

Evaluation of deep neural network architectures in the identification of bone fissures

Fredy Martínez, César Hernández, Fernando Martínez

Facultad Tecnológica, Universidad Distrital Francisco José de Caldas, Colombia

Article Info

Article history:

Received Aug 16, 2019

Revised Jan 4, 2020

Accepted Feb 13, 2020

Keywords:

Biomedical computing

Deep neural network

Fissures recognition

Image processing

ABSTRACT

Automated medical image processing, particularly of radiological images, can reduce the number of diagnostic errors, increase patient care and reduce medical costs. This paper seeks to evaluate the performance of three recent convolutional neural networks in the autonomous identification of fissures over two-dimensional radiological images. These architectures have been proposed as deep neural network types specially designed for image classification, which allows their integration with traditional image processing strategies for automatic analysis of medical images. In particular, we use three convolutional networks: ResNet (residual neural network), DenseNet (dense convolutional network), and NASNet (neural architecture search network) to learn information from a set of 200 images labeled half as fissured bones and half as seamless bones. All three networks are trained and adjusted under the same conditions, and their performance was evaluated with the same metrics. The final results consider not only the model's ability to predict the characteristics of an unknown image but also its internal complexity. The three neural models were optimized to reduce classification errors without producing network over-adjustment. In all three cases, generalization of behavior was observed, and the ability of the models to identify the images with fissures, however the expected performance was only achieved with the NASNet model.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Fredy Martínez,

Facultad Tecnológica,

Universidad Distrital Francisco José de Caldas,

Bogotá, Colombia.

Email: fhmartinezs@udistrital.edu.co

1. INTRODUCTION

In recent years there have been more and more advantages of the use of digital image processing is used as a tool to support the diagnosis from medical images [1, 2] and it's even proved very valuable to track throughout images (temporal quantification and growth) both damage and behavior of tissues [3, 4]. An automated system has the advantage of quickly identifying specific patterns in large volumes of images with a high degree of reliability. As a support tool for specialized medical personnel, this tool can not only reduce diagnostic time but also reduces confusing or misdiagnoses [5]. The great advantage of diagnostic imaging is its non-invasive character since most of these images are captured by resonance or radiography [6].

In general terms, these tools use a certain algorithm of classification on the image to determine if it possesses or not a certain characteristic, and thus to classify it [7, 8]. Images are normally pre-processed to maximize the ability to detect the characteristics of interest [9-11]. The classification algorithm, in general, is not applied to an image indiscriminately, on the contrary, a region of interest (ROI) is identified on the image,

which can be done manually, or even automatically in more advanced schemes, for example, through segmentation strategies [12-14]. It is also possible to use iterative searches on certain image structures to determine characteristics and ROI. Statistical methods are also used in which the image is navigated from previous information of characteristic behaviors [15].

Deep convolutional neural networks have become a powerful tool for image classification, with particular application to medical images [16, 17]. These correspond to regularized versions of the traditional multilayer perceptrons (fully connected forward layers) [18, 19]. Thanks to this regularization process, convolutional networks achieve complex structures with simple patterns that reduce the problem of network over-adjustment [20, 21]. We trained three models of deep neural networks to identify fissures on digitized radiological images. The images used for the training correspond to sections of bones in which ROI has been previously identified, but no morphological operation is applied to them [22, 23]. The types of deep nets selected correspond to the state of the art in convolutional nets for image classification [24, 25]. The following part of the paper is arranged in this way. Section 2 presents preliminary concepts and problem formulation. Section 3 illustrates the design profile and development methodology. Section 4 we present the preliminary results. And finally, in Section 5, we present our conclusions.

