Performance Comparison Analysis for Medical Images Using Deep Learning Approaches

Neha Adawadkar, Roshani Chavan, Chaitra Patwardhan, Advait Thakur, Prof. Dr. Mrs. Suhasini Itkar

Neha Adawadkar, Dept. of Computer Engineering, PES Modern College of Engineering, Maharashtra, India
Roshani Chavan, Dept. of Computer Engineering, PES Modern College of Engineering, Maharashtra, India
Chaitra Patwardhan, Dept. of Computer Engineering, PES Modern College of Engineering, Maharashtra, India
Advait Thakur, Dept. of Computer Engineering, PES Modern College of Engineering, Maharashtra, India
Prof. Dr. Mrs. Suhasini Itkar, Dept. of Computer Engineering, PES Modern College of Engineering, Maharashtra, India

---***---

Abstract - *The recent pandemic has enforced a lot of stress on our healthcare industry as well as the increase in population has brought to light that the work bestowed upon the healthcare specialists needs to be reduced. Deep learning models such as neural networks are efficient in finding hidden patterns, which assists the experts in the specified field. Medical images like chest x-rays are of utmost importance for the diagnosis of diseases, for instance, brain tumors, dementia, pneumonia, COVID-19, cardiovascular, and many more. Collaborating these medical images along with the deep learning techniques have paved the path for enormous applications leading to the reduction in stress given upon the health sector. In this paper, we propose a brief comparison and study of existing technologies for analyzing chest x-rays using deep learning.*

***Key Words*: Chest X-Ray (CXR), Convolutional Neural Network (CNN), Fully Convolutional Network (FCN), Deep Cascade of Convolutional Neural Network (DCCNN), Lookup based Convolutional neural networks (LCNN), Deep Convolutional Neural Network(DCNN)**

1. INTRODUCTION

Image segmentation is the method of dividing the images into various pixels or regions called as segments and processes images accurately. Diagnosis with the aid of imaging can help in better localization and staging of the disease. Often, Medical Images follow a traditional image segmentation life cycle to analyze them. [1] Highlights values of imaging in the diagnosis of a disease. The analysis of these image segmentation includes pre-processing of image followed by feature extraction and perform classification. These processes are accomplished by using various deep learning methodologies. The deep learning techniques from the following researches have displayed remarkable results for diagnosis of medical images. In the pre-processing stage, the most commonly used technique is normalization which helps in reducing the dimensionality of the image for making the image fit for segmentation. Classification of images can be done using Support Vector Machine (SVM) [3]. Other than that, the data augmentation [4] is used for increasing the robustness of deep learning. Along with these techniques, an Up-sampling method [4] is implemented which helps the network layers to propagate context data to higher resolution layers.

Feature extraction is a technique which is used to reduce a large input data set into relevant features. The images required for segmentation have a large number of variables. A lot of computing resources are required for these variables. Therefore, feature extractions help to attain best feature from those big data sets by selecting and integrating variables into features, effectively decreasing the amount of data. The final step is the classification, where we obtain the desired output. The pre-processed and selected features are provided to the model for training purposes which ultimately gives us the appropriate result. The classification task is a supervised learning approach in machine learning domain. Models such as SVM in [3], U-net in [4], Resnet18, Resnet50 and Xception in [8], LCNN and DCNN in [9], and Fully Convolutional Network (FCN) in [15] are used for this step in the examined research.

