaluated and compared for the classification performance of CT scans into fracture and non-fracture categories based on coronal, sagittal, and transverse views. The bone fracture detection algorithm incorporated fracture area matching using the speeded-up robust features (SURF) method, Canny edge detection, and contour tracing.

Results: Results showed that ResNet was comparable in accuracy (98%) to the VGG network for bone fracture classification but achieved better performance for involving a deeper neural network architecture. ResNet classification results were used as the input for detecting the location and type of bone fracture using SURF algorithm.

Conclusions: Results from real patient fracture data sets demonstrate the feasibility using deep CNN and SURF for computer-aided classification and detection of the location of calcaneus fractures in CT images.

1. Introduction

The calcaneus, also known as the heel bone, is the largest tarsal bone forming the rear part of the foot. While it is designed to withstand great force, the calcaneus is the most frequently fractured tarsal bone, representing ~2% of all fractures and 60% of tarsal bones fractures [1]. The primary cause of these fractures, which most commonly occur in males 30–50 years of age, is traumatic axial loading caused by falling, twisting, or motor vehicle accidents. Calcaneal fractures can be categorized into two types, on the basis of subtalar joint involvement: intra-articular and extra-articular fractures. Intra-articular fractures are more common as the posterior talar articular facet includes the calcaneus. Calcaneal injuries are often associated with a poor prognosis, with 40% of cases resulting in complications such as an unhealed wound, infection, or posttraumatic arthritis which can be exacerbated by pre-existing patient conditions such as diabetes or multiple open injuries [1]. In addition to pain and impaired mobility, inadequate and inaccurate initial treatment of calcaneal fractures can lead to serious long-term problems including neurovascular dysfunction, compartment syndromes, and even amputation [2].

Several new advances have been made in cross-sectional medical imaging in recent years [3], particularly in computed tomography (CT), which have assisted physicians in identifying and characterizing calcaneal fractures. Unlike conventional x-rays, tomographic slices consist of a sequence of thin-section 2-D images which provide voxels that are essentially isotropic, allowing for high-resolution multi-planar 3-D reconstructions. As such, CT images include significantly more information than conventional radiographs. In addition, CT images have been shown to be superior to routine radiographs for evaluating the complex structures of the axial skeleton including the osseous structures of the foot and ankle [4]. For instance, Grigoryan et al. (2003) showed that CT images allowed for more complete and detailed visualization for

* Correspondence to: Jia-Ching Wang, Department of Computer Science and Information Engineering; I-Hui Hsieh, Institute of Cognitive Neuroscience, National Central University, No.300, Jhongda Rd., Jhongli County, Taoyuan City 32001, Taiwan.
E-mail addresses: jcw@csie.ncu.edu.tw (J.-C. Wang), ihsieh@cc.ncu.edu.tw (I.-H. Hsieh).

https://doi.org/10.1016/j.cmpb.2019.02.006
0169-2607/

evaluating radial and tibial shaft fractures during healing compared with conventional X-rays [5]. Due to the significant amount of information contained in CT images, details such as skeletal structure, organ boundaries, and soft tissue properties may not be accurately assessed via visual inspection, depending on the experience of the radiologist. Manual detection of bone fractures in CT is also challenging due to low image resolution, the partial volume effect, and the potential presence of artifacts and noise in CT images [6]. The complex anatomy and corresponding soft tissue structure of the calcaneus further complicates this process, making it difficult to diagnose and treat. There is currently a lack of consensus on the use of operative or conservative fracture management plans, due in part to the lack of a standardized calcaneus fracture classification system [7]. In addition, the interpretation of CT images for diagnosing calcaneus fractures is difficult and studies have shown low intra-observer agreement and low intra-rater reliability in evaluating such scans. As such, a computer-assisted fracture detection methodology could be highly beneficial for assisting physicians in the timely and accurate diagnosis and treatment of these injuries [2].

While fracture detection in calcaneus CT images is an emerging area of research, several attempts have been made to detect bone fractures in other parts of the body using X-ray or CT images. For example, Anu et al. focused on bone fracture detection in X-ray images using gray-level co-occurrence matrices (GLCMs) [8]. However, the resulting classification using a decision tree or neural network for distinguishing fractures from normal bone was only ~50% accurate, and in some cases the area of the fracture could not be determined. Similarly, Kurniawan et al. proposed using Canny edge detection to identify fractured bones in X-ray images [9]. However, the accuracy of this method was limited by the quality of the images, and the algorithm required significant processing time. Preliminary efforts in automated CT detection were also conducted by Sharma et al. who detected the edges of lungs using morphological algorithms [10]. Wu et al. proposed the registered active shape method (RASM) for segmentation, combined with an adaptive window and stationary wavelet transform, to detect contour discontinuities indicative of potential traumatic fractures [6]. However, this dataset was limited to pelvic CT scans.

Convolutional neural networks (CNNs) are a type of deep learning algorithm which produce the most satisfying results for visual image recognition [11,12]. CNNs are multi-layer neural networks with the ability to imitate image recognition systems in the human visual cortex. CNNs belong to a class of multilayer perceptrons (MLPs), with the intended purpose of processing two-dimensional data, and have been widely applied in image processing. The original MLPs made of fully-connected layers are less suitable for image classification as they are unable to store spatial information from image data and consider each pixel to be an independent feature, thus producing poor results. CNNs have been shown to outperform other machine learning methods, such as support vector machines (SVMs), in image classification tasks. For example, Holger et al. used deep convolutional networks for the automated detection of posterior-element fractures in spinal CT, achieving good results with area-under-the-curve, AUC, of 0.857 corresponding to 81% sensitivity [13]. Speeded-up robust features (SURF) [14] is a scale- and rotation-invariant feature detector and descriptor used to identify regions of interest and commonly used in image recognition. It has been shown to be much faster than comparable techniques, producing highly-accurate results when provided appropriate reference images. For example, Du et al. used SURF for facial recognition and achieved better results in the detection step [15]. In the present study, we will first apply CNN in binary fracture classification of calcaneus CT images and use SURF for detecting the location and type of classified calcaneal fractured cases.

This study proposes an efficient method for the computer-assisted automated classification and detection of calcaneus bone fracture locations in CT images to reduce human error and reduce the amount of manpower required for the diagnosis of bone fracture. The current method combined CNN for classification and SURF for detection, which have not been applied previously in calcaneal fracture detection. Bone CT image data from clinical patients were classified into two groups, fracture and non-fracture, and the images were labeled by physicians accordingly to the Digital Imaging and Communications in Medicine (DICOM) standard. Fracture classification was then performed using the CNN to separate CT scans into these two classes (fracture and non-fracture) from coronal, transverse, and sagittal views. Different network architectures with different width or depth have been reported to have different effect on classification performance [16]. Among different CNNs, these distinct architectures shared most of the network architecture but differed in depth were investigated: a visual geometry group (VGG) and a residual network (ResNet), with modifications on the softmax layer and the fully connected layer right before the softmax layer of each architecture type. This results in the number of output being reduced from 1000 to 2 categories, fracture and normal for the classification. Fracture detection was performed using the SURF algorithm, including reference images in which the areas of calcaneus fractures were based on ResNet classifier results from coronal, sagittal, and transverse views. Results obtained demonstrated high accuracy and reduced runtime in classifying and detecting the location of calcaneus bone fractures in CT images.