2. PROBLEM FORMULATION

We evaluate models based on deep neural networks by identifying characteristics in bone structures as shown in Figure 1. In particular, we look for models that identify and classify bones with fissures and fractures in one category, and those healthy bones in a second category. In deep learning, a convolutional neural network (CNN) is a class of deep neural networks commonly applied to analyzing images. They have the great advantage that they require much less image pre-processing to identify the features of interest than any other digital processing strategy. They operate as classification algorithms in which an adjustable weight value is assigned to the characteristics of the image that make it distinguishable from others. With proper training and adjustment, a convolutional network can replicate the behavior of a sophisticated filter on the image. Besides, unlike traditional neural networks, a convolutional network can identify special and temporal dependencies in images.

The high performance of convolutional networks is due to the design of their network architecture. While their operation is still a black box, their high performance is attributed to characteristics such as network depth, network width (greater number of parameters), and skip connections (whether dense or residual, which increases the complexity of the network and its ability to represent information). Consequently, the networks selected for the visual categorization task of this performance test are ResNet (residual neural network), DenseNet (dense convolutional network), and NASNet (neural architecture search network).



Figure 1. Sample database of cracked bone images used for model training

3. METHODOLOGY

The training of the three models is performed with the same dataset, a custom set of X-ray images separated into two categories (fissured and seamless). The images corresponding to a category are stored in the same folder for easy identification by category (the name of the folder is the name of the category). We use 1000 images for each category keeping the balance of classes to avoid biases in the model. We use TensorFlow as the framework on which we run Keras. Numpy, Scikit Learn, Pandas, OpenCV and Matplotlib were also used as support libraries.

The images were randomly mixed in the data list to improve network performance. Besides, they are all resized to the same size (256*256 pixels) with the same goal. We do not consider the aspect ratio of the images when resizing them. In all three cases, the dataset was divided into two groups, a training group, and a test group. We used 70% of the data for training and 30% for performance evaluation.

The three models are compiled specifying the optimization function, the cost or loss function, and the metrics. We use the stochastic gradient descent optimization function, the categorical cross-entropy function, which can be used to reflect the accuracy of the predictions, and for the metrics, accuracy (or hit rate) and mse (mean of the quadratic errors).

3.1. ResNet (residual neural network)

This network mimics the structure of pyramidal cells in the cerebral cortex. This structure is achieved by jumping (double or triple) over some of the layers, which use ReLu (Rectified Linear Units) activation function as shown in Figure 2.

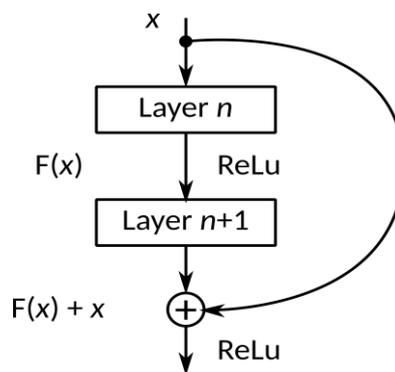


Figure 2. Building block (ResNet)

3.2. DenseNet (dense convolutional network)

The DenseNet structure also has similar jumps to the ResNet, but each layer receives input from the previous layers, and connects to the subsequent layers (each layer receives knowledge from the previous layers as shown in Figure 3).

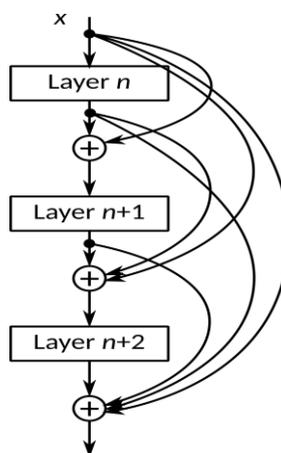


Figure 3. Building block (DenseNet)

3.3. NASNet (neural architecture search network)

The NASNet network consists of a specific block, the best convolutional structure for CIFAR-10, which is then generalized for ImageNet, and finally replicated as a block for large datasets as shown in Figure 4.

