A Machine Learning project for breast cancer detection and diagnosis employing multi-layer perceptron neural network(MLP) and Convolutional neural network(CNN) By implementing the Keras library

Fateme Amouei
4013074508
Master's student in computer science
Direction of decision sciences and knowledge engineering
Kharazmi University, Karaj branch, Faculty of Mathematics and Computer Science
machine learning
Dr. Keivan Borna
Winter 1401

*My reference is this article:

An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN)" Meha Desai, Manan Shah; 1-Department of Electronics and Instrumentation Engineering, School of Electronics and Electrical Engineering, Vellore Institute of Technology, Tamil Nadu, India, 2-Department of Chemical Engineering, School of Technology, Pandit Deendayal Petroleum University, Gandhinagar, Gujarat, India"

You can see my Project and Database and images in Data file with my git: https://github.com/Sofia-Amouei/ML-Methods-on-breast-cancer-detection-employing-MLP-CNN-with-Keras-

1-Abstract

This paper aims to review Artificial neural networks, Multi-Layer Perceptron Neural network (MLP) and Convolutional Neural network (CNN) employed to detect breast malignancies for early diagnosis of breast cancer based on their accuracy in order to identify which method is better for the diagnosis of breast cell malignancies. Deep comparison of functioning of each network and its designing is performed then analysis is done based on the accuracy of diagnosis and classification of breast malignancy by the network to decide which network outperforms the other. CNN is found to give slightly higher accuracy than MLP for diagnosis and detection of breast cancer. There still is the need to carefully analyse and perform a thorough research that uses both these methods on the same data set under same conditions in order identify the architecture that gives better accuracy. By implementing the methods of this article (CNN, MLP and Keras with different Python language libraries) on different test samples, I achieved close accuracies such as 96.49

2-Introduction

Medical science and health research are essential for the survival of all species. It includes researches and information of various diseases, medications, risks and most importantly diagnosis and treatments. Integrating and expanding technology in the field of medicinal science is of prime importance for increasing the capacity and accuracy of disease diagnosis, disease trend and other factors like treatment1,11. In 2018, WHO recorded 9.6 million deaths due to cancer, out of which 6,27,000 deaths were caused due to breast cancer. WHO also suggests early diagnosis as a strategic method for treatment and cure of breast cancer. Thus, The application of Neural Network in Breast cancer detection and diagnosis proves to be of immense importance in the field of Oncology. Artificial neural network (ANN) is a sophisticated system that functions closely like the human brain and its nervous system.59,43,70,16,26 As the brain is also a self-learning organ, so is the artificial neural network system. It performs thousands of iterations and learning to predict outputs based on them. The various nodes of the ANN, are analogous to that of neurons in the human brain.55,54,49,75 Each node output is taken as the input for the next node after adding a weight function to it.55,54,53 The learning rate can be modulated for a better accuracy of outputs. For a greater accuracy of output, ANNs perform back propagation, which means it basically performs one round of iterations with a certain set of weights and then it propagates in the backward direction and reduces the errors found in those weights, such that the accuracy of the output after several such forward and backward propagations is extremely high and reliable.58,57,46 For complex analysis of data and non-linearity between the input and the predicted output, ANN is a powerful tool of execution of various predictive-output tasks.76,6,48,50 ANN is a reliable and accurate source for solving problems that have no simple algorithmic solutions35 like disease prediction and detection. Healthcare has beamingly started applying ANN to various disease detections, analysis and predictions. Analysis of waveforms and signals like that of the Electrocardiogram can be performed using the Kohenen self-organising maps that work based on ANN.52,55,54 The output of the mapping helps in the interpretation of the waveform.12,35. An ANN was also developed for Ophthalmology, that captures images of eye as its input data set and interprets the dysfunctionalities.27,35 In cardiology, cardiac imaging is used as the input data set for Artificial intelligent network to diagnose cardiac disorders.

3-Compare

Comparison of accuracy of breast cancer detection from researches that used CNN or MLP architecture for breast cancer diagnosis and classification:

References	Architecture	Accuracy
Dabeer et al.8	CNN	99.86height
Sanap and Agrawal	MLP	Benign = 72.5 percent, Malignant = 89.00 percent
Chtihrakkannan et al	MLP	Average: 89.77percent, Highest: 96percent
Kathija et al	MLP	WBCD dataset: 98.99percent
Guan and Loew	CNN	90.5 ± 3.2
Nrea et al	CNN	91.86percent
Mojarad et al	MLP	65.21percent
Kiyan and Yildirim	MLP	98.8percent
Bardou et al.	CNN	between 96.15percent and 98.33percent
Iesmantas and Alzbutas	CNN	Cross validation accuracy of 87percent

4-Main Part

I researched the issue of diagnosis of this disease in three parts using CNN and MLP model and Cress library and using the data that I put on my GitHub to implement them in a way that Git notebooks done, paid And I spent about 3 months researching and investigating and implementing this project.