2. LITERATURE SURVEY

Over the past few years, the machine learning framework has immensely helped in analyzing the complex medical images. Tremendous advancements in medical image segmentation and its applications have motivated us to study such research techniques. The aim of [1] is to concisely present the issue of coronavirus and highlight the importance of medical imaging in its diagnosis. Inception inspired, a novel deep CNN architecture where depth-wise separable convolutions replace the Inception modules is being discussed in [2]. Here, an architecture named Xception which is similar to Inception V3 is developed which efficiently uses the model parameters. Emphasizing automatic tuberculosis detection [3] proposes a binary classifier using SVM which classifies a CXR as either normal or abnormal achieving an accuracy of 82.1 % and area within the ROC curve as 88.5%. In [4] a U-Net Network is trained with the use of data augmentation by using elastic deformations for effective use of annotated samples. The segmentation results for ISBI cell tracking challenge with average IOU ("intersection over union") for "PhC-U373" data set is 92% and for e "DIC-HeLa" dataset Is 77.5%. [5] In this paper a framework for dynamic sequences of 2-D cardiac Magnetic Resonance Images (MRI) reconstructed from Under-sampled data using DCCNN is proposed. With the help of this model, there is feasibility in the reconstruction of images with good quality through a network having interleaved data consistency stages. It can reconstruct most of the anatomical structures more precisely by previously learned data. In [6] the proposed system repurposes the networks which are trained on image classification. Semantic segmentation is done by using 'Atrous Convolution' by un-sampled filters for feature extraction. We can explicitly control two things: the resolution for feature response computation and effectively amplify the field of view of filters to integrate larger context without the need to increase the number of parameters nor the computational amount. In [7] this model shows that convolutional networks when trained by themselves exceed the existing semantic segmentation. The purpose of this research is to develop a FCN which input the arbitrary size and create a similar-sized output with efficient inference. This model adapts contemporary classification networks into FCN and then passes on their learned parameters by fine-tuning into the segmentation block. In [8] COVID-19 and Thorax diseases, from chest X-Ray images are being diagnosed by the system using residual networks 18, 50 and Xception model. This Resnet18 model gives an accuracy of 94% while resnet50 gives accuracy of 93% and the Xception model is having an accuracy of 95%. In FUIQA [9] LCNN and DCNN are used together for analysis wherein LCNN gives overall accuracy of 92% and the DCNN gives overall accuracy of around 97%. For the image quality assessment [10] CNN model is used to reconstruct all types of images i.e., assessing the quality of the image and also reconstructing it in a better way. The model involves a convolutional layer, maximum and minimum pooling followed by two fully connected layers, and an output node. An optimization process consisting of the network structure, feature learning and regression is integrated, this directs to a more efficient model for judging the image quality. In the model for image segmentation [11] U-net is used for the segmentation of biomedical images. In [12] for cardiovascular magnetic resonance (CMR) images, an automated analysis method is based on FCN. The network is trained and tested on an extensive dataset from the UK Biobank. Several technical metrics, such as the mean contour distance, Dice metric and Hausdorff distance, as well as some medical measures, including end-systolic volume (LVESV), LV mass (LVM); left ventricle (LV) end-diastolic volume (LVEDV) and right ventricle (RV) end-diastolic volume (RVEDV) and end-systolic volume (RVESV) have been used to evaluate the system. The system gives an accuracy of 0.93. In [13] convolutional recurrent neural network (CRNN) model is used for image quality assessment. High-quality CMR images are reconstructed by CRNN. This is done with the help of highly under sampled k-space data with exploiting the dependencies of the temporal sequences as well as the iterative nature of the traditional optimization algorithms. This model gives a minimum time of 3s. In [14] the authors have proposed the use of VGG16 for Breast Cancer Detection giving an overall accuracy of 94.77%. For Brain tumor analysis [15] use of deep neural networks like FCN, Res-net, U-net, Encoder/Decoder, and CNN is done giving an overall accuracy of 75% to 80%. Quality assessment of Echocardiograms in [16] regression model is enhanced to assess the standard and quality of echo images. A stochastic gradient descent algorithm is used to evaluate the loss function of the system. The model was designed by a convolution stage (Convolutional and Pooling Layer) and a fully connected stage. In the paper for image quality assessment [17] basic operations are done to eliminate known distortions from images. Distorted and reference signals are scaled and aligned then metrics of digital values need to convert into pixel intensity using nonlinear pointed transformations. Images are decomposed into channels that are selective for spatial and temporal orientation. Discrete Cosine Transform or Separable wavelet transform methods are used for quality assessment methods. Errors between decomposed references in each channel are normalized so that presence of one image component will decrease another image component proximate in orientation concerning the spatial or temporal location and frequency.

