ISSN 2522-9400 **European Modern Studies Journal** Vol 5 No 6
available at journal-ems.com

GMD Model Based on Multi-Label Classification for Detection and Diagnosis of Eye Diseases

Serhii Yaroshchak and Mahmoud Smaida
The National University of Water and Environmental Engineering, Rivne, Ukraine

Abstract. The diagnosis of eye diseases is still considered a challenge in the health care sector. This is mainly due to the fact that eye diseases are difficult to recognize even for the optometrist as they come in a variety of forms. Therefore, using of AI should be made available to help the ophthalmologist to detect and diagnose of eye diseases. In this work, we will introduce GMD model based on multi-label classification with additional techniques to detect and diagnosis eye diseases. Confusion matrix will be used as evaluation performance model.

In this study, we will compare three models of neural network, AlexNet, VGG16 and Inception-v3 with the GMD model in order to evaluate performance of the models. The dataset consists of four classes of eye diseases, Glaucoma, Myopia, Diabetic retinopathy and Normal. All these networks based on label classification deployed using GPU in Google Colab, according to all the experiments, we obtained results for each combination and observed that for multi-label classification, GMD model gives the best accuracy (95%) compared with the other models.

Keywords: multi-label classification, GMD model, eye diseases, confusing matrix, deep learning, diabetic retinopathy, glaucoma, myopia

Introduction

The significance of machine learning, artificial intelligence and deep learning algorithms have been significant in evolving of the modern world. These technological algorithms are however observed to facilitate the subjects having various eye diseases by means of accurate identification. The identification of early complications associated with eye diseases or ophthalmic diseases is of great eminence to protect the eyes from further disruptions and complexity. The diagnostic methods of diseases associated with the eye are generally observed by doctors but it may lead them towards the excess of duration and can be misdiagnosed due to human error, thus a deep implementation of machine learning and artificial intelligence algorithm is required for providing suitable clinical measures for the safety of further complexions (Malik, 2021). The field of artificial intelligence is designed to copy the style and nature of a human and is further partitioned under general, super, and narrow intelligence systems.

The whole artificial intelligence is a complete system that includes scheduling and automatic learning, conventional learning, natural language processor, smart sensing, and robotics (Meskó, Hetényi, & Győrffy, 2018). Although, machine learning and its algorithms have a significant role in detecting the causes of ophthalmic disorders by means of its applications such as writing recognition, speech learning, medical interventions, and machine reading. But both deep learning and conventional machine learning have been identified as the driving force of machine learning (Bengio, Courville, & Vincent, 2013). The field of ophthalmology has however been special in the implementation of driving artificial intelligence along with the deep and machine learning algorithms to simply assess the disease using analysing a variety of ophthalmic shots which provides the necessary data through computed algorithms (LeCun, Bengio, & Hinton, 2015). Based on the current trends of the automated significance and applications of artificial intelligence, deep and machine learning, it is time to predict that these interventions would be beneficial in the identification of diabetic retinopathy, glaucoma, and other related ophthalmic complexions, irrespective of the doctors, consultants and trained specialists. Thus, the study is merely focused on the available and modified clinical applications of machine learning or deep learning algorithms that can be extensively utilised with the ophthalmic imaging machines or modalities to diagnose any complications related to ophthalmic disorders.

Our goal in this paper is to create convolutional neural network model based on multi label classification more effective to detects and diagnoses of eye diseases, such as, Glaucoma, Myopia and Diabetic retinopathy (GMD) using fundus images. Moreover, authors added to the model some different functionalities to make it able to detect between variations features for all the classes.

Multi Label Classification and Label Smoothing Techniques

Multilabel classification is a predictive modelling task that assigns to each sample a set of target labels. Using label classification for the purpose of the possibility of having more than one disease in the input image. The major reasons multi-label classification is considered to be effective and applicable within the medical field is based on the fact that such models are considered to be efficient in terms of producing accurate results that are regarded to be conducive for detecting and treating the eye disease effectively. For all the multi label classification tasks sigmoid function is almost used (Lin, Cai, & Lin, 2021).

It has been observed that label smoothing is a regularisation technique that is used for classification problems and helps the model to prevent to predict labels confidently. There are problems that have been encountered such as overfitting and overconfidence therefore, the method of label smoothing has been used to eliminate any biasness. Zhang et al. (2019) state that label smoothing is used to provide gains across many tasks and it has been further observed that label smoothing is equivalent to confidence penalty if the order of KL divergence between uniform distributions and models output is reversed. Label smoothing has been used to improve the accuracy of deep learning models across different range of tasks such as image classification, speech recognition and machine translation (Ju et al., 2021). Label smoothing causes confidence of the model to decrease and it prevents the model from descending into deep crevices of loss of landscape where outfitting occurs. Label smoothing regularisation (LSR) was introduced as a regularization technique for classification problems to prevent the model from predicting the labels too confidently during training and generalizing poorly.

The technique that is adopted in the following research is the label smoothing where the study conducted by Tymchenko, Marchenko, and Spodarets (2020) has demonstrated that the label smoothing is an effective method in respect to early diagnosis along with detecting the eye diseases at the early stages. The label smoothing is associated with the use of diagnosing algorithms that concentrates on the image-level diagnosis. Moreover, the label classification is identified to be in high demand in respect to the medical systems which is also indicated as automate diagnostic processes. Furthermore, the label smoothing utilizes the colour fundus photography that produces high quality images that is highly useful for the detection of eye disease (Adaimi, Kreiss, & Alahi, 2019). Therefore, the use of label smoothing is highly important in the development of algorithms as it can support in early diagnosis of the eye diseases. Furthermore, a classification model is calibrated and cause probabilities to be identified and reflect the outcome of the accuracy (Hussain, Bird, & Faria, 2018). The calibration model is important for the model interpretability and reliability and helps to decide thresholds for application and the formula of label Smoothing is as follows:

$$y\_ls = (1 - \alpha) * y\_hot + \alpha / K \quad (1)$$

Where, K is the number of label classes and α is the hypermeter that is used to determine the amount of smoothing. It has further been observed that if $\alpha = 0$ the original one hot encoded *y_hot is* obtained. Label smoothing is used when there is less of function and it can cause the predictability of the model to decrease along with the confidence level (Li, Dasarathy, &

Berisha, 2020). The TensorFlow is already implemented in the label smoothing within different cross entropy loss functions. It has been observed that different algorithms are applied in this regard and it causes different thresholds to be established. Therefore, it causes segmentation of complex images to be considered and it is one of the most complicated computation processes (De Fauw et al., 2018). The process involves the labelling of each pixel in the complex image separately so that different segments can formed for simpler calculations of their visual characteristics. There are several methods are designed for the automatic segmentation of the complex images in the medical diagnosis which incorporates termed thresholding, edge-based segmentation, region-based segmentation, and the clustering-based segmentation methods (Pion-Tonachini, Kreutz-Delgado, & Makeig, 2019). In terms of medical diagnosis, the segmentation of the complex images makes the process of analysis and understand.