1-Neural network for breast cancer diagnosis and classification:

Application of neural networks in breast cancer detection has a major advantage over traditional methods in terms of time taken for examination. Where conventional methods take up a large amount of time in examination of one data at a time, ANN examines a large amount of data after a short training period. ANNs predict outputs with a high accuracy and are easy to code. Whereas traditionally, predicting outputs in medicine takes years of experience and knowledge in the particular field. Plus, various types of ANNs can be developed in order to diagnose breast cancer, which broadens the horizons for earlier and easier breast cancer detection. A Multilayer Perceptron Neural Network is a Neural Network of the feed forward type. It uses the Backpropagation technique for learning. It has an input layer of neurons that act as receivers, one or more hidden layers of neurons that compute the data and undergo iterations and then the output layer which predicts the output. Ayer et al.2 discussed a three-layer ANNs that examines256 mammogram images and classifies them. Micro calcification of samples into size ranges of 50-250 lm, 100-500 lm, 200-1,000 lm, and 400-2,000 lm was used to formulate the network. The classification based of malignancy size range was achieved using four ANNs. The ML-NN was found to be 84 percent sensitive at 75 percent specificity. Ting and Sim38 implemented a supervised multilayer neural network in classification of breast cancer in order to assist medical examiners. They considered 170 mammograms that have already been tested of malignancies as image input set. Upon using these in an ML-NN, the neural network successfully classified the breast malignancies. Based on the intensity of malignancy, the three classifications were, benign, malignant and normal. When compared with the known malignancy of the input data used, they found the accuracy of the ML-NN to be 90.59 percent and its sensitivity to be 90.53 percent. Fogel et al compared two construc-

tions of ML-NN for diagnosis of breast cancer by conducting two experiments using 400 image sets. Meinel et al.28 used a back propagation neural network to create a Computer Aided Diagnosis system that enhances Magnetic Resonance images of breasts. 80 lesions were used for the procedure, out of which 43 were malignant and 37 were benign. Image parameters for the NN were fixed and 13 input features were chosen from 42 different ones. The network was trained to distinguish and enhance the lesion from the background. They found that CAD systems improve the MRI by enhancing it, making it easier and increasing efficiency to detect disorders in it. Cedeño et al78 devised a neural network called, Artificial Metaplasticity Multi-Layer Perceptron. Artificial metaplasticity was used in order to minimise error. It was implemented in a Multilayer perceptron neural network to detect breast malignancy. The network was trained using 410 samples of breast images, of which 144 were malignant and 266 were benign. To test the system, 233 samples consisting of 95 malignant and 178 benign samples were used.The AMMLP network output accuracy and sensitivity was compared to the conventional methods of CAD that use Back propagation NN. The result was found to have 99.26 percent accuracy, 97.89 percent specificity and 100 percent sensitivity, proving it to be better than conventional back propagation NN.

2-Convolutional neural network for breast cancer diagnosis and classification:

Convolution Neural Network (CNN) is a neural network class mostly employed to examine, identify or classify images as it simplifies the images for better analysis. This network is advantageous as it needs fewer human efforts and pre-processing. Back propagation is also included in the learning process to make the network more accurate. In terms of design, it is closely related to MLP, as it consists of an input layer of neurons, multiple hidden layers and an output layer. Each neuron in one layer is connected to every neuron in the succeeding layer. The image (say, a flower) to be classified or analysed is passed through multiple layers. The convolution layer is used to filter the image upon performing convolution in order to enhance the features. Then, pooling layer down samples (reduces the sample size) of the sample of features extracted. This makes the processing faster, as the parameters decrease. Max pooling takes the maximum quantity obtained upon

pooling and Average pooling takes the average of the pooling output. In CNN, an activation function called Rectified Linear Unit is used to ensure non linearity. Then after passing through a fully connected layer, the output is predicted into classes (for classifying a flower, the classes may be daisy, windflower, sunflower, pansy).This network frame is increasingly being used in image processing as it solves the signal transition problem and accurately extracts features. The CNN architecture has been employed in breast cancer diagnosis and classification because of its ability of feature extraction that can be used to enhance and easily see malignancy in breast masses. Thus, aiding in the process of early detection of breast cancer, so that it can be treated at a lower stage, before it spreads more. Wang et al.39 devised a Max-pooling based convolution neural network to detect mitosis in breast images. The examination was performed on a small dataset. They considered a data set of 50 images and used a CNN with a series of convolution and pooling layers. 2D convolution was performed with a rectangular filter. This network was trained to pixelate the image and classify each pixel in it.

