PROJECT ABSTRACT

Dog Breed Classification using Transfer Learning

A transfer-learning classifier that identifies 120 dog breeds from photos, fine-tuning a pretrained ResNet on Stanford Dogs with notebook-measured validation accuracy.

01. Title

Dog Breed Classification using Transfer Learning

02. Introduction / Background

Distinguishing 120 dog breeds is a fine-grained classification challenge: breeds differ in subtle details of coat, ears, and muzzle that defeat shallow models. Training a deep network from scratch would demand enormous data, so this project uses transfer learning: a ResNet (He et al., 2016) pretrained on ImageNet already knows edges, textures, and object parts, and the included training notebook fine-tunes it on the Stanford Dogs dataset (Khosla et al., 2011), 20,580 annotated images across 120 breeds. A demo app serves top-3 breed predictions with confidence scores for uploaded photos.

03. Problem Statement

Fine-grained breed identification needs expert knowledge most people lack, and training a 120-way deep classifier from scratch is infeasible on student-scale data and hardware. The need is a transfer-learning pipeline that reuses pretrained visual knowledge, documents its augmentation and train/validation split, and computes accuracy, loss curves, and a confusion matrix during the builder's own training run — with per-breed error analysis for the viva.

04. Objectives

The objectives are to fine-tune an ImageNet-pretrained ResNet on the Stanford Dogs dataset; to build an augmentation pipeline and a reproducible train/validation split; to log accuracy and loss curves and compute a validation confusion matrix with per-breed breakdown during training; to serve top-3 predictions with confidence scores through a demo app; and to add Grad-CAM visualizations of decision regions.

- Fine-tune pretrained ResNet on Stanford Dogs (120 breeds).
- Build augmentation (crops, flips, color jitter, rotations) and a seeded split.

- Log accuracy/loss curves; compute validation confusion matrix with per-breed breakdown.
- Serve top-3 breed predictions with confidence scores in a demo app.
- Add Grad-CAM visualizations of decision regions.

05. Proposed System / Working

Stanford Dogs images are organized by breed and split into train and validation sets with a fixed seed. Training images pass through augmentation (random resized crops, horizontal flips, color jitter); validation images are only resized and normalized. An ImageNet-pretrained ResNet is fine-tuned by first training the replaced 120-way classifier head, then unfreezing deeper layers at a low learning rate. Each epoch logs training and validation accuracy and loss, the best checkpoint is saved, and a confusion matrix is computed on the validation split with a per-breed accuracy breakdown. The demo app loads the best weights at startup; an uploaded photo is preprocessed identically to validation images and softmax scores are ranked into top-3 breed predictions with confidence bars. Grad-CAM overlays show the attended regions.

- Seeded train/validation split of Stanford Dogs images.
- Augmentation pipeline for training images.
- Two-stage fine-tuning of pretrained ResNet.
- Per-epoch accuracy/loss logging and best-checkpoint saving.
- Validation confusion matrix with per-breed breakdown.
- Demo app: upload, top-3 predictions, Grad-CAM overlays.

06. Technologies / Components Used

The project uses Python 3.10 and PyTorch for training, validation, and inference, with torchvision supplying the pretrained ResNet, transforms, and augmentation. NumPy and scikit-learn compute metrics and the confusion matrix; Matplotlib and Seaborn plot accuracy and loss curves. A Flask app serves the demo, the Stanford Dogs dataset provides training data, and trained weights ship as a .pth file exported from the included training run.

- Python 3.10, PyTorch (training, validation, inference).
- torchvision (pretrained ResNet, augmentation transforms).
- NumPy, scikit-learn (metrics, confusion matrix).
- Matplotlib, Seaborn (accuracy and loss curves).
- Flask demo web app with photo-upload UI.
- Stanford Dogs dataset (20,580 images, 120 breeds).
- Trained weights as .pth, exported from the included training run.

07. Key Features

Key features include a ResNet classifier fine-tuned on 120 breeds via transfer learning, an ImageNet-pretrained backbone with a replaceable classifier head, a training notebook logging accuracy, loss curves, and a validation confusion matrix, an augmentation pipeline, a demo app with top-3 predictions and confidence scores, Grad-CAM visualizations, and a per-breed accuracy breakdown for error analysis.

- ResNet classifier fine-tuned on 120 dog breeds.
- ImageNet-pretrained backbone, replaceable classifier head.
- Training notebook with accuracy, loss curves, confusion matrix.
- Augmentation: crops, flips, color jitter, rotations.
- Demo app with top-3 predictions and confidence scores.
- Grad-CAM visualizations of decision regions.
- Per-breed accuracy breakdown for error analysis.

08. Applications / Use Cases

Breed classification applies to pet-adoption platforms (auto-tagging listings), veterinary intake prototypes, pet-insurance quoting assistants, and educational demos of fine-grained recognition. It also teaches transfer learning, residual networks, and confusion-matrix error analysis as core deep-learning concepts.

- Pet-adoption platforms auto-tagging listings.
- Veterinary intake and shelter prototypes.
- Pet-insurance quoting assistants.
- Educational demos of fine-grained recognition.
- Teaching transfer learning and residual networks.

09. Results & Scope

The expected outcome is a trained 120-way breed classifier with a design target of ~85%+ validation accuracy, where the actual accuracy, loss curves, and confusion matrix are computed by the training notebook during the buyer's training run — the report's numbers come from that run, not from a claim. The per-breed breakdown names the confused breed pairs explicitly, which doubles as viva error-analysis material. Scope covers the 120 Stanford Dogs breeds on well-framed photos; documented limits include occluded, multi-dog, and low-light photos. Future scope includes larger breed sets, mixed-breed scoring, and mobile deployment.

- Design target ~85%+ validation accuracy; actual metrics computed by the notebook during the buyer's training run.
- Per-breed breakdown naming confused breed pairs.
- Scope: 120 Stanford Dogs breeds, well-framed photos.
- Documented limits: occluded, multi-dog, low-light photos.
- Future scope: more breeds, mixed-breed scoring, mobile deployment.

10. Conclusion

Transfer learning turns 120-way breed classification from a data-impossible problem into a tractable fine-tuning task, and this project delivers it as a complete auditable experiment: dataset pipeline, two-stage ResNet fine-tuning, logged metrics, per-breed confusion analysis, Grad-CAM interpretability, and a demo app. Built to order, it ships with source code, trained weights, report, presentation, and viva preparation material.

11. References

The project cites the ResNet architecture, the Stanford Dogs dataset, and the interpretability method used.

- K. He, X. Zhang, S. Ren, and J. Sun, 'Deep Residual Learning for Image Recognition,' Proc. CVPR, 2016.
- A. Khosla, N. Jayadevaprakash, B. Yao, and L. Fei-Fei, 'Novel Dataset for Fine-Grained Image Categorization: Stanford Dogs,' Proc. CVPR Workshop on Fine-Grained Visual Categorization, 2011.
- R. R. Selvaraju et al., 'Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,' Proc. ICCV, 2017.