Related Work

The study of Gour and Khanna (2021) has evaluated the multi-class multi-label classification for eye disease detection. Fundus imaging has been considered the retinal image modality for getting the anatomical structures and abnormalities in the human eye. The primary tools for the observation and detection of a wider range of ophthalmological diseases. The changes in and around the anatomical structures such as the optic disc, fovea, blood vessels, and macula reflect the disease's presence, such as the diabetic retinopathy, age related macular degeneration, glaucoma, myopia, cataract, and hypertension. The patient would be suffering from the different ophthalmological diseases observed in one eye or both eyes. Two models have been proposed for the multi-class multi-label fundus images classification of the disease applying transfer learning based on the convolutional neural network approaches. The ocular disease intelligent recognition database with the fundus images of patients' right and left eyes for the eight categories has been used for experimentation. The study has obtained the effectiveness of the multi-label classification in eye disease detection.

In terms of multi-label classification, it is considered crucial to reiterate the views of Gour and Khanna (2021) who claimed that the notion of multi-label classification is regarded to be effective in terms of early diagnosis as well as the eye diseases are detected at an early stage. Since multi-label classification is considered to use an array of classification, He et al. (2021) state that it helps to reduce the eye disease from progress any further and thus prevents any irreversible damage that can be levied on the eyes. Moreover, it is stated by Fu et al. (2018) that the reason multi-label classification is considered to enjoy mainstream success is based on the fact that it employs the use of colour fundus photography and is widely considered to be an effective tool in terms of screening the eyes for detection purposes. In addition, the author’s further state that with the use of colour fundus photography, early symptoms of eye diseases are effectively detected, and for further progression of the disease, multi-label classification employs the use of diagnosing algorithms. Wang et al. (2020) further illustrate that the notion of colour fundus photography allocates its focus on concentrating on image-level diagnoses of the impending eye diseases and helps in establishing a better correlation between the right and left eyes.

Following on with the study conducted by Yin et al. (2019), it can be stated that multi-label classification also uses Convolutional Neural Networks since it is regarded as an effective measure for extracting a considerate amount of information for the purpose of establishing a correlation between the features involved. Moreover, the author’s further state that the notion of Convolutional Neural Networks works independently in terms of multi-label classification and serves its objective of extracting two data sets in their entirety to be used for colour fundus pictures.

In addition to the points presented above, Gargiulo et al. (2019) state that the methodology called Hierarchical Label Set Expansion (HLSE) used for ordering data labels, as well as an analysis of the impact of different word Embedding (WE) models that explicitly include grammatical and grammatical features. The authors evaluate the above methodologies in the PubMed cohort of scientific papers, where the problem of multi-class and multi-label text classification is defined using the Medical Subject Heading (MeSH) label set. The empirical evaluation proves the usefulness of the proposed HLSE methodology and also gives interesting results regarding the influence of different applications and WE model sets as input to a neural network in this type of application.

In accordance with the study conducted by Bali and Mansotra 2021), it can be stated that multi-label classification has been realized amid a large number of demands in the medical systems and one of the major demands is considered to automate diagnostic processes. Moreover, it is stated by Lin, Cai, and Lin (2021) that one of the major reasons multi-label classification is considered to be effective and applicable within the medical field is based on the fact that such models are considered to be efficient in terms of producing accurate results that are regarded to be conducive for detecting and treating the eye disease effectively. On the other hand, Jain et al. (2020) dictate that for the purpose of realizing early detection of eye diseases, multi-label classification similarly uses colour fundus photography for the purpose of producing visually high-quality images however, as an additional feature, multi-label classification realizes the objective of using contract sensitivity (Rezaei-Dastjerdehei, Mijani, & Fatemizadeh, 2020).

Material and Method

Datasets

The dataset is utilized from the open-source database “Kaggle” that is confirmed and classified by an Ophthalmologist (Katrina Demetri) works in Brytanskyy Oftalmolohichnyy Tsentr clinic in Kiev. This dataset has four different classes that incorporated 2050 images based on Glaucoma, Diabetic retinopathy, Myopia, and Normal images. In this work, the highest variance among each class is distinguished which is a primary intention of this work. Moreover, 2050 images are utilized in the training phase, 20% for validation from total dataset.

The dataset divided in two types, original fundus images and synthetic fundus images. The original dataset was collected from Kaggle and iChallenge-GON Comprehension that has high-resolution images. The second type of dataset is from synthetic fundus images they are generated through Deep Convolutional Generative Adversarial Network (DCGAN) (Smaida, Yaroshchak, & El Barg, 2021). This dataset is utilized for dealing with eye-based diseases including Glaucoma, Myopia, Diabetic Retinopathy, and Normal diseases. These four types of eye diseases are the major diseases and the datasets are obtained based on these diseases. In the following Figure 1 and Figure 2, the sample images from original and synthetic datasets are shown respectively. For the original dataset, 1395 images are collected based on Fundus images, and the synthetic datasets are obtained from DCGAN are 655 and they are differentiated in Table 2 below.

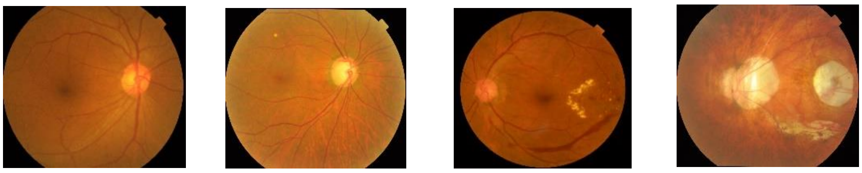


Figure 1. Normal Fundus, Glaucoma, Diabetic retinopathy and Myopia in the original dataset

Figure 2. Normal Fundus, Glaucoma, Diabetic retinopathy and Myopia in the synthetic dataset (Smaida, Yaroshchak, & El Barg, 2021)

As similar to original dataset, the synthetic of datasets are utilized in terms of size (64x64), colour, format (png), etc.

