Deep Neural Network Based Human Kidney Cortex Cell Image Classification

R. Taimoor (6716747)
Dept. Of Computer Sciences, University of Surrey, Guildford, Surrey, GU2 7JP

Background

Medical image classification based on a deep neural network has helped optimise the identification of early diseases. This is achieved via the classification of cell deformation and their impact on organs based on their spatial organisation. In this research, we will be classifying kidney cortex cells from an image set provided by the Broad Bioimage Benchmark Collection(BBBC). However, medical image classification poses certain challenges such as overfitting, data augmentation issues, and invariant class distributions. Consequently, there are many methods and approaches to overcome these challenges. According to research by Sahiner et al. (2018), there are various workarounds to overfitting, including the use of regularisation, early stopping and dropout layers[1]. Regularisers and dropout layers tend to penalise overly complex CNN architectures by imposing an additional term to the loss function during training, smoothing the solution and dropping units in the neural network, respectively.
The objective of this project is to classify a kidney cortex cell image set sourced through the Broad Bioimage Benchmark Collection (BBBC) into 8 classes of different cells. The images were taken of three deceased donor kidney nephrectomy tissues, and ”Image acquisition was performed in four separate consecutive channels using an upright Leica SP8 Confocal Microscope controlled by LAS X software (Germany)” (Woloshuk et al., 2021)[5]. Moreover, according to J. Yang et al., (2021) the images have been manipulated and resized to greyscale and image size of 28x28[6]. Furthermore, the dataset comprises 200,000 images consisting of 150,000 images for training and 50,000 images for testing. This project aims to identify and overcome overfitting and invariant image classes using various hyperparameters and model architecture modifications. The methods of overcoming these issues discussed in this report will encompass regularisation techniques, dropout layers, and data manipulation (data augmentation) to optimise the ratio of accuracy and validation accuracy.

Preprocessing and Initial Data Analysis

The image dataset is first given as a training set, a testing set, and a CSV file with image IDs and their corresponding classes. Following this, the training data is then divided into different folders with subfolders of each class with the help of the train.csv file. Once the data is divided into different sections, the following diagram depicts its class variations.

The training set is not divided into classes as it is given. Therefore, each image in the dataset is classified into folders individually with respect to its class in the train.csv file. Once the data is divided into different sections, Figure 1 depicts its class variations. As it can be visualised, there are a total of 8 classes, with great disparity in the data, with Types 0, 6, and 7 encompassing more than half of the data set. Therefore, there is a large class imbalance in the dataset.

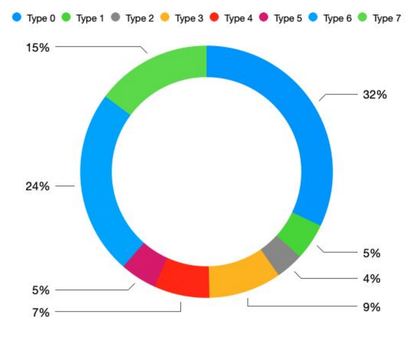


Figure 1: Initial Training Image Classification

From this initial data analysis, it can be inferred that there will be a rigorous need for data augmentation and class rebalancing. Furthermore, the training data is then divided into a training and validation set. Out of 150,000 training images, 120,000 images are used for training and 30,000 images are used for validation. These images are shuffled with the label_mode set as categorical, and the batch size is fixed at 128. This convention is constant in all CNN model runs.

Analysis and Approach

Initially, after careful analysis of the data, it is clear that overfitting and imbalanced classes are an issue. To tackle these challenges, I used various regularisation techniques. In each consecutive model, I have added each technique to the next model until I achieved a well-balanced model. The three significant techniques used include data augmentation, dropout Layers, batch normalization and Reduce LR On Plateau.

Data Augmentation (Model 1)

According to a study conducted by P. Thanapol et al. (2020), the application of data augmentation significantly reduces overfitting[3]. While the first solution this research looks into is data augmentation, it is important to note that medical images are very sensitive to data augmentation. This is due to the structural composition of the cells we are trying to classify; hence, only horizontal flip, vertical flip and rotation are used for augmentation. Figure 2 is a sample of an augmented image.

Figure 2: Data Augmentation (Flips and Rotations)

Augmentation drastically increases the dataset size and allows for more training params on new/seemingly new images.

Batch Normalization + Data Augmentation (Model 2)

In the second model, to further decrease the gap between the training and validation accuracy and increase the accuracy of the model, I have implemented batch normalisation layers on top of Model 1. Batch normalisation is a technique used to “standardise only the mean and variance of each unit in order to stabilise learning but allows the relationships between units and the nonlinear statistics of a single unit to change”(Goodfellow et al., 2016)[2]. Experimental studies by Ioffe et al. (2015) indicate that batch normalisation prevents overfitting when training deep networks, along with improving convergence and generalisation [4]. Therefore, the second model focuses solely on batch normalisation coupled with image augmentation. The notion of using this strategy came from P. Thanapol et al. (2020) discussion on “Utilising a Combination of data augmentation techniques together with batch normalisation in training the CNN” [3].

Dropout + Batch Normalization + Data Augmentation (Final Model)

The final model that I have implemented furthers the last two models by including dropout layers. Although it is unconventional to include dropout layers and batch normalisations in a model architecture, it has been successfully implemented before to reduce overfitting drastically.

