CAD-Based Machine Learning Project for Reducing Human-Factor Related Errors in Medical Image Analysis

Adegun AA, Adebiyi MO, Ogundokun RO, and Asani EO

Department of Computer Science, Landmark University, Omu- Aran Kwara-State. Nigeria.

Abstract. *Machine learning technique such as deep learning methods have produced promising results in medical images analysis. This work proposes a user-friendly system that utilizes deep learning techniques for detecting and diagnosing diseases using medical images. This includes the design of CAD-based project that can reduce human factor related errors while performing manual screening of medical images. The system accepts medical images as input and performs segmentation of the images. Segmentation process analyses and identifies the region of interest (ROI) of diseases from medical images. Analyzing and segmentation of medical images has assisted in diagnosis and monitoring of some diseases. Diseases such as skin cancer, age-related fovea degeneration, diabetic retinopathy, glaucoma, hypertension, arteriosclerosis and choroidal neovascularization can be effectively managed by the analysis of skin lesion and retinal vessels images. The proposed system was evaluated on diseases such as diabetic retinopathy from retina images and skin cancer from dermoscopic images.*

Keywords: Human actor, CAD, Machine learning, deep learning, Project, Diseases, Encoder, Decoder, Images, Segmentation.

1   Introduction

Automatic segmentation of the medical images is important in the detection of a number of diseases (Sharma, Anchal, & Shaveta Rani, 2016). Recently state-of-the-arts techniques such as deep learning have been applied in this segmentation processes. Computer-aided detection (CADe) and diagnosis

(CAD) project for analysis medical images has evolved and is rapidly growing (Doi, K., & Huang, H. K. ,2007). This includes the design of CAD techniques that ease the rigorous task of manual screening of medical images which is susceptible to human errors in disease diagnosis. Machine learning technique such as deep learning methods have produced promising results in medical images analysis. It has been shown that image analysis and segmentation carried out based on deep learning methods has produced improved results with a very high accuracy percentage as against the manual screening method that is characterized with human factor errors (Abràmoff, M. D., Garvin, M. K., & Sonka, M. ,2010). This chapter proposes a human-friendly project that utilizes deep learning techniques for detecting and diagnosing diseases using medical images. The system accepts medical images as input and performs segmentation of the images. Segmentation process analyses and identifies the region of interest (ROI) of diseases from medical images. The proposed system was evaluated on diseases such as diabetic retinopathy from retina images and skin cancer from dermoscopic images. Retina vessels images and dermoscopic images datasets were used to test and evaluate the performance of the system. The output gave a promising result. This paper proposes better human interfacing tools for the proposed project in the future work.

2   Related works

In the last decade, there have been a lot of research about the application of deep learning to medical image analysis. Some works have been particularly carried out in the segmentation process of medical image analysis state-of-the arts techniques. The performance of the deep learning projects has been compared with manual approach with so much human factors related errors. This section performs the review of related works in this aspect.

Deep learning method was utilized for detection and segmentation of colorectal liver metastases by (Vorontsov et al., 2019). They applied three-dimensional automated segmentations to resolve deficiencies of fully automated segmentation for small metastases and it was faster than manual three-dimensional segmentation. They compared the performance of fully automated and user-corrected segmentations with manual segmentations. Chen, L., Bentley, P., & Rueckert, D. (2017) proposed framework to automatically segment stroke lesions images. The framework was made up of two convolutional neural networks to evaluate the lesions detected in order to remove potential disease.

Vesal, S., Ravikumar, N., & Maier, A. (2018) proposed a convolutional neural network (CNN) project called SkinNet that employed dilated and densely block convolutions to incorporate multi-scale and global context information for skin

lesion segmentation. Baur, C., Wiestler, B., Albarqouni, S., & Navab, N. (2019). combined the advantages of supervised and unsupervised methods into a novel framework for learning from both labeled & unlabeled data for the challenging task of White Matter lesion segmentation in brain MR images. They proposed a semi-supervised setting for tackling domain shift which is a known problem in MR image analysis. Chlebus et al., (2018) developed a fully automatic method for liver tumor segmentation in CT images based on a 2D fully convolutional neural network with an object-based post-processing step. The system was compared with human performance.