2. Method

2.1. Dataset

Two sets of real-patient calcaneus bone CT image were obtained from the Show Chwan Memorial Hospital in Changhua, Taiwan. These two datasets, acquired at two different periods in time, consisted of one-channel grayscale DICOM images with a 2-mm slice thickness and dimensions of 512 × 512 pixels. Each data set included images acquired from three different views (sagittal, coronal, and transverse). The Sanders classification system [17] was used to classify the datasets into normal and eight classes of fractured calcaneus bone CT images. Fig. 1 shows the CT image datasets of calcaneus fractures classified as Type 1 to 4 according to Sanders classification system. This classification produced 255 fractured images and 732 normal images from the first dataset (a total of 987 images), 428 fractured images and 516 normal images from the second dataset (a total of 944 images). A combined total of 1931 images were categorized from these two datasets, which were separated into two folders for the training (80%) and testing steps (20%).

2.2. Classification

Deep learning has recently been applied to a broad range of tasks in image processing. Several studies have demonstrated the successful classification of natural images [18,19] as well as processed medical images [20,21], which have become increasingly important for assisting physicians in timely and accurate diagnosis and treatment planning. The proposed classification algorithm uses a CNN to differentiate calcaneal bone CT scans into fractured and normal images. This method involves pre-processing of CT scan images, classification of CNN bone fractures, and detection of bone fracture locations. These various stages are described in detail below.

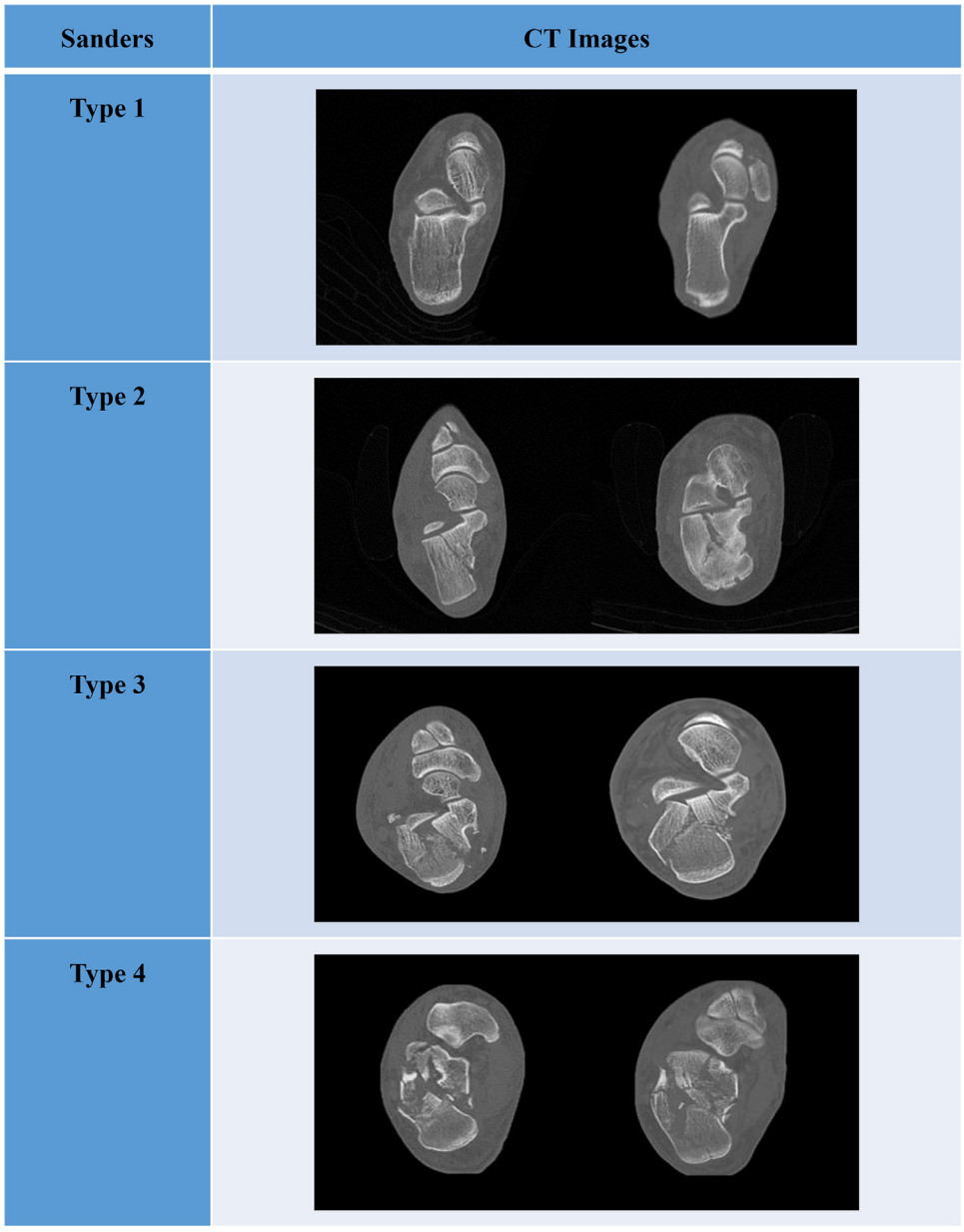

Fig. 1. Examples CT scan images of calcaneal fracture of Type 1–4 of Sanders classification system. Type 1 includes all intraarticular fractures that have less than 2 mm of articular displacement. Type 2: displaced with one fracture line in the posterior facet. Type 3: displaced with two fracture lines in the posterior facet. Type 4: comminuted with more than three fracture lines in the posterior facet.

