Automated Detection and Classification of Skin Lesion in Dermoscopy Images

[1]Mrs. Blessy Paul P, [2]Mrs. Surya Krishna S.L

[1]Assistant Professor, [2]PG Scholar
Department of Computer Science and Engineering,
Marthandam College of Engineering and Technology, Kuttakuzhi

***Abstract*: Automated skin lesion classification in dermoscopy images is an essential way to improve the diagnostic performance and reduce melanoma deaths. Automatic localization of skin lesions within dermoscopy images is a crucial step toward developing a decision support system for skin cancer detection. However, segmentation of the lesion image can be challenging, as these images possess various artifacts distorting the uniformity of the lesion area. Recently, deep convolution learning-based techniques have drawn great attention for pixel-wise image segmentation. These deep networks produce coarse segmentation, and convolutional filters and pooling layers result in segmentation of a skin lesion at a lower resolution than the original skin image. To overcome these drawbacks, the proposed system uses a superpixel-based fine-tuning strategy to effectively utilize the characteristics of the skin image pixels to accurately extract the border of the lesion. The proposed approach not only learns a global map for skin lesions, but also acquires the local contextual information, such as lesion boundary. It can, therefore, accurately segment lesions within a given skin image, even in the presence of fuzzy boundaries and complex textures.**

***Keywords*: Dermoscopy, Skin lesion, Super pixel**

1. INTRODUCTION

A skin lesion is a part of the skin that has an abnormal growth or appearance compared to the skin around it. Two categories of skin lesions exist: primary and secondary. Primary skin lesions are abnormal skin conditions present at birth or acquired over a person’s lifetime. Secondary skin lesions are the result of irritated or manipulated primary skin lesions. Change of recreational behavior together with the increase in ultraviolet radiation cause a dramatic increase in the number of melanomas diagnosed. The DCNN is used for the classification task has a remarkable localization ability that can highlight the discriminative regions in images, despite being trained with only image-level labels, instead of the bounding boxes of discriminative regions. Hence, strengthening the discriminative ability of a DCNN via taking advantage of its self-attention ability. Since the higher layers in a DCNN have a better ability for semantic abstraction than lower ones, it might be possible to use the feature maps obtained by higher layers as the attention mask of lower ones. Meanwhile, the residual network is more suitable for small-sample learning problems than other DCNNs, such as AlexNet , VGG, and GoogLeNet, since it uses “shortcut connections” to skip one or more layers, and thus enable the construction of a deeper network. The raw input image is transformed through a multi-scale convolutional network, which produces a set of feature maps. The feature maps of all scales are concatenated, and then the coarser-scale maps are upsampled to match the size of the finest image scale map. In parallel, a single segmentation technique using superpixels is devised to exploit the natural contours of the image. Finally, a supervised classifier is used to classify each superpixel by computing the average class distribution of the dense features within the superpixels. Yan et al. formulate the object detection as a multi-class superpixel labeling problem. Their proposed energy minimization has several terms such as data cost, smooth cost term and label cost. The data cost is learned through a convolutional neural network. Following that, a smooth cost term and label cost term is used to obtain the final labelling of the superpixels. The parameters in the labeling model are learned through a structural support vector machine.

The presence of artifacts such as fiducial markers and hairs on the skin may occlude the lesion, making accurate skin lesion segmentation difficult. Both of these methods combine CNN responses and superpixel prediction maps in a supervised fashion to obtain the final segmentation outcome. In this method, however, it smooth the initial segmentation of the FCN network by constructing an energy function on the image's superpixels to infer their labels. In particular, it refine the boundary of the skin lesion by analyzing the relationship of the superpixels inside the region of interest (ROI) obtained by the FCN.

II. RELATED WORKS

Jianpeng Zhang et al., proposed an attention residual learning convolutional neural network (ARL-CNN) model for skin lesion classification in dermoscopy images, which is composed of multiple ARL blocks, a global average pooling layer, and a classification layer. Each ARL block jointly uses the residual learning and a novel attention learning mechanisms to improve its ability for discriminative representation. The attention mechanism is an effective technique that helps a model pay more attention to important information. It has made great progress in the cross fields of computer vision and natural language processing, such as image/video caption and visual question answering.