CNN architecture was used for blood vessel segmentation of fundus images (Xiancheng, et al.,2018) The architecture was derived from U-Net architecture that implemented an encoder-decoder architecture. A trainable Frangi-Net project was developed to perform segmentation of retinal images (Fu et al.,2017). A multi-scale convolutional neural network structure and label processing approach was applied for retinal vessel segmentation (Li et al., 2017). The method used two different scale image segments to generate input for two deep convolutional networks. Lastly, deep neural network (DNN) that uses max-pooling layers (MPCNNs) instead of subsampling or down sampling was also used for retina images segmentation (Melinščak, M., Prentašić, P., & Lončarić, S. 2015). The MPCNN mapped input samples into output class probabilities using hierarchical layers to extract features and fully connected layers to classify extracted features.

3 Methodology

a. Dataset

We evaluate our method on some publicly available databases namely:

i. Retinal image vessel segmentation dataset: DRIVE. The DRIVE (Digital Retinal Images for Vessel Extraction) (Staal et al.,2004) dataset consists of 400 color fundus photographs. The set of 400 images has been divided into a training set and a test set, both containing 200 images.

ii. Skin lesion images ISIC dataset: The dataset used is repository in ISIC 2018 archive (Codella et al., 2017). The dataset contains 2600 samples in total. It consists of 2000 training samples in JPEG format and 600 testing samples. The original size of each sample was 700x900, which was rescaled to 256x256 for this implementation. The training samples include the original images and the corresponding ground truth label in PNG format.

b. Model Implementation

The software used for the model implementation includes:
- Python Version 3.5
- Keras 2
- Tensor flow backend
- Scikit-image Version 0.14.1

Running Environment

The experimental platform is a PC equipped with an Intel Core i7 processor with ten (10) 3.4GHZ cores, 16GB memory and NVIDIA Tesla K40c GPU, and the operating system is Ubuntu 16.04.

c. General Architecture of the proposed system

The proposed project utilizes Deep Convolutional Encoder-Decoder Model Architecture which adapts the popular U-Net model. The U-Net model implements a contemporary classification network into fully convolutional networks. It operates by learning from pixels to pixels in an end-to –end manner. Training image dataset and corresponding ground truth image dataset serves as the input into the system. These will be taken through the other sections of the system such as data preparation section and the learning and training section. Supervised form of training is employed here whereby the target goal in the form of the ground truth images is supplied together with the input images. After a period of training, the validation image data set will now be fed into the system for the expected predictions. The output data from the prediction will finally be compared with our target goals for accuracy and the performance metrics evaluated.

The architecture is composed of the following major sections:

i. Data Preparation section

The training image data set and the corresponding ground truth image data set are used to train the model simultaneously. In this section, the images are first resized into the appropriate image size. The image sets are then resampled to remove every form of noise. The ground truth serves as the expected output and the training set serves as the input to train the system.

The sample training image data set and ground truth images for both the retina images and the skin lesion are illustrated in Fig 1, 2, 3 and 4 below.

Fig. 1. A Sample Skin Lesion Training Image Data Set

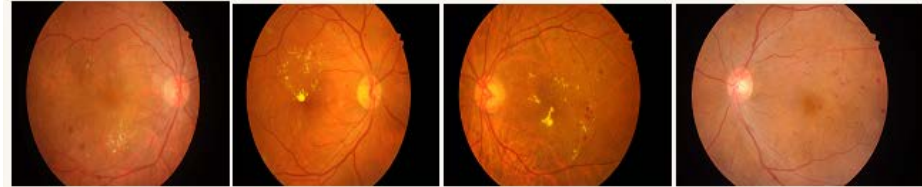**Fig. 2.** The Corresponding Skin Lesion Ground Truth images