Table 1. Eye diseases samples in the original dataset and synthetic dataset

Diseases	Original Dataset	Synthetic data using DCGAN	Total images
Glaucoma	323	161	484
Diabetic retinopathy	407	133	540
Myopia	208	278	486
Normal	457	83	540
Total	1742	655	2050

Model Building

The main objective of this work is to create a model (GMD model) based on label classification model capable of detecting and diagnosing eye diseases. In this work, we will explain GMD model, and therefore we will compare the results in a comparison with popular models such as, Alex net, VGG16 net and Inception v3 net.

GMD model

The following research intends to adopt the Glaucoma, Myopia and Diabetic retinopathy (GMD) model for comparing the accuracy. GMD model is a predicted model based on multi label classification, it is a set of convolution layers, pooling layers and fully connected layers as it shown in Figure 3. This model utilized additional functions in order to enhance the accuracy to classify eye diseases. Figure 3 shows that the input image is set to 64 x 64 pixels with 3 RGB channels. Data augmentation layer has been used to increase the amount of data by adding slightly modified copies of existing data. Features were extracted from the images used seven convolution layers: the first convolutional layer is 32 filters with (3 x 3) convolution filter. Zero padding has been used to allow for more space for the kernel to cover the image and keeping information at the borders. Next, convolutional layer 32 filters with (3 x 3) convolution filter, followed by max pooling size of 2x2 pixels. Then convolutional layer 64 filter with (3 x 3) convolution filter, followed by Zero padding and drop out. Another convolutional layer 64 filter with (3 x 3) convolution filter has been used followed by max pooling and batch normalization to speed up the learning of the model. convolutional layer 128 filters with (3 x 3) convolution filter added to the network followed by zero padding. Then, convolutional layer 128 filters, max pooling a 2x2 pixel window was used and drop out to avoid overfitting. The last convolutional layer 256 filters used (3 x 3) convolution filter, then flatten layer to flattens the multi-dimensional input tensors into a single dimension. Dense 300 units were used as fully connection followed by drop out, finally, output layer (4 units) to predict eye cases. CNNs adjust their filter weights through backpropagation, which means that after the forward pass, the network is able to look at the loss function and make a backward pass to update the weights.

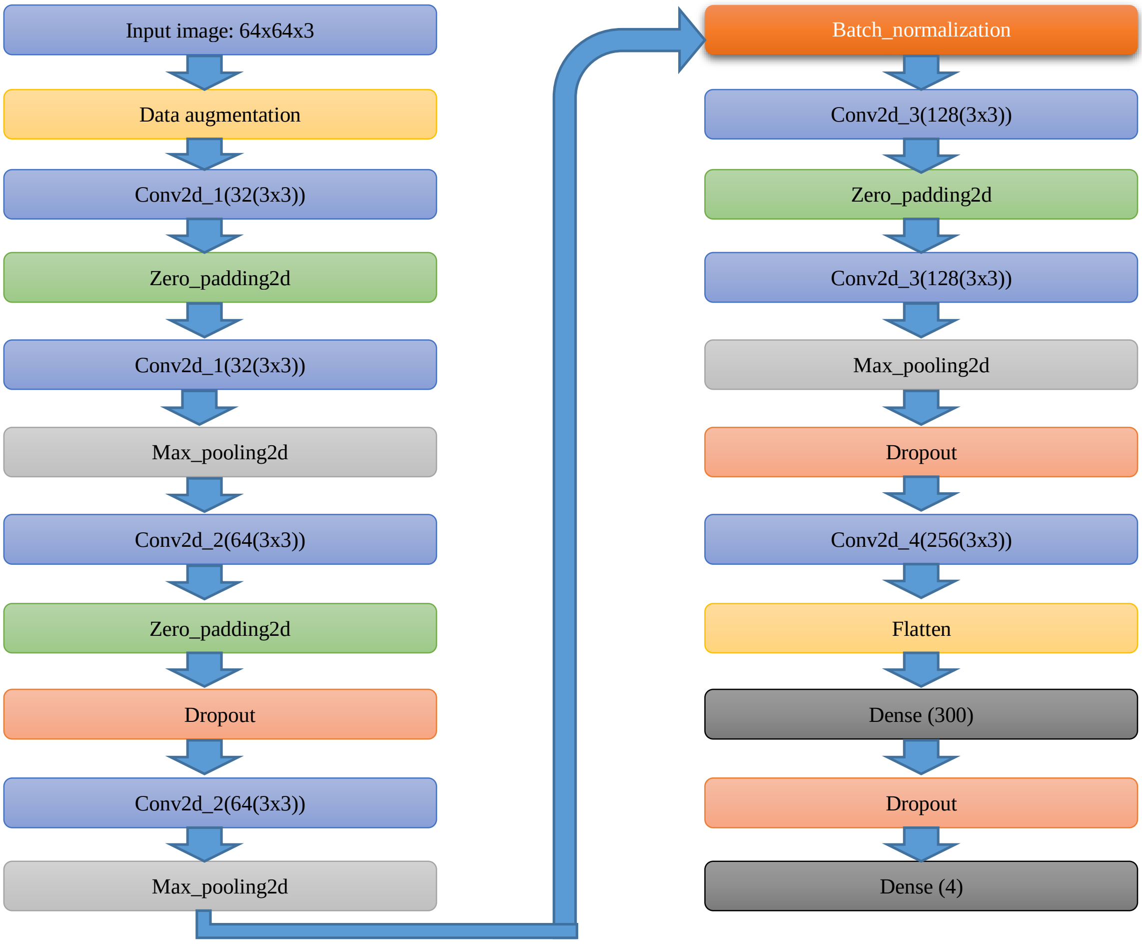


Figure 3. Block diagram of GMD network

GMD model is a type of neural probabilistic classifier, we proposed some additional techniques to improve the training model and speed up the training model.

One-hot-encoding. One-hot-encoding: is Per-Processing dataset techniques to vectorize a class as vector representation and all the classes present in the way below:

{Normal= [1,0,0,0], Myopia= [0,1,0,0], Diabetic retinopathy= [0,0,1,0], Glaucoma= [0,0,0,1]}. This technic also called binarize label.

Activation function. For all the multi label classification tasks sigmoid function is almost used. In our case, we used softmax() function for reason that makes the model give a probability for each label in the image input and using sigmoid for the output. Sparse_categorical_Crossentropy() function has been used for GMD model to categorical class.

By applying these techniques, the accuracy was increased and multi label classification is done. Our goal is to make GMD model more predictable to use in real-world task, here another problem occurs, which is the training of the model being slowly and sometimes get overfitting when the model training in high number of epochs.

To solve the problems which we mentioned above (speed up the training model and pass overfitting), we proposed label smoothing for overfitting and overconfidence, and Batch Normalization for speed up the training Model.

