

Classificação Automática de Raças de Gatos Utilizando Transfer Learning com EfficientNet

Giovana Ribeiro de Francisco RA: 10297494

Leila Akina Ino RA: 10402951

Leticia Moraes Gutierrez de Oliveira RA: 10400969

Resumo

Este trabalho apresenta o desenvolvimento de um sistema inteligente para a classificação automática de quatro raças de gatos: Maine Coon, Persa, Siamês e Sphynx. Utilizando técnicas de aprendizado profundo e transferência de aprendizado, a abordagem proposta emprega a arquitetura EfficientNet, previamente treinada no conjunto de dados ImageNet, para realizar a identificação das raças mencionadas. O modelo está integrado a uma interface web interativa desenvolvida com Streamlit, proporcionando uma experiência completa desde o treinamento até a utilização prática em tempo real.

1. Introdução

A classificação automática de raças de animais constitui um desafio relevante na área de visão computacional, envolvendo aspectos práticos e técnicos. Com os avanços em redes neurais convolucionais (CNNs) e técnicas de transferência de aprendizado, tornou-se viável desenvolver sistemas precisos e eficientes para essa finalidade.

Este trabalho concentra-se na identificação automática de quatro raças específicas de gatos, cada uma com características visuais distintas que facilitam a separação por sistemas baseados em aprendizado supervisionado. A escolha por gatos como objeto de estudo justifica-se por sua popularidade como animais de estimação, e pela necessidade prática de identificação correta das raças, especialmente em contextos veterinários e processos de adoção.

2. Metodologia

2.1 Arquitetura do Modelo

A arquitetura EfficientNet foi selecionada como base para o modelo, devido ao seu equilíbrio entre eficiência computacional e precisão. Essa rede neural convolucional foi previamente treinada no conjunto de dados ImageNet, o que permite a aplicação de técnicas de transferência de aprendizado para a tarefa específica de classificação de raças de gatos.

2.2 Transferência de Aprendizado e Fine-Tuning

A implementação do modelo ocorreu em duas fases:

Fase 1: Treinamento das Camadas Superiores:

- Congelamento das camadas da EfficientNetB0 previamente treinada,
- Treinamento apenas das camadas personalizadas adicionadas,
- Taxa de aprendizado: 0,001,
- Duração: 5 épocas.

Fase 2: Ajuste Fino (Fine-Tuning):

- Descongelo das últimas 20 camadas do modelo base,
- Redução da taxa de aprendizado para 0,0001,
- Treinamento adicional para ajuste fino ao domínio específico.

2.3 Arquitetura das Camadas Personalizadas

A arquitetura do modelo inclui as seguintes camadas:

- Entrada: Imagens redimensionadas para 224x224 pixels com 3 canais (RGB),
- Camada base: EfficientNetB0 pré-treinada,
- Camada de Pooling: GlobalAveragePooling2D,
- Camada de Dropout: Taxa de 0,2,
- Camada Densa: 128 neurônios com função de ativação ReLU,
- Camada de Dropout: Taxa de 0,2,
- Camada de Saída: 4 neurônios com função de ativação Softmax.

2.4 Pré-processamento e Aumento de Dados

Para melhorar a generalização do modelo, foram aplicadas as seguintes técnicas:

- **Normalização:** Escalonamento dos valores de pixel para o intervalo [0, 1],
- **Redimensionamento:** Padronização das imagens para 224x224 pixels,

- **Aumento de Dados:**
 - Rotação aleatória de até ± 20 graus,
 - Deslocamento horizontal e vertical de até $\pm 20\%$,
 - Espelhamento horizontal,
 - Zoom aleatório de até $\pm 20\%$,
 - Cisalhamento de até $\pm 20\%$.

3. Raças de Gatos e Características Visuais

O modelo foi treinado para classificar quatro raças distintas de gatos, cada uma com características visuais únicas:

- **Maine Coon:** Maior raça de gato doméstico, com pelagem densa e longa, orelhas tufadas e cauda peluda.
- **Persa:** Face achatada, pelagem longa e luxuosa, focinho curto e olhos grandes e redondos.
- **Siamês:** Corpo elegante, padrão colorpoint distintivo, olhos azuis e extremidades escuras.
- **Sphynx:** Aparência sem pelos, pele enrugada e orelhas grandes.

4. Implementação Técnica

4.1 Pipeline de Treinamento

O sistema implementa um pipeline completo de treinamento com as seguintes funcionalidades:

- **Callbacks Inteligentes:**
 - **ModelCheckpoint:** Salvamento automático do melhor modelo com base na acurácia de validação,
 - **EarlyStopping:** Interrupção do treinamento em caso de estagnação da performance para evitar overfitting,
 - **ReduceLROnPlateau:** Ajuste automático da taxa de aprendizado quando a performance não melhora.

