ISSN 2278-3091
Volume 9, No.1, January – February 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse119912020.pdf
https://doi.org/10.30534/ijatcse/2020/119912020

Abnormal X-Ray Detection System using Convolution Neural Network

Shubhangi Tirpude[1], Naman Vidyabhanu[2], Hashir Sheikh[3] Shoeb Pathan[4], Zeeshan Ali syed[5], Shivam Singh[6]

[1]Assistant Professor, Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, India, tirpudes@rknec.edu
[2] Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, vidyabhanuna@rknec.edu
[3] Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, sheikhhn@rknec.edu
[4] Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, pathansa@rknec.edu
[5] Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, syedzi@rknec.edu
[6] Shri Ramdeobaba College of Engineering & Management, Nagpur, Maharashtra, singhss_1@rknec.edu

ABSTRACT

Musculoskeletal conditions become very serious health problems in the life of people that affect muscles, bones, and joints. Musculoskeletal disorders include tendinitis, carpal tunnel, syndrome, osteoarthritis, rheumatoid arthritis (RA),fibromyalgia and bone fractures. Many methods have been developed for this purpose. In this paper, we detect and localize abnormalities in X-rays using densely connected convolutional neural network. The convolutional neural network are trained and tested on the large dataset of musculoskeletal radiographs (MURA). We train a 169-layer densely connected convolutional network on this dataset to detect and highlight abnormalities in the X ray using concepts of supervised learning. The results is presented in terms of localizing the abnormality, achieving good amount of accuracy & comparing it's predictions to those of radiologists. We find that on finger, hand, and wrist studies, the model achieves good amount of accuracy but on elbow, forearm, humerus, and shoulder studies our model finds it difficult to detect the abnormality for some cases but still results in decent amount of accuracy.

Key words: Musculoskeletal Radiographs, Convolutional Neural Network, Abnormality Detection, machine

1. INTRODUCTION

More than 1.7 billion people worldwide [BMU, 2017], affected with musculoskeletal conditions which causes severe, long- term pain and disability [16]
The abnormality detection task is a critical radiological task i.e. it determines whether a radiographic study is normal or abnormal. A study interpreted as normal rules out disease and thus eliminates the requirement of patients to endure further diagnostic procedures. We develop an abnormality detection model on MURA which takes as input one or more views for a study of an upper extremity. The 169-layer convolution neural network predicts the probability of abnormality on each view. Then the per-view probabilities are averaged to determine the probability of abnormality for the study.

The main objectives of the system is to
• Design a model which detects and highlights the abnormalities in X-ray images.
• Train the model over 50,000 training images avoiding over-fitting and under- fitting problems.
• Build the abnormality detection process efficient so that it is much faster than manual detection.
• Formulate the model by achieving good amount of accuracy in abnormalities detection.

2. LITERATURE REVIEW

Andreas Kanavos & Panagiotis Pintelas [1] implements automatic detection of lung abnormalities from digital chest X-rays They proposed a new semi-supervised learning algorithm for the classification of lung abnormalities from X-rays based on an ensemble philosophy. The efficacy of the presented algorithm is demonstrated by numerical experiments, illustrating that reliable prediction models could be developed by incorporating ensemble methodologies in the semi-supervised framework.

Pranav Rajpurkar, Awni Y. Hannun, Masoumeh Haghpanahi, Codie Bourn, Andrew Y. Ng[2] developed an algorithm which detects a wide range of heart arrhythmias from electrocardiograms recorded with a single-lead wearable monitor which exceeds the performance of board certified cardiologists. They build a dataset with more than 500 times the number of unique patients than previously studied corpora. They train a 34-layer convolution neural network which maps a sequence of ECG samples to a sequence of rhythm classes. They compare the performance of model to that of 6 other individual cardiologists and exceed the average cardiologist performance in both recall (sensitivity) and precision (positive predictive value).

William Gale, Luke Oakden-Rayner, Gustavo Carneiro, Andrew P. Bradley, Lyle J. Palmer [3] suggested an automated deep learning system to detect hip fractures from frontal pelvic x-rays, an important and common radiological task. The system was trained on a decade of clinical x-rays (~53,000 studies) and can be applied to clinical data, automatically excluding inappropriate and technically unsatisfactory studies.. Translated to clinical practice, such a system has the potential to increase the efficiency of diagnosis, reduce the need for expensive additional testing, expand access to expert level medical image interpretation, and improve overall patient outcomes.

