Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

MDCF: MULTI-DISEASE CLASSIFICATION FRAMEWORK ON FUNDUS IMAGE USING ENSEMBLE CNN MODELS

E. Sudheer Kumar [1] and C. Shoba Bindu [2]

[1]Research Scholar, Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India, sudheerkumar.e@gmail.com

[2]Professor, Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India, shobabindhu.cse@jntua.ac.in

Abstract

The purpose of fundus imaging is to examine the anomalies related to diseases that affect the eye. A fundus image plays a crucial role in the observation and detection of various ophthalmological diseases. The majority of earlier researchers have concentrated their approaches on the identification of individual diseases from fundus image. But, simultaneous detection of multi-disease from fundus image is still facing a great challenge. To the patient, there is a chance of having more than one disease while diagnosing either or both the eyes. So, it is planned to address this challenge by designing a framework which can detect multi-disease from fundus image with improved accuracy. This paper proposes a novel Multi-Disease Classification Framework (MDCF) by incorporating ensemble neural architectures. In the proposed framework, the initial task is to perform preprocessing on the dataset with certain steps like: contrast enhancement, oversampling, resizing, and normalization. The MDCF will be conceded in two stages: the first stage is to detect whether the fundus image is at disease risk or not and the second stage is to classify multi-disease on fundus image. Two convolutional neural networks Densenet201 and EfficientNetB4 were used for disease risk detection and in addition to these two networks ResNet105 is added for multi-disease classification. Retinal Fundus Multi-disease Image Dataset (RFMiD) is used for training and validation of the proposed work. The MDCF is tested on Ocular Disease Intelligent Recognition (ODIR) 2019 dataset and the output demonstrated that the proposed work is performing well compared to the other state-of-the-art results.

Keywords— Fundus Imaging; Multi-Disease Classification; Convolutional Neural Network; Ensemble Approach; Deep Learning; Ophthalmological Diseases;

1. INTRODUCTION

A retinal tissue with thick layer which can draw the posterior of the eye inwards can be converted in to neural signals from incident light. The signals can be managed with the visual cortex of a brain in order to identify the object/scene. The overall health of a person can also be assessed by the analysis of retinal tissue. In the process of analysis of retina image involves various imaging techniques such as optical coherence tomography (OCT), fundus and fluorescein angiography. But, OCT imaging technology

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

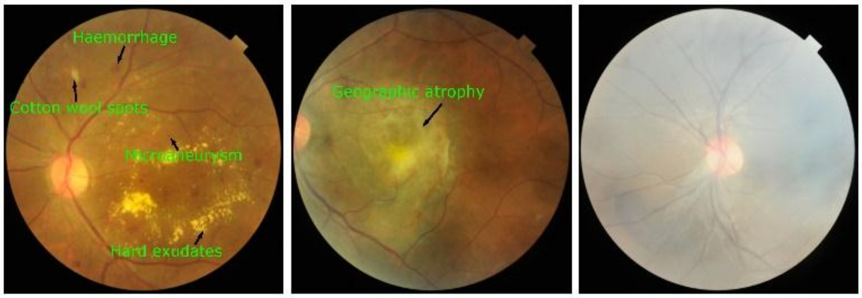
is quite expensive; whereas fundus imaging was an invasive and economical way for diagnosing eye diseases. Ophthalmologists use fundus imaging as a primary modality to diagnose “Diabetic Retinopathy (DR), Glaucoma, Age-related Macular Degeneration (AMD), Cataracts, Hypertension, and Myopia” diseases [1]. There is a dearth of timely and effective cure in the short terms, which will origin irreversible visual impairment.

Therefore, primary recognition and treatment of fundus diseases is very crucial. Fast growing technology like AI can help ophthalmologists based on comprehensive medical data while diagnosing and also offers new approaches to improve the level of analysis and treatment of eye disease in major hospitals. A mixture of AI and ophthalmology medical therapies can encounter the practical needs of patients with diseases. In ophthalmological image analysis most of the research is restricted to specific diseases. The main objective of this paper is to design and develop a Multi-Disease Classification Framework (MDCF) which is competent of discovering a wide variety of eye diseases from fundus image for multi-class classification of ophthalmic diseases. To the best of our knowledge this is one of the best solution in terms of classification accuracy achieved so far when compared with earlier works on RFMiD dataset.

However, still challenges are there in using convolutional neural network (CNN) for fundus image research. Firstly, fundus image classification with multi-label was further usual and concrete problem, as the actual images in the real life involves multiple diseases. Secondly, it is tough to acquire adequate accurate fundus images, exclusively on some special fundus diseases. Thirdly, in the face of inadequate data and unavoidable image noise, training a single model to productively attain high diagnostic accuracy is challenging.

