Image Classification with ImageNette

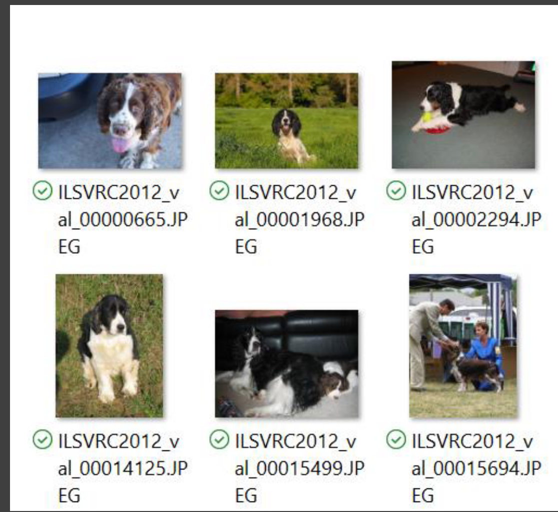
Caspar Chan (z5206252), Fenyi Shen (z5323706), Henry Ho (z5312521), Rachael Yu (z5164297), William Wong (z5205428)

Problem Statement

- Image classification aiming high accuracy
  - ImageNette
- Evaluate effectiveness of different approaches

Data Sources

- From Imagenette (Not imagenet)
- Easy to access
- Sufficiently non-trivial
- Size is not too big
- Project limitations



Data Exploration

- Similar number of images across classes (858 to 993)
- Images have different sizes (between 200 and 500 pixels)
- Premade training and validation sets (9469 training, 3925 validation)
- Slight noisy labelling (~5%)

Data pre-processing

Image Transformations

- Resize
- RandomCrop
- RandomHorizontalFlip
- RandomErasing
- ToTensor
- Normalize

Methods

Approach

- Tested performance of popular networks used for image classification such as GoogLeNet and various versions of ResNet (ResNet34, ResNet50, ResNet50-d)
- ResNet architecture displayed overall better performance
- Modifications were made based on ResNet's architecture incorporating features from other networks and alternative components proposed by recent research papers
- Final model is a composition of all adjustments that demonstrated increase in peak accuracy or general performance

Proposal for new block

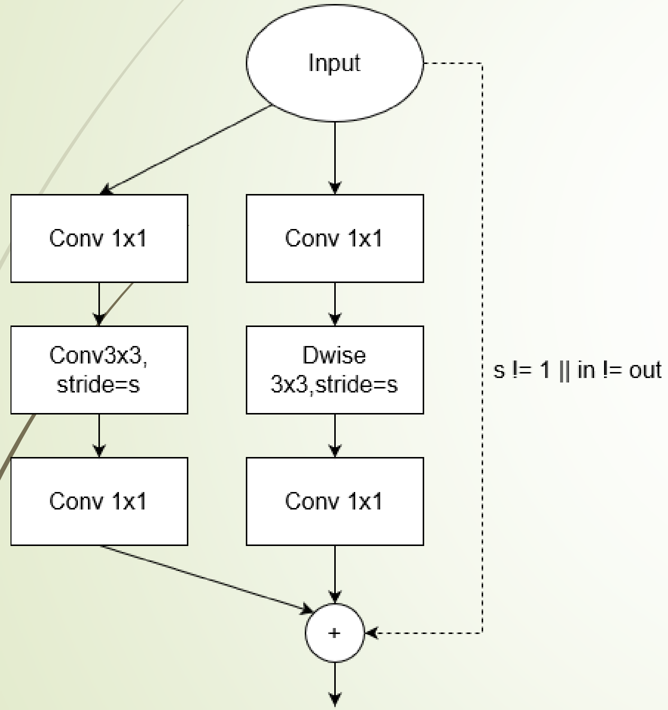
- How would we gather more information?
  - By performing multiple passes on the same input!

Problems

- Computationally more expensive
  - *n* times more expensive for each block added
- High contrast blocks may pick up completely different things
  - Might not pick up any new information at all

Model Architecture

Block Structure



Main Structure of our Model

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	

Experimentation

- Mish -> Faster convergence
- MaxBlurPool -> Decrease in accuracy
- Dropout -> Drastic increase in training time without showing significant effect in reducing overfitting or increasing accuracy
- Adam optimiser -> Major increase in learning speed compared to SGD
- Learning rate scheduler -> Significant improvement in learning speed up to 80% accuracy with Cosine annealing scheduler demonstrating better performance than Exponential scheduler