Monika Grewal, Muktabh Mayank Srivastava, Pulkit Kumar, Srikrishna Varadarajan[8] describes a deep learning approach for automated brain hemorrhage detection from computed tomography (CT) scans. This model emulates the procedure followed by radiologists to analyze a 3D CT scan in real-world. Similar to radiologists, the model sifts through 2D cross-sectional slices while paying close attention to potential hemorrhagic regions. They uses Recurrent Attention DenseNet (RADnet).. RADnet demonstrates 81.82% hemorrhage prediction accuracy at CT level that is comparable to radiologists. Further, RADnet achieves higher recall than two of the three radiologists, which is remarkable.

3. PROPOSED APPROACH

3.1 Data Set Used

MURA (musculoskeletal radiographs) is a large dataset of bone X-rays and it is one of the largest public radiographic image datasets, containing 14,863 studies from 12,173 patients, with a total of 40,561 multi-view radiographic images. The dataset contains 9,067 normal and 5,915 abnormal musculoskeletal radiographic studies including the shoulder, humerus, elbow, forearm, wrist, hand, and finger. Each belongs to one of seven standard upper extremity radiographic study types: elbow, finger, forearm, hand, humerus, shoulder, and wrist and it were manually labeled as normal or abnormal by board-certified radiologists from the Stanford Hospital. The dataset is freely available at https://stanfordmlgroup.github. Io/projects/mura.

3.2 Process

MURA dataset comes with train, valid and test folders containing datasets, train.csv and valid.csv contain paths of radiographic images and their labels. These radiographic images are labeled as 1 (abnormal) or 0 (normal) based on whether its corresponding study is negative or positive.

We visualize the parts of the radiograph which contribute most to the model's prediction of abnormality by using Class Activation Mappings (CAMs) [17]. The radiograph X is inputted to the fully trained network to obtain the feature maps output by the final convolution layer. To compute the CAM M(X), we take a weighted average of the feature maps using the weights of the final fully connected layer. Denote the kth feature map output by the network on image X by fk(X) and the kth fully connected weight by wk. formally

$$\mathbf{M(X)} = {}_{k}\sum^{n} \boldsymbol{k} * \mathrm{Fk}()$$

To highlight the salient features in the original radiograph, we upscale the CAM M(X) to the dimensions of the image and overlay the image.

3.3 Building the Data Pipeline

The model takes as input one or more views for a study of an upper extremity. Then 169-layer convolution neural network predicts the probability of abnormality on each view. We compute the overall probability of abnormality for the study by taking the arithmetic mean of the abnormality probabilities output by the network for each image. The model makes the binary prediction of abnormal if the probability of abnormality for the study is greater than 0.5. Figure 2 illustrates the model's prediction pipeline. So we need to predict the probability of abnormality at study level. Now we need a study level data pipeline, one which returns all images of a study to be fed to the model and respective label of the study.

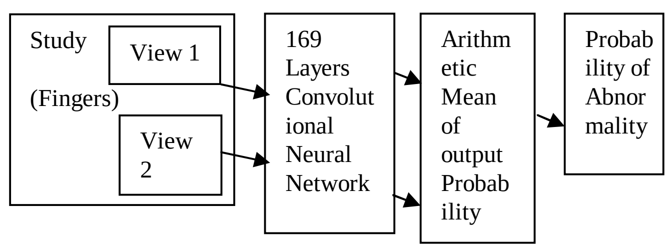


Figure 1: Flow of Model

3.4 Preprocessing

As the radiographs are of different sizes, so before feeding images into the network, we normalized each image to have the same mean and standard deviation of images in the Image Net training set. The training data set get normalized to 224×224 pixel size. Then we flipped or augmented the images in all the directions by applying rotation & inversion so as to increase the training data set. The purpose of these is to get the accurate weights which results in more accurate curve. PyTorch provides easy to use data pipeline and data augmentation modules dataset and data loader.

