Dog vs Cat vs Bird Classifier Report

Musab bin Jamil (29409)[*] Hussain Diwan (29410)[†]

December 30, 2024

Contents

[*]
[†]

1 Dataset Preprocessing and Exploration

The dataset for this project includes images of dogs, cats, and birds. Key preprocessing steps included:

- Loading and normalizing the image data.
- Resizing images to a uniform size of 32×32 pixels.
- Applying transformations such as flipping, rotation, and normalization to improve model robustness.

The dataset was visualized to ensure balanced representation across the three classes.

2 Baseline Model Development

A simple Convolutional Neural Network (CNN) was used as the baseline model. The architecture included:

- Two convolutional layers with ReLU activation and max-pooling.
- Fully connected layers for classification into three categories.

The model was trained using the Cross-Entropy Loss function and optimized using the Adam optimizer with a learning rate of 0.001. This model produced a prediction with an accuracy of 71.466% on Kaggle.

3 Optimization Techniques

Several techniques were implemented to improve the model's performance:

- **Data Augmentation**: Added transformations like color jittering and random cropping to increase dataset diversity.
- **Regularization**: Used dropout layers to prevent overfitting.
- **Learning Rate Scheduling**: Adjusted the learning rate dynamically during training.

After this fo some reason the accuracy dropped to 61.7%.

4 Transfer Learning

To further enhance performance, transfer learning was employed using a pre-trained ResNet model. The final classification layer was replaced with a custom layer for three-class classification. The pre-trained layers were frozen during initial training and later fine-tuned. Using this model accuracy improved by roughly 10%. The resulting accuracy was 82.233%.

5 Optimization of Transfer Learning

Fine-tuning the transfer learning model involved:

- Unfreezing selected layers of the pre-trained model.
- Experimenting with different learning rates for fine-tuning.
- Incorporating batch normalization to stabilize training.

Finally, after these steps a final accuracy of 82.233% was achieved. This had no significant changes to previous ResNet model.

6 Results, Metrics, and Insights

The following metrics were used to evaluate the model:

- **Accuracy**: Achieved a training accuracy of approximately 72% with the baseline model and over 83% with transfer learning.
- **Loss**: Observed a steady decrease in loss over epochs. Starting off with 92.69% and ending on 27.49% by the 10th Epoch.

Insights:

- Data augmentation significantly improved model generalization, by decreasing the accuracy. This can be further tested with more Epochs, to improve accuracy.
- Transfer learning did not have that much of an impact on accuracy compared to the baseline model.
- Regularization techniques like dropout effectively reduced overfitting, as we can see with the decrease in accuracy.

7 Conclusion

Through this project it is clear that pre-trained models that have a wider pattern recognition spectrum are more accurate as compared to a homemade model with problem specific data. Secondly, augmentation of data can result in a overall better generalization and lower chances of over fitting.