INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 03 | MARCH - 2024 SJIF RATING: 8.176 ISSN: 2582-3930

Retinal Disease Classification using Deep Learning

S.A. Shivarkar
Assistant Professor
Department of CSE
JSPM's JSCOE,Pune, India

Aishwaryaraj Chandel
Student
Department of CSE
JSPM's JSCOE, Pune, India

Sumit Maradkar
Student
Department of CSE
JSPM's JSCOE, Pune, India

Umakant Dodtalle
Student
Department of CSE
JSPM's JSCOE, Pune, India

Abstract- Retinal disease, which consist a varying range of eye disorders, can be a serious and alarming threat to vision and ocular health. Timely diagnosis is the most important as in early stages it is easier to plan the treatment and intervention. In recent years, the involvement of the deep learning methodologies into the field of ophthalmology has been realized through medical imaging that has revolutionized the ophthalmology study. This research investigates the use of deep learning methods that help in classification of retinal diseases through high-resolution cadaver images. The focus of this work is providing an entirely operational and cutting-edge system for retinal disease classification using deep learning covering the main technical challenges and issues on ethics. The made model is a major advancement in the process of detecting retinal diseases with the aim of improving the quality and efficiency of the eye care. The model is also expected to improve patient outcomes.

Keywords: Re-winding, health classification, deep learning, RNN.

I. INTRODUCTION

There are absolutely no place barriers, or economic barriers, or cultural barriers for the spread of blindness. Modern imaging technology has assisted these medical practitioners establish accurate diagnosis by ascertaining the symptoms. There exists a variety of diagnostic modalities, namely angiofluoroscopy and ocular coherence tomography. Thermal imaging that has been used lately for breast cancer detection could also be used o diagnose diabetic foot or a number of eye diseases.

It focuses on retinal images, like fundus photographs or OCT scans, and picks out intricate patterns or tiny abnormalities which might be impossible to the naked eye identified. Replacing doctors as frontline diagnostic screening with deep learning based algorithms learns from huge data sets by creating highly precise computer-aided diagsnostics that would determine the type of several retinal diseases as diabetic retinopathy (DRD), age-related maculodepression (AMD), or glaucoma (glaucoma) with potential life-changing implications for the affected patients. This technology increases the possibility of early detection, and better diagnosis can lead to vision impairment prevention through the whole world.

This work can be summarised as follows:This work can be summarised as follows:

The multi dimensional dataset (MuReD Dataset) is generated in a personalized way and consists of twenty classes of the diseases collected from the state of the art sources. It is cleaned automatically using quality scores which are given according to the sharpness and brightness of the image.

A new approach is used for the first time and that is of transformer-based model which has been optimized by very thorough experimentation to detect and classify some retinal diseases.

The categorization techniques for fundus imaging which currently exist are presented in (section III). Also, a review of traditional approaches to balancing class imbalance is done.

The issue of building a transformer model is addressed in section IV.

Finally, we mention some fundus datasets available for public in part V.

Finally, the article concludes and presents the most important contributions of the research and also suggesting new ways of research that might be useful for the future research.

II. EASE OF USE

A. Accuracy: These models have the ability to analyze large datasets and identify complex patterns in retinal pathology, resulting in more precise disease classification

B. Accessibility: Accessibility allows healthcare professionals to utilize these models regardless of their location or available resources.

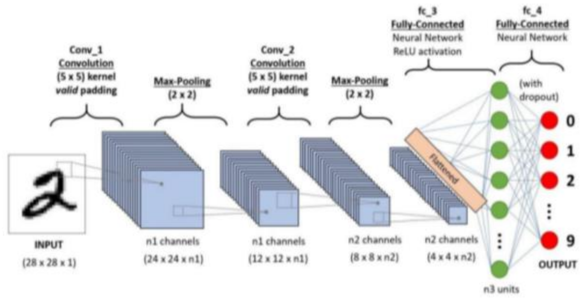
C. Integration: Smooth integration allows healthcare workflows and researchers to adopt these systems without disrupting their current practices.