3.5 Building the Model

The network uses Dense Convolution Network architecture detailed in [7, 12] which connects each layer to every other layer in a feed-forward fashion to make the optimization of deep networks tractable. We replaced the final fully connected layer with one that has a single output, after which we applied a sigmoid nonlinearity. The weights of the network were initialized with weights from a model pre trained on ImageNet [20].

By default PyTorch has DenseNet implementation, but so as to replace the final fully connected layer with one that has a single output and to initialize the model with weights from a model pretrained on ImageNet, we need to modify the default DenseNet implementation.

For each image X of study type T in the training set, we optimize the weighted binary cross entropy loss.

$$L(X, y) = -w_{T,1} \cdot y \log p(Y = 1|X) - w_{T,0} \cdot (1 - y) \log p(Y = 0|X),$$

Where y is the label of the study, $p(Y = i|X)$ is the probability that the network assigns to the label i, $w_{T,1} = |N_T|/(|A_T| + |N_T|)$, and $w_{T,0} = |A_T|/(|A_T| + |N_T|)$ where $|A_T|$ and $|N_T|$ are the number of abnormal images and normal images of study type T in the training set respectively

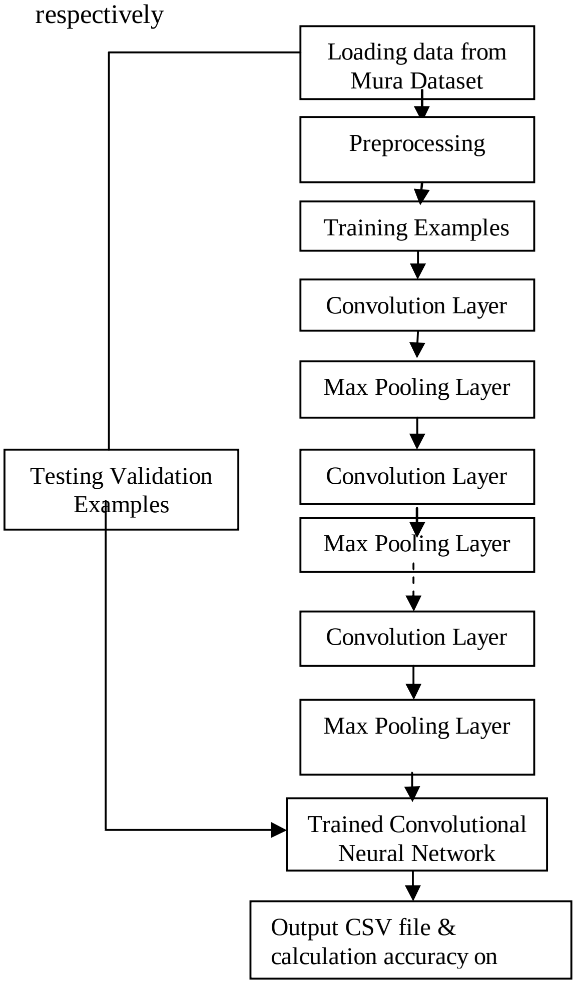


Figure 2: System Architecture

The weights of the network are initialized with weights from a model pre trained on Image Net (20). The network is trained end-to-end using default parameters β1 = 0.9 and β2 = 0.999 (Kingma & Ba, 2014). We use an initial learning rate of 0.0001 that is decayed by a factor of 10 each time the validation loss plateaus after an epoch, and pick the model with the lowest validation loss.

3.6 Structure of System Model

Dense Net connects each layer to every other layer in a feed-forward fashion. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers. Dense Nets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