3-Comparison of MLP and CNN:

Übeyli compared the classification accuracies for breast cancer for different ANN algorithms namely multilayer perceptron neural network (MLPNN), combined neural network (CNN), probabilistic neural network (PNN), recurrent neural network (RNN) and support vector machine (SVM) in order to identify the best and the optimum diagnostic approach. The Wisconsin breast cancer dataset was studied and 683 records were analysed under each algorithm. Malignancy was confirmed by performing biopsy on the breast tissue. Between MLP and CNN, MLP had a sensitivity of 91.19 percent, specificity of 92.34 percent and accuracy of classification was 91.92 percent, whereas, CNN showed much higher rates of sensitivity (96.86 percent), specificity (87.81 percent) and accuracy (97.46 percent). This showed that CNN was better than MLP as a breast cancer detection algorithm. At 30 features are extracted from each image, 100 batch size, 0.03 learning rate and 1000 epoch size, the accuracy of CNN was the highest at 96.49 percent, while the highest MPL achieved was 97.891 at five layers,100 training time and 0.03 learning rate. This again shows that CNN is better at breast cancer detection than MLP.

5-Conclusion

The healthcare system is an integral part of the society that ensures that every individual gets effective diagnosis and treatment. It also performs researches to combat new diseases, viruses and other ailments. Health is a crucial factor for deciding the capabilities of any individual and thus, a healthy life is a need for everyone. The field of medical science aids this ubiquitous need of good health. In this technologically advancing, fast paced world, integrating new technology in the field of healthcare has become essential and inevitable. Technology helps in making ailment diagnosis, treatment, medicine prescription, etc more efficient and less time consuming. It also reduces the need of trained work force. Thus, implementing various technologies in healthcare is a crucial step in advancement of medical sciences. This paper examines various approaches that use ANN to diagnose breast cancer, and compares Multi-Layer Perceptron Neural Network and Convolutional Neural Network based on their accuracy of diagnosis and breast cancer classification. Convolutional Neural network is explained in terms of architecture and its working. Then various researches that use CNN for breast cancer detection are examined. Then, Multi-Layer Perceptron Neural Networks are explained with its architecture and working, followed by an elaborate examination of various researches that use MLP-NN for breast cancer diagnosis and classification. Convolutional neural network has convolution and pooling layers that make it the better choice for complex image classifications, however, upon comparison from examination of various experiments, CNN shows higher accuracy than MLP in diagnosing and classifying breast cancer cells. Majority of results propose that CNN gives a high accuracy as various layers like convolutional layer, pooling layer, and fully connected layer are incorporated. The numbers of hidden layers need to be minimised and using artificial metaplasticity in MLP can help minimise error. MLP is closely as efficient as CNN, but CNN shows results with higher accuracy. Yet, the most accurate result can be obtained upon elaborate experiment of breast cancer diagnosis and classification using same dataset in both CNN and MLP-NN under similar conditions of testing and training. But,implementing ANN in breast cancer diagnosis and classification is the need of the hour, in order to reduce work load on doctors who each have to perform diagnosis on several patients per day,improve efficiency and also help women diagnose breast cancer themselves safely by extending this technology to make safe and handy mobile applications.

6-References

1. Araujo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, et al. Classification of breast cancer histology images using Convolutional Neural Networks. PLoS ONE. 2017;12(6). https://doi.org/10.1371/journal.pone.0177544.
2. Ayer T, Chen Q, Burnside ES. Artificial Neural Networks in Mammography Interpretation and Diagnostic Decision Making. Comput Math Methods Med. 2013;2013:1-10. https://doi.org/10.1155/2013/832509.
3. Bengtsson E, Malm P. Screening for Cervical Cancer Using Automated Analysis of PAP-Smears. Comput Math Methods Med. 2014;2014:1-12. https://doi.org/ 10.1155/2014/842037.
4. Chen Y, Wang Y, Yang B. Evolving Hierarchical RBF Neural Networks for Breast Cancer Detection. Lect Notes Comput Sci. 2006;137-144.
5. Chtihrakkannan R, Kavitha P, Mangayarkarasi T, Karthikeyan R. Breast Cancer Detection using Machine Learning. International Journal of Innovative Technology and Exploring Engineering.. 2019;8(11):3123-3126.
6. Chunekar VN, Ambulgekar HP. Approach of Neural Network to Diagnose Breast Cancer on three different Data Set2009 International Conference on Advances in Recent Technologies in Communication and Computing. https://doi.org/10.1109/ artcom.2009.225.
.....

Thanks for your attention