Catarina Barata et al., proposes two systems for the automatic classification of melanocytic skin lesions. The first system uses global methods to classify skin lesions. This approach evolves in three sequential steps. First, the lesion is segmented using an automatic segmentation method. Then, a set of features from the ABCD rule (color and texture features) is extracted and used to train a classifier to perform binary classification as melanoma or benign.

Jianpeng Zhang et al., proposes a synergic deep learning (SDL) model to address this issue. Specifically, a dual deep convolutional neural network with a synergic signal system is designed to mutually learn image representation. The synergic signal is used to verify whether the input image pair belongs to the same category and to give the corrective feedback if a synergic error exists. The SDL model can be trained 'end to end'. In the test phase, the class label of an input can be predicted by averaging the likelihood probabilities obtained by two convolutional neural network components.

Emre Celebi et al propose the systems allow the use of a computer as a second independent diagnostic method, which can potentially be used for the prescreening of patients performed by non-experienced operators and for aiding clinicians. Although computerized analysis techniques cannot provide a definitive diagnosis, they can be used to improve biopsy decision-making, which some observers feel is the most important use for dermoscopy.

X. Yang et al propose a multi-task deep neural network is proposed for skin lesion analysis. The proposed multitask learning model solves different tasks (e.g., lesion segmentation and two independent binary lesion classifications) at the same time by exploiting commonalities and differences across tasks.

F. Xie et al proposes a novel method for classifying melanocytic tumors as benign or malignant by the analysis of digital dermoscopy images. The algorithm follows three steps: first, lesions are extracted using a self-generating neural network (SGNN); second, features descriptive of tumor color, texture and border are extracted; and third, lesion objects are classified using a classifier based on a neural network ensemble model.

Liang et al's work address the task of semantic image segmentation with Deep Learning and make three main contributions First, it highlight convolution with upsampled filters, or `atrous convolution', as a powerful tool in dense prediction tasks. Atrous convolution allows to explicitly control the resolution at which feature responses are computed within Deep Convolutional Neural Networks. It also allows us to effectively enlarge the field of view of filters to incorporate larger context without increasing the number of parameters or the amount of computation.

Z. Zhou et al Intense interest in applying convolutional neural networks (CNNs) in biomedical image analysis is wide spread, but its success is impeded by the lack of large annotated datasets in biomedical imaging. Annotating biomedical images is not only tedious and time consuming, but also demanding of costly, specialty - oriented knowledge and skills, which are not easily accessible. To dramatically reduce annotation cost, this paper presents a novel method called AIFT (active, incremental fine-tuning) to naturally integrate active learning and transfer learning into a single framework. AIFT starts directly with a pre-trained CNN to seek "worthy" samples from the unannotated for annotation, and the (fine-tuned) CNN is further fine-tuned continuously by incorporating newly annotated samples in each iteration to enhance the CNNs performance incrementally.