4. IMPLMENTATION

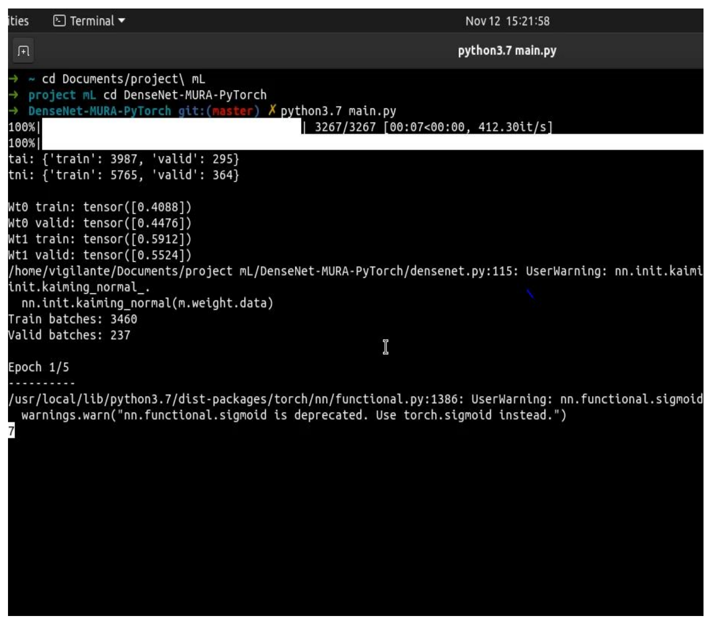


Figure 3: Training Dense Net

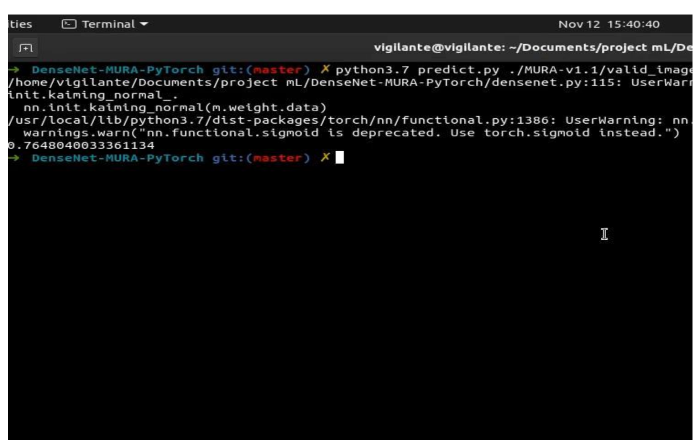
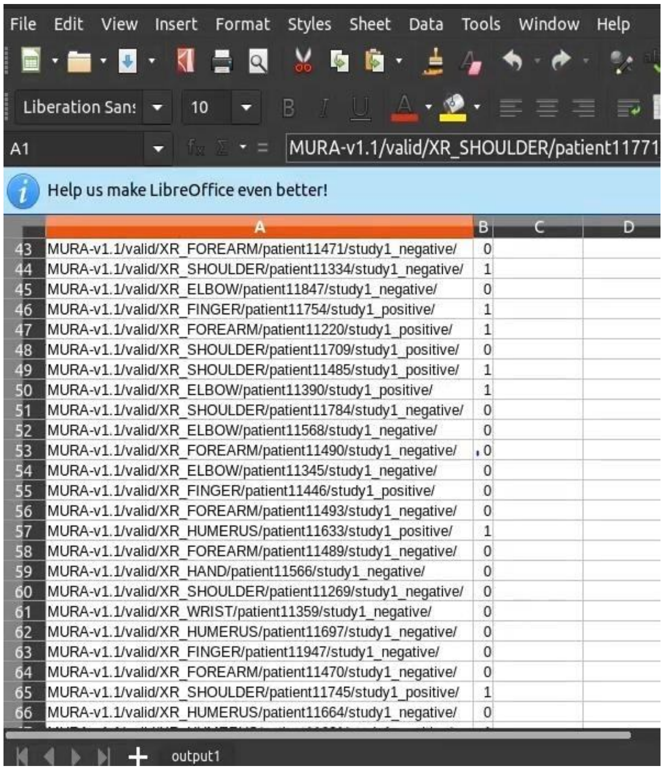