Label smoothing. It has been observed that label smoothing is a regularization technique that is used for classification problems and helps the model to prevent to predict labels confidently. There are problems that have been encountered such as overfitting and overconfidence therefore, the method of Label Smoothing has been used to eliminate any biasness.

$$\boldsymbol{L}(\boldsymbol{\theta}) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{k}\boldsymbol{q}(\boldsymbol{k}\backslash\boldsymbol{x}_n)\log[\boldsymbol{P}_{\boldsymbol{\theta}}(\boldsymbol{k}\backslash\boldsymbol{x}_n)] \quad (2)$$

Where $\boldsymbol{L}(\boldsymbol{\theta})$ is Cross-entropy cost function, $\boldsymbol{q}(\boldsymbol{k}\backslash\boldsymbol{x}_n)$ is Label distribution, $\boldsymbol{P}_{\boldsymbol{\theta}}(\boldsymbol{k}\backslash\boldsymbol{x}_n)$ is Output distribution.

Label smoothing: Smoothing one-hot label distribution with strength ∝.

For example: in the case of k=3 classes classification, if ∝=0.2 then:

(0, 0, 1) ⟶ Label Smoothing ⟶ (0.1, 0.1, 0.8) where 0.8=1-∝ and 0.1= $\frac{\propto}{\boldsymbol{k-1}}$

Where (0, 0, 1) One-hot label distribution and (0.1, 0.1, 0.8) Smoothing label distribution (Li, Dasarathy, & Berisha, 2020).

Batch Normalization. It is a normalization technique that is applied between layers of a neural network. This is done in small batches rather than a full dataset. It speeds up learning and uses higher learning rates, which makes learning easier (Ioffe, & Szegedy, 2015).

Advantages of using BN:

- Speeds up training: It helps with the flow of data between the middle layers of the neural network, which means we can use a higher learning rate. and this will lead to increase the speed of training.
- Decreases importance of initial weights: Reduces the dependence of gradients on the scale of parameters or their initial values.
- Regularizes the model: It has a regularizing effect which means we can often remove dropout.

AlexNet

AlexNet is a CNN architecture, designed by Alex Krizhevsky in collaboration with Ilya Sutskever and Geoffrey Hinton, who was Krizhevsky's Ph.D. advisor. It is competed in the ImageNet Large Scale Visual Recognition Challenge on September 30, 2012 (Deng et al., 2019). AlexNet consists of eight layers, five convolutional layers and three fully-connected layers as it shown in Figure 4. This architecture used ReLU Nonlinearity instead of tanh function to increase the speed of training. Also, Alex used overlapping Pooling to reduce the error and to avoid overfitting (Wei, 2019).

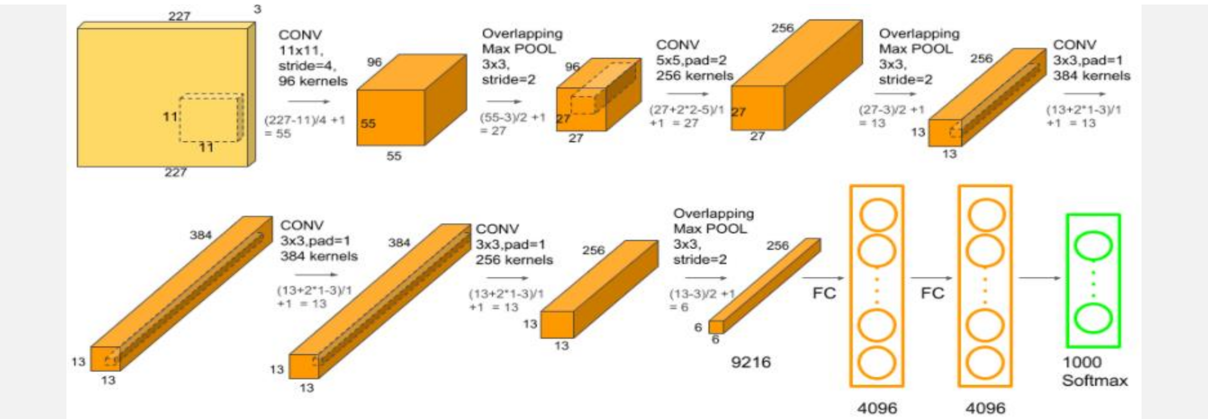


Figure 4. Architecture of AlexNet (Wei, 2019)

VGG16Net

VGG16 is a CNN architecture developed by Visual Geometry Group from oxford university in 2014. This model used as a transfer learning loads a set of weights pre-trained used 16 layers network. The size of input image is 224x224 RGB. It consists of 5 blocks, where each block consists of a set of convolutional layers and max pooling layer over 2x2 windows with stride 2. The stride in VGG16 is 1 and the convolutional layer input is padded. The last five blocks are followed by three fully connected layers (FC). The last layer is softmax layer which represented the output. Figure 5 shows the full form of VGG16 architecture (Tindall, Luong, & Saad, 2015).

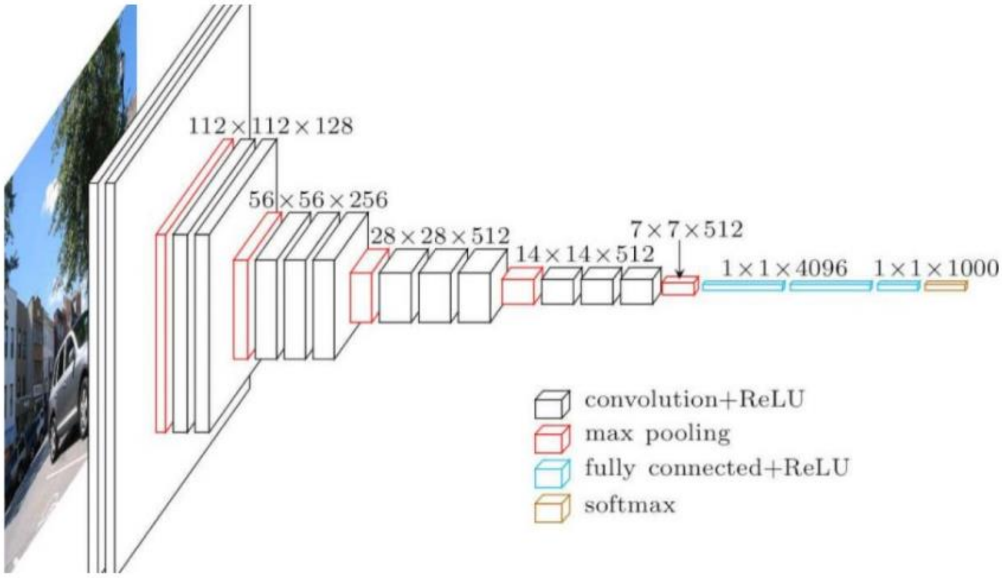


Figure 5. Architecture of VGG16 (Tindall, Luong, & Saad, 2015)

Inception-V3 (GoogleNet)

Inception-v3 is CNN's 48-layer architecture. It has trained over a million images in the ImageNet database. It can classify images into 1000 categories of objects (Szegedy et al., 2015).

