

PyTorch Capstone Project

Cat Breed Classification System

Wiwat Pholsomboon

April 8, 2025

Deep Learning with Pytorch - INFO-6147-(01)-25W

Introduction

- Study on using PyTorch models to classify cat breeds
- Implementation of Grad-CAM for visual explanation
- Integration with ChatGPT-4o-mini for image identification
- Interactive quiz system to compare model accuracy with human performance

Dataset

- **Source:** Kaggle 'Geno Cat Breed Image Collection'
- **Size:** 15 cat breeds with 375 photos each (5,625 total)
- **Preprocessing pipeline:**
 - Resize to 256×256 pixels
 - Random crop to 224×224 pixels
 - Random horizontal flip for augmentation
 - Random rotation for variety
 - Conversion to PyTorch tensors
 - Normalization using ImageNet statistics

Model Architecture

Custom CNN:

- 5 Convolutional layers
- 2 Fully connected layers
- Dropout for regularization
- ReLU activation functions

Transfer Learning Models:

- ResNet18
- EfficientNetB2
- VGG16
- All pre-trained with ImageNet weights

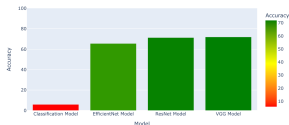
Training Methodology

Consistent hyperparameters across all models:

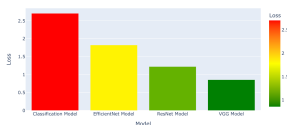
- Batch size: 128
- Learning rate: 0.001
- Momentum: 0.9
- Epochs: 5
- Loss function: Cross Entropy
- Optimizer: SGD

Results and Evaluation

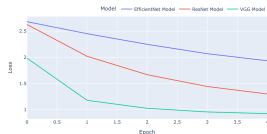
Model Accuracy



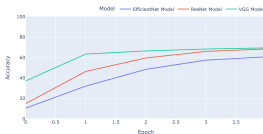
Model Loss



Model Training Loss

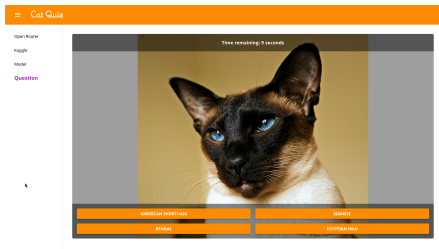


Model Training Accuracy



- Comparison of model accuracy and loss
- Interactive quiz implementation for human vs AI performance analysis





Visualization and Interpretability



- **Grad-CAM visualization** highlights regions of interest
- **ChatGPT-4o-mini** provides reasoning for predictions
- **Interactive interface** for user engagement

Evaluation Results

- Side-by-side comparison of accuracy:
 - Human participants
 - ChatGPT-4o-mini
 - PyTorch models (CNN, ResNet, EfficientNet, VGG)

<p>Your Answer</p> <p>British Shorthair ✔ Correct!</p>	<p>ChatGPT-4o-mini Answer (1.53 seconds)</p> <p>British Shorthair ✔ Correct!</p> <p>The cat appears to have a stocky build and round face with short fur, which are characteristics typical of the British Shorthair breed.</p>	<p>Classification Model Answer (0.10 seconds)</p> <p>I don't know ✖ Incorrect!</p> <p>The correct answer was: British Shorthair</p> 
<p>ResNet Model Answer (0.11 seconds)</p> <p>Persian ✖ Incorrect!</p> <p>The correct answer was: British Shorthair</p> 	<p>EfficientNet Model Answer (0.43 seconds)</p> <p>British Shorthair ✔ Correct!</p> 	<p>VGG Model Answer (0.13 seconds)</p> <p>I don't know ✖ Incorrect!</p> <p>The correct answer was: British Shorthair</p> 

You : 2 ChatGPT : 2 Classification : 0 ResNet : 1 EfficientNet : 2 VGG : 1

[Next question](#)

Challenges and Solutions

Challenges:

- Different target layers for Grad-CAM across models
- Low accuracy with custom CNN model

Solutions:

- Manual selection of appropriate CNN layers
- Transfer learning to leverage pre-trained weights

Future Improvements

Potential Enhancements

- User management system to track quiz performance over time
- Enhanced data augmentation techniques for better generalization
- Additional model architectures for comparison
- More comprehensive evaluation metrics beyond accuracy
- Fine-tuning of transfer learning models

Conclusion

- **ChatGPT-4o-mini** demonstrates superior accuracy due to advanced image understanding capabilities
- **Transfer learning models** perform moderately well with limited training
- **Custom CNN model** shows potential but requires more extensive training
- **Visualization tools** provide valuable insights into model decision-making

Thank you!