Alternative Models Tested

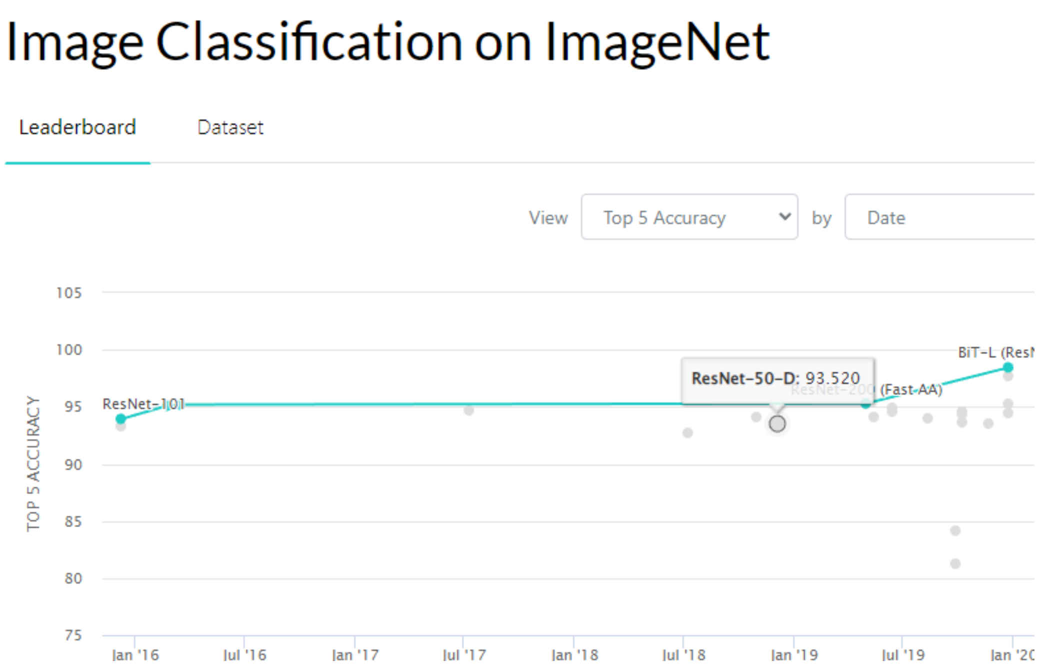
- Current model with Dropout
- ResNet with Dropout
- ResNet34
- ResNet50
- ResNet50-d
- GoogLeNet

Results

Merging features from different models can bring more insights on new model

Class Accuracy	Modified MobileNetV2	Raw Resnet50
Tench	94.8 %	93.0 %
English Springer	93.9 %	95.9 %
Cassette Player	91.9 %	81.2 %
Chain Saw	80.8 %	74.4 %
Church	90.7 %	90.2 %
French Horn	88.1 %	91.9 %
Garbage Truck	90.7 %	89.5 %
Gas Pump	85.0 %	77.8 %
Golf Ball	91.7 %	84.7 %
Parachute	90.0 %	89.7 %
Overall	90%	87%

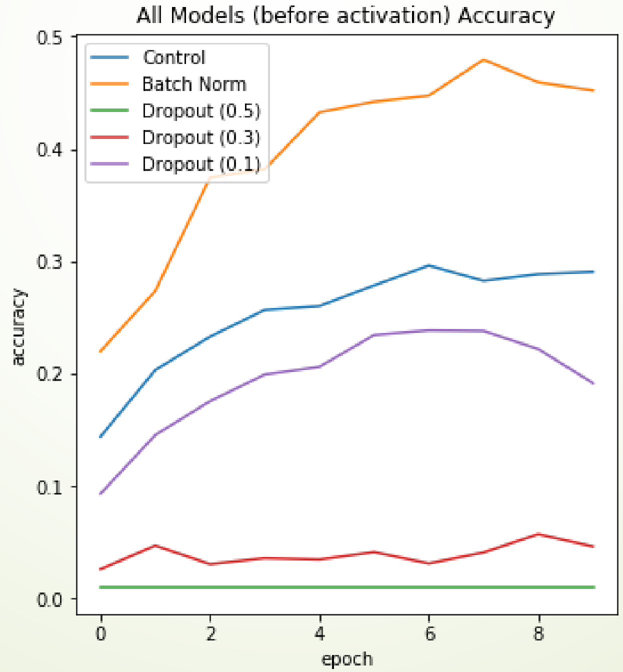
The size and purity of the training dataset can have an impact on the accuracy

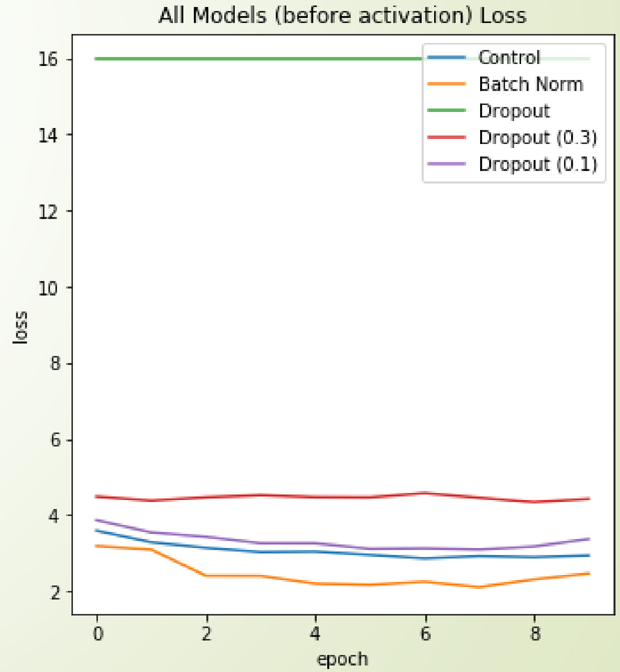


- Raw Resnet50 only achieves an 87 % accuracy
- Insufficient dataset size
- 5% noisy data
  - Increases the model complexity
  - Increases time of learning

Dropout layers do not necessarily improve the model performance

- Batch normalization already has a regularizing effect
- Partial dropout can be harmful in CNN





Fine-tuning hyperparameters is essential to achieve a high accuracy

- ResNet50 with a consistent 0.001 learning rate
  - Barely reach 80%

- CosineAnnealingLR scheduler
  - Improve to 87%

Mish activation and BlurPool

- Activation change: ReLU6 to Mish
  - Improved training speed, faster in reaching 80%
- Blurpool
  - Accuracy decreased

Limitations

- GPU limitation
- Training process is time consuming
- Time constraint

Conclusion

- Merging features from different models can bring insights on new model architecture
- Effectiveness of different techniques and approaches in model construction and training
  - Mish
  - BlurPool
  - Scheduler
  - Dropout