2.2.1. Pre-processing

Image pre-processing was needed to satisfy the input size requirements for the classification step using the CNN with fully connected layers [22]. This process involved resizing, vector conversion, and normalization. Each CT image was first resized from 512 × 512 pixels to 224 × 224 pixels and then transformed into a vector matrix. We resized the images to satisfy the standard input image size required for using pre-trained VGG and ResNet networks, which were very deep convolutional neural networks pre-trained on ImageNet images with required input image size of 224 × 224 pixels. In the normalization procedure, the strong pixel correlation in each image caused by image downsizing and high variance within the image set was reduced by applying zero-phase component analysis (ZCA) whitening (i.e., sphering) [23]. This enhanced raw input detectability by lowering adjacent pixels, reducing their correlation and producing better classification results than in images without whitening [24]. The ZCA whitening procedure uses mean-centered random vectors z_c and x_c with $E(z_c)=0$ and $E(x_c)=0$, as described in the following expressions:

$$E\left((z_c - x_c)^T (z_c - x_c)\right) = tr(\mathrm{I}) - 2E\left(tr\left(z_c x_c^T\right)\right) + tr(\boldsymbol{\Sigma})$$
$$E\left((z_c - x_c)^T (z_c - x_c)\right) = d - 2tr(\Phi) + tr\ (\boldsymbol{V}) \quad (1)$$

where E represents the expected value, d represents dimension, Φ represents cross-covariance matrix, and V represents diagonal matrix contains variance of x_i. z_c and x_c were the transformed and original pixel locations, respectively [25]. These normalized images were then used in the CNN classification step.

2.2.2. Classification with the convolutional neural network

Once the normalized image vectors were acquired, they were input to the CNN for bone fracture classification training. A pre-trained model was implemented to reduce the required runtime and avoid the need to train the model from scratch [26]. Another

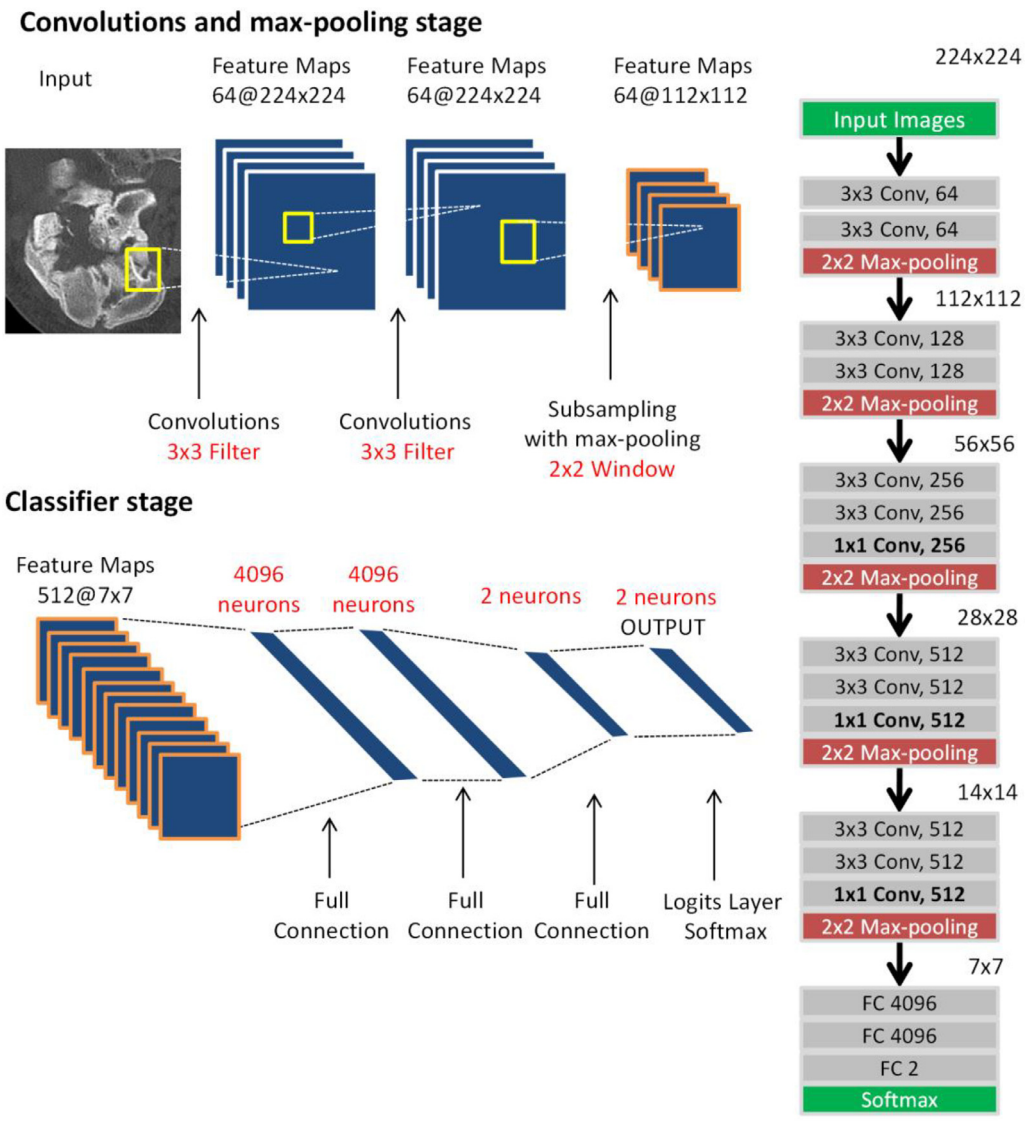

Fig. 2. A schematic diagram of the VGG network architecture.

benefit of using pre-trained model was its better classification accuracy compared to the model trained from scratch [27], resulting from training with 15 million labeled images across 22,000 categories, ranging from animal, fruit, to geological formation and artificial data sets from the ImageNet Large-Scale Visual Recognition Competition (ILSVRC) [19,28]. The pre-trained model was then fine-tuned by updating the weighting of this network with domain-specific category of images [26], using our pre-processed image data, to provide bone fracture classification output. Two CNN pre-trained models, VGG and ResNet, were compared using our proposed method, the results of which are discussed in the following sections.

2.2.3. VGG network architecture and training

VGG is a neural network architecture developed by the Visual Geometry Group for large-scale image recognition [29]. It was proposed to investigate CNN performance on object recognition with different depths of CNN architectures, significantly improving this accuracy by increasing its depth to 16–19 layers [29]. VGG applies multiple convolutions using the same 3×3 filter, as opposed to a larger filter which involves fewer parameters. Such parameters are typically implemented in the neural network by applying a max-pooling for image size reduction. However, the number of network parameters is still too high, making VGG a computationally expensive model [30].

The VGG network was implemented and trained to classify calcaneus CT images into fracture and normal categories using sagittal, coronal, and transverse views. Fig. 2 shows a schematic diagram of the VGG neural network architecture.