III. PROPOSED SYSTEM

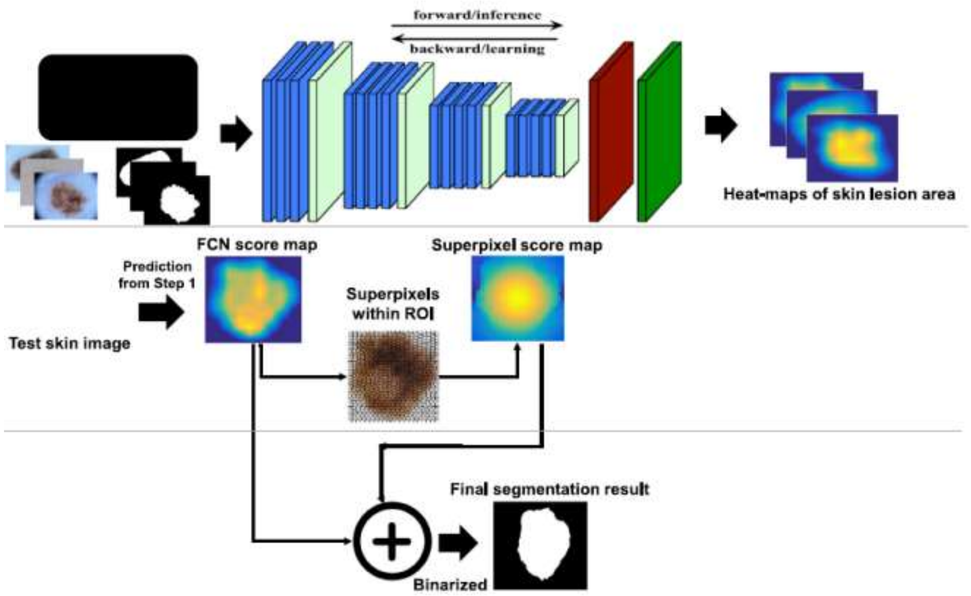


Fig 1: Block Diagram

The proposed method consists of modules including FCN-based initial segmentation and superpixel based finetuning which the raw input image is transformed through a multi-scale convolutional network, which produces a set of feature maps. The feature maps of all scales are concatenated, and then the coarser-scale maps are upsampled to match the size of the finest image scale map. The convolutional networks trained by themselves, trained end-to-end, pixels to-pixels, exceed the state-of-the-art in semantic segmentation. The key insight is to build “fully convolutional” networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning. It define and detail the space of fully convolutional networks, specifies their application to spatially dense prediction tasks, and draw connections to prior models. It then define a skip architecture that combines semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations. In parallel, a single segmentation technique using superpixels is devised to exploit the natural contours of the image. Finally, a supervised classifier is used to classify each superpixel by computing the average class distribution of the dense features within the superpixels. Superpixel segmentation is the process of partitioning an image into multiple segments called superpixels, which are homogeneous as in pixels inside every portion are comparable concerning certain attributes, for example, shading and surface. In spite of the fact that superpixel segmentation as a rule yields over-sectioned results instead of item level fragments, it radically diminishes the quantity of picture primitives with insignificant loss of data and offers a simple approach to separate the probably picture objects with as few portions as could be expected under the circumstances. Likewise, since superpixel segmentation gives a more characteristic and perceptually significant representation of the info picture, it is more helpful and powerful to concentrate area based visual elements utilizing superpixels.

IV. METHODOLOGY

A.Training

Training a neural network typically consists of two phases:

- A forward phase, where the input is passed completely through the network.
- A backward phase, where gradients are backpropagated (backprop) and weights are updated.