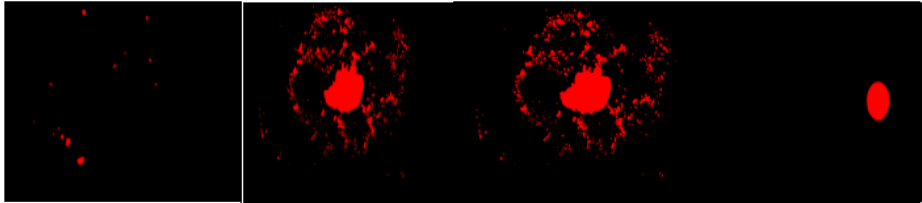**Fig. 3.** A Sample Retina Training Image Data Set

Fig. 4. The Corresponding Retina Ground Truth images

ii.   Learning and Training Section:

In this section, we have the encoder units and the decoder units.

Encoder:

The encoder is made up of set of convolution layers and pooling layers. Features extraction from the input image takes place at the convolution layer and the pooling layers reduce the resolution of the image feature maps. In this model, the encoder applies rectified linear unit (RELU) activation function.

Each encoder uses the Max-pooling to translate invariance over small spatial shifts in the image and combine this with Subsampling to produce a large input image context in terms of spatial window. This method reduces the feature map size and this leads to image representation that is noisy with blurred boundaries. The restoration is done by decoder as the output image resolution must be the same as input image.

Decoder:

The decoder ensures that the image resolution of image set from the encoder units is increased to the initial resolution status. It is also made up of set of convolution layers. Each of the layers in the decoder stage corresponds to the layer in the encoder i.e. for each of the encoders there is a corresponding decoder which up samples the feature map using the already stored max-pooling indices. Sparse feature maps of higher resolutions are then produced. These are fed through the training section to produce dense feature maps.

Prediction:

The prediction section of this model is performed by predicting pixel-wise labels for an output which has the same resolution as the input image. The last part of the decoder is connected to a softmax classifier which classifies each pixel.

The sparse feature maps restored from the original resolution are then fed to the softmax classifier to produce the final segmentation.

Output:

The final segmented images are generated at this section. These are stored with preds as the output name. The performance of the model is also evaluated at this section. The model is described by the diagram in Fig 5

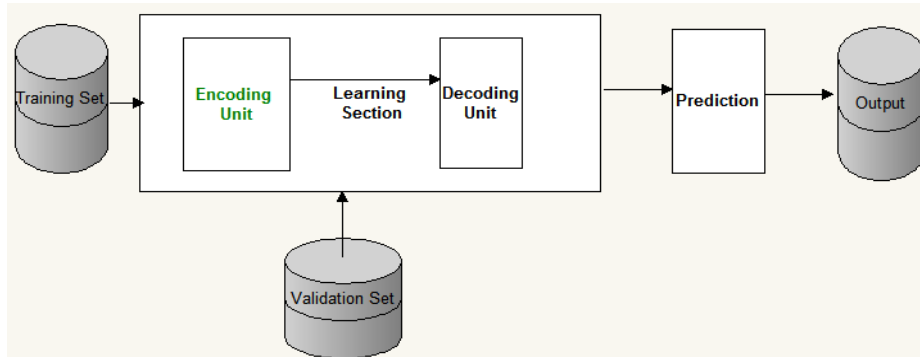The general layout of encoder-decoder architecture is illustrated with Fig 5

Fig. 5. General Layout Diagram of A proposed segmentation method for Medical Images analysis

d. Experimental Results and Analysis

The performance of the proposed model was evaluated and the results displayed below. The segmentation accuracy has been assessed by comparing the predicted results with the manual version. The final results of the validation image set used for testing segmentation are compared with the ground truth.

The metrics can be described as stated below:

• Intersection of Union: also known as Jaccard similarity coefficient, it can be described as similarity measure over two image sets.

