Automated Detection of Diabetic Retinopathy (DR) Stages in Retinal Fundus photographs Using Deep Learning

by

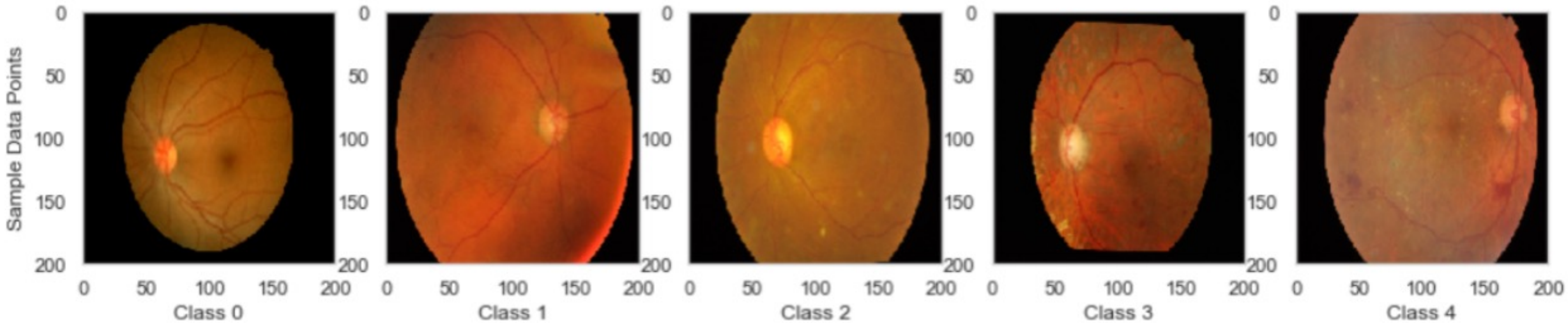
Muniba Shaikh

CS 767 – Machine Learning

This project is an application of deep learning model that screens retinal fundus images for diabetic retinopathy and provide information on how severe the condition is on a scale of 0 to 4:

- 0 - No DR - Healthy
- 1 - Mild
- 2 - Moderate
- 3 - Severe
- 4 - Proliferative DR

Here is the URL to my research paper (https://tinyurl.com/DR-classification) and source code https://github.com/munibas/Diabetic_Retinopathy/blob/main/Diabetic_Retinopathy.ipynb



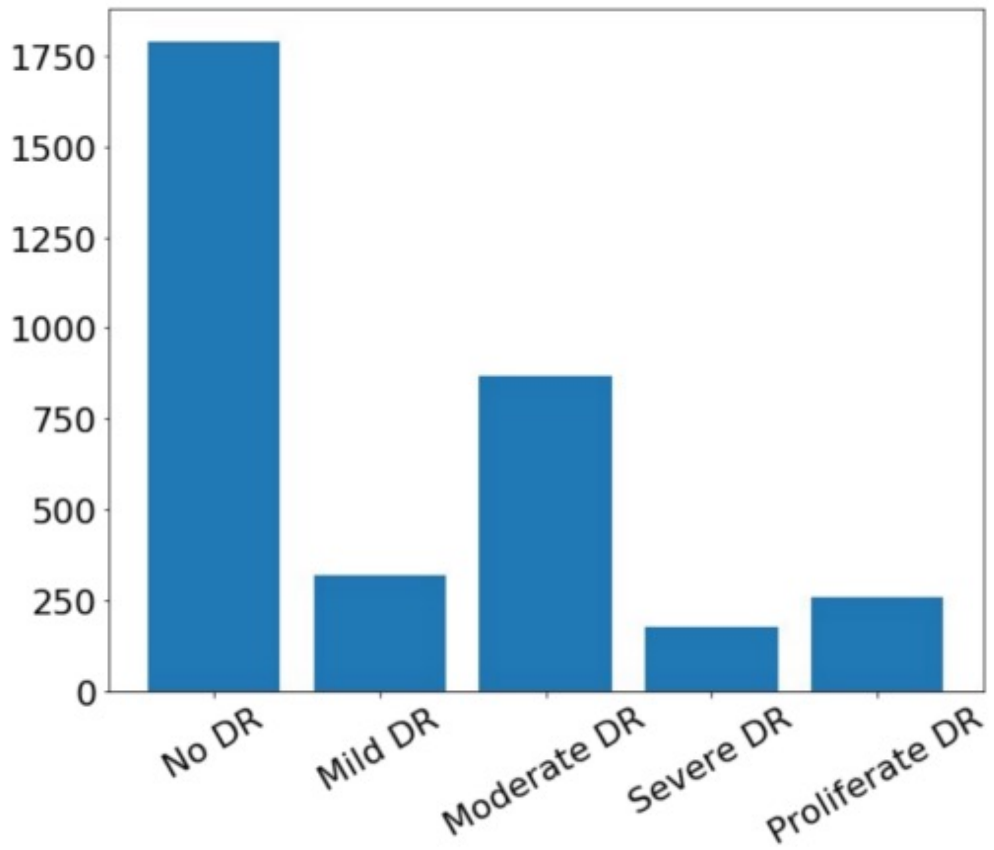
Eye images corresponding to each Blindness class severity (0–4)