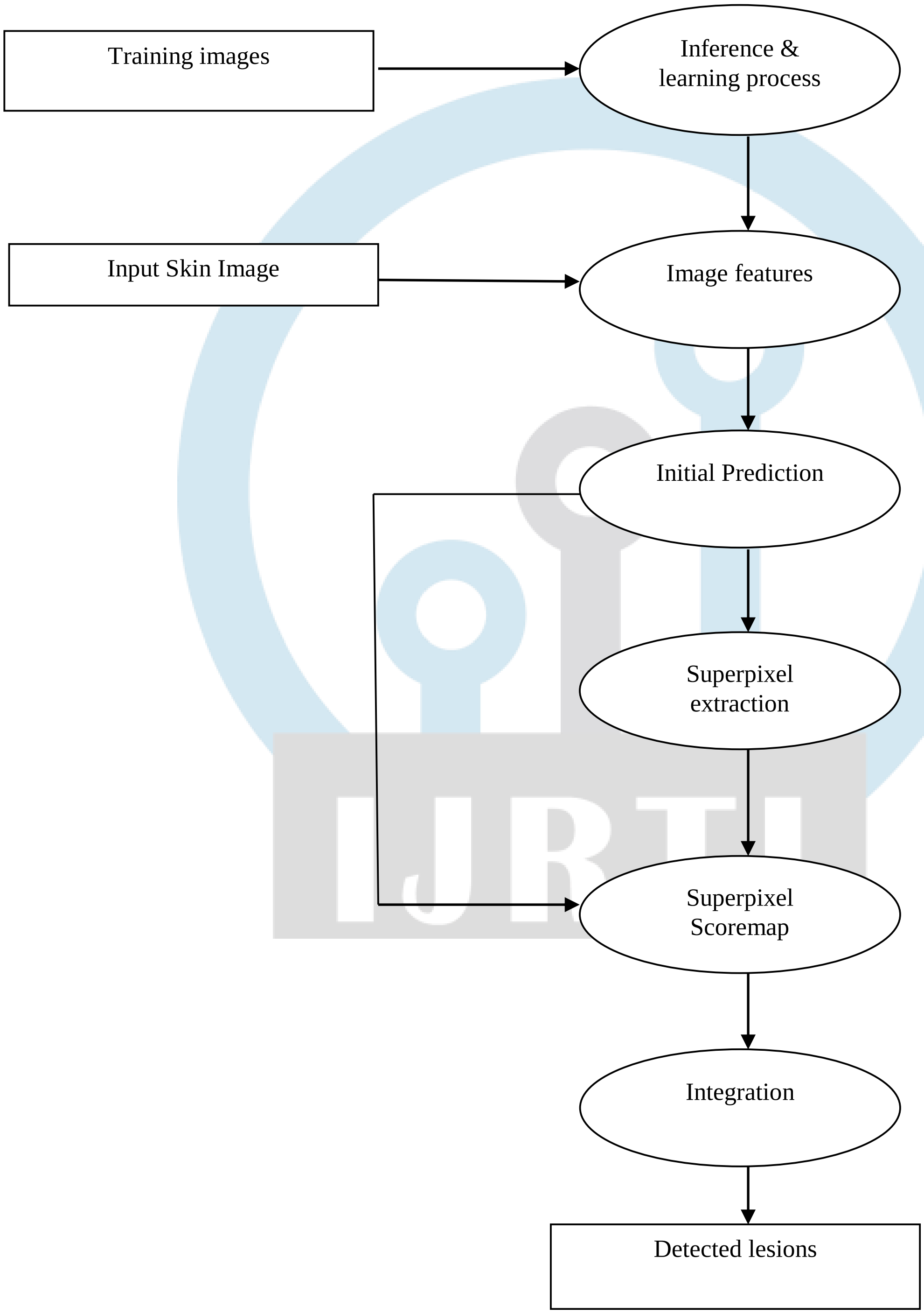


Fig 2: Process flow

During the forward phase, each layer will cache any data (like inputs, intermediate values, etc) it'll need for the backward phase. This means that any backward phase must be preceded by a corresponding forward phase.

During the backward phase, each layer will receive a gradient and also return a gradient. It will receive the gradient of loss with respect to its outputs (∂L / ∂out) and return the gradient of loss with respect to its inputs (∂L / ∂in).

B. Initial segmentation

A fully convolutional network can be trained end-to-end and performs pixels-to-pixels mapping for semantic segmentation. Fully convolutional versions of existing networks predict dense pixel-wise outputs from arbitrary sized input images. Both learning and inference are performed for whole-image in an end-to-end manner through dense feed-forward computation and backpropagation. Typical image recognition nets. It take fixed-sized inputs and produce non-spatial outputs. The fully connected layers of these nets have fixed dimensions. However, these fully connected layers can also be viewed as convolutions with kernels that cover their entire input regions. Doing so casts them into an FCN that takes any image size as input and outputs a pixel-wise label prediction map. In the proposed method, it adopts the FCN architecture, the proposed system includes a deconvolution layer to upsample the last convolution layer predictions back into the image pixels. As skin lesions can potentially be imaged from a variety of camera rotations, it augment the training data by rotating the images with respect to the lesion centroid.