• Sensitivity: also known as true positive rate (TPR) Sensitivity: This proportion of actual positives which are predicted positive. It can be defined as the proportion of examples which were predicted to belong to a class with respect to all of the examples that truly belong in the class.

• Dice Coefficient: also shows similarity measure and it is related to Jaccard index

Table 1. PERFORMANCE EVALUATION OF THE PROPOSED MODEL
BASED ON THE SPECIFIED METRICS EARLIER DISCUSSED

Performance Metrics	Dice Coefficient (%)	Intersection of Union (%)	Positive Predictive value (%)	Sensitivity (%)
Proposed model	82.23	72.23	96.5	96.5

Evaluation was done using Dice's coefficient and some other metrics, which are calculated as:

$$\text{Dice } (A, B) = \frac{|A \cap B|}{(|A| \cap |B|) \div 2} \qquad (1)$$

$$\text{PPV } (A, B) = \frac{|A \cap B|}{|A|} \qquad (2)$$

$$\text{SENSITIVITY} = \frac{|A \cap B|}{|B|} \qquad (3)$$

$$\text{IOU } (A, B) = \frac{|A \cap B|}{|A| \cap |B|} \qquad (4)$$

Where A denotes the segmented region and B denotes the manually labelled region, $|A \cap B|$ denotes the overlap area between A and B, AND $|A|$ and $|B|$ represent the areas of A and B respectively.

It can be shown that the results of the performance evaluation of the proposed model as stated above gave a reasonably high percentage of similarity and low level of diversity between the predicted results and ground truth results.

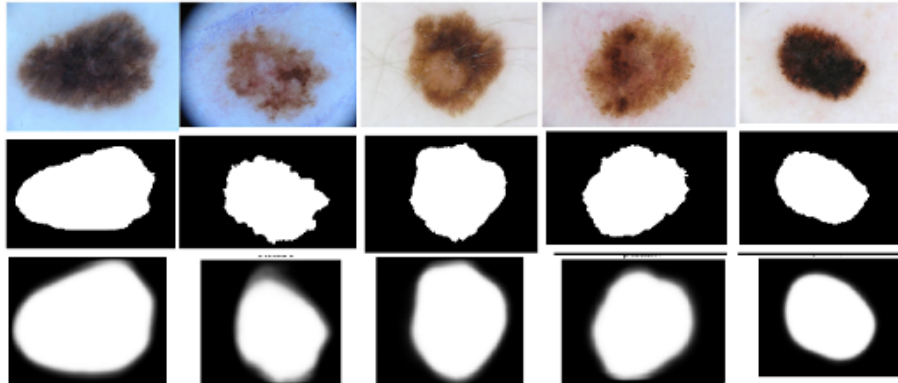The final segmented output for the skin lesion images is presented in Fig 6 below.

Fig. 6. The figure contains the project validation image set followed by the ground truth labels and the predicted skin lesion segmented images results from the proposed model.

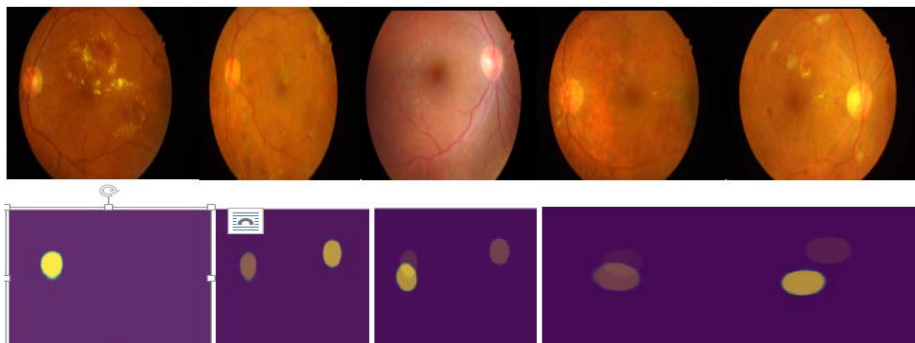The final segmented output for the retina images is presented in Fig 7 below.