The VGG very deep network [29] was adopted in this study. The channel numbers for the last fully-connected layer and the softmax layer were reduced to two for binary (true or false) determination of the presence of bone fractures. The pre-trained VGG model was already pre-trained using images from the ILSVRC dataset. Training images for fine-tuning the neural network were collected from the two DICOM datasets (see Section 2.1 for details) were resized to 224×224 pixels, the VGG very deep 16 neural network size requirement, prior to the fine-tuning phase. Hyper-parameters used for fine-tuning the pre-trained model for bone fracture classification were the same as those used in pre-training the model [29]. These parameters included learning rate, weight decay, and batch size, to determine which combinations resulted in the highest accuracy for bone fracture classification. Table 1 displays the five different combinations of classification parameters tested in the

Table 1
VGG network architecture classification parameters.

Combination		A	B	C	D	E
Learning Rate		0.005 and 0.00005	0.0001 and 0.00001	0.00005 and 0.000005	0.00001 and 0.000001	0.005 and 0.0005
Weight Decay		0.0001	0.0002	0.0003	0.0004	0.0005
Batch Size		50	30	10	5	1
Accuracy		98 %	96 %	95 %	93 %	92 %
#Iteration	Training	25	42	124	248	1236
	Validation	7	11	31	62	308
Computational Time		1,120 seconds (18.67 minutes)	1,227 seconds (20.45 minutes)	1,326 seconds (22.1 minutes)	1,423 seconds (23.72 minutes)	2,915 seconds (48.58 minutes)

study. Among these, combination A resulted in the highest classification accuracy and required the least number of iterations with the VGG network.

2.2.4. Deep residual learning network architecture and training

The deep residual network (ResNet) is a new 152-layer network architecture generating unprecedented results in classification, detection, and localization studies. The network architecture is significantly deeper yet required less computational complexity than conventional VGG networks [29]. In addition to its unparalleled layer quantities, ResNet recently won the ILSVRC 2015 classification task competition with a surprisingly low error rate of 3.6% [31]. ResNet solve the problem of vanishing gradients in stacked layers by using a residual mapping structure. ResNet and VGG network architectures mostly exhibit the same primary components, including convolution layers, subsampling, fully connected layers, and activation for backpropagation. However, ResNet are nearly 8 times deeper than VGG nets. It has been shown that deeper neural network architecture generally performed better classification than networks with similar classification outcome [11]. The pre-trained imagenet-resnet-50-dag model [31] was used in this study. The number of channels in the last two layers was modified to determine whether fractures can be detected in CT images. Fig. 3 shows a schematic diagram of the ResNet architecture.

The hyperparameters used for fine-tuning the pre-trained model for bone fracture classification were the same as those used in training the pre-trained model [31]. The combination of hyperparameters selected for fine-tuning included: the number of sub batches, weight decay, and batch size, with the aim to accurately classify bone fractures. Three different combinations of hyperparameters were tested, with combination A resulting in the highest classification accuracy and required the least number of iterations (Table 2).

2.3. Detection of bone fracture locations in calcaneal CT images

A multi-stage process was implemented, after bone fracture classification using the CNN, to determine the exact location of the fracture in calcaneal CT scans. The included detection process, based on the SURF algorithm, consisted of the following steps: pre-processing, feature detection, feature extraction, feature matching, and contour identification. Fig. 4 provides the detection flowchart for this process.

2.3.1. Detection pre-processing

SURF is often used for object recognition in which reference images are included with a set of test images. This process was used to identify bone fracture locations in patient CT images through comparison with a reference bone fracture CT image. In this step, reference images were produced based on the results of the previous classification stage. A red square marker was manually added to indicate the fractured regions in pre-processed coronal, sagittal, and transverse calcaneal CT images, as shown in Fig. 5. This red area served as a reference image in the feature detection stage.

2.3.2. Feature detection

Test images must be converted into features prior to object recognition. As such, a fast-Hessian matrix was applied in the SURF feature detection process. This matrix served as a detector for locating regions of interest (RoIs) within an image, leading to reduced computational runtimes from 700 to 160 ms [14]. However, not all detected features are useful for object recognition and threshold filtering is commonly applied to exclude extraneous information. The threshold for the determinant of fast-Hessian detector was set to 700 for leaving stable interest points with its determinant higher than the threshold to achieve the best feature detection results. The corners of the object were detected after feature identification and the located frames were converted to grayscale. The number of detected features was then output and the RoI in the images was used for the feature extraction process described in the following section.

2.3.3. Feature extraction

The application of SURF ensures that extracted features are scale and translation invariant, but not necessarily rotation invariant, in which objects can be robustly located when placed at different angles. The feature descriptor in SURF is then constructed from features detected in previous results, calculated using the Haar wavelet response and standard SURF procedure [14].

2.3.4. Feature matching and mismatch reduction

After feature extraction, detected features and partial images rounded by red squares in the reference image were compared by calculating the Euclidean distance between the input image and reference image to find the best matched bone-fracture cases. The Euclidean distance of descriptors was calculated from the two images using the nearest neighbor matching diagram, to reduce the

