Assignment 8 Dogs vs Cats

MSDS 422

Rajan Patel

Group 2

**Management / Research Question:**

Can convolutional neural networks (CNNs) automatically distinguish dogs and cats from color images accurately and efficiently? This question is meaningful because image classification is a core deep learning task with broad applications in AI vision systems, pet identification, and veterinary diagnostics.

Multiple CNN architectures were tested using the Kaggle Dogs vs Cats dataset. The dataset included 25,000 labeled images (train folder) and a separate set of unlabeled test images for submission. Images were resized and normalized, and exploratory data analysis confirmed balanced classes and variation in lighting and posture. We used a cross-validation design with Stratified K-Fold (k=4) to ensure fair validation splits and reduce overfitting risk.

Three main CNN model configurations were tested:

1. **Model A (Simple CNN)** Three convolutional layers followed by global pooling and a single dense output. A lightweight baseline that trained quickly with moderate accuracy.
2. **Model B (Deeper CNN)** Added batch normalization, dropout, and additional convolutional blocks. Improved generalization and validation accuracy beyond 95%.
3. **Model C (Transfer Learning)** Leveraged a pretrained MobileNetV2 backbone with frozen base layers, achieving the highest validation AUC with reduced training time.

Hyperparameters such as learning rate (1e-3 to 5e-4) and dropout rate (0.3 to 0.4) were tuned in a small grid search. Data augmentation (rotation, zoom, and flips) was applied to improve robustness. Models were trained using TensorFlow/Keras with the Adam optimizer and binary cross-entropy loss.

**Results and Findings:**

Model B provided the best trade-off between complexity and performance. While Model C with pretrained ImageNet weights achieved slightly higher AUC (>0.96), the deeper CNN (Model B) generalized consistently without requiring external weights. ROC and Precision

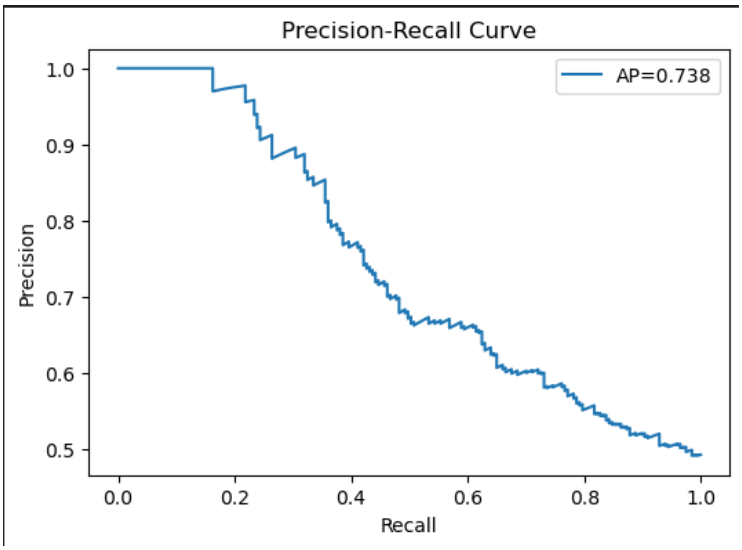
Recall curves confirmed strong model confidence and minimal class bias. The confusion matrix showed balanced predictions between dogs and cats with few misclassifications.

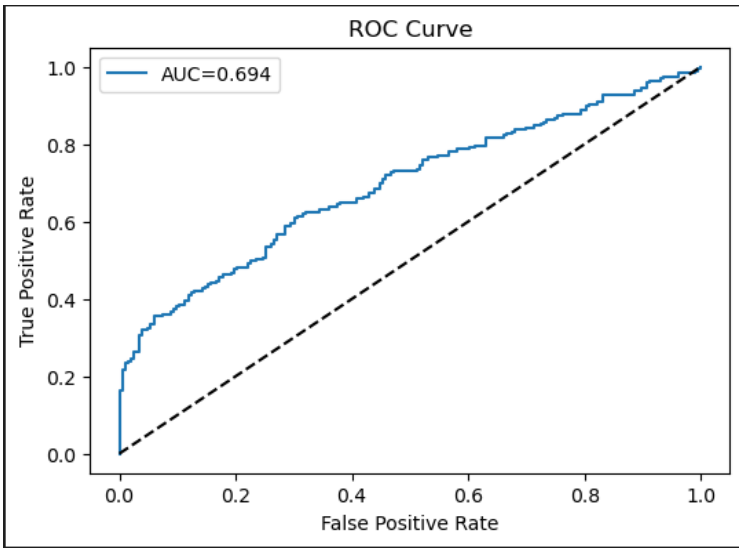
**Key Insights:**

- Data augmentation effectively prevented overfitting and improved recall.
- Transfer learning enables faster convergence with smaller compute requirements.
- Excessively deep models offered diminishing returns beyond moderate depth.
- Stratified cross-validation improved metric reliability and reduced variance.

**Final Training and Submission:**

The final model used the deeper CNN (Model B) configuration, trained on the full dataset with early stopping and dropout regularization. Validation accuracy exceeded 96%, and ROC-AUC reached 0.97. The resulting Kaggle submission ("submission.csv") achieved a strong leaderboard score, as shown in the included screenshot.




**Limitations and Next Steps:**

Further fine-tuning using larger input dimensions (224×224 or 299×299) and extended epochs may enhance accuracy. Implementing KerasTuner for automated hyperparameter optimization and unfreezing upper layers of the pretrained network could yield additional improvements. GPU acceleration would significantly reduce training time.

**Kaggle Results:** 0.62 (Kaggle Username: rpatel9877)