APTOS2019 dataset

- The dataset used in this research was taken from Kaggle (https://www.kaggle.com/competitions/aptos2019-blindness-detection/data), collected by Aravind Eye Hospital from rural areas in India.
- The full dataset consists of 3662 retinal fundus photographs for left and right eyes, which are divided into 2747 training, 755 validation, and 160 testing images.

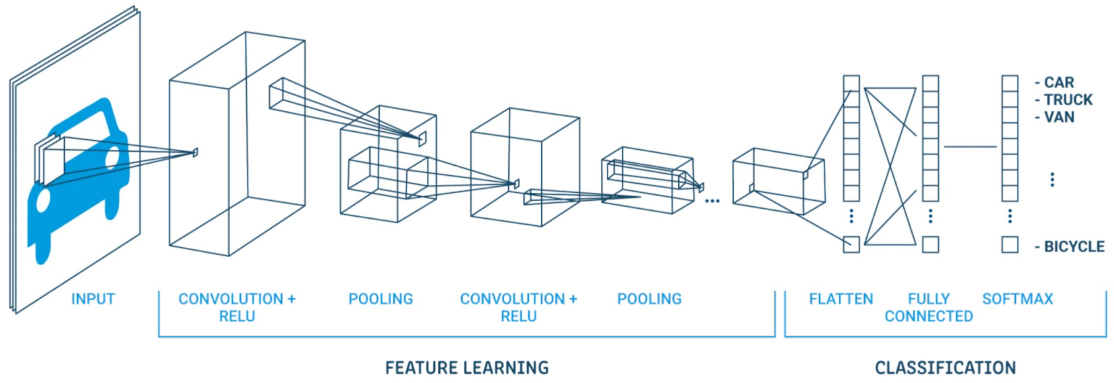
Class distribution in APTOS2019 dataset

- Class distribution in APTOS2019 training data is imbalanced with most images classified as 0 and least number of images in class 3 and 4.
- As different diabetic retinopathy datasets have a similar class distribution, it is considered as a fundamental property of this type of data.
- Like any real-world data set, this data set may have noise in both the images and labels.
- Images may be out of focus, underexposed, or overexposed. The images were gathered from multiple clinics using a variety of cameras.
- The images are labelled by clinicians based on the severity of diabetic retinopathy on a scale of 0 to 4, so possibility of human error while labelling the images.