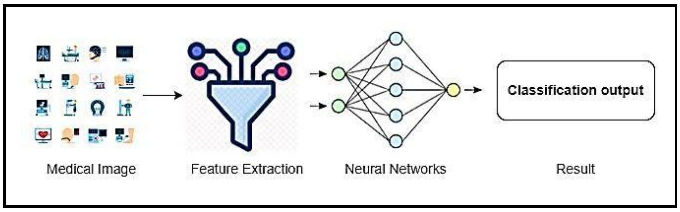

Fig-1: The general steps for evaluating images using deep learning techniques

3. RELATED WORKS

Considering the literature review in this section we will focus upon deep learning models SVM [3], U-net [4], Resnet18, Resnet50 and Xception [8], LCNN and CCNN [9] and CNN [14]. In the upcoming five sections, the models will be studied in detail and will then be compared according to their accuracy discussed in the last section.

3.1 Abnormal CXR with tuberculosis detection using SVM

The research presents an automated system for the detection of tuberculosis. In [3] the performance of the system is measured on two datasets: The MC dataset from the Department of Health and Human Services of Montgomery County (MC), Maryland, and the Shenzhen dataset, from Shenzhen No.3 Hospital in Shenzhen, Guangdong providence, China. The system uses a graph cut-based for lung segmentation and selects training masks according to the horizontal intensity projections of the histogram equalized images and Bhattacharyya coefficient for similarity projections between input and training CXR images. Further, the subtle structures in a CXR are picked by features computed through feature computation techniques. Object Detection Inspired Features are taken from the MC Dataset and CBIR-based Image Features are taken from the Shenzhen dataset where, the 594 dimensioned feature vectors which is greater than three times the feature vector of the MC dataset. The computation of a set of edge, shape, and texture features is then fed as a binary classifier input. Thus, the SVM classifier system can classify a given input image into classified into normal or abnormal. For the MC dataset, the classification accuracy is 78.3% and the area under the ROC curve (AUC) is 86.9%, for the Shenzhen dataset the classification accuracy is 82.5% and the AUC is 88%.

Model	Dataset	Accuracy	AUC
SVM	MC	78.3	86.9
	Shenzhen	82.5	88

Table-1. Abnormal CXR with tuberculosis detection

3.2 U-Net: Convolution Network for segmentation of image

The proposed architecture contains contracting paths to capture context and symmetric expanding paths for precise localization. This [4] model uses a weighted loss, large weight in loss function allocated by separating labels between touching cells, giving accurate segmentation. Data augmentation is used for desired robust properties and invariance when a handful of training samples are available. The important modification in the architecture is the up-sampling method where a large number of feature channels are present which helps the network to propagate sensitive information to high-resolution layers. As a result, the expansive path and contracting path are nearly symmetric, producing a u-shaped architecture. The proposed model does not contain any fully connected layers and uses only associated parts of each convolution layer which allows smooth segmentation of arbitrary large images with the help of overlap-tile strategy. In this research, the model is evaluated using two datasets. The first data-set "PhC-U373" contains Glioblastoma-astrocytoma U373 cells on a polyacrylamide substrate recorded by phase-contrast microscopy. Here, the model achieves an average IOU (Intersection Over Union) of 92%, which is significantly better than the prior best algorithm (Sliding-window Convolutional Network) with an IOU of 83%. The second data-set "DIC-HeLa" are HeLa cells on a flat glass recorded by differential interference contrast (DIC) microscopy. Here the achieved IOU average of 77.5% which is significant in front of the prior best algorithm with 46%.