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 03 | MARCH - 2024 SJIF RATING: 8.176 ISSN: 2582-3930

III. RELATED WORK

A. *CNN*

The rapid growth of Artificial Intelligence surpasses all other technologies in terms of capability to diminish the gap between the human potential and the artificial capacity. As such, the researchers alongside the fans have been engaged in diverse spheres of the subject, pushing for more desirable things. Paper considers the subject with other examples that are the typical cases of computer vision. The field here is to enable machines to view the world akin to that of humans, think like they do, and use differentially. Though computer vision with deep learning advances, mostly throughout time, the main concern has been about a single algorithm – Convolutional Neural Network. A deep learning algorithm called ConvNet/CNN operates as follows: it takes an image as input and weights the objects/aspects within it in terms of their significance and distinguishes them from one another. It is the stimuli within only the small visual field segmented to receptive areas that triggers the neurons to react. Since the receptive fields focussing on the same points overlap, they cover the whole visual field.



B. *Class Imbalance*

Multi-label datasets have a disability of class imbalance frequently where minority of classes have majority of examples and on the other hand, majority classes have fewer examples compared to the minority classes. This is known as the long-tail distribution problem.There are three main approaches to address the class imbalance, with their effectiveness being dependent on the particular application: Resampling Techniques: These methods weigh on pre-processing of the multi-label dataset making it classifier independent. The methods can be classified into two categories: oversampling techniques that create new samples from the classes which have fewer number of samples and undersampling methods that remove samples from the classes which have more number of samples. Due to their model-independence property, this category of methods is the one that remains the most popular in this field.

Classifier Adaptation: The model has to be built to manage the skewedness of the data set. It is not as well known as it because it needs both the adeptness of classifier and the ability of the domain problem and usually produces the more complex and specialized ones training pipelines.

Ensemble Approaches: That is to say these methods utilize two or more models, each of which learns a different list of labels, and ultimately, the predictions of each model are combined to form the complete list of labels.

C. *Transformers*

The transformers are a large class of pre-trained neural parameters that effectively considering the positional contexts of the retinal images and provide correct diagnosis.

.

Attention Mechanism: Attention mechanisms enable transformers to focus on the distinguished parts of the input data, while suppressing the other parts. This is possible in retina disease classification by making the model to detect specific areas of the retina that are most telling of the presence or severity of the disease.

Feature Extraction: The transformers efficiently use feature extraction from images of high dimension. Instead of decomposing the input image into a sequence of small tokens and concatenating them sequentially, transformers parse the helpful semantic segments and reconstitute them repeatedly through multiple layers, which enables the extraction of hierarchical representations. The depiction of the local as well as the global features (that are important in eye disease classification) is obtained in these representations.

Model Interpretability: Transformers' attention mechanism makes the model interpretable by displaying a map of the areas in the retinal image that carry the most relevance when making classification decisions.

Transfer Learning and Pre-training: The pre-training of transformers on datasets such as ImageNet has been shown to be effective in terms of capturing general visula representations that can later be fine-tuned for a specific purpose such as retinal disease identification.

IV. METHODS.

i. **Data Collection and Preprocessing**: Collect a large cohort of retinal images with labels specifying whether retinal diseases are present or not.Prepare the images for analysis by uniformly resizing, resolving and formatting them. This might imply changing image resolution, trimming, and scaling.

ii. **Transformer Architecture Selection**: Pick a transformer model that can be applied to image categorization processes. The most often selected ones include Vision Transformer (ViT), Convolutional Transformer (CT), and the other kinds of transformers customized for image processing.

iii. **Data Augmentation**: Expand the training dataset to make the diversity and reliability through its diversification. Common augmentation techniques are

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 03 | MARCH - 2024 SJIF RATING: 8.176 ISSN: 2582-3930

image rotation, flipping, scaling, and adding noise to the images.

iv. **.Model Training**:Use random or pretrained weights from a model that was built to recognize common objects in a large amount images of ImageNet dataset for the transformer model initialization.

v. **Fine-Tuning and Transfer Learning**:Lastly, fine-tune the pre-trained transformer model on the retinal disease classification task with the target dataset using the optionally. With this step, the model can conform to the peculiar features of the retinal images.

vi. .**Model Evaluation**:Assess the trained model on a separate test dataset to obtain in classifying retinal diseases' performance. Compute evaluation metrics such as accuracy, precision, recall, F1-score, and area-under-the-curve (AUC) to quantify the model's efficacy.