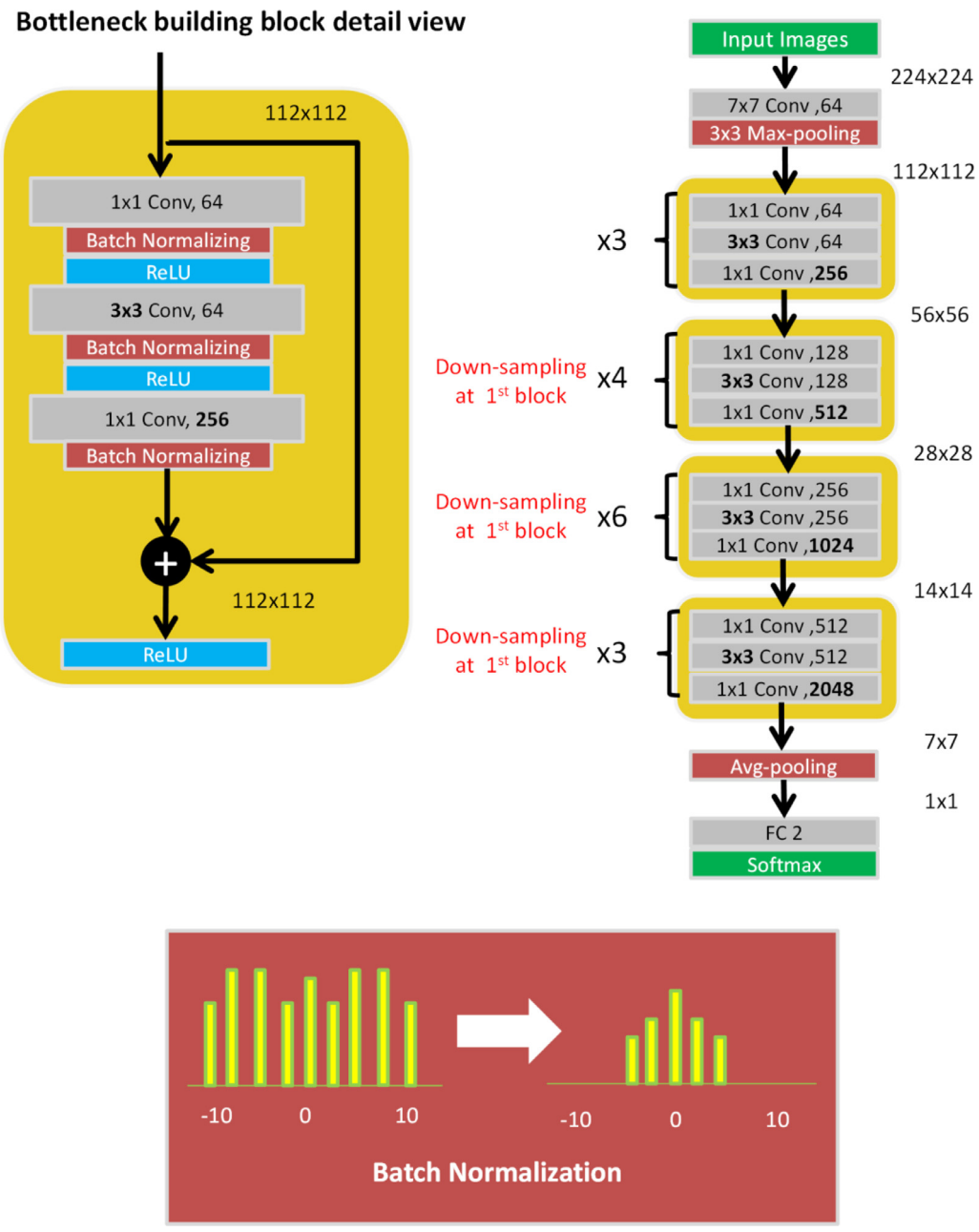

Fig. 3. A schematic diagram of the ResNet.

probability of a mismatch across the large reference set. Fast library for approximate nearest neighbors (FLANN) [32], an algorithm used for nearest neighbor calculation, was applied to match extracted features from test and reference images. Matches due to outliers were removed [33] and robust fitting methods were utilized based on epipolar geometry information, which are the geometric relationships among different CT scanner views on the same object and the interest points of the same object on the CT scans. It is almost impossible for radiologists to always scan from the same angle for different patients because of the differences of patients' foot orientation and casting degrees of the tube [17]. Random sample consensus (RANSAC) [34] was implemented for mismatch reduction with epipolar geometry consistency [35]. The areas which best matched those rounded by the red squares in the reference images were cropped and used as input for the edge detection step.

2.3.5. Labeling potential bone fractures through contour detection

The complexity of contour detection is correlated with image size, as smaller images require less runtime. As such, the width and height of the input images was reduced by considering only the areas matched with the red squares in the reference image. Canny edge detection was then applied to the reduced image to identify the edge of a bone fracture. Noise was then removed by applying image smoothing and a Gaussian filter. Potential bone fractures were outlined using the 'findContours' function in OpenCV, which generated red lines around potential fracture sites. The cropped images containing rounded contours were then superimposed with the original images from which they were cropped. The resulting combined images displayed the locations of calcaneus fracture areas in patient CT scans.

Table 2
Residual network classification parameters.

Combination		A	B	C
Number of Sub Batches		10	10	5
Weight Decay		0.0002	0.0001	0.00005
Batch Size		30	10	5
Accuracy		98 %	80 %	95 %
# Iteration	Training	42	124	248
	Validation	11	31	62
Computational Time		1,575 seconds (26.25 minutes)	4,510 seconds (75 minutes)	778 seconds (13 minutes)

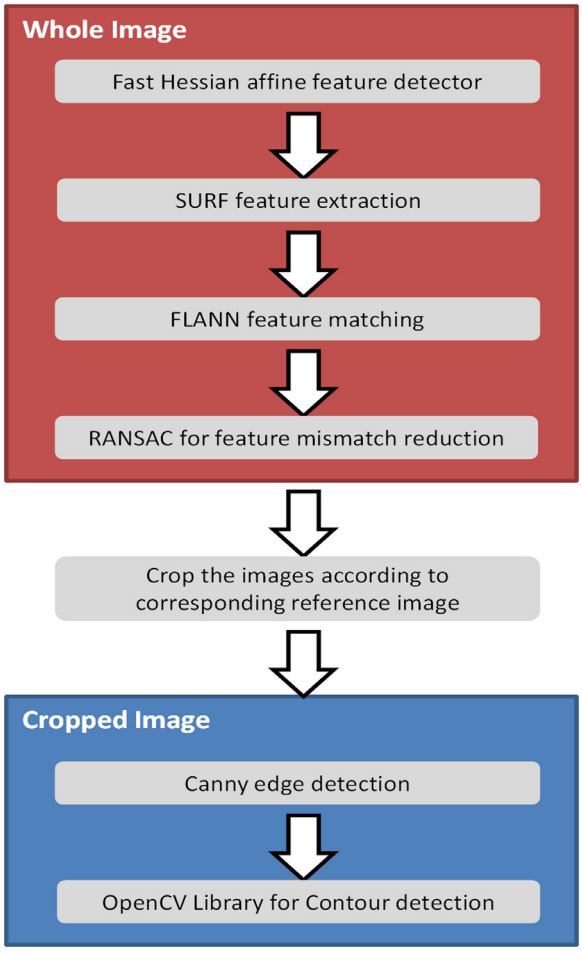

Fig. 4. A schematic flowchart of calcaneal bone fracture detection.

3. Results

3.1. Dataset

The two datasets included in the validation tests were the same set of DICOM images used in fine-tuning the pre-trained model (see Section 2.1). We tested the networks using 1931 CT images from these DICOM datasets.

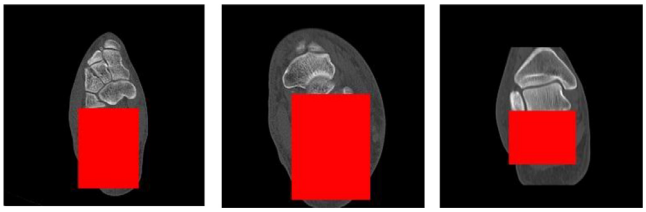

Fig. 5. Pre-processing results from fractured calcaneus CT images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Fine-tuning hyperparameters for VGG and ResNet pre-trained models

The hyperparameters used for fine-tuning the pre-trained model include learning rate, weight decay, and batch size. The ranges of batch size, weight decay and learning rate are usually 1–50, 0.0001–0.0005, and 0.000001–0.005, respectively. In this experiment, five parameter combinations were chosen empirically based on our previous work [36]. These parameter combinations were used in the training step to create the best model in the classification step. We analyzed the relationship between these three variable (learning rate, weight decay, and batch size) in terms of the accuracy result, number of iterations, and computational time for each parameter combination to determine the best parameter combination to use.

