Ice Cream VS Pizza Classification (Transfer Learning)

Introduction:

This project aims to address the challenge of classifying images into two categories: Ice Cream and Pizza. The motivation behind this task is to develop a model capable of distinguishing between these two visually distinct classes. The primary technologies used include VGG16 architecture as the base model and the Keras Tuner for hyper parameter optimization.

Data Collection:

I collected the dataset from kaggle, comprising images of Ice Cream and Pizza.
Dataset Link: https://www.kaggle.com/datasets/hemendrasr/pizza-vs-ice-cream

Data Preprocessing:

To prepare the dataset for training, several preprocessing steps were applied:

- Image augmentation techniques, such as rotation, horizontal and vertical flips, shear, and were employed to increase the diversity of the training set.
- All images were rescaled to a common size of (417,676) pixels to ensure uniformity in input dimensions.

Model Building (Transfer-Learning):

The VGG16 architecture was chosen as the base model due to its proven effectiveness in image classification tasks. The model was modified by freezing specific layers (block1) to retain pre-trained features while allowing fine-tuning on subsequent layers.

Hyper parameter tuning was performed using the Keras Tuner, focusing on the dense layer. The following hyper parameters were optimized:

- Units in the dense layer.
- Activation function (sigmoid, tanh, relu).
- Weight initialization (glorot_uniform, glorot_normal, he_uniform, he_normal).

Additionally, batch normalization and dropout were introduced to enhance model generalization.

Compile the model:

- The model was compiled using binary crossentropy as the loss function and dynamically choosing an optimizer (sgd, Adam, rmsprop, adadelta) based on hyper parameter tuning. Accuracy was chosen as the evaluation metric.

Training and Evaluation:

- The model was trained for 15 epochs, with early stopping to prevent over fitting. Model Checkpoint and CSVLogger callbacks were employed for model saving and logging, respectively. Training and validation performance were visualized using loss and accuracy plots.

Prediction:

The trained model was utilized to make predictions on a dedicated test dataset. This phase is essential to evaluate the model's performance on unseen data. The following steps outline the prediction process and the subsequent evaluation:

1. **Making Predictions:** The trained model was applied to the images in the dataset using the **'predict'** method. This method produces probability scores for each class (pizza and ice cream). The class with the highest is considered the predicted class for a given image.
2. **Evaluation Metrics:** To quantitatively assess the model's performance on the test dataset, standard evaluation metrics were employed. The primary metric used is accuracy, which measures the proportion of correctly classified samples over the total number of samples in the test set.
3. **Confusion Matrix:** A confusion matrix was generated to provide a more detailed analysis of the model's performance. The confusion matrix illustrates the number of true positive, true negative, false positive, false negative predictions.

This comprehensive evaluation ensures a thorough understanding of the model's performance and facilitates informed decision-making regarding potential model refinements or deployment.

Model Performance:

- After training and evaluating the model, the final accuracy achieved on the test set was 58%. This accuracy achieved metric represents the proportion of correctly classified instances out of the total number of samples in the test set. The achieved accuracy provides valuable insight into the effectiveness of the model in distinguishing between ice cream and pizza images.

Conclusion:

In conclusion, the project successfully tackled the task of classifying ice cream and pizza images, achieving an accuracy of 58%. The combination of transfer learning, hyper parameter tuning, and comprehensive model evaluation demonstrates a systematic approach to building an effective image classification model.