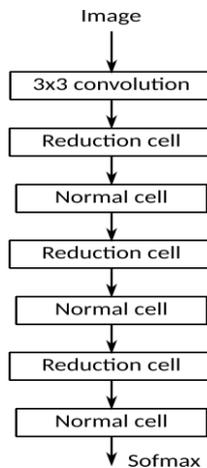


Figure 4. ImageNet architecture (NASNet)

4. FINDINGS

To evaluate the performance of the three models, in addition to loss and accuracy with training and validation data, we have used precision, recall, and F1-score as performance metrics. The results show superior NASNet performance over ResNet and DenseNet. ResNet had the poorest performance, not only are its metrics very low, but its accuracy does not increase significantly with loss reduction, and the model is the most complex (over 23 million parameters). DenseNet has similar performance but with only 7 million parameters, but with still very low metrics. NASNet is the only one that gets an acceptable performance and with a lower number of parameters (a little over 4 million). Summary of the model: ResNet (residual neural network as shown in Figures 5, 6 and 7):

- Total params: 23,591,810
- Trainable params: 23,538,690
- Non-trainable params: 53,120

Summary of the model: DenseNet (Dense Convolutional Network as shown in Figures 8, 9 and 10):

- Total params: 7,039,554
- Trainable params: 6,955,906
- Non-trainable params: 83,648

Summary of the model: NASNet (Neural Architecture Search Network, Figures 11, 12 and 13):

- Total params: 4,271,830
- Trainable params: 4,235,092
- Non-trainable params: 36,738

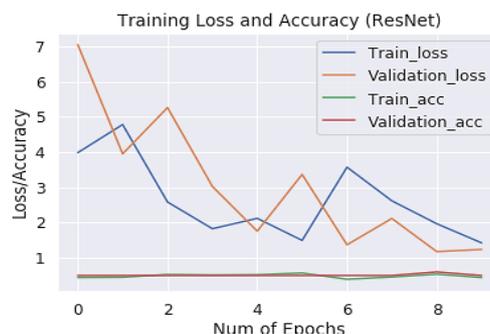


Figure 5. Training loss and accuracy (ResNet)

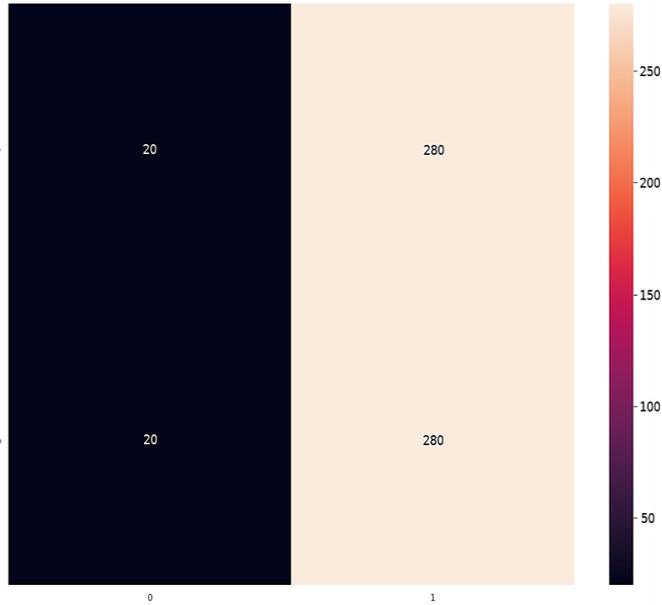
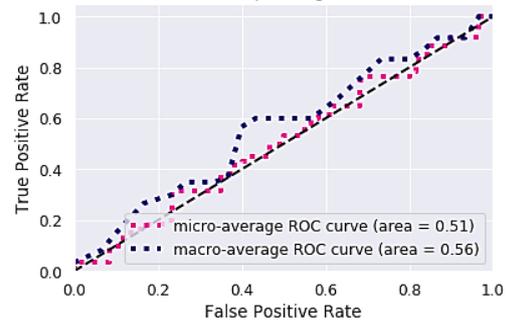


Figure 6. Confusion matrix (ResNet)

	precision	recall	f1-score	support
0	0.50	0.07	0.12	300
1	0.50	0.93	0.65	300
micro avg	0.50	0.50	0.50	600
macro avg	0.50	0.50	0.38	600
weighted avg	0.50	0.50	0.38	600