The Figure 1 shows sample fundus image from Retinal Fundus Multi-Disease Image Dataset (RFMiD) [2]. The proposed work aims at developing a Multi Disease Classification Framework (MDCF) in two stages. The first stage is to detect whether the fundus image is at disease risk or not. If the image is at risk, then in the second stage it will classify diseases on single image if it contains. To increase the classification accuracy pre-processing is also adopted in this framework. The proposed approach includes high enactment and simplification competencies of deep learning methods for classification. Ensemble based approach with pre-trained CNN is suggested to classify images into Twenty Eight kinds of diseases.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C



(a) (b) (c)

Figure 1: Fundus image from RFMiD dataset (a) DR (b) AMD (c) Media Haze (MH) [2].

The rest of the paper is structured as follows: Section 2 discusses the literature and Section 3 enlightened about proposed transfer learning with ensemble based approach for classification. Results and evaluation criteria were discussed in Section 4. Finally, the work is concluded in Section 5.

2. LITERATURE

A multi-class multi-label classification method [3] is suggested for the identification of ophthalmological diseases. This approach is based on transfer learning approach. There are two models proposed for the disease identification. The method uses four different CNN architectures along with dual dissimilar optimizers. Upon experimentation it is confined that the multi-class classification performance is increased with VGG16 along with SGD optimizer. This method uses ODIR 2019 dataset for experimentation. This work [4] performs multi-label classification ensemble model with the help of EfficientNet architecture. This approach has divided in to two parts: feature extraction process is the first part with the help of EfficientNet network and in the second part with the help of custom neural network architecture multi-label classification is performed. This method also uses ODIR 2019 dataset for experimentation. By considering multiple models as ensemble the recognition result is obtained from the fused output probabilities of various models.

To achieve correlations between paired fundus photography a Dense Correlation Network (DCNet) [5] is suggested. DCNet is comprised of CNN as a backbone along with spatial correlation module and a classifier to perform disease classification. The spatial correlation module is responsible to fuse the extracted features from corresponding fundus image in pixel wise manner. The decisive convolution network [6] with some pre-processing steps along with cross-entropy based loss function is used for better classification accuracy. This network classifies eight types of ophthalmological diseases from fundus image.

An attention map mechanism with shallow CNN architecture [7] is proposed as a deep CNN which consists of significant features emphasized through CNN at several stages. Heatmaps are assembled using gradient class activation charting and expended for

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

recognition of ocular diseases such as "AMD, Diabetic Retinopathy, and Glaucoma on the privately collected database". The similar kind of approach [8] has been proposed with dual pre-trained networks. VGG16 and InceptionV3 are used as ensemble approach for disease classification. A transfer learning based approach [9] is investigated for the association amongst the no. of classes and fully-connected layers for ophthalmology diseases.

Multiple pre-trained networks were used as a transfer learning method for fundus image classification [10, 11] and also framework based approaches on fundus images [12, 13]. Generally for diagnosing retinal fundus images, there are certain basic steps [14] which can be performed like Optic Disc segmentation [15], Optic Cup segmentation [16], Lesions segmentation [17] and finally disease classification. But this is a single disease classification which will encounter in the last step. Whereas multi disease classification is a challenging task.

3. Proposed System And Methodology

The aim of this work is to develop a multi disease classification framework. This framework consists of three stages: pre-processing, disease risk detection and multi disease classification. The pre-processing stage includes augmentation, upsampling, resizing and normalization. Disease risk identification can be done with the help of two CNN architectures namely: Densenet201 and EfficientNetB4. Multi-disease classification stage is an ensemble approach consisting of three CNN architectures Densenet201, EfficientNetB4 and ResNet105. The above mentioned stages will be discussed elaborately in the further sub sections.

3.1 Dataset

The dataset considered in this work is the new "Retinal Fundus Multi-Disease Image Dataset (RFMiD) which consists of 3200 fundus images for which 1920 is available for training purpose. These images were captured by three different fundus cameras having a resolution of 4288x2848 (277 images), 2048x1536 (150 images) and 2144x1424 (1493 images), respectively.

Figure 2: Class Distribution of RFMiD Dataset

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

These images were captured by a retinal specialist at an Eye Clinic and public screening camp organized at the Center of Excellence in Signal and Image Processing, SGGS Institute of Engineering and Technology both located in Nanded, India. The dataset was published associated to the Retinal Image Analysis for Multi-Disease Classification (RIADD) challenge from the ISBI 2021" [18]. The annotation frequency of each class can be seen in the below Table 1 and in the Figure 2.