Inception-v3 can be used for transfer learning, it allows you to retrain the last layers of the model, resulting in a significant reduction in training time. As mentioned above, inception-v3 is trained on over a million images from the ImageNet database, which means you can keep the knowledge gained by the model during the initial training and apply it to a smaller data set.

As it shown in Figure 6 Inception Layers are a combination of a set of layers (namely, 1×1 Convolutional layer, 3×3 Convolutional layer, 5×5 Convolutional layer) with their output filter banks concatenated into a single output vector forming the input of the next stage (ImageNet, n.d.).

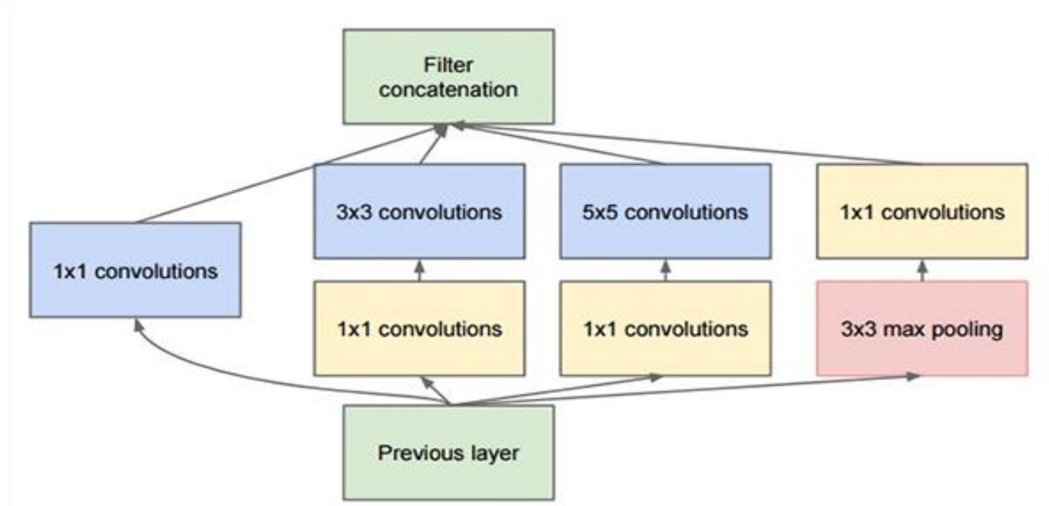


Figure 6. Architecture of Inceotion-v3 (Smaida & Yaroshchak, 2019)

Experiments and Results

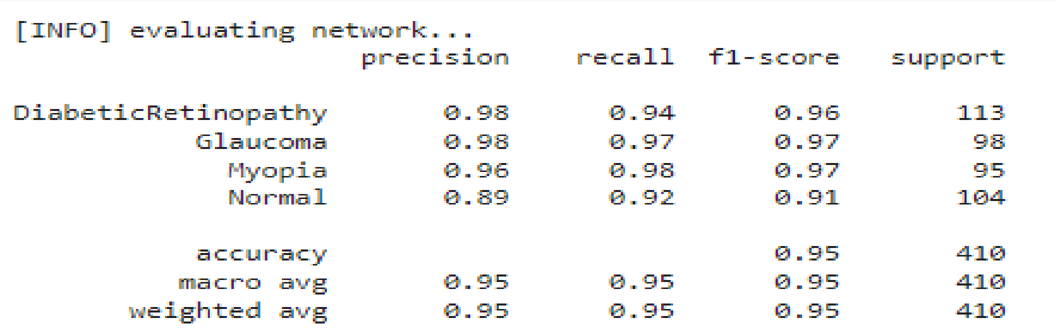
According to the models we explained above, all these models are implemented by Python language (Google Colab) using dual Graphics Processing Unit (GPU). The main goal of this work is to improve the accuracy of GMD model for detection and diagnosis of eye diseases. Our dataset has been explained in the GMD Results section, it consists of four types of eye disease, Glaucoma, Myopia, Diabetic retinopathy and Normal eyes. The number of samples of each class are shown in Table 1.

GMD Results

GMD structure as it shown in Figure 3, has been applied using our dataset which consists of 1640 samples of fundus images size 64x64 coloured for training set, 410 samples used for validation set. The results have been addressed in Table 2 below, Figure 7 and Figure 8. The accuracy performances of GMD model based on label classification and label smoothing were evaluated.

Table 2. GMD confusion matrix

		Predicted			
		Precision	Recall	F1-score	Support
Actual	Diabetic retinopathy	0.98	0.94	0.96	113
	Glaucoma	0.98	0.97	0.97	98
	Myopia	0.96	0.98	0.97	95
	Normal	0.89	0.92	0.91	104
	Accuracy			**0.95**	410
	Macro avg	0.95	0.95	0.95	410
	Weighted avg	0.95	0.95	0.95	410



```
[INFO] evaluating network...
                      precision    recall  f1-score   support

DiabeticRetinopathy        0.98      0.94      0.96       113
           Glaucoma        0.98      0.97      0.97        98
             Myopia        0.96      0.98      0.97        95
             Normal        0.89      0.92      0.91       104

           accuracy                            0.95       410
          macro avg        0.95      0.95      0.95       410
       weighted avg        0.95      0.95      0.95       410
```

