Optimizing Cats & Dogs Classification with Baseline, Expanded Dataset, and Pretrained Network Approaches

Introduction The objective of this project was to develop an accurate and generalizable classifier for the Cats & Dogs dataset. By evaluating multiple architectures, dataset sizes, and optimization techniques, this study aimed to determine the best strategy for achieving high performance. The final phase included using a VGG16 pretrained model with additional optimizations to compare results across different sample sizes and validate the most effective model configurations.

Question 1: Baseline Model Exploration

Objective:
To establish a foundation for performance, four models were developed with distinct architectural features, providing a baseline for evaluation.

Methodology:
Each model was built and trained from scratch, incorporating different layers and regularization techniques:

- **Model 1**: Basic convolutional model using MaxPooling layers with filters increasing from 32 to 256 across 5 layers.
- **Model 2**: Added data augmentation, MaxPooling with the same filter structure as Model 1, and a dropout rate of 0.5 to reduce overfitting.
- **Model 3**: Expanded filter depth from 32 to 512 across 6 layers and added an additional convolutional layer. Data augmentation and dropout were included to improve robustness.
- **Model 4**: Same structure as Model 2 but with a higher dropout rate of 0.6, aiming to reduce overfitting.

Results:
Among these, **Model 3 achieved the best performance**, with a training accuracy of 80% and a validation accuracy of 72%. However, it showed signs of overfitting, highlighting potential for further improvement through increased data and optimizations. **Model 3** was selected for subsequent trials due to its promising baseline performance.

Question 2: Expanding the Training Dataset

Objective:
To assess the impact of a larger training dataset on generalization and accuracy, the dataset size was increased to 4,000 images while keeping the validation and test sets constant.

Methodology:
Model 3 was retrained with 4,000 images. By increasing the data volume, the model had the opportunity to learn a broader range of features, thereby enhancing its capacity to generalize to unseen images.

Findings:
With the expanded dataset, Model 3 achieved:

- **Validation Accuracy:** 81%
- **Validation Loss:** 0.4114

The expanded data allowed the model to better capture underlying patterns, improving generalization without significant overfitting.

Conclusion:
Increasing the dataset size demonstrated notable benefits, as it enabled the model to recognize more nuanced features and achieve higher validation accuracy. This experiment confirmed that more data can enhance model stability and generalization.

Question 3: Optimizing Sample Size for Best Performance

Objective:
To identify the most effective training sample size, balancing training efficiency and generalization capability.

Methodology:
Experiments were conducted with different sample sizes, specifically **1,500, 2,000, and 2,500 images**. The model's performance was evaluated across each size to determine the optimal configuration.

Results:
The **2,000-sample size** yielded the best validation performance, with:

- **Validation Accuracy:** 78.2%
- **Validation Loss:** The lowest among the tested sizes.

Insight:
The results suggested an optimal point at 2,000 samples, where the model demonstrated strong generalization while avoiding excessive training time and overfitting. This optimal size provided a balance between the amount of data and model performance.

Question 4: Using a Pretrained Network with Optimizations

Objective:
Leverage a VGG16 pretrained network to improve the model's feature extraction capabilities, applying targeted optimizations to further enhance performance.

Methodology:
Using VGG16 as a base model, various techniques were applied to adapt it to the Cats & Dogs classification task:

1. **Data Augmentation**: Advanced data augmentation techniques were used, including:

    - **Rotations, Shifts, Shear, Zoom with fill mode**: This allowed the model to encounter a wider range of image variations, improving its ability to generalize to unseen data.
2. **Learning Rate Scheduling**: The **ReduceLROnPlateau** callback reduced the learning rate once the validation loss plateaued, allowing finer adjustment of weights as the model approached convergence.
3. **Regularization**:
    - **Dropout**: Set to a rate of 0.5 to prevent overfitting, encouraging the model to learn simpler, more generalizable patterns.
    - **L2 Penalty**: This regularization technique penalized complex models, pushing the model towards a more robust architecture.
4. **Fine-Tuning**: The last four layers of the VGG16 base were unfrozen, enabling them to learn dataset-specific features that enhanced the model's overall accuracy and stability.

Results Across Sample Sizes:

- **Smallest sample size (1,000 images)**: Achieved moderate improvement but was limited by data scarcity, resulting in a validation accuracy of around 73%.
- **Intermediate sample sizes (1,500 and 2,000 images)**: These sizes led to incremental accuracy gains, with reduced validation loss and stronger generalization ability.
- **Largest sample size (2,500 images)**: Produced the best results, achieving nearly 75% validation accuracy and the lowest validation loss.

Conclusion:
The optimizations, especially the fine-tuning of VGG16's final layers, contributed to a more stable and high-performing classifier. By applying advanced data augmentation, learning rate scheduling, dropout, and fine-tuning, the model maximized the use of available data, achieving optimal accuracy and generalization.

Overall Conclusion

The project demonstrated a clear path from baseline model exploration through dataset expansion, sample size optimization, and pretrained network adaptation. Each phase provided valuable insights:

1. **Baseline Performance**: Model 3 served as an effective starting point, revealing potential for improvement.
2. **Increased Data**: Expanding the dataset improved generalization, highlighting the impact of data volume on performance.
3. **Sample Size Optimization**: Testing various sizes identified the optimal sample size, achieving strong accuracy without requiring excessive data.
4. **Pretrained Network and Fine-Tuning**: The final phase using VGG16 with optimizations provided the best performance, effectively balancing complexity and generalizability.

In summary, this report illustrates the importance of combining data volume, architecture enhancements, and pretrained models with robust optimization techniques. The combination of these strategies allowed for the development of a highly effective model for the Cats & Dogs classification task, underscoring the impact of well-designed data augmentation, learning rate scheduling, and regularization on model performance.