

DSCI 552 Group Report

Group 14 -
Sian Jyun Kuo,
Renqi Huang,
Nemil Shah
23/04/2023

Abstract

This study investigates the impact of pre-processing techniques and freezing layers on model accuracy. The initial model achieved a validation accuracy of 17%, but it was found to be overfitting the data. To address this issue, the group implemented pre-processing techniques on the data and applied freezing layers to prevent the model from overfitting. The pre-processing technique involved using image pre-processing, and the freezing layer method involved freezing the layers. As a result of these techniques, the model's validation accuracy was improved from 17% to 96%. The study provides insights into the benefits of using pre-processing and freezing layers for enhanced model accuracy and performance. This research can be useful for practitioners and researchers interested in improving model performance and reducing overfitting in image identification.

Introduction

Deep learning has become increasingly important in many fields, such as image recognition, natural language processing, and predictive modeling. However, one common challenge in deep learning is overfitting, where the model becomes too complex and performs well on the training data but poorly on new, unseen data. To address this issue, various techniques have been proposed, including pre-processing techniques and freezing layers.

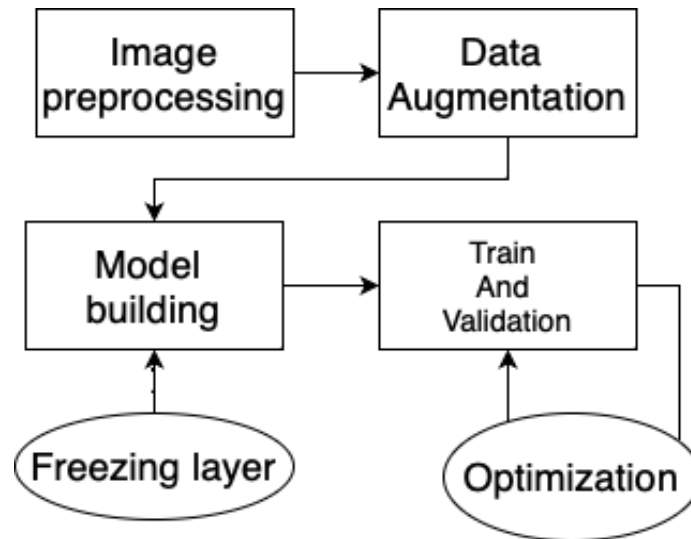


Figure 1. The process workflow of the project.

In this study, we investigate the impact of pre-processing techniques and freezing layers on a model's classification accuracy. Specifically, we focus on image recognition, where pre-processing techniques such as photo pre-processing can be applied to enhance the applicability of data for our own purpose. Additionally, we apply freezing layers to prevent the model from overfitting by restricting the training of certain layers in the network.

The dataset is made up of a selection of pictures of well-known (or lesser-known) sites. A two-level data hierarchy has been used to organize the collection. The categories for the landmarks are on the first level, and the specific landmarks are on the second level. There are 6 categories, including Gothic, Modern, Pagodas, Pyramids, Neoclassical, and Mughal architecture. There are a total of 30 landmarks, 5 for each category. There are 14 photographs for each landmark.

The initial model achieved an accuracy of 17% on validation data accuracy with training accuracy 96%, which was found to be overfitting the data. To improve the validation accuracy and solve the overfitting problem, we implemented pre-processing techniques on the data and applied freezing layers. Our results show that these techniques led to a significant improvement in the model's performance, achieving an accuracy of 96% on a validation set.

This study contributes to the ongoing efforts to improve model accuracy and reduce overfitting in machine learning. The findings can be useful for practitioners and researchers interested in applying pre-processing techniques and freezing layers to improve model performance in image recognition tasks.

Methods

In the beginning, we just tried VGG16 and EfficientNetB0 to train and test with some data augmentation by applying a range of image transformations to the original images, including horizontal flipping, rotation, shear, and zoom. we get around 17% validation accuracy.

We applied methods, which included adjusting brightness, saturation, hue, and gamma to make data more suitable for our needs. We also tried the Sobel filter for edge detection and converted the images to grayscale for easier processing. After pre-processing, the validation accuracy greatly increased from 17% to around 80%. We found out that our training accuracy is pretty high around 95% and validation accuracy is around 80%, then we realized that we still have the overfitting problem in this model.

To prevent overfitting, we improved our model training method with freezing layer technique. In the EfficientNet, the model we used, has 237 layers, but not all of these layers need to be trained. Since it is already a powerful network for image identification, further learning of the model is very likely to cause overfitting. We only retained the last 20 layers which can be learned from training. This greatly improved the validation accuracy to around 96%.

Data Augmentation

For augmenting the dataset, we used image transformation techniques such as horizontal flipping, rotation, shear and zoom.

Pre-processing dataset

We tried the following three methods for preprocessing our images:

1. Adjusting: gamma, brightness, saturation, hue



Figure 2. The example building after transformation.

2. Apply Sobel filter to get edge detection effect



Figure 3. The example building after edge detection transformation.