Figure 7. GMD confusion matrix

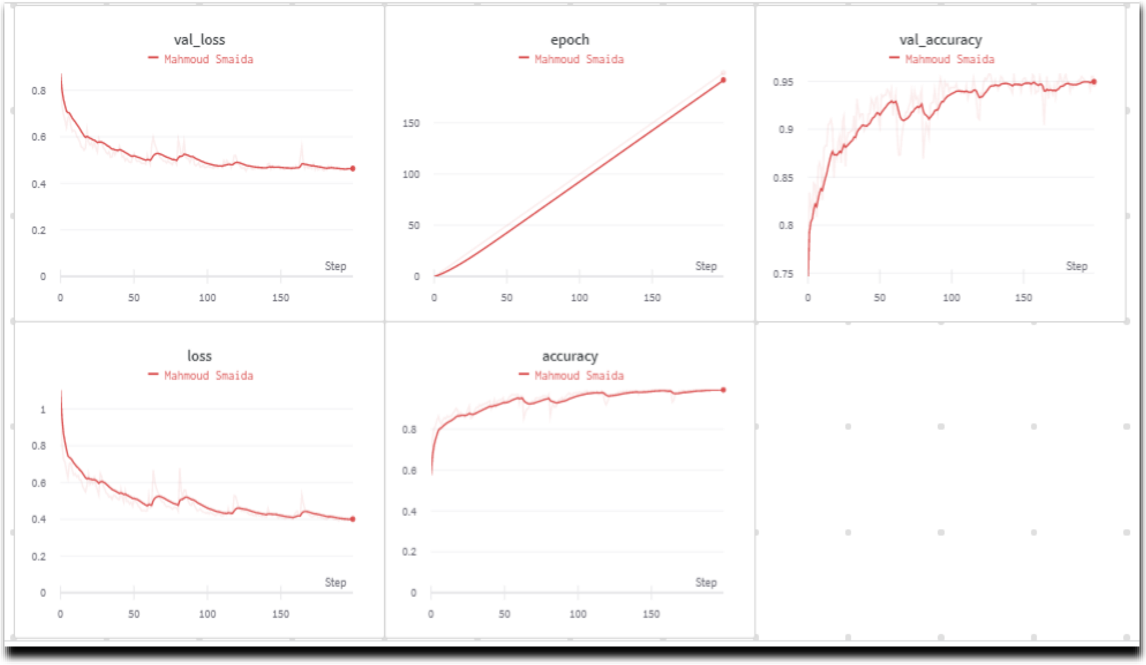


Figure 8. Losses and accuracy of GMD model

AlexNet Results

AlexNet has been applied with all the additional techniques used in GMD model. The results have been reported in Table 3 below, Figure 9 and Figure 10. The accuracy performances of Alex model based on label classification and label smoothing were evaluated.

Table 3. AlexNet confusion matrix

		Predicted			
		Precision	Recall	F1-score	Support
Actual	Diabetic retinopathy	0.91	0.94	0.93	106
	Glaucoma	0.92	0.93	0.92	108
	Myopia	1.00	1.00	1.00	88
	Normal	0.87	0.83	0.85	108
	Accuracy			**0.92**	410
	Macro avg	0.93	0.93	0.93	410
	Weighted avg	0.92	0.92	0.92	410

```
[INFO] evaluating network...
                      precision    recall  f1-score   support

DiabeticRetinopathy        0.91      0.94      0.93       106
           Glaucoma        0.92      0.93      0.92       108
             Myopia        1.00      1.00      1.00        88
             Normal        0.87      0.83      0.85       108

           accuracy                            0.92       410
          macro avg        0.93      0.93      0.93       410
       weighted avg        0.92      0.92      0.92       410
```

Figure 9. AlexNet confusion matrix

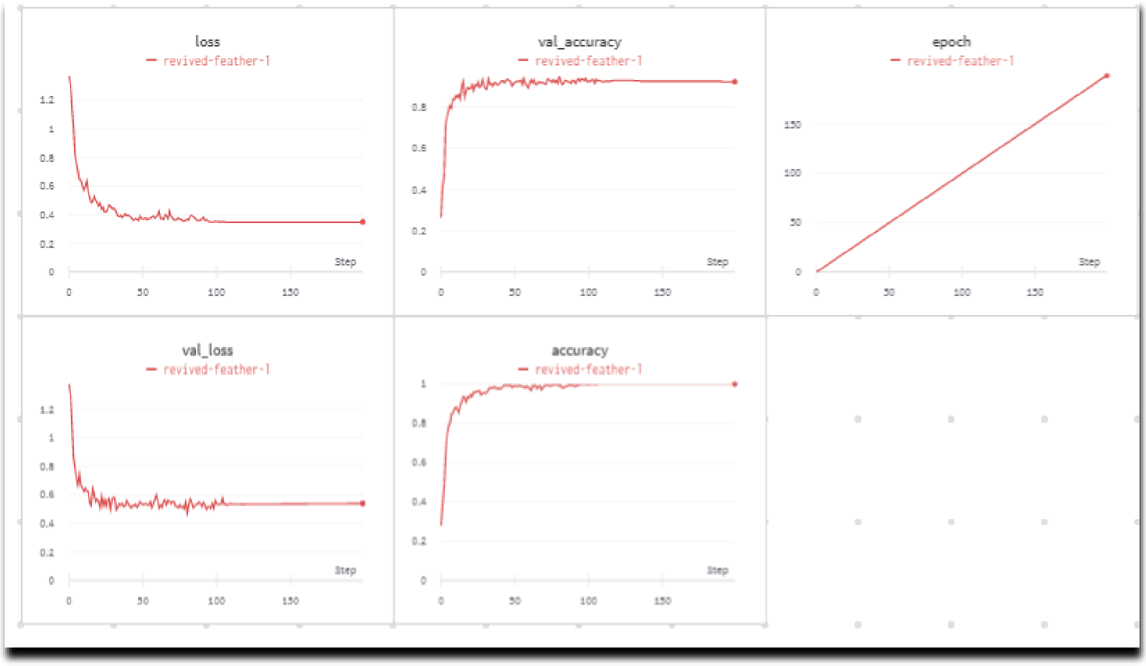


Figure 10. Losses and accuracy of AlexNet

VGG16 Results

The Network of VGG16 Net has been applied with all the additional techniques used in GMD model. The results have been obtained, and reported in Table 4 below. The accuracy performances of VGG16 model based on label classification and label smoothing were evaluated.