(a)

Some extension of Receiver operating characteristic to multi-class



(b)

Figure 7. Performance metrics (ResNet):
 (a) Classification report (ResNet), (b) ROC curve and ROC area (ResNet)

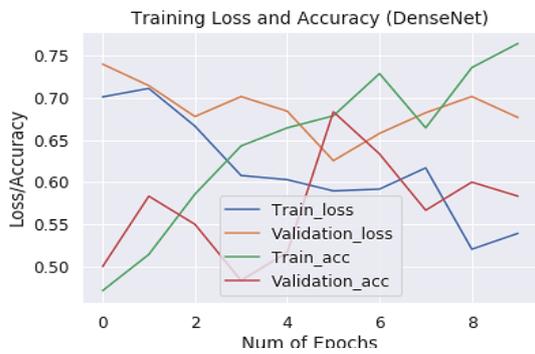


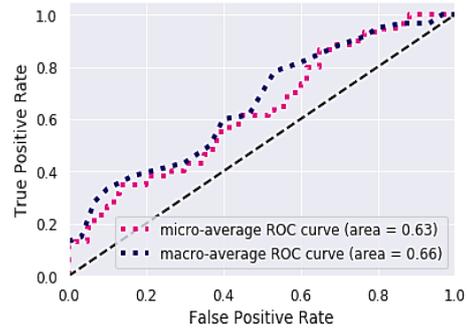
Figure 8. Training loss and accuracy (DenseNet)



Figure 9. Confusion matrix (DenseNet)

	precision	recall	f1-score	support
0	0.59	0.53	0.56	300
1	0.58	0.63	0.60	300
micro avg	0.58	0.58	0.58	600
macro avg	0.58	0.58	0.58	600
weighted avg	0.58	0.58	0.58	600

Some extension of Receiver operating characteristic to multi-class



(a)

(b)

Figure 10. Performance metrics (DenseNet):
 (a) Classification report (DenseNet), (b) ROC curve and ROC area (DenseNet)

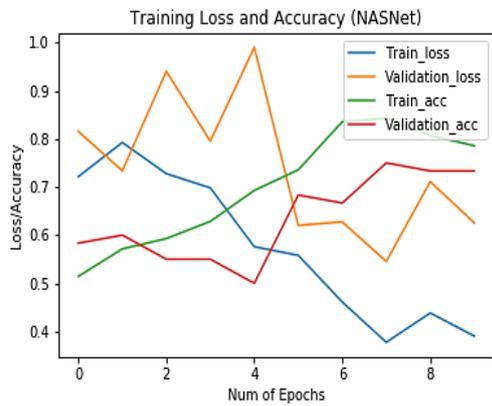
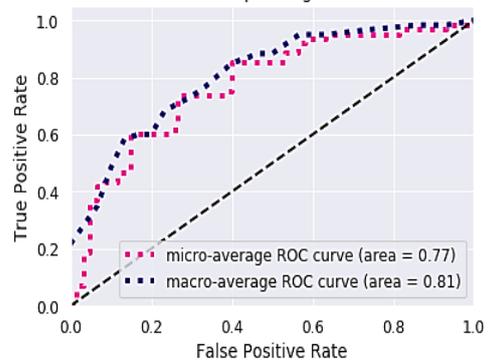


Figure 11. Training loss and accuracy (NASNet)

Figure 12. Confusion matrix (NASNet)

	precision	recall	f1-score	support
0	0.82	0.60	0.69	300
1	0.68	0.87	0.76	300
micro avg	0.73	0.73	0.73	600
macro avg	0.75	0.73	0.73	600
weighted avg	0.75	0.73	0.73	600

Some extension of Receiver operating characteristic to multi-class



(a)

(b)

Figure 13. Performance metrics (NASNet):
 (a) Classification report (NASNet), (b) ROC curve and ROC area (NASNet)

The convolutional network models used have an optimized structure for image classification. The ResNet network-based model achieves a considerable reduction compared to the number of adjustable parameters for a deep network, however the number of parameters remains high, and the best optimisation still shows an under-adjustment of the data (50% accuracy). The DenseNet model with a much denser architecture achieves higher accuracy than ResNet (58%), but with a much higher number of parameters. Finally, the optimized NASNet architecture achieves the highest metric values (75% accuracy) with a much lower number of parameters, becoming the right solution to the problem.