Deep learning

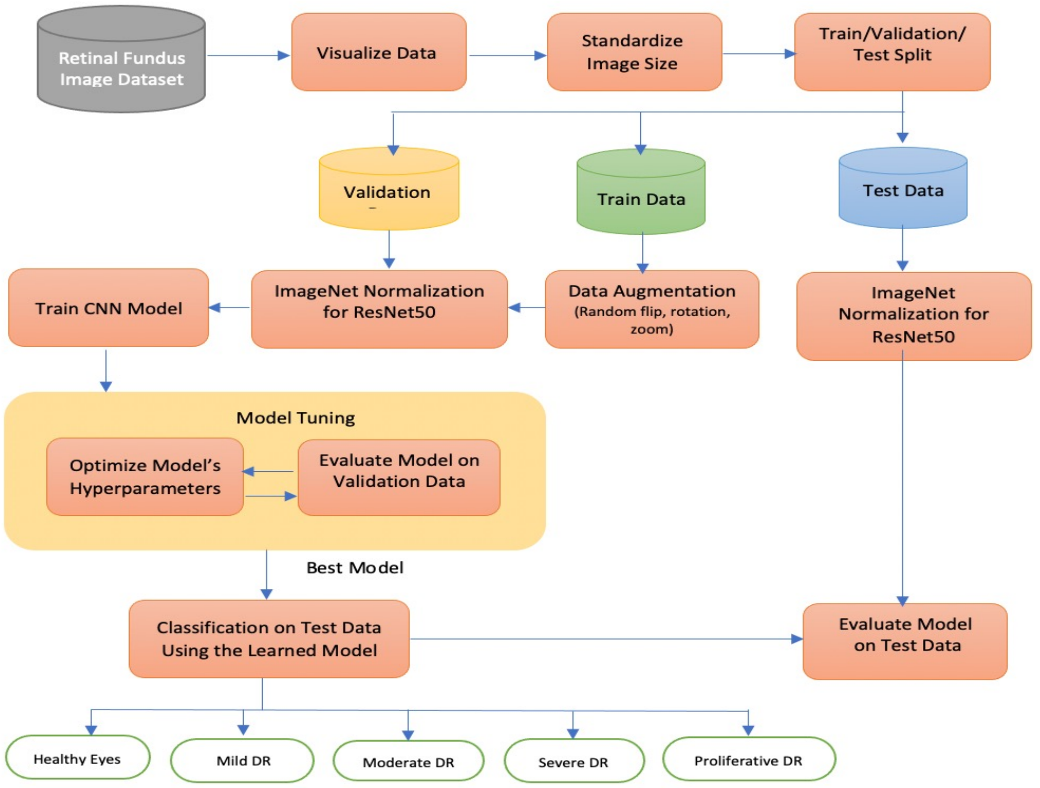
- Deep learning image processing has become one of the most common techniques in machine learning that have achieved better performance in medical image analysis, classification and automated diagnosis.
- Convolutional neural networks are proven to be a highly effective deep learning method in image analysis because they preserve the spatial relationship between pixels by learning internal feature representations using small squares of input data. Features are learned and used across the whole image, allowing for the objects in the images to be shifted or translated in the scene and still detectable by the network.
- Implemented a deep convolutional neural network (CNN) model in Python using TensorFlow and Keras API that screens retinal fundus images for diabetic retinopathy.



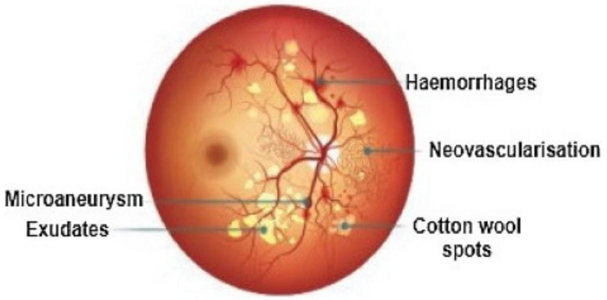
ML System Design

- **Convolutional neural network (CNN)** applies to the classification problem of diabetic retinopathy in retinal fundus images as it can identify the intricate features involved in the classification task such as micro-aneurysms, exudate, and hemorrhages in the retina and consequently provide an automatic diagnosis.
- Images are standardized to a fixed image size 224 × 224 pixels.
- Image data is split into train, and validation sets using 75/25 split. Then validation set is further split into validation and test sets using 80/20 split. The train set is used for model fitting. The validation set is used during the model fitting to evaluate the loss and metrics. The test set is used at the end to evaluate how well the model generalizes to new data.
- Additional training data is generated from existing examples by augmenting them using random transformations such as horizontal flipping, random rotation (0.2), and random zoom (0.1).This helps expose the model to more aspects of the data and generalize better.