Model	Dataset	IOU
U-Net	Phc-U373	92%
	DIC-HeLa	77.5%

Table-2. U-Net based image segmentation

3.3 Image Quality Assessment for fetal ultrasound with Deep Convolutional Neural Networks

This paper [8] proposes the use of LCNN and CCNN for classification purposes for the image quality assessment for fetal ultrasounds. The performance of the system is evaluated based on data taken from Shenzhen Maternal and Child Healthcare Hospital from 6 September 2012 to November 2013. The range of fetal ages were between 16 to 40 weeks. In the proposed paper, uses two deep convolutional neural networks namely LCNN and CCNN with an accuracy of 96%. The L-CNN purpose is to find the region of interest (ROI) of the fetal abdominal in the image. According to the ROI discovered by the L-CNN, the C-CNN assesses the quality of an image by determining its goodness of deciding the essential structures of the stomach bubble and umbilical vein. Thereafter, for increasing the performance of the L-CNN, augmentation of data is done between the input of the neural network containing the features of local phase with the original data collected. The experimental evaluations of the scheme are carried out in three parts. In the first part, the L-CNN and C-CNN implementation issues are investigated. In the second part, the performance of the C-CNN model before and after training with visualization of the model is illustrated quantitatively. To illustrate the performance of the C-CNN and LCNN, a comparison of the results of the region of interest (ROI) identification and the assessment of SB (stomach bubble) and UV (umbilical vein) structures from FUIQA scheme is done. Assessment metrics such as accuracy, sensitivity and specificity are selected for comparing quantitatively. In this method, the outcomes according to 3 doctors E1 E2 and E3 are considered

Doctors	ROI	SB	UV
	Accuracy	Accuracy	Accuracy
E1	0.93	0.99	0.98
E2	0.91	0.99	0.97
E3	0.92	0.98	0.98

Table-3. Fetal ultrasound image quality assessment

3.4 Image Segment-based Classification for Chest X-Ray

The system [9] uses U-net with Xception encoder and Feature Pyramid network with Xception encoder and Resnet50 encoder for segmentation of images and Xception, Resnet50, ResNet18 for classification purposes with an accuracy of 94%, 93%, and 95% respectively. This system is based on Deep Learning and AI Summer/Winter School (DLAI) Hackathon Phase 3—Multiclass COVID-19 Chest X-ray Challenge dataset and was collected from various sources. The dataset comprises of three classes that are "COVID-19", "Thorax-Diseases", and "Clear". In order to tune with the CNN model, the original RGB X-Ray images were converted into monochrome images along with floating-point pixels. The dataset was randomly split into training set and test sets with an 80–20 split. Also, augmentation increased the number of the training set and enabled the model to become robust and aids in a better comprehension uncommon CXR images. A random selection of one or more of the following operations: random cropping of CXR with concentration on lung area, horizontal and vertical flipping, 90-degree rotation, and histogram equation were implemented on the original image. To segment the lung part a deep learning model for semantic segmentation was used. A CNN model with ResNet18, Resnet 50, and Xception was used. ResNet is residual learning model which skips connections to resolve the gradient vanishing problem along the increasing depth of CNN layers. The overall accuracy was enhanced by Xception: A CNN model, established upon depth-wise separable convolution layers. The cross-channel mapping and spatial correlations in the feature maps of CNN can be entirely decoupled and was linearly assembled with ResNet connections.

Method	**Cross Entropy**			
	Accuracy	**Precision**	**Recall**	**F1**
ResNet18	94	90	89	89
ResNet50	93	87	90	88
Xception	95	94	90	92
Average	94	90.3	89.7	89.7

Table-4.a. Chest X-Ray Classification(Cross Entropy)

Method	**Focal Loss**			
	Accuracy	**Precision**	**Recall**	**F1**
ResNet18	94	91	91	91
ResNet50	93	87	90	89
Xception	94	91	93	92
Average	93.7	89.7	91.3	90.7

Table-4.b. Chest X-Ray Classification(Focal Loss)

3.5 Breast Cancer Medical Image Classification using CNN

In this system [14], CNN Model is used primarily. Model pre-processes the images with specified parameters of medical images required to fit in the network. Images are in 224x224 pixel PNG format with rescaled boundaries of images. For normalizing the images, reducing complexity of topology function and increasing optimization of learning of model, implementation of new level of orthogonality between the layer is introduced. GPU performs elimination of over-fitting, zooming, and translation with zero computation cost. For given 15 layered CNN model, input used is VGG16 conventional specified images with dimension 224x224x3 with RGB color channels. There is batch size of 8 is used due to GPU requirement. The model will have epoch set to 100. The preferred optimizer used in work is stochastic gradient descent (SGD) because of faster and frequent updates. The activation function used is ReLU as having good computational efficiency and ability to work in over fitting case. The class mode used is Binary i.e., Malignant and benign. Learning rate being 0.1 and momentum set to 0.9. Strides fixed t 1 pixel. The convolutional layer has 64 filters of 3x3 dimension outputting 224x224x64. The fully connected layer has 2 class output and dropout probability up to 50% to work properly with overfitting in the model.