5. CONCLUSION

In this paper, we have evaluated the performance of three convolutional neural networks in the identification of fissures on bones. The aim of the research is to find an automatic model that is capable of processing radiological images and giving a preliminary diagnosis of possible bone fissures, in the hope of reducing the probability of misdiagnosis, increasing the percentage of patients attended and improving the quality of medical service. The selected networks were: ResNet (residual neural network), DenseNet (dense convolutional network), and NASNet (neural architecture search network). The performance of each of the models was evaluated by calculating the precision, recall, and F1-score metrics. The models were also used to evaluate loss and accuracy with training and validation data. Details of the number of parameters of each model, confusion matrices and ROC curve were also shown. After analyzing the behavior of the models, it was found that only the NASNet network produces an acceptable classification for the problem. The precision values of the NASNet model were higher than the other two models. Similar behavior was observed in the other calculated metrics. In addition, the NASNet model is the smallest of the three, requiring a little more than 4 million trainable parameters, compared to 7 million in the DenseNet model and more than 23 million in the ResNet model. These results are important for the correct selection of an automated diagnostic model, and show that it is possible to improve the performance of this model through a larger set of training images and better tuning of parameters. Future work will focus on improving the fit of this network by altering its depth and using images with more visual information.

ACKNOWLEDGEMENTS

This work was supported by the Universidad Distrital Francisco José de Caldas, in part through CIDC, and partly by the Facultad Tecnológica. The views expressed in this paper are not necessarily endorsed by District University. The authors thank the research group ARMOS for the evaluation carried out on prototypes of ideas and strategies.

REFERENCES

- [1] M. Hussain, A. Bhuiyan, A. Turpin, C. Luu, R. Smith, R. Guymer, and R., "Kotagiri. Automatic identification of pathology-distorted retinal layer boundaries using sd-oct imaging," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1638–1649, 2017.
- [2] O. Zenteno, F. Zvietcovich, D. Zapata, H. Maruenda, B. Valencia, A. Llanos, J. Arevalo, M. Montero, R. Lavarello, and B. Castaneda, "An integrated protocol for the research and monitoring of cutaneous leishmaniasis," *IEEE Latin America Transactions*, vol. 15, no. 11, pp. 2164–2170, 2017.
- [3] S. Makrogiannis, K. Fishbein, A. Moore, R. Spencer, and L. Ferrucci, "Image-based tissue distribution modeling for skeletal muscle quality characterization," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 805–813, 2016.
- [4] A. Anand, I. Moon, and B. Javidi, "Automated disease identification with 3-d optical imaging: A medical diagnostic tool," *Proceedings of the IEEE*, vol. 105, no. 5, pp. 924–946, 2017.
- [5] J. Zhou, T. Zhong, and X. He, "Auxiliary diagnosis of breast tumor based on pnn classifier optimized by pca and pso algorithm," *In 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC 2017)*, vol. 2, pp. 222–227, 2017.
- [6] N. Tsai, J. Goodwin, M. Semler, R. Kothera, M. Van Horn, B. Wolf, and D. Garner, "Development of a non-invasive blink reflexometer," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 5, no. 1, pp. 1–4, 2017.
- [7] O. Bertel, C. Moreno, and E. Toro, "Aplicación de la transformada wavelet para el reconocimiento de formas en visión artificial," *Tekhnê*, vol. 6, no. 1, pp. 3–8, 2009.
- [8] Y. Guo, L. Jiao, S. Wang, S. Wang, F. Liu, and W. Hua, "Fuzzy superpixels for polarimetric sar images classification," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 5, pp. 2846–2860, 2018.
- [9] J. Castañeda and Y. Salguero, "Adjustment of visual identification algorithm for use in stand-alone robot navigation applications," *Revista Tekhnê*, vol. 14, no. 1, pp. 73–86, 2017.
- [10] P. Joris, W. Develter, W. Voorde, P. Suetens, F. Maes, D. Vandermeulen, and P. Claes, "Preprocessing of heteroscedastic medical images," *IEEE Access*, vol. 6, no. 1, pp. 26047–26058, 2018.

