Weather Image Classification

ACM40960-Projects in Maths Modelling

Department of Mathematics and Statistics

Sarvesh Naik(22204841)
Neha Sharma(22201277)

University College Dublin

17th Aug 2023

Contents

- Weather Image Classification holds substantial importance within the fields of meteorology and computer vision.
- Its implications extend across a diverse spectrum, encompassing activities such as weather prediction, environmental surveillance, and disaster preparedness.
- The conventional techniques predominantly rely on human interpretations, potentially leading to inaccuracies and uncertainties. This underscores the need for more sophisticated methodologies.
- As a solution, the integration of Convolutional Neural Networks (CNNs) in weather image classification has emerged as a robust and pragmatic approach.

- The weather image classification dataset comprises 6862 images representing various weather conditions.
- These images are categorized into 11 distinct classes, including dew, fog/smog, frost, glaze, hail, lightning, rain, rainbow, rime, sandstorm, and snow.

Figure 1: Different types of Weather Images

Model 1

- CNN with 4 convolution layers and 4 max-pooling layers then followed by 2 fully connected layers.
- The first convolution layer is set with 32 filters and a 3×3 kernel with strides 1 (default).
- Next 3 convolution layers are set with 64 and 128 filters with 3×3 kernels. The fully connected layer uses 512 units and ReLU activation function.

Model 2

- CNN model with 4 layers and the number of filters has been increased from 64 to 512.
- Dropout regularization is used to help prevent overfitting.
- Batch normalization layer is added after each convolutional layer or dense layer.
- Different kernel sizes is used in convolutional layers, such as 5x5 or 7x7. In one layer, 3x3 kernel size is also used.

Model 3

- Data Augmentation is used to increase the number of images.
- Four convolution layers with increased filter count (64 to 512).
- Dropout for overfitting prevention.
- Batch normalization after each convolutional or dense layer.
- Varying kernel sizes (5x5, 7x7, and 3x3) for convolutions.

Model Network Architecture

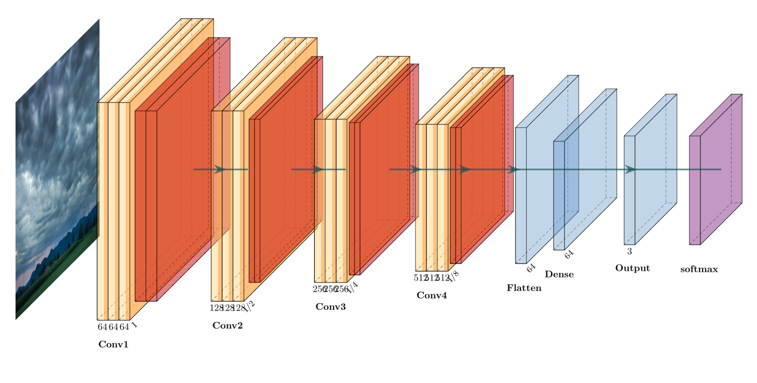

Figure 2: Model 2 Network Architecture

Model 1 Simulation Result

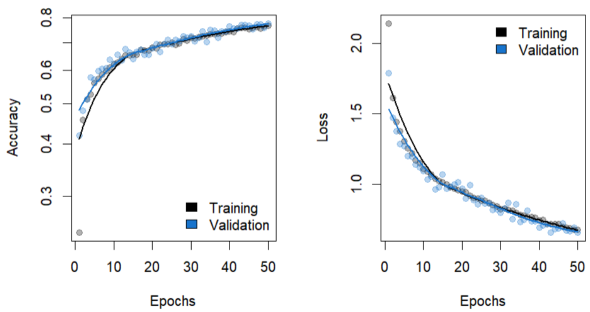

Figure 3: Model 1 Accuracy, Loss curves

The accuracy curve flattens early or before reaching a high level, it indicates that this could be due to model limitations or insufficient complexity to capture the data's patterns.