Fig. 7. The figure contains the project validation image set and the predicted segmented retina images results from the proposed project model.

e. Conclusion

In this work, a CAD-Based Machine Learning Project for Reducing Human-factor Errors in Medical Image Analysis is proposed. A novel deep learning model for segmentation of medical images in order to detect diseases such as melanoma skin cancer, diabetic retinopathy, glaucoma, hypertension has thus been investigated. The review of the existing works was also carried out. The proposed model was evaluated and analyzed. This paper shows that the proposed project model gave a better performance over the manual approach of medical image analysis. This is as a result of errors due to human factors that has been eliminated in the automated approach. The encoder–decoder mechanism in this model can be improved on in the future works where ensemble methods with some other state-of-the arts techniques can also be explored. A more human friendly system can be developed in the future works.

References

Abràmoff, M. D., Garvin, M. K., & Sonka, M. (2010). Retinal imaging and image analysis. IEEE reviews in biomedical engineering, 3, 169-208.

Almotiri, Jasem, Khaled Elleithy, and Abdelrahman Elleithy. "Retinal Vessels Segmentation Techniques and Algorithms: A Survey." Applied Sciences 8, no. 2 (2018): 155.

Baur, C., Wiestler, B., Albarqouni, S., & Navab, N. (2019, May). Fusing Unsupervised and Supervised Deep Learning for White Matter Lesion Segmentation. In International Conference on Medical Imaging with Deep Learning (pp. 63-72).

Chen, L., Bentley, P., & Rueckert, D. (2017). Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. NeuroImage: Clinical, 15, 633-643.

Chlebus, G., Schenk, A., Moltz, J. H., van Ginneken, B., Hahn, H. K., & Meine, H. (2018). Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. Scientific reports, 8(1), 15497.

Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... &

Doi, K., & Huang, H. K. (2007). Computer-aided diagnosis (CAD) and image-guided decision support. Computerized Medical Imaging and Graphics, 4(31), 195-197.

Fu, W., Breininger, K., Würfl, T., Ravikumar, N., Schaffert, R., & Maier, A. (2017). Frangi-Net: A Neural Network Approach to Vessel Segmentation. arXiv preprint arXiv:1711.03345.

Halpern, A. (2018, April). 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). In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 168-172). IEEE.

Li, M., Ma, Z., Liu, C., Zhang, G., & Han, Z. (2017). Robust retinal blood vessel segmentation based on reinforcement local descriptions. BioMed research international, 2017.

Melinščak, M., Prentašić, P., & Lončarić, S. (2015, January). Retinal vessel segmentation using deep neural networks. In 10th International Conference on Computer Vision Theory and Applications (VISAPP 2015).

Sharma, A., & Rani, S. (2016, April). An automatic segmentation & detection of blood vessels and optic disc in retinal images. In 2016 International Conference on Communication and Signal Processing (ICCSP) (pp. 1674-1678). IEEE.

Staal, J., Abràmoff, M. D., Niemeijer, M., Viergever, M. A., & Van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. IEEE transactions on medical imaging, 23(4), 501-509.

Vesal, S., Ravikumar, N., & Maier, A. (2018). SkinNet: A Deep Learning Framework for Skin Lesion Segmentation. arXiv preprint arXiv:1806.09522.

Vorontsov, E., Cerny, M., Régnier, P., Di Jorio, L., Pal, C. J., Lapointe, R., ... & Tang, A. (2019). Deep Learning for Automated Segmentation of Liver Lesions at CT in Patients with Colorectal Cancer Liver Metastases. Radiology: Artificial Intelligence, 1(2), 180014.

Xiancheng, W., Wei, L., Bingyi, M., He, J., Jiang, Z., Xu, W., ... & Zhaomeng, S. (2018). Retina blood vessel segmentation using a U-net based Convolutional neural network. In Procedia Computer Science: International Conference on Data Science (ICDS 2018), Beijing, China, 8-9 June 2018.