Mini Project 001:

Cải thiện độ chính xác của bộ dataset ImageWoof

# 1. Mô tả bài toán

## Ngữ cảnh / bài toán:

Image classification là một trong những bài toán cơ bản và quan trọng của deep learning và computer vision. Những giải thuật, mô hình deep learning có kết quả tốt cho các bài toán image classification gần như sẽ đều được đem ứng dụng vào các bài toán khác trong computer vision: ResNet, EfficientNet, .... ImageNet là một bộ dataset standard để test các giải thuật và mô hình mới cho image classification. Tuy nhiên, do kích cỡ tương đối lớn của bộ dataset này, không phải ai cũng tiếp cận được nguồn computing đủ lớn để có thể thử nghiệm giải thuật và phát triển mô hình dùng ImageNet.

Để hỗ trợ việc thử nghiệm các ý tưởng nhanh hơn, chuyên gia Jeremy Howard (former president and chief scientist of Kaggle) đã phát triển 3 bộ dataset nhỏ Imagenette, Imagewoof và Imagewang với độ khó tăng dần với hy vọng là, những giải thuật, mô hình chạy tốt trên những bộ này có thể cũng sẽ chạy tốt khi chạy trên bộ dataset ImageNet.

## Task cụ thể cần thực hiện:

Tìm cách tăng độ chính xác của mô hình trên bộ Imagewoof. Bộ Imagewoof tương đối khó, còn bộ ImageWang tuy khó hơn nhưng thiên về test các giải thuật self-supervised learning. Chỉ có thể dùng bộ train của ImageWoof để optimize mô hình. Bộ validation của Imagewoof sẽ dùng để report độ chính xác của mô hình.

Độ chính xác của mô hình state-of-the-art (SOTA) cho bộ này theo anh biết thì đang thuộc về các tác giả của bài báo [“Data Efficient Stagewise Knowledge Distillation”](https://github.com/IvLabs/stagewise-knowledge-distillation)

với kết quả là 93.4% cho ResNet26.

## Metric đánh giá:

Độ chính xác = (Number of correct predictions) / (Total number of predictions)

Ngoài ra, chương trình cũng cần theo một số yêu cầu liệt kê trên [trang của Imagewoof](https://github.com/fastai/imagenette)

Cuối cùng, điểm đánh giá cũng sẽ phụ thuộc vào the size of the model (hai mô hình có độ chính xác tương đương nhau, mô hình nhỏ hơn sẽ được điểm cao hơn) và sự mới mẻ (novelty) trong các phương pháp dùng để tăng độ chính xác của mô hình.

# 2. Mô tả dữ liệu

Imagewoof là một tập con của ImageNet, gồm 10 lớp (classes). 10 lớp này là 10 giống chó khác nhau: Australian terrier, Border terrier, Samoyed, Beagle, Shih-Tzu, English foxhound, Rhodesian ridgeback, Dingo, Golden retriever, Old English sheepdog.

Chúng ta có thể download bộ Imagewoof với 3 kích cỡ khác nhau: ảnh full size, 320x320 và 160x160. Chúng ta có thể dùng cả 3 kích cớ này để training và testing.

Với mỗi kích cỡ, bộ Imagewoof bao gồm 2 thư mục train và val dùng để training và validating mô hình. Hai thư mục này bao gồm các thư mục con, tương ứng với 10 classes. Ngoài ra, còn có file “noisy\_imagewoof.csv” chứa label của bộ dataset. File label này gồm 7 cột. Mỗi dòng tương ứng với một ảnh trong bộ này. 7 cột này là:

1. path: đường dẫn tới ảnh
2. noisy\_labels\_0: label của ảnh đó
3. noisy\_labels\_1, noisy\_labels\_5, noisy\_labels\_25, noisy\_labels\_50: cột labels của ảnh đó nhưng labels được randomly changed to an incorrect label theo probability là 1%, 5%, 25% và 50% respectively.
4. is\_valid: ảnh đó thuộc về bộ val hay bộ train

# 3. Yêu cầu đầu ra:

1. Code + và chương trình để chạy training và testing
2. 1 báo cáo (từ 4 trang trở lên) theo mẫu của bài báo khoa học. Một ví dụ chính là bài báo [“Data Efficient Stagewise Knowledge Distillation”](https://arxiv.org/abs/1911.06786)

# Tài liệu tham khảo

1. Data Efficient Stagewise Knowledge Distillation
   1. <https://arxiv.org/abs/1911.06786>
   2. https://github.com/IvLabs/stagewise-knowledge-distillation
2. Sách “Deep Learning for Coders with fastai and PyTorch” viết bởi Jeremy Howard và Sylvain Gugger: https://github.com/fastai/fastbook
3. Sách “Deep Learning with PyTorch”: <https://pytorch.org/assets/deep-learning/Deep-Learning-with-PyTorch.pdf>
4. Sách “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow” 2nd Edition
5. Và nhiều nguồn tham khảo khác trên Internet như paperswithcode, semanticscholar.