Model	**Training Accuracy (%)**	**Val. Accuracy (%)**	**Training Loss (%)**	**Val. Loss (%)**
VGG16	94.77	96.33	13.94	10.90
CNN	**98.02**	**98.50**	**5.84**	**4.81**

Table-5. Brest cancer medical image classification

4. PERFORMANCE COMPARISION

For biomedical images segmentation, the deep learning approach surpasses the accuracy. In section E, the CNN model outperforms other models in terms of performance. The combination of CNN classification and image segmentation produces a finegrained detection model. The results show that training for multiple iterations on a designed model, accuracy is obtained are best among all. Implementation of a new level of orthogonality between the layers is introduced to increase learning in the process. The use of zero computational cost GPU for image augmentation purposes. In sections A and B we spot that the accuracy of models varies as the dataset differs. For section A, SVM has been applied to two datasets (MC and Shenzen), it is seen that from Table 1 the accuracy for the dataset Shenzen is more than that of MC. Similarly in section B, U-net has worked on two different datasets. From Table2 we acknowledge that the accuracy for the dataset PhC-U373 is greater than that for DIC-HeLA. There are multiple reasons for the change in accuracy on different datasets irrespective of the same technique applied. In section D, Xception encoder which is a CNN model, using depth-wise separable convolutional layers with cross channel mapping, feature maps of CNN can be entirely uncoupled and were stacked linearly with residual connections causing an increase in the accuracy. In Section C, LCNN and CCNN for image classification of fatal ultrasound image quality. According to the ROI discovered by the L-CNN, the C-CNN evaluates the quality of an image by determining its goodness of depiction for the key structures of the image data. Experiments conducted using different deep learning models, Xception, LCNN+CCNN, Resnet18 and Resnet50 achieve performs well after CNN. Fine-tuning avoiding losses give higher accuracy when used on a larger dataset.

5. EXPERIMENTAL RESULTS

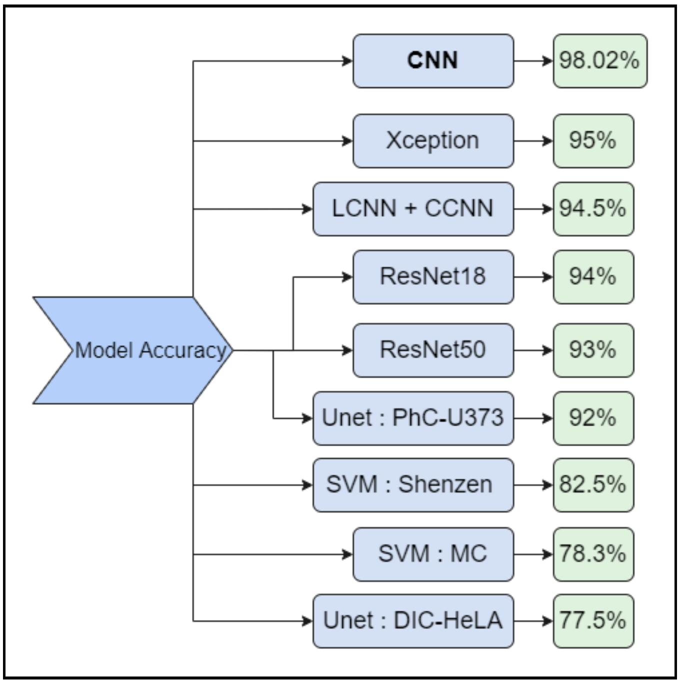

Fig-2. Models according to increasing accuracy

From this study, we have deduced that the 15-layer CNN model most effectively classifies and analyzes the medical images using relevant CNN classification techniques. Xception model along with the CNN model gives the second-best accuracy. Further, we propose the implementation of CNN models with different activation functions and focus on removing the noise for better segmentation of the images hence improving the performance.