3. Turn RGB to Gray-scale

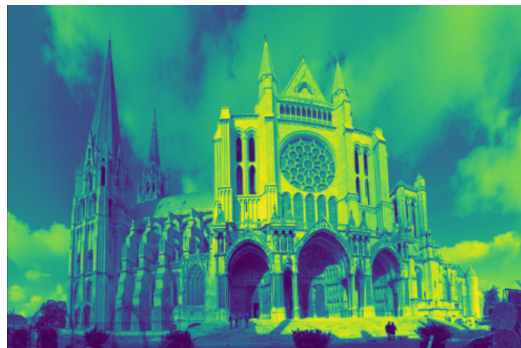
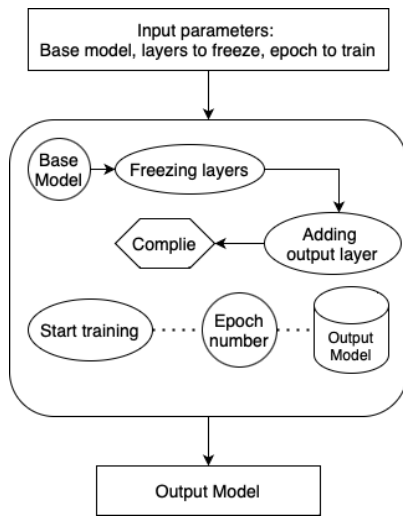


Figure 4. The example building turned into gray-scale.

Out of these methods, we found out adjusting gamma, brightness, hue and saturation to be the most effective. It helped us increase our validation accuracy from 17% to 80%.

Model Building



For the parameters which we used to build the model: we choose EfficientNetB0 as our base model, which has 237 layers in total. We apply the freezing layer technique on the first 217 layers of the base model. For the epoch number, we chose 30 for our training progress.

Figure 5. Model Building approach

Freezing layer

Freezing Layer: first 217 layers, unfreeze the last 20 layers.

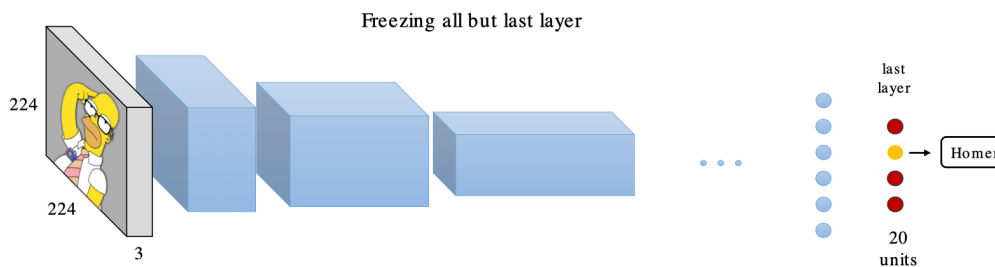


Figure 6 . The presentation of the freezing layers theory.

Results

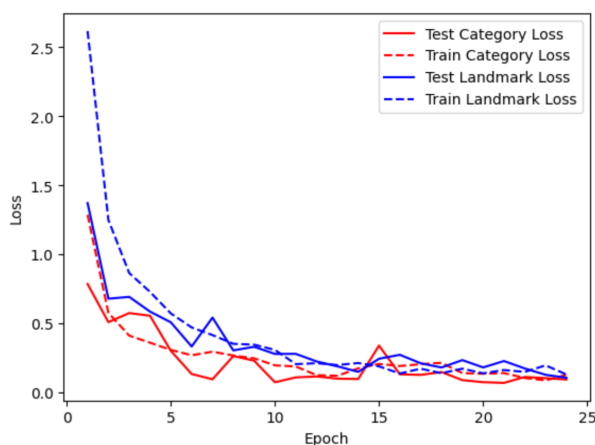


Figure 6 . The curve of the loss function on the training data set and validation data set.

F1 score for Landmark: 0.98(validation)

F1 score for Category: 0.91(validation)

To evaluate our results, we used the loss curve and F1 score. We were able to achieve F1 scores of over 0.9 for both category and landmark classification, indicating that our models were performing well on both tasks. The loss curves of the validation set and training set exhibit similar decreasing trends, and they almost reached convergence at the same time which imply the model is not overfitting. Overall, our project highlights the potential for deep learning to be applied to image classification tasks, and the importance of addressing overfitting in model development.

Conclusions

In this study, we have shown that pre-processing techniques and freezing layers can significantly improve model accuracy and reduce overfitting. Specifically, our results demonstrate that the application of photo pre-processing and freezing convolutional layers led to a considerable improvement in model accuracy, from 17% to 96%.

These findings have important implications for machine learning practitioners and researchers who seek to improve model performance and reduce overfitting in image recognition tasks. By utilizing pre-processing techniques and freezing layers, they can enhance the quality of the data and prevent the model from overfitting, leading to more accurate and reliable predictions.

Overall, our study contributes to the ongoing efforts to improve machine learning algorithms' performance and highlights the potential of pre-processing techniques and freezing layers in achieving this goal. Future research can explore additional pre-processing techniques and freezing layer methods to further improve model accuracy and performance.

Future Directions

There are three potential directions to improve the performance and application of machine learning models: refining and optimizing machine learning models, exploring multi-task learning, and applying the models to real-world use cases.

Refining and optimizing machine learning models involves experimenting with different model architectures, hyperparameters, and training techniques. Multi-task learning can help models learn more generalizable features by training a single model to perform both category and landmark classification. Applying machine learning models to real-world use cases such as automatically categorizing and tagging images or identifying landmarks for tourists would involve integrating the models into a larger system and testing their performance in real-world scenarios. These three directions have the potential to enhance the usefulness and effectiveness of machine learning models.

Bibliography

- TensorFlow API documentation. (n.d.). Retrieved April 25, 2023, from https://www.tensorflow.org/api_docs
- Stack Overflow. (n.d.). Retrieved April 25, 2023, from <https://stackoverflow.com>