Figure 4: Testing any 3000 images on our trained dense net



	A	B	C	D
43	MURA-v1.1/valid/XR_FOREARM/patient11471/study1_negative/	0		
44	MURA-v1.1/valid/XR_SHOULDER/patient11334/study1_negative/	1		
45	MURA-v1.1/valid/XR_ELBOW/patient11847/study1_negative/	0		
46	MURA-v1.1/valid/XR_FINGER/patient11754/study1_positive/	1		
47	MURA-v1.1/valid/XR_FOREARM/patient11220/study1_positive/	1		
48	MURA-v1.1/valid/XR_SHOULDER/patient11709/study1_positive/	0		
49	MURA-v1.1/valid/XR_SHOULDER/patient11485/study1_positive/	1		
50	MURA-v1.1/valid/XR_ELBOW/patient11390/study1_positive/	1		
51	MURA-v1.1/valid/XR_SHOULDER/patient11784/study1_negative/	0		
52	MURA-v1.1/valid/XR_ELBOW/patient11568/study1_negative/	0		
53	MURA-v1.1/valid/XR_FOREARM/patient11490/study1_negative/	0		
54	MURA-v1.1/valid/XR_ELBOW/patient11345/study1_negative/	0		
55	MURA-v1.1/valid/XR_FINGER/patient11446/study1_positive/	0		
56	MURA-v1.1/valid/XR_FOREARM/patient11493/study1_negative/	0		
57	MURA-v1.1/valid/XR_HUMERUS/patient11633/study1_positive/	1		
58	MURA-v1.1/valid/XR_FOREARM/patient11489/study1_negative/	0		
59	MURA-v1.1/valid/XR_HAND/patient11566/study1_negative/	0		
60	MURA-v1.1/valid/XR_SHOULDER/patient11269/study1_negative/	0		
61	MURA-v1.1/valid/XR_WRIST/patient11359/study1_negative/	0		
62	MURA-v1.1/valid/XR_HUMERUS/patient11697/study1_negative/	0		
63	MURA-v1.1/valid/XR_FINGER/patient11947/study1_negative/	0		
64	MURA-v1.1/valid/XR_FOREARM/patient11470/study1_negative/	0		
65	MURA-v1.1/valid/XR_SHOULDER/patient11745/study1_positive/	1		
66	MURA-v1.1/valid/XR_HUMERUS/patient11664/study1_negative/	0		

Figure 5: Output for Abnormities for Shoulder

5. CONCLUSION

The system successfully uses a convolution neural network to detect abnormalities in X-Rays with the help of "labeled" MURA- Dataset. With reasonably decent amount of accuracy, advancements of this model could provide a great substitute to radiologists in areas where access to skilled radiologists is limited

The network was trained end-to-end using Adam with default parameters $\beta1 = 0.9$ and $\beta2 = 0.999$ (Kingma & Ba, 2014). We trained the model using mini batches of size 8. We used an initial learning rate of 0.0001 that is decayed by a factor of 10 each time the validation loss plateaus after an epoch, and chose the model with the lowest validation loss.

6. FUTURE SCOPE

Abnormality detection in musculoskeletal radiographs has important clinical applications. In this scenario, the studies detected as abnormal could be moved ahead in the image interpretation workflow, allowing the sickest patients to receive diagnoses more quickly. Furthermore, the examinations identified as normal could be automatically assigned a preliminary reading of "normal"; this could mean

- More rapid results can be conveyed to the ordering provider (and patient) which would improve disposition in other areas of the healthcare system (i.e., discharged from the ED more quickly)
- A radiology report template for the normal study can be served to the interpreting radiologist for more rapid review and approval.