6. CONCLUSIONS

Medical Image analysis is mainly used for the detection and diagnosis of diseases which helps in proper and accurate treatment given to patients. In this paper we have taken an overview of various deep learning models focusing on the classification task. This study has concluded that a union of traditional analysis approaches along with deep learning framework upshots in an effective diagnosis of brain tumor, fatal ultrasound, COVID19, thorax, etc. Amongst the following models U-net, SVM, Residual networks, CNN, and Xception, the highest accuracy is achieved by the CNN model of 98.0.2%. The study illuminates that the drawbacks of a single model are overcome by encouraging a combination of deep learning models. In future works, this study can assist in enhancing the system using advanced deep learning frameworks.

REFERENCES

[1] Wenjing Yang, Arlene Sirajuddin, Xiaochun Zhang, Guanshu Liu, Zhongzhao Teng, Shihua Zhao, and Minjie Lu. 2020. The role of imaging in 2019 novel coronavirus pneumonia (COVID-19). European Radiology 30, 9 (April 2020), 4874–4882. https://doi.org/10.1007/s00330-020-06827-4.

[2] F. Chollet. 2017. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017). IEEE, Honolulu, HI, 1800–1807. https://doi.org/10.1109/CVPR.2017.195.

[3] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z.Xue, K. Palaniappan, R. K. Singh, S. Antani, G. Thoma, Y. Wang, P. Lu, and C. J. McDonald. 2014. Automatic Tuberculosis Screening Using Chest Radiographs. IEEE Transactions on Medical Imaging 33, 2 (Feb 2014), 233–245.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). Springer, Munich, Germany, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[5] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," IEEE transactions on Medical Imaging, vol. 37, no. 2, pp. 491– 503, 2017.

[6] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. 2018. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence 40, 4 (April 2018), 834–848.

[7] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proc. CVPR, 2015, pp. 3431–3440.

[8] Phongsathorn Kittiworapanya, Kitsuchart Pasupa, "An Image Segmentbased Classification for Chest XRay Image" in

ACM November 19, 2020, https://doi.org/10.1145/3429210.3429227.

[9] L. Wu, J.-Z. Cheng, S. Li, B. Lei, T. Wang, and D. Ni, "FUIQA: Fetal ultrasound image quality assessment with deep convolutional networks," IEEE Transactions on Cybernetics, vol. 47, no. 5, pp. 1336–1349, 2017.

[10] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1733– 1740

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation" in MICCAI. Springer, 2015, pp. 234–241.

[12] W. Bai, M. Sinclair, G. Tarroni, O. Oktay, M. Rajchl, G. Vaillant, A. M. Lee, N. Aung, E. Lukaschuk, M. M. Sanghvi, "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks" in Journal of Cardiovascular Magnetic Resonance, vol. 20, no. 1, p. 65, 2018.

[13] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolu�tional recurrent neural networks for dynamic MR image reconstruction," IEEE transactions on medical imaging, vol. 38, no. 1, pp. 280–290, 2018.

[14] Yongbin Yu, Ekong Favour, Pinaki Mazumder, "Convolutional Neural Network Design for Breast Cancer Medical Image Classification ", 2020 IEEE 20th International Conference on Communication Technology

[15] MAHNOOR ALI, SYED OMER GILANI, ASIM WARIS, KASHAN ZAFAR, AND MOHSIN JAMIL, "Brain Tumour Image Segmentation Using Deep Networks" IEEE Open Access Journal Volume 8, 2020.

[16] A. H. Abdi, C. Luong, T. Tsang, G. Allan, S. Nouranian, J. Jue, D. Hawley, S. Fleming, K. Gin, J. Swift, et al., "Automatic quality assessment of apical four-chamber echocardiograms using deep convolutional neural networks," in Proc. SPIE, vol. 10133, 2017, pp. 101 330S– 1.

[17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600– 612, Apr. 2004.