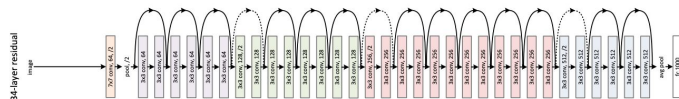# Breaking the CIFAR-10

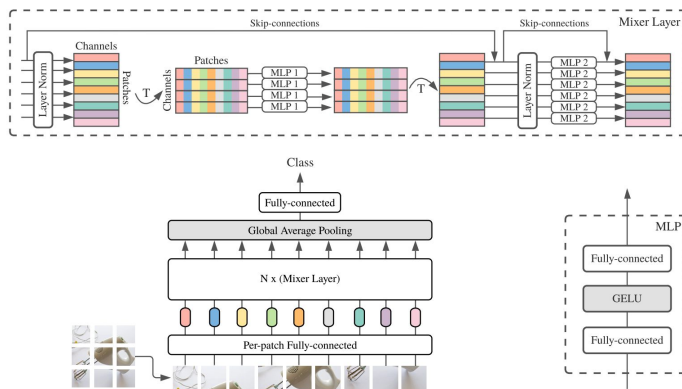Guglielmo Camporese

Models

ResNets

Visual Transformer (ViT)

MLP-Mixer

Results

Results

	ResNet-18 [1]	ResNet-34 [1]	ResNet-50 [1]	ResNets Ens [1]	MLP-Mixer [2]	ViT-S/16 [3]	ViT-B/16 [3]
Accuracy	95.91 %	96.92 %	95.46 %	97.53 %	94.67 %	96.05 %	<u>98.67 %</u>

Training Methods

- Train/Val **stratified split** over classes 0.8 train / 0.2 val → for maintaining the same class distribution,
- **Augmentation**: Resize(256), RandomCrop(224), RandomHorizontalFlip()
- **Finetuning** of a pretrained model from a large dataset (→ ImageNet) ,
- Model **ensemble**,
- **SGD Optimizer:** lr = 0.01, momentum = 0.9 → (SGD sometimes generalizes better than Adam & friends...),
- **Regularization**: dropout = 0.2, L2 regulariz. weight decay = 5e-4
- **LR scheduling**: cosine annealing with linear warmup,
- **Label smoothing** = 0.1,
- **Early stopping** (keep the model checkpoint that maximized the validation accuracy during the learning process).

→ Code for reproducing the results: https://github.com/guglielmocamporese/break_cifar10

→ A Recipe for Training Neural Networks: http://karpathy.github.io/2019/04/25/recipe/