Table 1 shows the training (80%) and testing results for the five different parameter combinations in terms of accuracy, number of iterations and computational time for the VGG network. Results showed that combination A produced the highest classification accuracy (98%) and required the fewest number of iterations (training and validation) compared to the other combinations using the same datasets. Combination A also required the shortest computational time compared to the other parameter combinations. In addition, when the values of learning rate, weight decay and batch size were bigger than parameter value used in combination A, the method became overfitting. Similar results are shown for ResNet network (Table 2).

3.3. Loss function and error rate of bone fracture classification

Fig. 6 shows training and validation results using VGG and ResNet for bone fracture classification both using hyperparameter combination A (for details, see Tables 2 and 3). A total of 30

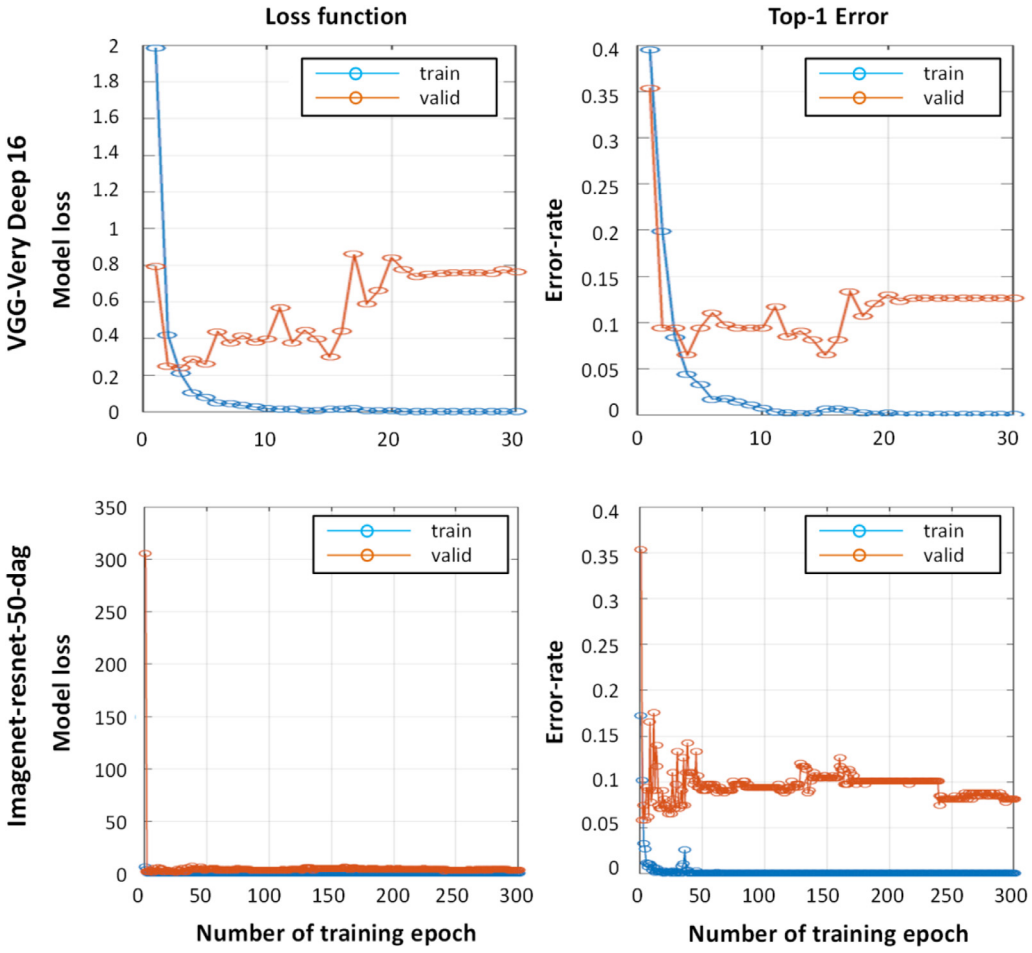

Fig. 6. The loss function and top-1 error of classification using VGG and ResNet.

Table 3
Classification accuracy and computational runtimes for VGG and ResNet.

	Accuracy	Training and validating time	Layer
VGG 16	98%	1120 s	23
ResNet-50	98%	1575 s	168

epochs were required to achieve convergence for the VGG network, with an acceptable error rate of 0.1% and a total runtime of 5–6 min each epoch for the training and validation steps. Approximately 300 epochs were required to achieve convergence for the ResNet, with an acceptable error rate of 0.1% and a total runtime of ~10 min each epoch for training and validation. The last epoch was used to assess classification results for the VGG network and ResNet, with the two algorithms achieving comparably high accuracy (98%) across 1931 images. In addition, the number of iterations required for training and validation of ResNet was larger than VGG network because the network architecture of ResNet is parallel while for the VGG network is serial. Although ResNet required a larger elapsed runtime (Table 3), we considered the ResNet to outperform VGG due to the eight times deeper number of layers involved in its network architecture compared to VGG.

3.4. Detection results using ResNet

ResNet classification results were used as input in the detection step, along with CT images of calcaneus bone fractures. The SURF method can be used to sequentially detect fractures in the classification results from CNN and ResNet with the pre-trained model. Fig. 7 shows results obtained at various stages of the fracture detection process, including (a) the original image, (b) the cropped image using SURF matching to a reference image, (c) the image after grey scaling, (d) the image after Canny edge detection, and (e) the image after contour detection to determine fracture location. These results demonstrate the method can successfully detect bone fractures over a relatively short time period of 10 s. The algorithm also produces good matching results between the reference and testing CT images.