Table 4. VGG16 Net confusion matrix

		Predicted			
		Precision	Recall	F1-score	Support
Actual	Diabetic retinopathy	0.95	0.92	0.93	106
	Glaucoma	0.92	0.90	0.91	108
	Myopia	0.97	1.00	0.98	88
	Normal	0.84	0.87	0.85	108
	Accuracy			**0.92**	410
	Macro avg	0.92	0.92	0.92	410
	Weighted avg	0.92	0.92	0.92	410

```
[INFO] evaluating network...
                     precision    recall  f1-score   support

DiabeticRetinopathy       0.95      0.92      0.93       106
           Glaucoma       0.92      0.90      0.91       108
             Myopia       0.97      1.00      0.98        88
             Normal       0.84      0.87      0.85       108

           accuracy                           0.92       410
          macro avg       0.92      0.92      0.92       410
       weighted avg       0.92      0.92      0.92       410
```

Figure 11. VGG16 confusion matrix

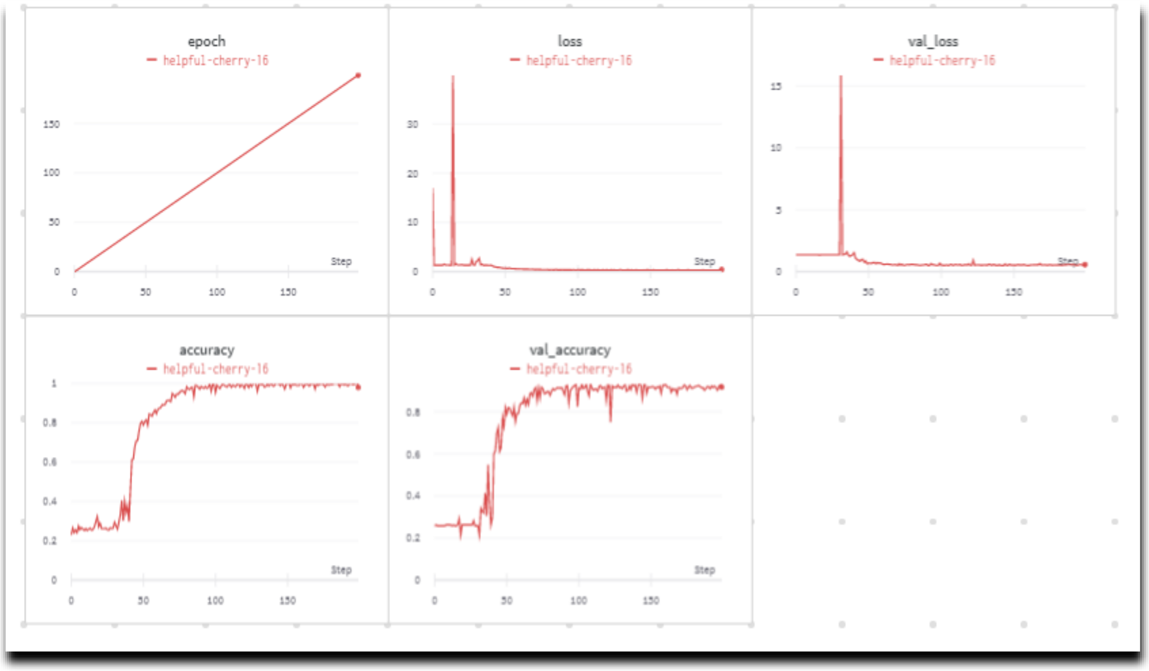


Figure 12. Losses and accuracy of VGG16 Net

Incerption-V3 Net

As shown in Figure 6, the structure of Inception-v3 has been implemented. All additional techniques which we mentioned added to this network, and the results documented as the following in Table 5, Figure 13 and Figure 14. The accuracy performances of Inception-v3 model based on label classification and label smoothing were evaluated.

Table 5. Inception-V3 Net confusion matrix

		Predicted			
		Precision	Recall	F1-score	Support
Actual	Diabetic retinopathy	0.99	0.93	0.96	106
	Glaucoma	0.92	0.94	0.93	108
	Myopia	1.00	1.00	1.00	88
	Normal	0.88	0.92	0.89	108
	Accuracy			**0.94**	410
	Macro avg	0.95	0.94	0.94	410
	Weighted avg	0.94	0.94	0.94	410

```
[INFO] evaluating network...
                      precision    recall  f1-score   support

DiabeticRetinopathy        0.99      0.93      0.96       106
           Glaucoma        0.92      0.94      0.93       108
             Myopia        1.00      1.00      1.00        88
             Normal        0.88      0.91      0.89       108

           accuracy                            0.94       410
          macro avg        0.95      0.94      0.94       410
       weighted avg        0.94      0.94      0.94       410
```

Figure 13. Inception-V3 confusion matrix

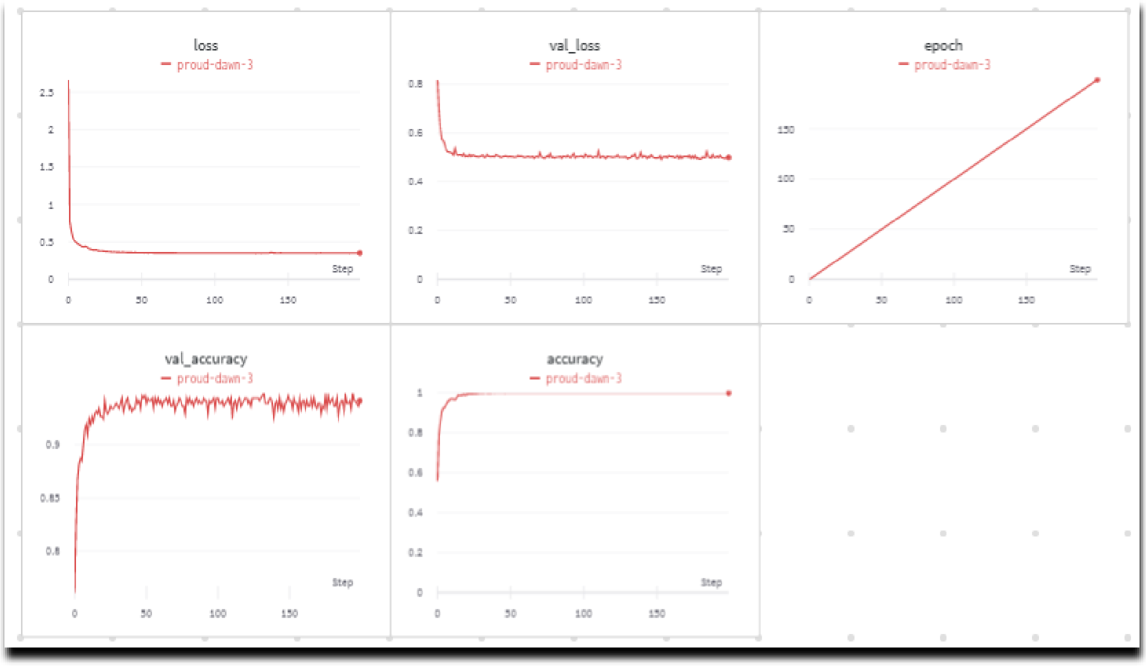


Figure 14. Losses and accuracy of Inception-V3 Net

Four multi-label classification accuracies have been obtained as shown in Table 6 according to the models mentioned above. All these accuracies are graphically represented in Figure 8, 10, 12, 14. Where each model structure shown with epochs and accuracies.