A. Multi-Label Classification Model :

C-Trans architecture (Convolutional Transformer), the composition of CNNs and transformers, to carry out the multilabel classification tasks, such classifying the retinal diseases in medical images.

Convolutional Layers: The architecture starts with a number of convolutional layers. The layers performs the feature extraction from the input images in a hierarchical fashion. Each convolutional level is evaluated with a set of filters on the input image to detect local patterns and features at different spatial levels.

Transformer Encoder: Coming after the convolution layers, the C-Trans system takes a transformer encoder. The encoder of the transformer is created using multiple blocks of transformers placed on top of each other. The self-attention mechanisms and feedforward neural networks comprise each transformer block. The self-attention mechanism encourages the model to remember long-range information and finally, comprehensively understand the input image.

.**Classification Head**: On top of the architecture, the classification head is added to output the last many-label classification predictions. The output of the transformer decoder usually is summarized and aggregated to obtain a fixed size of the whole image representation. Eventually, this representation is provided to fully connected layer or to a series of fully connected layers and softmax activation function to go on with the training.

Loss Function:The C-Trans network is optimized during the training to cater to multi label classification tasks using an appropriate loss function.

Training Process: The C-Trans model is trained on a labelled dataset of retinal images with the associated multi-label annotations that either suggest the presence or absence of different types of retinal diseases. During training, the model attains the capability of extracting useful features of the input images and to correctly predict the disease label for each disease class.

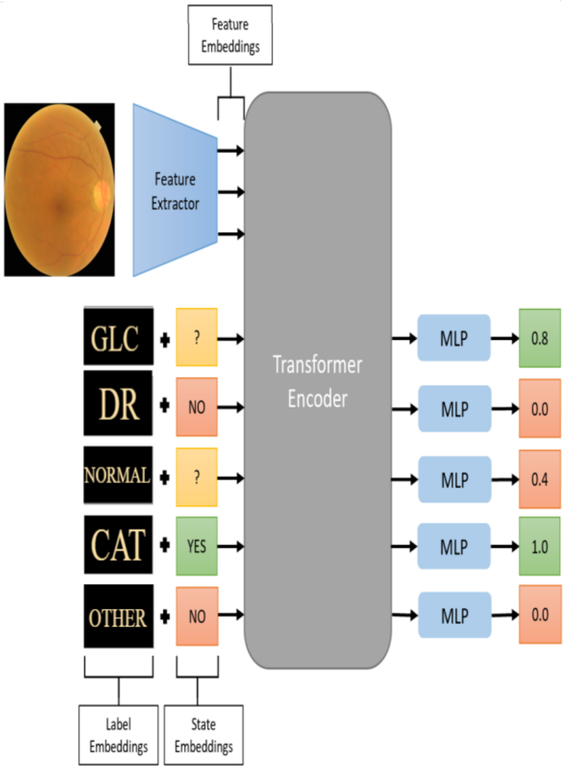


Fig: C-Transformer

B. Equations

Convolution Operation:

$$z_{ij} = \sum_{m=0}^{f-1} \sum_{n=0}^{f-1} x_{i+m,j+n} \cdot w_{m,n} + b$$

Pooling Operation (e.g., Max Pooling):

$$y_{ij} = \max_{m=0}^{k-1} \max_{n=0}^{k-1} x_{i \times s+m, j \times s+n}$$

V. DATASET

Implementing a deep learning model to detect the diseases within the retina through a collection of images of eye patients database. There are several thousand images in the training dataset. Testing dataset composed of images of retinal diseases which are used for image processing.