C. Superpixel based fine-tuning

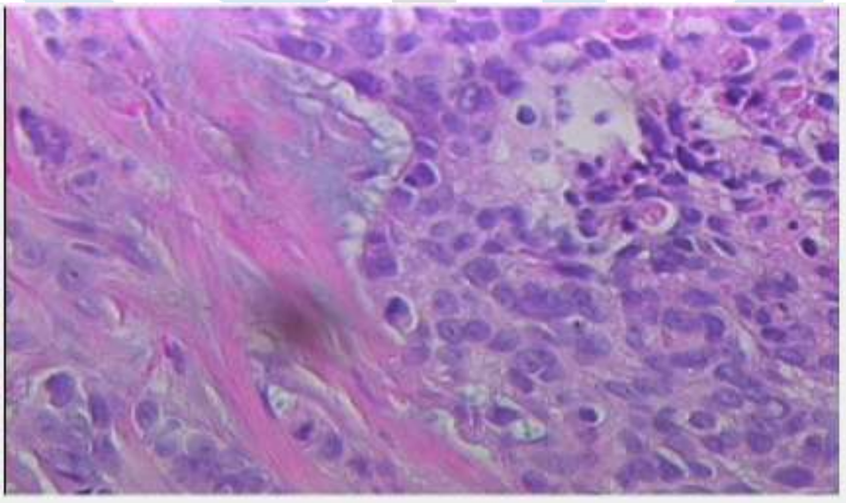
Since the output segmentation score map obtained by FCN is not homogeneous, it formulate the pairwise relationship between neighboring image patches, by imposing the graph structure. The prediction image pixels scores from a fully convolutional network provide a region of interest (ROI) that shows the probabilities of image pixels or superpixelsbelonging to the lesion. In our proposed method, the main idea is to propagate the label of those highly confident image patches (superpixels) and infer the labels of the less confident superpixels based on the similarities between each two superpixels.

D. Fusion

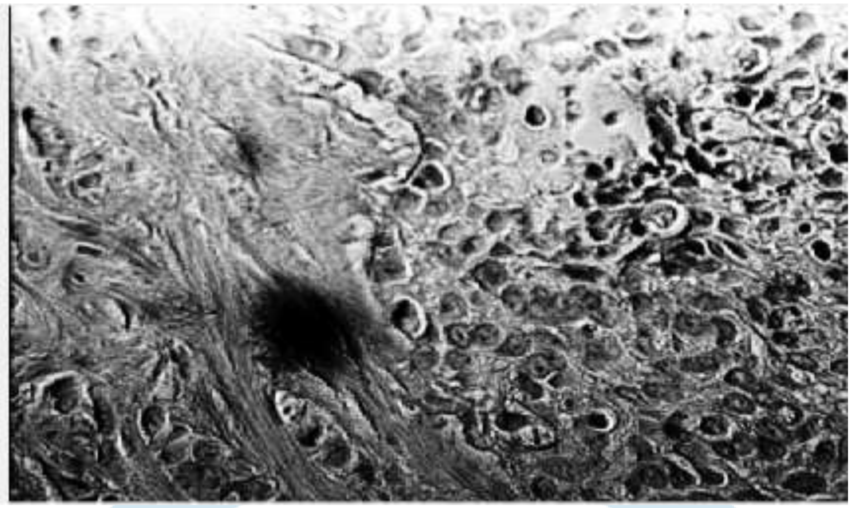
In the proposed method, both global and local segmentations models are exploited. Since the local lesion confidence map based on the image superpixels tends to recognize fine details such as the lesion fuzzy boundary, and the coarse lesion confidence map obtained by FCN concentrates on the lesion's overall shapes, these two are combined to generate the final saliency map The final prediction map is then binarized using the Otsu threshold to obtain the final segmentation result.