4. Discussion

CT image classification is one of the most important stages for disease diagnosis, treatment planning, and disease progression or healing assessment. However, there is no simple solution for early, timely and accurate detection and identification of diseases on CT images due to technical constraints or human considerations. The proposed computer-assisted approach demonstrates the feasibility of using deep learning approach combined with SURF and Canny edge detection method to efficiently classify and detect the location of calcaneal fractures in CT images. As shown in our results, fractures on calcaneal CT images can be automatically classified and detected with relatively high accuracy and short computational time. ResNet perform better than VGG network in classifying fracture images for involving a deeper number of layers in its network architecture, which has been shown to produce better

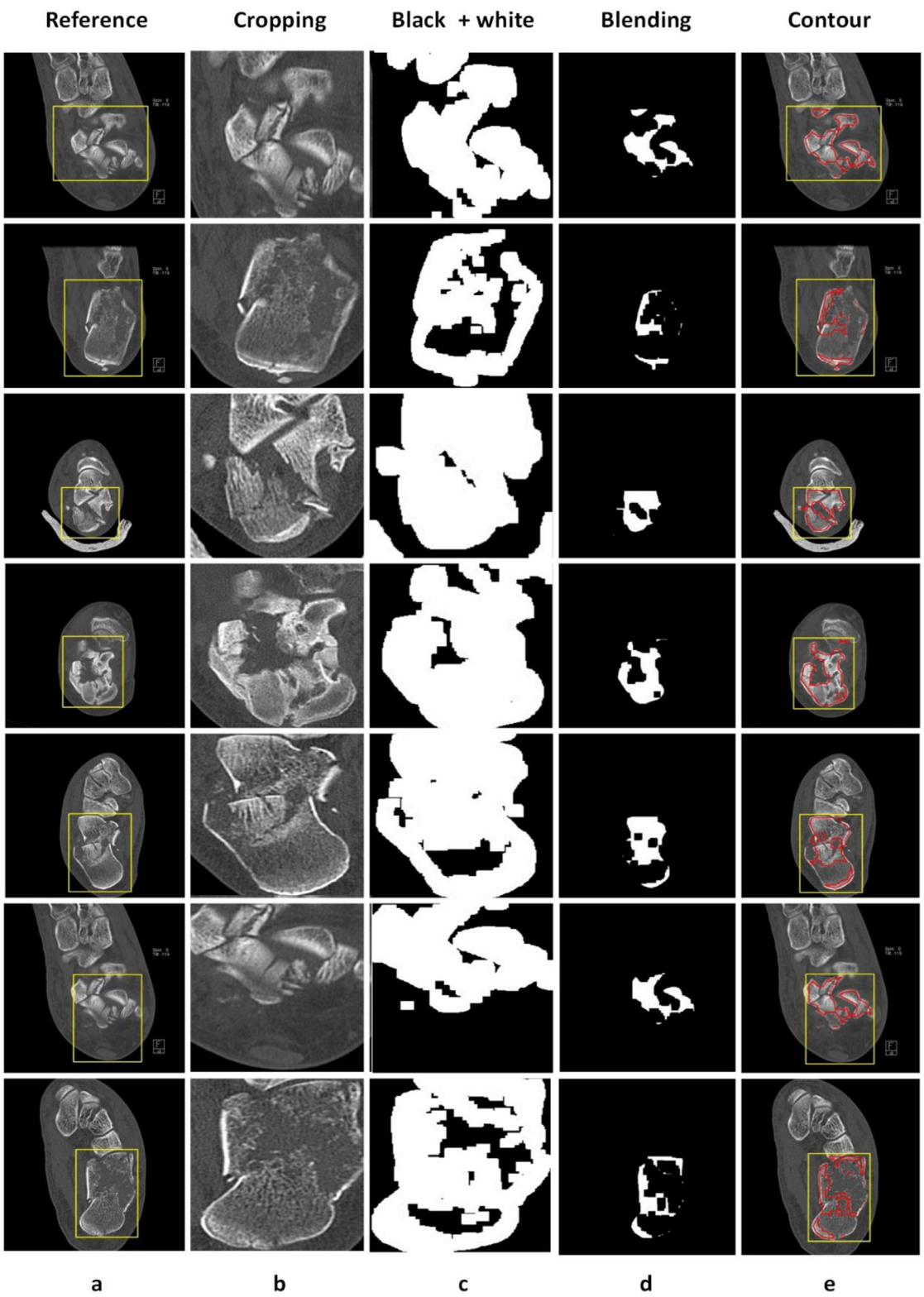

Fig. 7. Detection results for fractured calcaneus CT images at various processing steps. (a) the original image, (b) the cropped image using SURF matching to a reference image, (c) the image after grey scaling, (d) the image after Canny edge detection, and (e) the image after contour detection.

classification results than network architectures that involve fewer layers [11]. Comparison of reference and testing CT images produced good matching outcome, indicating that the fracture location was determined accurately using SURF method. Additionally, the proposed method use Canny edge detection and contour tracing to delineate the boundary along the fractured areas on CT images. Canny edge detector for edge detection has been shown to be one of the best edge detection techniques when comparing with other commonly used edge detection methods in almost all scenarios, especially under noisy conditions [37,38]. Future works can

implement deep learning algorithm within the traced contour to further determine the type and quantify the severity of calcaneal bone fractures.

Our trained ResNet neural network can distinguish between fractured and normal calcaneus bone in CT-image from coronal, transversal, and sagittal views with relatively high accuracy (98%). Recent studies have examined the calcaneal bone microarchitecture using texture analysis methods to assess the risk of osteoporotic fractures. For instance, Harrar et al. [39] used oriented texture analysis (OTA) with anisotropic piecewise fractional Brownian motion model (ap-fBm) and piecewise Whittle estimator (ap-WhE) to extract the fractal dimension of trabecular bone microarchitecture on calcaneus X-ray images and successfully differentiated osteoporotic patients from control subjects [39]. Similarly, Vokes et al. [40] applied radiographic texture analysis (RTA) to extract features from calcaneal X-ray images obtained using a portable densitometer specially equipped with a high resolution camera and differentiated between patients with and without prevalent vertebral fractures [40]. These texture analysis methods can assess the minute details of bone microstructure and may be implemented in our future works to evaluate potential bone fracture risks.

The proposed algorithm, once verified with additional data, can be implemented in the development of computer-aided decision systems to support physicians in making timely diagnosis and treatment. In addition, the proposed automated fracture detection system can be used as a training tool for less-trained radiologists or for use in areas without a radiologist available at all times. Automated detection scheme has the advantages of searching the entire field of view and weighting all the regions with equal importance, thus reducing the number of human errors associated with missed readings on radiography. In particular, mild and small fractures such as stress fractures may only be partially visible on CT images and often require multiple reads by the radiologists. Future studies can focus on automated fracture detection with slight fractures and implement other steps in the algorithm such as detection of the corners or edge discontinuity of fractured bones to improve detection sensitivity.

5. Conclusions

This study demonstrated that CNNs can achieve 98% accuracy in the automated classification of calcaneus bone fractures in CT images. Results showed the feasibility of locating fractured regions using a CNN with a VGG or ResNet as a pre-trained model. The ResNet exhibited 98% accuracy and 5-minute runtime in classifying the fractures and was therefore used as the input for the detection step in the SURF method. The 98% accuracy and 10-s runtimes achieved with this SURF detection technique make it a viable candidate for future use in the computer-assisted diagnosis of calcaneus bone fractures.

Conflicts of interest

The authors declare that there are no conflicts of interest related to this manuscript.