No	**OCT Images Table**			
	Labels	***No of Training Images***	***No of Testing Images***	***Image Size (Pixels)***
1	GLC	8000	2000	64x64
2	DR	8000	2000	64x64
3	NORMAL	8000	2000	64x64
4	CAT	8000	2000	64x64

A. Dataset Cleaning

To make sure of data quality and reliability, data sets are error-free. Firstly, preprocessing is carried out based on the techniques of normalizing and rescaling to standardize retinal photographs, ensuring the uniformity of the levels of pixel intensity and its size. One of the approaches on dealing with the imbalance of classes is the use of intellectual strategies to create a balanced representation of classes of diseases by oversampling, undersampling or original synthetic data generation.

For this, outlier detection and removal are carried out because abnormal data points can corrupt the neural network training. Feature selection and dimensionality reduction techniques, performed through the selection and retaining of only the most suitable/relevant features, help to improve the ease of data handling by reducing the computational complexity and overfitting risks. Finally, by increasing the diversity of the dataset more data augmentation methods like geometric modifications and color adjustments are used to achieve better model generalization.

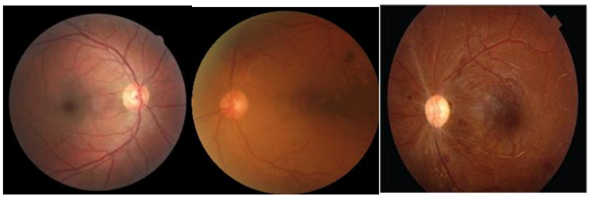


Fig: Dataset Images of retinal diseases

B. Normalization

Data input standardization is vital as it is the key to efficient model training, and this is where the concept of dataset normalization comes into the picture. One of the normalization techniques that is frequently used is scaling the retinal image pixel intensity values in a uniform range, usually 0-1. This normalization guarantees that in the pictures contrast or brightness differences won't make bad the model's work.

C. Feature Extraction

Feature extraction for detecting pretentious patterns diagnostic of different disorders is a must-do task in the process of the assessment of retinal images. One well-known and efficient feature extraction technique is based on employing pre-trained convolutional neural networks (CNNs), such as the VGG, which have already been trained on large image datasets like ImageNet. Networks of convolutional layers have the ability to approximate a hierarchy of characteristics starting from raw image input where features can be anything from low-level edges, textures to high-level semantics structures or shapes.

VI. EXPERIMENTAL SETUP AND RESULTS

A. Metrics :

i.**Accuracy Score**: When we were conducting the performance of deep learning model for retinal disease classification, accuracy score was the main metric that we used. The accuracy defined as the quotient of the correct predictions to the total number of provided input samples is among the most convenient metrics of the model's overall classification success. It is calculated as :It is calculated as :

Accuracy = (Total correct Prediction)/(Total Prediction).

In our study, the deep learning model was tasked with classifying retinal images into one of four possible conditions: Anabolism, Jerky, Diabetic Retinopathy, or Normal. The model's performance was then measured by its accuracy score which reflected its ability to classify these conditions correctly. The higher the accuracy score, the more correct predictions the model makes, and vice versa, implying that the model is better able to classify retinal diseases.

ii.**Precision**: It is given as the number of correctly classified positive samples (True Positives) divided by the sum of the correctly classified positive ones (True Positives) and the number of false positives (False Positives). It attempts to answer the question: What was the proportion of pre-emptive identifications that turned out to be correct? Mathematically, precision is defined as follows:Mathematically, precision is defined as follows:

A model with a false positive score of 0 produced the precision of 1.0.

iii. **Recall** :, It is equivalent to sensitivity, measures the ratio of correctly classified positive samples (the True Positives) to the total number of actually positive samples (the Total True Positives and False Negatives). It attempts to answer the question: In terms of how many actual positives are found successfully? Mathematically, recall is defined as follows:Mathematically, recall is defined as follows:

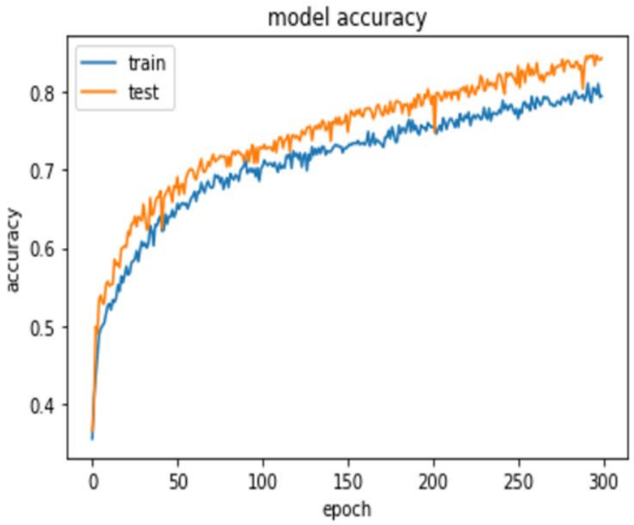


Figure : Model accuracy for normal eyes

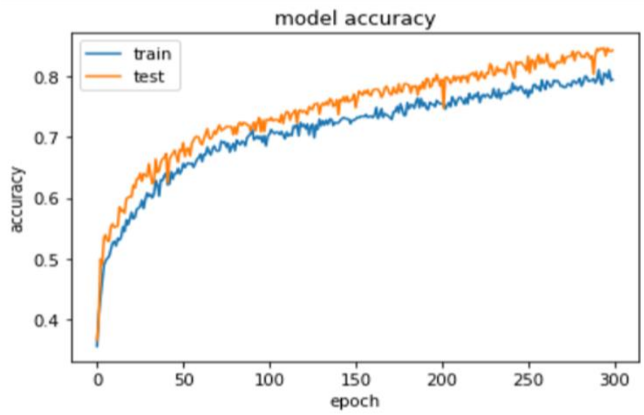


Figure : Model accuracy for Diabetic retinopathy

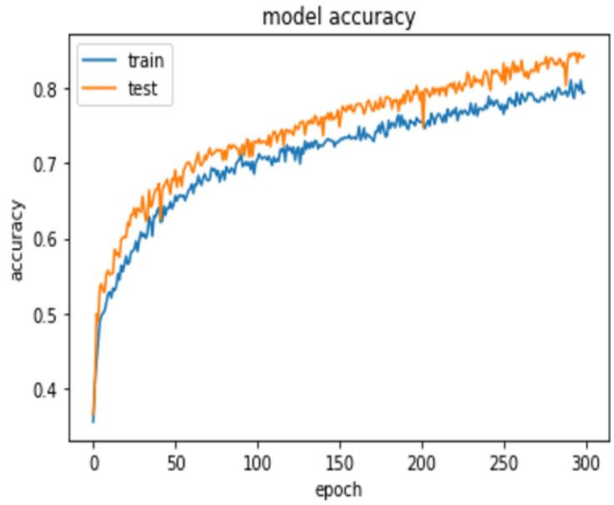


Figure : Model accuracy for cataract

V. CONCLUSION

Innovation in retina image processing introduced is exploiting the combination of different methods for segmentation, and cognitive model of a swarm of nanoparticles for swarm control possibility in retina vessels at future. Being data processing is done using segmentation and denoising of retina images, a proposed bit combination of methods is an option provided to increase the quality of images which will be a framework for a good diagnostic. Machine learning experienced using deep learning to classify retina retinal-fundus images has the advantages of accuracy and being useful as support for doctors, thereby, what is required to have an automated diagnosis of retina disease is just possible.

VI. REFERENCES

[1] K. Mittal, V. Mary, and A. Rajam, "Computerized retinal image analysis - a survey," Multimedia Tools Appl., vol. 79, no. 31/32, pp. 22389–22421, Aug. 2020.

[2] M. Badar, M. Haris, and A. Fatima, "Application of deep learning for retinal image analysis: A review," Comput. Sci. Rev., vol. 35, Feb. 2020, Art. no. 100203.

[3] J. Han, M. Kamber, and J. Pei, Data Mining Concepts and Techniques (Morgan Kaufmann Series in Data Management System), 3rd ed. Amsterdam, The Netherlands: Elsevier, 2011.

[4] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," IEEE Rev. Biomed. Eng., vol. 3, pp. 169–208, 2010.

[5] M. M. Fraz et al., "Blood vessel segmentation methodologies in retinal images–a survey," Comput. Methods Programs Biomed., vol. 108, no. 1, pp. 407–433, 2012.

