

HW3 Technical Report

Bryce Grover

Introduction

The data set is a collection of images of either cancerous or noncancerous skin. The goal is to train a convolutional neural network (CCN) to predict if an image is of cancerous or noncancerous skin. This is a binary classification problem. A model that can predict whether an image of skin is cancerous or not is valuable in the medical industry for cancer diagnosis. Such a model can be helpful in speeding up diagnosis or it may add another layer of surety for a medial provider, all of which increases the likelihood of successful cancer treatment.

Analysis

Cleaning and Pre-processing

To prepare the data for analysis, the image data were preprocessed using tensorflow and keras packages. The data was split into training and validation sets, image sizes were set to 256 x 256, and batch sizes were set at 32 images per batch.

Methods

Building a custom Deep Convolution Neural Network

To address the challenge of classifying skin cancer from images, a custom deep convolutional neural network was developed in addition to a ResNet50 model. The input layer took took in 256 x 256 images (to match the preprocessing) with three channels (RGB).

The main convolutional architecture is made up of five convolutional layers and three max-pooling layers to reduce dimensionality while preserving spatial hierarchies. The secifics of the model are that an initial convolutional layer with 32 filters of size 3x3 and ReLU activation

captures basic visual features. The next convolutional layers have 64 filters and 128 filters respectively. The increasing in depth is to capture more complex features. The final convolution layer has 256 filters to train the model to understand deep relationships in the spatial data. All the layers are followed by max-pooling to continuously reduce dimensionality. A flattening layer makes the two dimensional data to one dimensional data to feed into dense layers

The dense network includes an initial dense layer with 100 nodes followed by dropout for regularization to prevent overfitting (drop out rate of 50%). The next dense layers have 10 and 5 nodes respectively. ReLU activation was used to consider non-linearity in the data.

The architecture concludes with a dense output layer utilizing a sigmoid activation function, which is used when making a binary classification.

The model was compiled using the Adam optimizer and binary crossentropy loss function. This optimizer and loss were chosen due to the binary nature of the skin cancer classification problem and for the most optimal gradient descent.

The model was trained on training data pulled from a dataset split of 80/20. 80% of the split was used for training and the other 20% was reserved for model validation.

Early stopping was also utilized to prevent over training and therefore overfitting.

Transfer Learning with ResNet50 from Keras

In addition to my custom CNN, a transfer learning approach was used. The model used was the ResNet50 model pre-trained on the ImageNet dataset. Using this pretrained model allowed the leveraging of its robust feature extraction capabilities. The ResNet50 base was integrated without its top classification layer, adapting it to the skin cancer classification task by freezing the weights of the ResNet50 base to retain learned features. Additionally, a GlobalAveragePooling2D layer was used to reduce dimensionality. Finally, a dense layer with 1024 neurons and ReLU activation was used followed by dropout for regularization. The transfer learning model ultimately culminated with a sigmoid-activated dense layer for binary classification.

This transfer model was also compiled using the Adam optimizer and binary crossentropy, matching the custom model's setup. Training and validation followed an identical data split and preprocessing strategy, ensuring the conditions for performance evaluation were comparable.

Results

Both models were evaluated on their ability to classify skin cancer using accuracy and recall metrics. Using these metrics allows one to determine if a model is effective when presented with unseen data.

The evaluation of the neural network and decision tree models focused on three metrics: accuracy, precision, and recall.

Accuracy reflects the overall correctness of the models. Recall measures how well the models can identify all actual instances of cancer. It is the true positive rate. This specific metrics were chosen for this classification task because when determining whether a patient has cancer, correct answers (accuracy) are very important in making sure the patient either receives the necessary treatment to fight the cancer OR they are not subjugated to the often long and painful cancer treatment process when the patient did not have cancer in the first place. The first principle of the hippocratic oath is to do no harm.

Recall is an important metric because it is better to correctly predict the presence of cancer rather than the absence of it

These results indicate that while the neural network generally performs better in terms of overall accuracy and recall, the decision tree model was better suited in precision.

Reflection

I have a bit of conceptual difficulty with CNNs because of the nature of the data. Each pixel in an image is numeric which can be manipulated with kernels but the image as a whole is difficult to transform and understand what is happening to them in the model. In this assignment I came to better understand the inputs and the outputs of the models on a conceptual level.