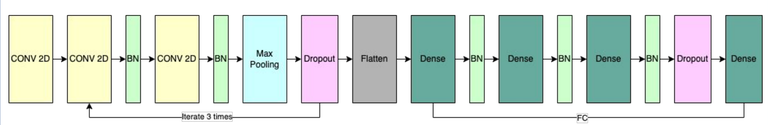


Figure 3: CNN Architecture for Final Model

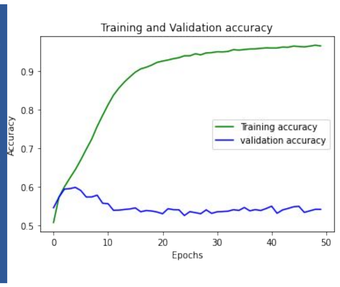
Figure 3 depicts the architecture of the final model and the placement of batch normalisation layers and dropout layers. Dropout layers reduce interdependent learning amongst the neurons by randomly setting input units to 0 with a frequency of rate at each step during training time, whereas inputs not set to 0 are scaled to keep the sum constant. By doing this, dropouts minimise overfitting by improving the model’s ability to generalise.

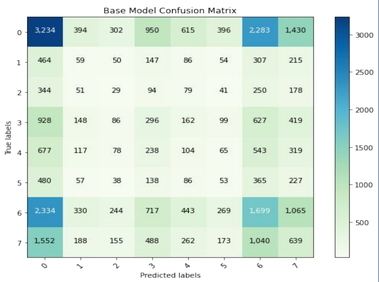
Implementation and Evaluation of Methods

All the models were implemented with a training set of 120,000 images and a validation set of 30,000 images. The loss function used was “Categorical Cross-Entropy” with the optimiser “Adam”. The learning rate is set to 0.01 initially, and Reduce LR on Plateau is used to reduce the learning rate every 3 epochs of increase in loss. It can be deduced from the accuracy-validation acc graphs below that the consequent models improve validation accuracy, and there is a significant decrease in the gap between the two accuracies, which reduces overfitting overall.

Base Model:

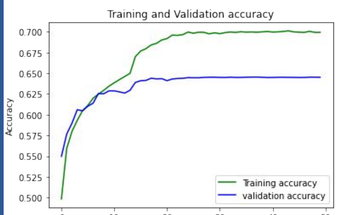
The base model has a large overfitting issue. The graph depicts a large difference between the accuracies. Furthermore, the validation accuracy is only 55%. In this model, there is no convergence. Moreover, the confusion matrix also depicts that there is significant discrepancy.

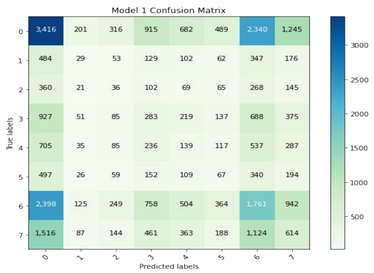




Model 1:

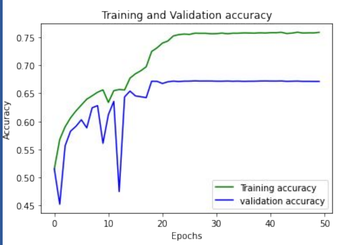
In this model, an augmentation layer is added, and a significant jump is noticed in the validation accuracy. There is also notable decrease in the overfitting as there is convergence. The validation accuracy is 63%. The confusion matrix also shows improvement as the predicted and true values converge.

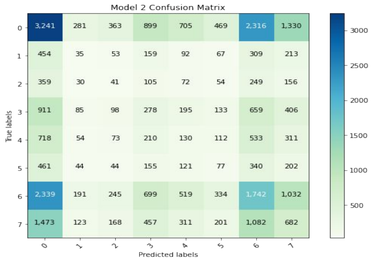




Model 2:

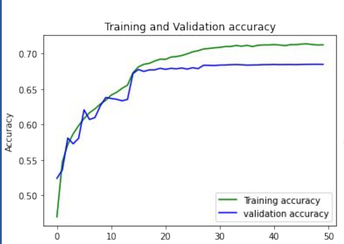
By adding batch normalization, this model further smooths convergence between the accuracies. Furthermore, the validation accuracy increases to 65%. Overfitting further decreases, and the confusion matrix further converges.

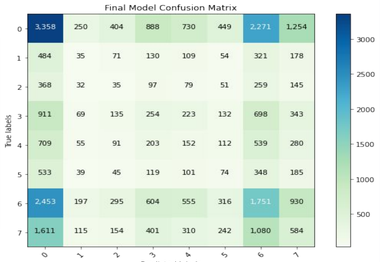




Final Model:

Finally, we include Dropout layers, and the training and validation accuracy is almost converged to one. The validation accuracy further increases to 68%, and in the confusion matrix true and predicted values match significantly more. This model reduces overfitting drastically compared to the base model, from 55% to 68%.





References

[1] Sahiner, Berkman et al. “Deep learning in medical imaging and radiation therapy.” Medical physics vol. 46,1 (2019): e1-e36. doi:10.1002/mp.13264
[2] Deep Learning (Ian J. Goodfellow, Yoshua Bengio and Aaron Courville), MIT Press, 2016.
[3] P. Thanapol, K. Lavangnananda, P. Bouvry, F. Pinel and F. Leprévost, "Reducing Overfitting and Improving Generalization in Training Convolutional Neural Network (CNN) under Limited Sample Sizes in Image Recognition," 2020 - 5th International Conference on Information Technology (InCIT), 2020, pp. 300-305, doi: 10.1109/InCIT50588.2020.9310787.
[4] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In ICML, 2015.
[5] Woloshuk, A, Khochare, S, Almulhim, AF, et al. In Situ Classification of Cell Types in Human Kidney Tissue Using 3D Nuclear Staining. Cytometry. 2021; 99: 707– 721. https://doi.org/10.1002/cyto.a.24274
[6] J. Yang et al. "Medmnist v2: A large-scale lightweight benchmark for 2d and 3d biomedical image classification." arXiv:2110.14795, 2021.