Table 1: Annotations Available in the Training Dataset.

Disease	No. of Samples	Disease	No. of Samples
Disease Risk	1519	AH	16
DR	376	ODP	65
ARMD	100	ODE	58
MH	317	ST	5
DN	138	AION	17
MYA	101	PT	11
BRVO	73	RT	14
TSLN	186	RS	43
ERM	14	CRS	32
LS	47	EDN	15
MS	15	RPEC	22
CSR	37	MHL	11
ODC	282	RP	6
CRVO	28	OTHER	34
TV	6	Normal	401

3.2 Pre-Processing

The purpose of pre-processing is to make the dataset flexible and suitable to the models considered in the training process. As an initial step in this stage, Image augmentation [19] has been performed with rotation and flipping operations. In turn can help the model to understand the data in different ways. When there is a un-uniformity in the distribution of classes, there exists a problem of class imbalance. In order to address this issue, upsampling [20] is the technique used in this pipeline to balance the class. After

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

performing the upsampling method the dataset has been increased from 1920 samples to 3215 samples.

As the developed framework consists of various CNN models for detection and classification tasks. It is necessary to resize the original images to the size of each individual network. So in order to address this issue, resizing has been performed on original samples as per the model requirement. Before sending the image to the CNN, it is been applied value intensity normalization as a final pre-processing step. "The intensities were zero-centered via the Z-Score normalization approach based on the mean and standard deviation computed on the ImageNet dataset" [21].

3.3 Ensemble Architectures

After completing the pre-processing step, the next task is to identify whether the fundus image is having disease risk or not. The disease risk detector is composed of two CNN architectures: Densenet201 [22] and EfficientNetB4 [23]. The two neural networks were trained separately with disease risk class labels from the dataset ignoring the other class labels. The two architectures were trained rigorously for 100 epochs which can be capable to detect whether the fundus image is at disease risk or not. Binary Cross-Entropy (BCE) loss [26] is used as loss function for this detector neural networks, because these networks needs to detect whether disease present or not. The BCE loss is defined as in equation 1.

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \times \log\big(p(y_i)\big) + (1 - y_i) \times \log(1 - p(y_i)) \quad □□□□$$

where *y* is the label (1 for having disease risk and 0 for not having disease risk) and *p(y)* is the predicted probability of the point being having disease risk for all *N* points.

Once the detection stage has performing well for identifying the diseased image then the classifier stage will perform its operations. The multi-disease classifier consists of Densenet201, EfficientNetB4 and ResNet105 [22-24] architectures. The three neural networks were trained individually for 28 classes (excluding disease risk) of ophthalmology diseases. These three architectures were trained rigorously for 100 epochs which can be capable enough to classify multi-disease from fundus image. Focal loss [25] is used as loss function for this classifier, because these networks need to detect multiple diseases simultaneously. Multi-class classification with focal loss is the best option in case of imbalanced datasets. The focal loss is defined as in equation 2.

$$FL(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t) \quad □□□$$

where, p_t is the probability for the correct ground truth class *t*, γ a tunable focusing parameter and α_t the associated weight for class *t*. The complete framework has shown in Figure 3.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

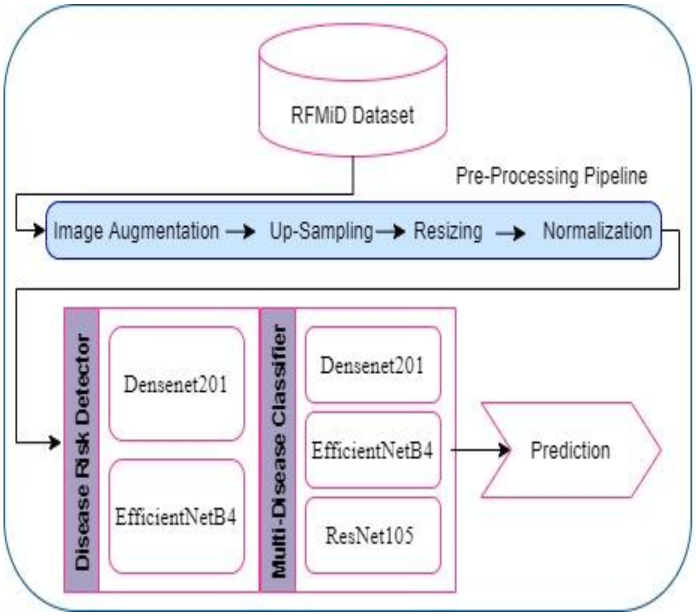


Figure 3: Proposed Multi-Disease Classification Framework

4 Results And Discussion

The CNN architectures used for disease risk detector and classifier is ensemble with Stacking technique. Stacking is apprehensive with connecting multiple classifiers generated from different CNNs on a single dataset which consists of pairs of feature vectors and their classifications. The final prediction result is of Multi-disease classification on a fundus image. The chosen models were employed by using training set of 3215 samples and evaluation set of 640 samples. For assessing the performance of models the developed framework is tested on ODIR 2019 dataset. The predicted labels were evaluated with the help of area under curve (AUC), precision, recall and F1 Score metrics. All metrics were defined in the equations 3, 4, 5 respectively.