REFERENCES

1. Andreas Kanavos Panagiotis Pintelasa **Detecting Lung Abnormalities From X-rays Using an Improved SSL Algorithm** https://doi.org/10.1016/j.entcs.2019.04.008
2. Rajpurkar, Pranav, Hannun, Awni Y, Haghpanahi, Masoumeh, Bourn, Codie, and Ng, Andrew Y. **Cardiologist-level arrhythmia detection with convolutional neural networks**. arXiv preprint arXiv:1707.01836, 2017a
3. Gale, W., Oakden-Rayner, L., Carneiro, G., Bradley, A. P., and Palmer, L. J. **Detecting hip fractures with radiologist-level performance using deep neural networks**. ArXiv e-prints, November 2017.
4. Berlin, Leonard. **Liability of interpreting too many radiographs**. American Journal of Roentgenology, 175(1):17–22, 2000.
5. Bhargavan, Mythreyi and Sunshine, Jonathan H. **Utilization of Radiology services in the United States: levels and trends in modalities, regions, and populations**. Radiology, 234(3):824–832, 2005
6. Rajpurkar, Pranav, Irvin, Jeremy, Zhu, Kaylie, Yang, Brandon, Mehta, Hershel, Duan, Tony, Ding, Daisy, Bagul, Aarti, Langlotz, Curtis, Shpanskaya, Katie, et al. Chexnet: **Radiologist-level pneumonia detection on chest x-rays with deep learning**. arXiv preprint arXiv:1711.05225, 2017b
7. Gertych, Arkadiusz, Zhang, Aifeng, Sayre, James, Pospiech-Kurkowska, Sylwia, and Huang, H.K. **Bone age assessment of children using a digital hand atlas. Computerized Medical Imaging and Graphics**, 31(4):322–331, 2007.
8. Gulshan, Varun, Peng, Lily, Coram, Marc, Stumpe, Martin C, Wu, Derek, Narayanaswamy, Arunachalam, Venugopalan, Subhashini, Widner, Kasumi, Madams, Tom, Cuadros, Jorge, et al. **Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs**. Jama, 316(22):2402–2410, 2016 https://doi.org/10.1001/jama.2016.17216
9. Hannun, Awni, Case, Carl, Casper, Jared, Catanzaro, Bryan, Diamos, Greg, Elsen, Erich, Prenger, Ryan, Satheesh, Sanjeev, Sengupta, Shubho, Coates, Adam, et al. **Deep speech: Scaling up end-to-end speech recognition**. arXiv preprint arXiv:1412.5567,2014
10. Huang, Gao, Liu, Zhuang, Weinberger, Kilian Q, and van der Maaten, Laurens. **Densely connected convolutional networks**. arXiv preprint arXiv:1608.06993, 2016.
11. Jaeger, Stefan, Candemir, Sema, Antani, Sameer, Wang, Yi-Xiang J, Lu, Pu- Xuan, and Thoma, George. **Two public chest x-ray datasets for computer-aided screening of pulmonary diseases**. Quantitative imaging in medicine and surgery, 4(6): 475, 2014.

12. Shiraishi, Junji, Katsuragawa, Shigehiko, Ikezoe, Junpei, Matsumoto, Tsuneo, Kobayashi, Takeshi, Komatsu, Ken-ichi, Matsui, Mitate, Fujita, Hiroshi, Kodera, Yoshie, and Doi, Kunio. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology, 174(1):71–74, 2000.
13. Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: **Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases**. arXiv preprint arXiv:1705.02315, 2017.
14. Woolf, Anthony D and Pfleger, Bruce. **Burden of major musculoskeletal conditions.** Bulletin of the World Health Organization, 81(9):646–656, 2003.
15. Zhou, Bolei, Khosla, Aditya, Lapedriza, Agata, Oliva, Aude, and Torralba Antonio. **Learning deep features for discriminative localization.** In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929, 2016.
16. Krupinski, Elizabeth A, Berbaum, Kevin S, Caldwell, Robert T, Schartz, Kevin M, and Kim, John. **Long radiology workdays reduce detection and accommodation accuracy**. Journal of the American College of Radiology, 7(9):698–704, 2010.
17. Demner-Fushman, Dina, Kohli, Marc D, Rosenman, Marc B, Shooshan, Sonya E, Rodriguez, Laritza, Antani, Sameer, Thoma, George R, and McDonald, Clement J. **Preparing a collection of radiology examinations for distribution and retrieval.** Journal of the American Medical Informatics Association, 23 (2):304–310, 2015.
https://doi.org/10.1093/jamia/ocv080
18. Deng, Jia, Dong, Wei, Socher, Richard, Li, Li-Jia, Li, Kai, and Fei-Fei, Li. Imagenet: **A large-scale hierarchical image database**. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 248–255. IEEE, 2009.
19. Mohammad Tariqul Islam, Md Abdul Aowal, Ahmed Tahseen Minhaz, Khalid Ashraf **Abnormality Detection and Localization in Chest X-Rays using Deep Convolutional Neural Networks** Computer Vision and Pattern Recognition https://doi.org/10.1016/j.entcs.2019.04.008
20. Rajesh R Karhe, Dr. S.N Kale, **Neural Network Classifiers for Cardiac Arrhythmiya From Scanned ECG IJATCSE,** doi.org/10.30534/ijatcse/2019/8181.62019