- [11] V. Jaouen, J. Bert, N. Boussion, H. Fayad, M. Hatt, and D. Visvikis, "Image enhancement with PDEs and nonconservative advection flow fields," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 3075–3088, 2019.
- [12] B. Lassen, E. van Rikxoort, M. Schmidt, S. Kerkstra, B. van Ginneken, and J. Kuhnigk, "Automatic segmentation of the pulmonary lobes from chest ct scans based on fissures, vessels, and bronchi," *IEEE Transactions on Medical Imaging*, vol. 32, no. 2, pp. 210–222, 2013.
- [13] P. Moeskops, M. Viergever, A. Mendrik, L. de Vries, M. Benders, and I. Isgum, "Automatic segmentation of mr brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
- [14] G. Wang, W. Li, M. Zuluaga, R. Pratt, P. Patel, M. Aertsen, T. Doel, A. David, J. Deprest, S. Ourselin, and T. Vercauteren, "Interactive medical image segmentation using deep learning with image-specific fine tuning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1562–1573, 2018.
- [15] J. Sedlar, M. Bajger, M. Caon, and G. Lee, "Model-guided segmentation of liver in CT and PET-CT images of child patients based on statistical region merging," *In International Conference on Digit. Image Comput. Tech. Appl. (DICTA 2016)*, pp. 1–8, 2016.
- [16] H. Tang, C. Zhang, and X. Xie, "Automatic pulmonary lobe segmentation using deep learning," *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, vol. 1, no. 1, pp. 1225–1228, 2019.
- [17] J. Ming, W. Kim, and K. Ryoung, "Finger-vein recognition based on deep DenseNet using composite image," *IEEE Acces*, vol. 7, no. 1, pp. 66845–66863, 2019.
- [18] V. Neagoe, A. Ciotec, and G. Cucu, "Deep convolutional neural networks versus multilayer perceptron for financial prediction," *In International Conference on Communications (COMM 2018)*, pp. 201–206, 2018.
- [19] S. Cui, Y. Luo, H. Tseng, R. Haken, and I. El Naqa, "Artificial neural network with composite architectures for prediction of local control in radiotherapy," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 242–249, 2019.
- [20] H. Tong and Z. and Zhi, "Bag of tricks for image classification with convolutional neural networks," *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, pp. 558–567, 2019.
- [21] X. Zhao, T. Zhang, H. Liu, G. Zhu, and X. Zou, "Automatic windowing for mri with convolutional neural network," *IEEE Access*, vol. 7, no. 1, pp. 68594–68606, 2019.
- [22] H. Montiel, E. Jacinto, and F. Martínez, "Recognition of fissures in bony structures through image processing," *International Journal of Engineering and Technology*, vol. 10, no. 4, pp. 1223–1229, 2018.
- [23] T. Savithri and S. Devi, "Nodule detection from posterior and anterior chest radiographs with different methods," *In Future Technologies Conference (FTC 2016)*, pp. 504–515, 2016.
- [24] G. Huang, S. Liu, L. Maaten, and K. Weinberger, "Condensenet: An efficient densenet using learned group convolutions," *In IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2752–2761, 2018.
- [25] J. Sen and B. Neil, "Eml-net: An expandable multi-layer network for saliency prediction," *ArXiv 1805(01047)*, pp. 1–10, 2018.