Acknowledgments

We would like to thank Dr. Pei-Yuan Lee, head of the Orthopedic Department at Show Chwan Memorial Hospital in Changhua, Taiwan for providing valuable inputs to this manuscript and for providing valuable patient CT data. We gratefully acknowledge financial support from the Ministry of Science and Technology of Taiwan (MOST106-2410-H-008-039-MY2 and MOST106-2218-E-008-002). This research is partially supported by the Ministry of Science and Technology under Grant #108-2634-F-008-004 through Pervasive Artificial Intelligence Research (PAIR) Labs, Taiwan.

References

[1] A. Daftary, A.H. Haims, M.R. Baumgaertner, Fractures of the calcaneus: a review with emphasis on CT, Radiographics 25 (5) (2005) 1215–1226.
[2] E. Guerado, M.L. Bertrand, J.R. Cano, Management of calcaneal fractures: what have we learnt over the years? Injury 43 (10) (2012) 1640–1650.
[3] W.E. Brant, Diagnostic imaging methods, in: W.E. Brant, C.A. Helms (Eds.), Fundamentals of Diagnostic Radiology, Lippincott Williams & Wilkins, Philadelphia, United States, 2012, pp. 3–25.
[4] A.L. Aagesen, M. Melek, Choosing the right diagnostic imaging modality in musculoskeletal diagnosis, Primary Care 40 (4) (2013) 849–861.
[5] M. Grigoryan, J.A. Lynch, A.L. Fierlinger, A. Guermazi, B. Fan, D.B. MacLean, A. MacLean, H.K. Genant, Quantitative and qualitative assessment of closed fracture healing using computed tomography and conventional radiography, Acad. Radiol. 10 (11) (2003) 1267–1273.
[6] J. Wu, P. Davuluri, K.R. Ward, C. Cockrell, R. Hobson, K. Najarian, Fracture detection in traumatic pelvic CT images, Int. J. Biomed. Imaging 2012 (2012) 1–10.
[7] M.S. Dhillon, K. Bali, S. Prabhakar, Controversies in calcaneus fracture management: a systematic review of the literature, Musculoskelet. Surg. 95 (3) (2011) 171–181.
[8] T.C. Anu, M.S. Mallikarjunaswamy, R. Raman, Detection of bone fracture using image processing methods, Int. J. Comput. Appl. (0975–8887) (2015) 6–9.
[9] S. Febrianto Kurniawan, I.K.G. Darma Putra, A.A. Kompiang Sudana, Bone fracture detection using opencv, J. Theor. Appl. Inf. Technol. 64 (1) (2014) 249–254.
[10] A. Sharma, P. Sharma, Rashmi, H. Kumar, Edge detection of medical images using morphological algorithms, Int. J. Sci. Emerg. Technol. Latest Trends 4 (1) (2012) 1–6.
[11] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[12] M.I. Razzak, S. Naz, A. Zaib, Deep learning for medical image processing: overview, challenges and the future, in: N. Dey, A.S. Ashour, S. Borra (Eds.), Classification in BioApps: Automation of Decision Making, Springer, 2017, pp. 323–350.
[13] H.R. Roth, Y. Wang, J. Yao, L. Lu, J.E. Burns, R.M. Summers, Deep convolutional networks for automated detection of posterior-element fractures on spine CT, Medical Imaging 2016: Computer-Aided Diagnosis, International Society for Optics and Photonics, 2016.
[14] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (SURF), Comput. Vis. Image Understanding 110 (3) (2008) 346–359.
[15] G. Du, F. Su, A. Cai, Face recognition using SURF features, MIPPR 2009: Pattern Recognition and Computer Vision, International Society for Optics and Photonics, 2009.
[16] S. Hoo-Chang, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imaging 35 (5) (2016) 1285.
[17] R. Sanders, Displaced intra-articular fractures of the calcaneus, J. Bone Joint Surg. Am. 82 (2) (2000) 225–250.
[18] L. Kang, J. Kumar, P. Ye, Y. Li, D. Doermann, Convolutional neural networks for document image classification, in: Pattern Recognition (ICPR), 2014 22nd International Conference on Pattern Recognition, IEEE, 2014, pp. 3168–3172.
[19] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[20] D.C. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2013, pp. 411–418.
[21] Z. Yan, Y. Zhan, Z. Peng, S. Liao, Y. Shinagawa, D.N. Metaxas, X.S. Zhou, Body-part recognition using multi-stage deep learning, in: International Conference on Information Processing in Medical Imaging, Springer, 2015, pp. 449–461.
[22] K. He, X. Zhang, S. Ren, J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition, in: European Conference on Computer Vision, Springer, 2014, pp. 346–361.
[23] A. Krizhevsky, Learning Multiple Layers of Features from Tiny Images, Citeseer, 2009.
[24] H. Jégou, O. Chum, Negative evidences and co-occurrences in image retrieval: the benefit of PCA and whitening, in: Computer Vision–ECCV 2012, Springer, 2012, pp. 774–787.
[25] A. Kessy, A. Lewin, K. Strimmer, Optimal whitening and decorrelation, Am. Stat. 72 (2018) 1–6.
[26] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans. Med. Imaging 35 (5) (2016) 1299–1312.
[27] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587.
[28] J. Deng, W. Dong, R. Socher, L. Li, L. Kai, F.-F. Li, ImageNet: a large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.
[29] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Proceedings International Conference on Learning Representations, 2015.

[30] A. Canziani, A. Paszke, E. Culurciello, An analysis of deep neural network models for practical applications, arXiv:1605.07678 (2016).
[31] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[32] M. Muja, D.G. Lowe, Fast approximate nearest neighbors with automatic algorithm configuration, in: Proceedings VISAPP, 2009, pp. 331–340.
[33] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[34] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.
[35] E. Oyallon, J. Rabin, An analysis of the SURF method, Image Process. Line 5 (2015) 176–218.
[36] J.-H. Li, Object Mask and Boundary Guided Recurrent Convolution Neural Network, National Central University, 2016.
[37] R. Maini, H. Aggarwal, Study and comparison of various image edge detection techniques, Int. J. Image Process. 3 (1) (2009) 1–11.
[38] N. Johari, N. Singh, Bone fracture detection using edge detection technique, in: Soft Computing: Theories and Applications, Springer, 2018, pp. 11–19.
[39] K. Harrar, R. Jennane, K. Zaouchi, T. Janvier, H. Toumi, E. Lespessailles, Oriented fractal analysis for improved bone microarchitecture characterization, Biomed. Signal Process. Control 39 (2018) 474–485.
[40] T. Vokes, D. Lauderdale, S.-L. Ma, M. Chinander, K. Childs, M. Giger, Radiographic texture analysis of densitometric calcaneal images: relationship to clinical characteristics and to bone fragility, J. Bone Miner. Res. 25 (1) (2010) 56–63.