V. RESULTS

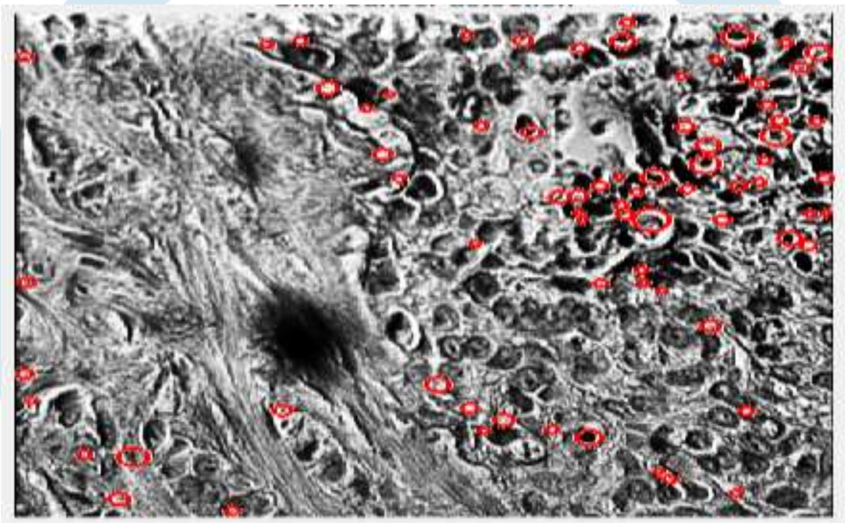
The following are the snapshots of results on execution of the program considering different cases are shown below. The snapshot in figure 1 shows the input skin image. The figure 2 shows the detection process. In next figure 3, shows the segmentation of lesions from the input skin image.



a) Input Image



b) Detection process



c) Segmentation process

CONCLUSION

The proposed system is a skin lesion segmentation algorithm using a deep convolution network uided by superpixels fine-tuning. It uses a superpixel-based fine-tuning strategy to effectively utilize the characteristics of the skin image pixels to accurately extract the border of the lesion. It not only learns a global map for skin lesions, but also acquires the local contextual information, such as lesion boundary. It can, therefore, accurately segment lesions within a given skin image, even in the presence of fuzzy boundaries and complex textures. It use a fully convolution neural network to obtain the coarse global segmentation map of the lesion. The local lesion map is then obtained by a superpixel-based ranking strategy on the region defined by the global segmentation map. The local lesion map and the global maps are then combined together to generate the final segmentation result.

REFERENCES

[1] R.L. Siegel, K.D. Miller, and A. Jemal. "Cancer statistics, 2016," CA: A Cancer Journal for Clinicians. vol. 66, no. 1, pp. 7-30, 2016.
[2] H. W. Rogers, M. A. Weinstock, S. R. Feldman, B. M. Coldiron. "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012." JAMA Dermatol. vol. 151, no. 10, pp. 1081-1086, 2015.
[3] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder. "Diagnostic accuracy of dermoscopy," Lancet Oncol. vol. 3, no. 3, pp. 159165, 2002.
[4] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images," in Proc. MICCAI, 2015, pp. 118-126.
[5] N. C. F. Codella, D. Gutman, M. Emre Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, A. Halpern. (Oct. 2017) "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)," [Online].
[6] J. Weese, and C. Lorenz. "Four challenges in medical image analysis from an industrial perspective," Med. Image Anal., vol. 33, pp. 44-49, Oct. 2016.

[7] Y. Song, W. Cai, H. Huang, Y. Zhou, D. D. Feng, Y. Wang, M. J. Fulham, and M. Chen. "Large Margin Local Estimate With Applications to Medical Image Classification," IEEE Trans. Med. Imag., vol. 34, no. 6, pp. 1362-1377, Jun. 2015.
[8] Y. LeCun, Y. Bengio, and G. Hinton. "Deep learning," Nature, vol. 521, pp. 436-444, 2015.
[9] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 1717-1724.
[10] Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1285-1298, May 2016.