[6] S. Stolte and R. Fang, "A survey on medical image analysis in diabetic retinopathy," Med. Image Anal., vol. 64, 2020, Art. no. 101742.

[7] J. Cheng et al., "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," IEEE Trans. Med. Imag., vol. 32, no. 6, pp. 1019–1032, Jun. 2013.

[8] M. A. Omar, M. A. Tahir, and F. Khelifi, "Multi-label learning model for improving retinal image classification in diabetic retinopathy," in Proc. 4th Int. Conf. Control, Decis. Inf. Technol., 2017, pp. 0202–0207.

[9] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. T. Salles, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks," Comput. Biol. Med., vol. 116, 2020, Art. no. 103537.

[10] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, and W. Qian, "An interpretable ensemble deep learning model for diabetic retinopathy disease classification," in Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2019, pp. 2045–2048.

[11] Y. N. Yan, Y. X. Wang, L. Wang, J. Xu, W. B. Wei, and J. B. Jonas, "Fundus tessellation: Prevalence and associated factors: The beijing eye study 2011," Ophthalmol., vol. 122, no. 9, pp. 1873–1880, 2015.

[12] L.-P. Cen et al., "Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks," Nature Commun., vol. 12, no. 1, Aug. 2021, Art. no. 4828.

[13] Diabetic retinopathy detection, [Online]. Available: Accessed: Jun. 2022. https://www.kaggle.com/c/diabetic-retinopathy-detection

[14] L. Ju, X. Wang, Z. Yu, L. Wang, X. Zhao, and Z. Ge, "Long-tailed multilabel retinal diseases recognition using hierarchical information and hybrid knowledge distillation," 2021, arXiv:2111.08913.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.

[16] Peking university international competition on ocular disease intelligent recognition (odir-2019), [Online]. Available: Accessed: Jun. 2022. https: //odir2019.grand-challenge.org [

17] J. He, C. Li, J. Ye, S. Wang, Y. Qiao, and L. Gu, "Classification of ocular diseases employing attention-based unilateral and bilateral feature weighting and fusion," in Proc. IEEE 17th Int. Symp. Biomed. Imag., 2020, pp. 1258–1261.

[18] 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," in Proc. IEEE 17th Int. Symp. Biomed. Imag., 2020, pp. 1–4.

[19] A. N. Tarekegn, M. Giacobini, and K. Michalak, "A review of methods for imbalanced multi-label classification," Pattern Recognit., vol. 118, 2021, Art. no. 107965.

[20] F. Charte, A. J. Rivera, J. MarïaJ. del, and F. Herrera, "Addressing imbalance in multilabel classification: Measures and random resampling algorithms," Neurocomputing, vol. 163, pp. 3–16, 2015.

[21] F. Charte, A. Rivera, María José del Jesus, and F. Herrera, "A first approach to deal with imbalance in multi-label datasets," in Proc. Int. Conf. Hybrid Artif. Intell. Syst., Jeng-Shyang Pan, Marios M. Michał Polycarpou Wo´zniak, C. P.L.F. André de Carvalho, Héctor Quintián, and E. Corchado, Eds., Berlin, Heidelberg: Springer, 2013, pp. 150–160.

[22] F.-F. Luo, W.-Z. Guo, and G.-L. Chen, "Addressing imbalance in weakly supervised multi-label learning," IEEE Access, vol. 7, pp. 37463–37472, 2019.

[23] K.W. Sun and C. H. Lee, "Addressing class-imbalance in multi-label learning via two-stage multi-label hypernetwork," Neurocomputing, vol. 266, pp. 375–389, 2017.

[24] M. A. Tahir, J. Kittler, and F. Yan, "Inverse random under sampling for class imbalance problem and its application to multi-label classification," Pattern Recognit., vol. 45, no. 10, pp. 3738–3750, 2012.

[25] M. A. Tahir, J. Kittler, and A. Bouridane, "Multilabel classification using heterogeneous ensemble of multi-label classifiers," Pattern Recognit. Lett., vol. 33, no. 5, pp. 513–523, 2012.

[26] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 2980– 2988.