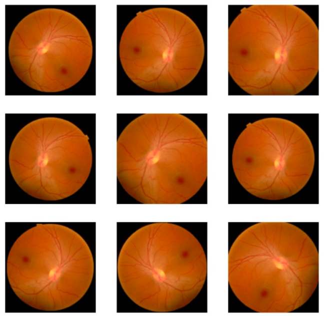
ML System Design

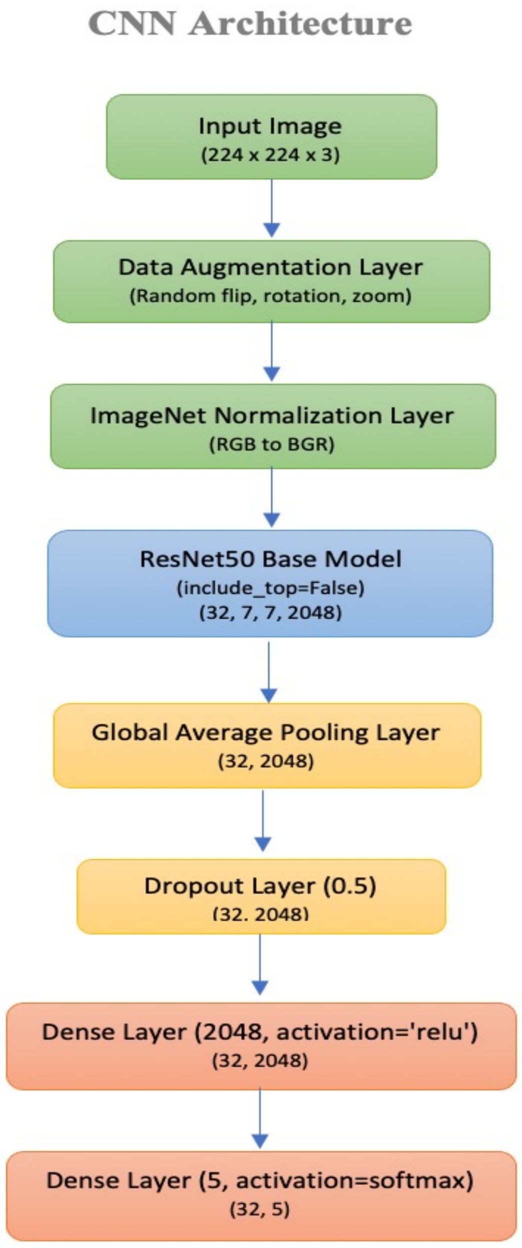


Diabetic Retinopathy



Data Augmentation





Convolutional Neural Network

- The custom CNN includes seven layers: one data augmentation layer, one ImageNet normalization layer, one base model (Resnet50), one max-pooling layer, one dropout layer and two FC layers.
- Since the dataset is small so I implemented transfer learning to take advantage of features learned by a pre-trained *ResNet50* on a much larger ImageNet dataset, as the base model to initialize the weights of my CNN models.
- A *Pooling layer* is used to down-sample feature maps from the previous convolutional layers to compress or generalize feature representations.
- *Dropout (0.5*) method is used to reduce overfitting.
- The *SoftMax* function is used as a classifier in the last FC layer.

Hyperparameters for Model Training

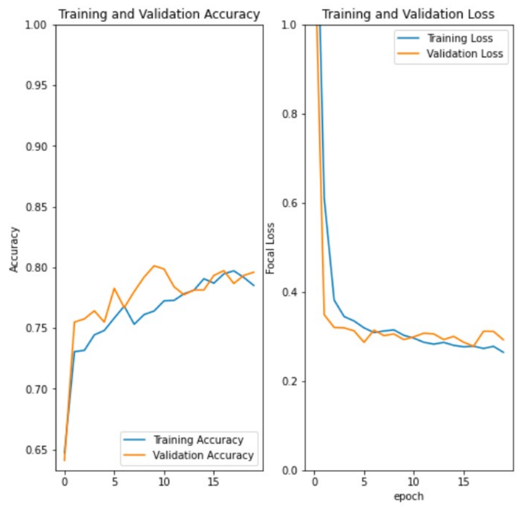
Sr. No.	Hyperparameters	Value
1.	Learning Rate	0.001
2.	Batch Size	32
3.	Epochs	20
4.	Activation Function	ReLu
5.	Optimizer	Adam
6.	Metric	Accuracy
7.	Loss Function	SparseCategoricalFocalLoss(gamma=2) SparseCategoricalCrossentropy

Loss function for Imbalanced Data Classification

Sparse categorical focal Loss and sparse categorical cross entropy loss functions are used in model training to address the issue of class imbalance of different stages of DR in the APTOS2019 dataset. Two models are trained with different loss functions to investigate and compare their performance on DR classification:

- **SparseCategoricalFocalLoss** function for multiclass classification with integer labels: *CNN Model_1* is based on the multi-class *SparseCategoricalFocalLoss*. It down-weights the loss assigned to well-classified examples. This loss function generalizes multiclass softmax cross-entropy by introducing a hyperparameter γ (gamma), called the *focusing parameter*, that allows hard-to-classify examples to be penalized more heavily relative to easy-to-classify examples. *CNN Model_1* predicted with 76.88% accuracy on test data.
- **SparseCategoricalCrossentropy** function for multiclass classification with integer labels: *CNN Model_2* is based on the multi-class *SparseCategoricalCrossentropy* loss function . Class weights are used in model training to address the imbalance of different classes for diabetic retinopathy classification. *CNN Model_2* predicted with 75.63% accuracy on test data.