$$Precision = \frac{TP}{FP+TP} \qquad □□□$$

$$Recall = \frac{TP}{TP+FN} \qquad □□□$$

$$F1score = 2 \times \left[\frac{Precesion \times Recall}{Precesion+Recall}\right] \qquad □□□$$

The developed model is evaluated on RFMiD dataset and the attained AUC values were shown in Table 2. Whereas the proposed framework is tested on ODIR 2019 dataset and the projected performance metrics has shown in the Table 3.

ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

Table 2: Model and Its Performance on RFMiD Dataset.

Model Type	Architecture	AUC
Detector	DenseNet201	0.97
Detector	EfficientNetB4	0.98
Classifier	DenseNet201	0.97
Classifier	EfficientNetB4	0.96
Classifier	ResNet150	0.97
Ensembler	**Stacking**	**0.98**

The proposed framework is built on minimal pre-processing steps and stacked with prominent CNN architectures in order to perform disease detection and classification task. The prediction of multi-disease on fundus image with ensemble of deep learning models has attained better performance compared with existing approaches. Still there is a scope to improve this work by including other pre-processing steps like class weighting, padding and many more.

Table 3: Performance Comparison of Proposed Framework with Existing Approaches on ODIR Dataset.

Author	Method	AUC	F1score
Neha Gour et al. [3]	Two VGG16	84.93%	85.57%
J. Wang et al. [4]	EfficientNetB3	74.23%	89.21%
C. Li et al. [5]	ResNet-101	93.01%	91.32%
M. T. Islam et al. [6]	Shallow CNN	80.50%	85.00%
C.C. Jordi et al. [8]	VGG16 multi-class	88.71%	81.76%
Proposed Approach	**Stacking**	**97.42**%	**94.32**%

Considered neural network architectures are proved efficient in classification tasks. The consideration of these CNNs has proven it can perform well either individually or stacked. Further we can consider even better architectures for this task and also we can make use of other ensemble techniques like bagging and any other. Cross validation is also another technique to validate the proposed model among the bagged architectures.

Still there exist better loss functions and optimizers while training the network. The parameter fine tuning is also another future task which may increase the framework performance.

5 CONCLUSION

This paper main objective is to design and develop a framework to attain disease detection and classification tasks. To address the difficulties facing by the ophthalmologists while diagnosing multiple diseases on fundus image this approach will stand as a second reference for them. MDCF comprises of pre-processing, ensemble detectors and classifiers stages. Later based on stacking technique prediction has been produced with multi-disease classification. The projected performance is quite superior to existing literature. Further this can be improved with additional pre-processing steps, other neural network architectures and with different ensemble methods.

REFERENCES

[1] Peking university international competition on ocular disease intelligent recognition (ODIR-2019), 2020, (accessed April 8, 2021). URL https://odir2019.grandchallenge.org/.

[2] Pachade S, Porwal P, Thulkar D, Kokare M, Deshmukh G, Sahasrabuddhe V, Giancardo L, Quellec G, Mériaudeau F. Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research. Data. 2021; 6(2):14. https://doi.org/10.3390/data6020014

[3] Neha Gour, Pritee Khanna, "Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network", Biomedical Signal Processing and Control, Volume 66, 2021, 102329, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2020.102329.J

[4] J. Wang, L. Yang, Z. Huo, W. He and J. Luo, "Multi-Label Classification of Fundus Images With EfficientNet," in IEEE Access, vol. 8, pp. 212499-212508, 2020, doi: 10.1109/ACCESS.2020.3040275.

[5] C. Li, J. Ye, J. He, S. Wang, Y. Qiao and L. Gu, "Dense Correlation Network for Automated Multi-Label Ocular Disease Detection with Paired Color Fundus Photographs," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 1-4, doi: 10.1109/ISBI45749.2020.9098340.

[6] M. T. Islam, S. A. Imran, A. Arefeen, M. Hasan and C. Shahnaz, "Source and Camera Independent Ophthalmic Disease Recognition from Fundus Image Using Neural Network," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), 2019, pp. 59-63, doi: 10.1109/SPICSCON48833.2019.9065162.

ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

[7] Q. Meng, Y. Hashimoto, S. Satoh, “Fundus image classification and retinal disease localization with limited supervision”, in: Asian Conference on Pattern Recognition, Springer, 2019, pp. 469–482.

[8] C.C. Jordi, N.D.R. Joan Manuel, V.R. Carles, “Ocular Disease Intelligent Recognition Through Deep Learning Architectures”, Universitat Oberta de Catalunya, 2019, pp. 1–114.

[9] A. Ram, d.C.C. Reyes-Aldasoro, “The relationship between fully connected layers and number of classes for the analysis of retinal images”, arXiv preprint arXiv:2004.03624.

[10] Wahyudi Setiawan, Moh. Imam Utoyo, Riries Rulaningtyas, “Transfer learning with multiple pre-trained network for fundus classification,” TELKOMNIKA Telecommunication, Computing, Electronics and Control, Vol. 18, No. 3, June 2020, pp. 1382-1388, ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v18i3.14868

[11] Wahyudi Setiawan, Moh. Imam Utoyo, Riries Rulaningtyas, “Reconfiguration layers of convolutional neural network for fundus patches classification,” Bulletin of Electrical Engineering and Informatics, Vol.10, No.1, February 2021, pp. 383-389, ISSN: 2302-9285, DOI: 10.11591/eei.v10i1.1974

[12] Pearl Mary S., Thanikaiselvan V., “Unified adaptive framework for contrast enhancement of blood vessels,” International Journal of Electrical and Computer Engineering (IJECE), Vol.10, No.1, February2020, pp. 767-777, ISSN: 2088-8708, DOI: 10.11591/ijece.v10i1.pp767-777

[13] Deepashree Devaraj, Prasanna Kumar S.C., “Development of a hybrid framework to characterize red lesions for early detection of diabetic retinopathy,” Indonesian Journal of Electrical Engineering and Computer Science, Vol. 13, No. 3, March 2019, pp. 962-973, ISSN: 2502-4752, DOI: 10.11591/ijeecs.v13.i3.pp962-973

[14] Sudheer Kumar .E, Shoba Bindu .C, “Medical Image Analysis Using Deep Learning: A Systematic Literature Review”. In: Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics (ICETCE 2019). Communications in Computer and Information Science, vol 985, pp 81-97, 18th May 2019, Print ISBN: 978-981-13-8299-4, Online ISBN: 978-981-13-8300-7, Springer, Singapore. https://doi.org/10.1007/978-981-13-8300-7_8.

[15] Kumar, E.S., Bindu, C.S. Two-stage framework for optic disc segmentation and estimation of cup-to-disc ratio using deep learning technique. J Ambient Intell Human Comput (2021). https://doi.org/10.1007/s12652-021-02977-5

[16] E. S. Kumar and C. S. Bindu, "An Efficient Approach to Accomplish Automatic Segmentation of Optic Cup Using Modified U-Net," 2020 11th International

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 40 Issue: 09-2021
DOI 10.17605/OSF.IO/ ZHA9C

Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-5, doi: 10.1109/ICCCNT49239.2020.9225537.

[17] E. Sudheer Kumar, C. Shoba Bindu, “Segmentation of Retinal Lesions in Fundus Images: A Patch Based Approach Using Encoder-Decoder Neural Network,” 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS).

[18] IEEE International Symposium on Biomedical Imaging (ISBI-2021), Nice Acropolis Convention, and Exhibition Centre, France. On April 13th, 2021, https://riadd.grand-challenge.org/ .

[19] Shorten, C., Khoshgoftaar, T.M. “A survey on Image Data Augmentation for Deep Learning,” J Big Data 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

[20] Dumitrescu D, Boiangiu C-A. “A Study of Image Upsampling and Downsampling Filters,” Computers. 2019; 8(2):30. https://doi.org/10.3390/computers8020030

[21] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[22] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” Proc- 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017 - January, pp. 2261–2269, Aug. 2016.

[23] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” 36th Int. Conf. Mach. Learn. ICML 2019, vol. 2019-June, pp. 10691–10700, May 2019.

[24] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Dec. 2016, vol. 2016-December, pp. 770–778, doi:10.1109/CVPR.2016.90.

[25] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 2, pp. 318–327, Aug. 2017, Available: http://arxiv.org/abs/1708.02002.

[26] Wang, B., Chen, L., Sun, W., Qin, K., Li, K., & Zhou, H. “Ranking-Based Autoencoder for Extreme Multi-label Classification,” NAACL-HLT. (2019).