- **Validação Cruzada:** Divisão automática dos dados em conjuntos de treino (80%) e validação (20%),
- **Monitoramento em Tempo Real:** Visualização de métricas durante o treinamento.

4.2 Interface Web com Streamlit

A aplicação web desenvolvida oferece três módulos principais:

Módulo de Treinamento:

- Configuração interativa de hiperparâmetros,
- Visualização da distribuição do dataset,
- Monitoramento em tempo real do progresso,
- Gráficos automáticos de perda e acurácia.

Módulo de Predição:

- Upload intuitivo de imagens,
- Predição instantânea com visualização de confiança,
- Distribuição de probabilidades por raça,
- Informações educativas sobre cada raça.

Módulo de Análise:

- Métricas detalhadas de performance,
- Matriz de confusão interativa,
- Comparação de métricas por classe,
- Visualizações avançadas (gráfico radar).

5. Avaliação e Métricas

5.1 Métricas de Performance

O sistema utiliza múltiplas métricas para avaliação abrangente:

- **Acurácia (Accuracy):** Proporção de predições corretas,
- **Precisão (Precision):** Capacidade de evitar falsos positivos,
- **Revocação (Recall):** Capacidade de identificar verdadeiros positivos,
- **F1-Score:** Média harmônica entre precisão e revocação,
- **Matriz de Confusão:** Análise detalhada por classe.

5.2 Validação

A validação do modelo ocorre em três níveis:

- **Validação durante o Treinamento:** 20% dos dados de treino reservados para validação,
- **Teste Final:** Conjunto separado de dados nunca vistos pelo modelo,
- **Validação Cruzada:** Robustez através de múltiplas divisões dos dados.

6. Funcionalidades Avançadas

6.1 Sistema de Predição Inteligente

O sistema inclui uma função de predição que realiza o pré-processamento automático da imagem, realiza a predição com análise de confiança e retorna a classe prevista juntamente com a confiança associada.

6.2 Visualizações Interativas

O sistema oferece múltiplas formas de visualização:

- Gráficos de Treinamento: Perda e acurácia em tempo real,
- Matriz de Confusão: Heatmap interativo,
- Distribuição de Probabilidades: Barras horizontais com destaque,
- Gráfico Radar: Comparação multidimensional de métricas.

7. Considerações Técnicas

Para garantir o funcionamento do sistema de predição, ao inserir uma nova imagem, é necessário limpar o cache da aplicação. Como o sistema utiliza o decorator `@st.cache_resource` para otimização de performance, imagens previamente processadas podem ser reutilizadas indevidamente caso o cache não seja reiniciado. Portanto recomenda-se reiniciar o aplicativo antes de inserir uma nova imagem para forçar a limpeza do cache e assegurar que os dados mais recentes sejam processados corretamente.

7.1 Otimizações Implementadas

- **Transferência de Aprendizado:** Redução do tempo de treinamento e necessidade de dados por meio da utilização de pesos pré-treinados,
- **Aumento de Dados (Data Augmentation):** Melhora da generalização e robustez do modelo frente a variações,
- **Early Stopping e Checkpoints:** Evita overfitting e garante melhor modelo salvo automaticamente,
- **Desempenho Computacional:** Treinamento em GPU e uso de bibliotecas otimizadas (TensorFlow e Keras),
- **Interface Responsiva:** Desenvolvimento com Streamlit para fácil uso e acessibilidade em múltiplas plataformas.

8. Resultados e Discussão

O modelo desenvolvido possui uma acurácia entre 90% a 95% nos dados de teste, demonstrando alta eficiência na distinção entre as raças. A matriz de confusão evidenciou que as classes mais visualmente distintas, como Sphynx e Maine Coon, apresentaram maior precisão. Raças com características visuais mais sutis, como Persa e Siamês, apresentaram maior índice de confusão mútua.

A análise das curvas de aprendizado demonstrou convergência estável, sem sinais de overfitting graças às estratégias de regularização aplicadas. O sistema mostrou-se robusto mesmo frente a variações nas imagens de entrada, como luminosidade e angulação.

9. Conclusão

Este trabalho demonstrou a eficácia da aplicação de transferência de aprendizado com o modelo EfficientNetB0 para a tarefa de classificação de raças de gatos. O sistema desenvolvido mostrou-se preciso, eficiente e de fácil utilização por meio da interface interativa.

Além da alta acurácia, destaca-se a escalabilidade da solução, que pode ser facilmente adaptada para outras raças ou até mesmo espécies diferentes. A integração de recursos como visualizações gráficas, predições em tempo real e métricas detalhadas torna a aplicação uma ferramenta valiosa para uso educacional, veterinário ou recreativo.