Model 2 Simulation

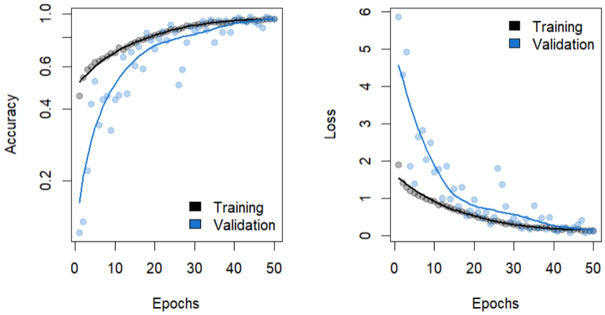

Figure 4: Model 2 Accuracy, Loss curves

The loss curve's downward trend for both training and validation data indicates successful learning. Moreover, the validation loss starts to decrease while training loss keeps decreasing, meaning the model is not overfitting.

Model 3 Simulation

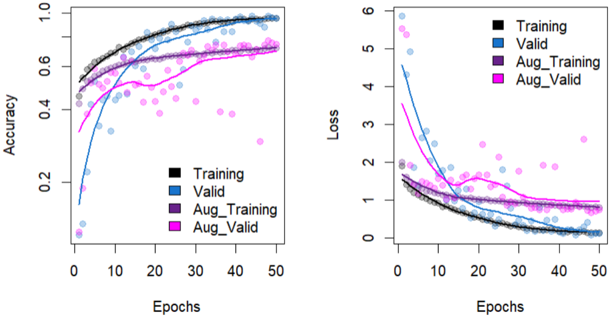

Figure 5: Model 3 Accuracy, Loss curves compared with Model 2

In the accuracy curve, a significant gap between training and validation accuracy indicates overfitting as the validation accuracy lags behind or remains stagnant.

Best Model macro-averaged precision, recall, and F1 scores

```
Class 0 Precision: 0.926      Recall: 0.855   F1 score: 0.889
Class 1 Precision: 0.792      Recall: 0.919   F1 score: 0.851
Class 2 Precision: 0.73       Recall: 0.632   F1 score: 0.677
Class 3 Precision: 0.668      Recall: 0.634   F1 score: 0.651
Class 4 Precision: 0.801      Recall: 0.725   F1 score: 0.761
Class 5 Precision: 0.852      Recall: 0.924   F1 score: 0.887
Class 6 Precision: 0.895      Recall: 0.682   F1 score: 0.774
Class 7 Precision: 0.865      Recall: 0.647   F1 score: 0.741
Class 8 Precision: 0.739      Recall: 0.917   F1 score: 0.818
Class 9 Precision: 0.895      Recall: 0.856   F1 score: 0.875
Class 10 Precision: 0.693     Recall: 0.65    F1 score: 0.671

Macro-averaged Precision: 0.805       Macro-averaged Recall: 0.767    Macro-averaged F1 score: 0.781
```

Figure 6: Best Model scores

Quantification of the comparative study

	Model 1	Model 2	Model 3
Test set accuracy	0.7751	0.952	0.7344
Macro-averaged Precision	0.404	0.805	0.353
Macro-averaged Recall	0.371	0.767	0.302
Macro-averaged F1 Score	0.280	0.781	0.189

Table 1: Quantification of the comparative study

We remark that

- **Model 2 performs best with test set accuracy of** 95.2%
- **Model 3 (with data augmentation) performs worse with accuracy of** 73.4%**, this could be due to overfitting.**
- **Model 2 is able to precisely classify all 11 categories of weather conditions with acceptable precision, recall and F1 scores.**

We conclude that:

Conclusion

- **Model 2 stood out with an accuracy of 0.952, showcasing the power of increased filters, dropout, and batch normalization.**
- **Varying kernel sizes in Model 2 compared to Model 1 highlighted the adaptability of CNNs in capturing weather features.**
- **Using a more complex model (such as Model 3) alongside data augmentation could potentially have introduced more opportunities for the model to memorize the augmented training examples, rather than truly learning the underlying patterns in the data.**
- **Despite successes, macro-averaged precision and recall indicate room for further enhancement.**

Thank You !