Table 6. Models accuracy

Models	Recall				Precision				F-Score				Accuracy
	Diabetic Retinopathy	Glaucoma	Myopia	Normal	Diabetic Retinopathy	Glaucoma	Myopia	Normal	Diabetic Retinopathy	Glaucoma	Myopia	Normal	Overall
Alex	0.94	0.93	1.000	0.83	0.91	0.92	1.00	0.87	0.93	0.92	1.00	0.85	92%
VGG16	0.92	0.90	1.00	0.87	0.95	0.92	0,97	0.84	0.93	0.91	0.98	0.85	92%
Inception V3	0.93	0.94	1.000	0.92	0.99	0.92	1.00	0.88	0.96	0.93	1.00	0.89	94%
GMD	0.94	0.97	0.98	0.92	0.98	0.98	0.96	0.89	0.96	0.97	0.97	0.91	**95%**

Conclusion

Multi-label classification has been realized amid a large number of demands in the medical systems and one of the major demands is considered to automate diagnostic processes.

In this study, we have applied GMD model based on multi-label classification with additional techniques as mentioned earlier to increase the model accuracy, improve the training model and speed up the training model. 2050 samples of our dataset have been used, 20% as validation.

In conclusion, we compared three popular networks, AlexNet, VGG16 and Inception-V3 with the GMD model in order to measure the accuracy using confusion matrix. All these networks based on label classification deployed using GPU in Google Colab. According to the all experiments, we obtained results for each combination and observed that for multi-label classification, GMD model gives the highest accuracy (95%) compared with the other models as it shown in Table 6. Therefore, the use of confusion matrix shows all classification models in varying proportions. We suggest strongly, that GMD model can be implement as web server or web application to use it in ophthalmology clinics or at general practitioner clinics to automate diagnostic processes.

References

Adaimi, G., Kreiss, S. & Alahi, A. (2019). Rethinking person re-identification with confidence. *arXiv preprint arXiv:1906.04692*.

Bali, A. & Mansotra, V. (2021). Proposed transfer learning-based convolutional neural network and one versus rest classifier for multi-class multi-label ophthalmological disease prediction in fundus images. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(14), 161-171.

Bengio, Y., Courville, A. & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, *35*(8), 1798-1828.

De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., Askham, H., Glorot, X., O'Donoghue, B., Visentin, D. & van den Driessche, G. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, *24*(9), 1342-1350.

Deng, J., et al. (2019). Imagenet large scale visual recognition competition. *ilsvrc2012*.

Fu, H., Cheng, J., Xu, Y., Wong, D.W.K., Liu, J. & Cao, X. (2018). Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *IEEE transactions on medical imaging*, *37*(7), 1597-1605.

Gargiulo, F., Silvestri, S., Ciampi, M. & De Pietro, G. (2019). Deep neural network for hierarchical extreme multi-label text classification. *Applied Soft Computing*, *79*, 125-138.

Gour, N. & Khanna, P. (2021). Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. *Biomedical Signal Processing and Control*, *66*, 102329.

He, J., Li, C., Ye, J., Qiao, Y. & Gu, L. (2021). Multi-label ocular disease classification with a dense correlation deep neural network. *Biomedical Signal Processing and Control*, *63*, 102167.

Hussain, M., Bird, J.J. & Faria, D.R. (2018, September). A study on cnn transfer learning for image classification. In *UK Workshop on computational Intelligence* (pp. 191-202). Springer, Cham.

ImageNet. http://www.image-net.org/,11/2019

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning*. PMLR.

Jain, S., Thakur, P.S., Bharti, K., Khanna, P. & Ojha, A. (2020, December). A Lightweight Multi-label Image Classification Model Based on Inception Module. In *International Conference on Computer Vision and Image Processing* (pp. 225-236). Springer, Singapore.

Ju, L., Wang, X., Wang, L., Mahapatra, D., Zhao, X., Harandi, M., Drummond, T., Liu, T. & Ge, Z. (2021). Improving medical image classification with label noise using dual-uncertainty estimation. *arXiv preprint arXiv:2103.00528*.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). *Deep learning.nature*, *521*(7553), 436-444.

Li, W., Dasarathy, G., & Berisha, V. (2020, June). Regularization via structural label smoothing. In *International Conference on Artificial Intelligence and Statistics* (pp. 1453-1463). PMLR.

Lin, J., Cai, Q. & Lin, M. (2021). Multi-Label Classification of Fundus Images With Graph Convolutional Network and Self-Supervised Learning. *IEEE Signal Processing Letters*, *28*, 454-458.

Malik, T.G. (2021). Artificial Intelligence in Ophthalmology. *Pakistan Journal of Ophthalmology*, *37*(1).

Meskó, B., Hetényi, G. & Győrffy, Z. (2018). Will artificial intelligence solve the human resource crisis in health-care?. *BMC health services research*, *18*(1), 1-4.

Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181-197.

Rezaei-Dastjerdehei, M.R., Mijani, A. & Fatemizadeh, E. (2020, November). Addressing Imbalance in Multi-Label Classification Using Weighted Cross Entropy Loss Function. In *2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME)* (pp. 333-338). IEEE.

Smaida, M., & Yaroshchak, S. (2019). Comparative Study of Image Classification Algorithms for Eyes Diseases Diagnostic. *International Journal of Innovative Science and Research Technology*, *4*(12).

Smaida, M., Yaroshchak, S. & El Barg, Y. (2021). DCGAn for enhancing eye diseases classification. *CEUR Workshop Proceedings*.

Szegedy, C., et al. (2015). Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Tindall, L., Luong, C., & Saad, A. (2015). Plankton classification using vgg16 network.

Tymchenko, B., Marchenko, P. & Spodarets, D. (2020). Deep learning approach to diabetic retinopathy detection. *arXiv preprint arXiv:2003.02261*.

Wang, J., Yang, L., Huo, Z., He, W. & Luo, J. (2020). Multi-Label Classification of Fundus Images With EfficientNet. *IEEE Access*, *8*, 212499-212508.

Wei, J. (2019). AlexNet: The Architecture that Challenged CNNs. https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951.

Yin, P., Wu, Q., Xu, Y., Min, H., Yang, M., Zhang, Y. & Tan, M. (2019, October). PM-Net: Pyramid multi-label network for joint optic disc and cup segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 129-137). Springer, Cham.

Zhang, K., Li, T., Liu, B. & Liu, Q. (2019). Co-saliency detection via mask-guided fully convolutional networks with multi-scale label smoothing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3095-3104).