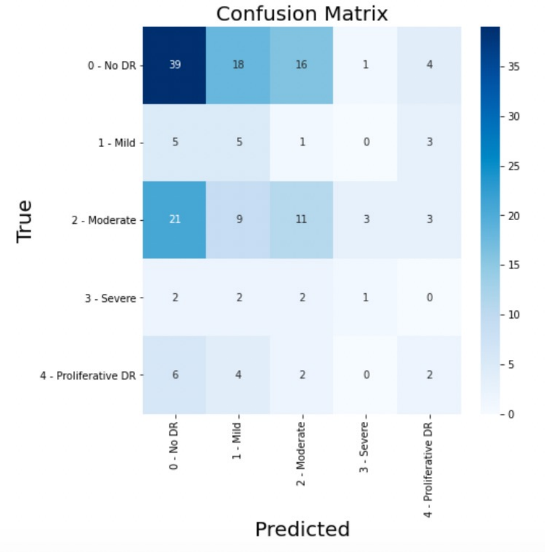
CNN Model_1

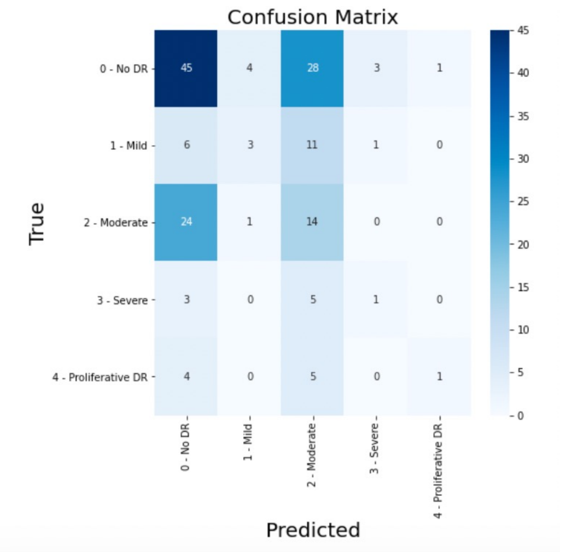


CNN Model_2

Performance Measures

- The model's performance is evaluated using accuracy, precision, recall and F1-score and confusion matrices are computed.
- Performance measures of both the models are almost similar.





Classification report of CNN Model_2

	precision	recall	f1-score	support
0 – No DR	0.53	0.50	0.52	78
1 – Mild	0.13	0.36	0.19	14
2 – Moderate	0.34	0.23	0.28	47
3 – Severe	0.20	0.14	0.17	7
4 – Proliferative DR	0.17	0.14	0.15	14

Classification report of CNN Model_1

	precision	recall	f1-score	support
0 – No DR	0.55	0.56	0.55	81
1 – Mild	0.38	0.14	0.21	21
2 – Moderate	0.22	0.36	0.27	39
3 – Severe	0.20	0.11	0.14	9
4 – Proliferative DR	0.50	0.10	0.17	10

Conclusion

- Class imbalance of APTOS2019 dataset is addressed by using *multi-class focal loss* function and *SparseCategoricalCrossentropy* function along with class weighting . Both the models are predicting all 5 classes of diabetic retinopathy although not with high sensitivity for the minority classes but with the overall accuracy of 76.88% and 75.63% respectively.
- Besides class imbalance, the dataset was small and consisted of only 3662 images, so it could not predict class 1, 3 and 4 with high sensitivity.
- In future work, I propose to investigate the performance of CNN model using large scale data for better training which would hopefully result in better classification and prediction of the different stages of diabetic retinopathy.

References

I. "Tf.keras.utils.image_dataset_from_directory : TensorFlow Core v2.9.1," *TensorFlow*. https://www.tensorflow.org/api_docs/python/tf/keras/utils/image_dataset_from_directory.

II. Rubina Sarkis, Sandra Michalska, K. A. H. W. Y. Z. (2019). Convolutional neural networks for mild diabetic retinopathy detection: an experimental study. bioRxiv.

III. "Image classification : Tensorflow Core," *TensorFlow*. https://www.tensorflow.org/tutorials/images/classification.

IV. M. D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, and M. Niemeijer, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of Deep Learning," *Investigative Ophthalmology & Visual Science*, 01-Oct-2016. https://iovs.arvojournals.org/article.aspx?articleid=2565719.

V. W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through Deep Learning Techniques: A Review," *Informatics in Medicine Unlocked*, 20-Jun-2020. https://www.sciencedirect.com/science/article/pii/S2352914820302069#bib12.

VI. B. Tymchenko, P. Marchenko, and D. Spodarets, "Deep Learning Approach to diabeticretinopathy detection," *arXiv.org*, 03-Mar-2020. https://arxiv.org/abs/2003.02261.

VII. M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, "Transfer learning-based model for diabetic retinopathy diagnosis using retinal images," *MDPI*, 22-Apr-2022. https://www.mdpi.com/2076-3425/12/5/535/htm.