AIMI LAB1 Pneumonia Classification

1. Introduction

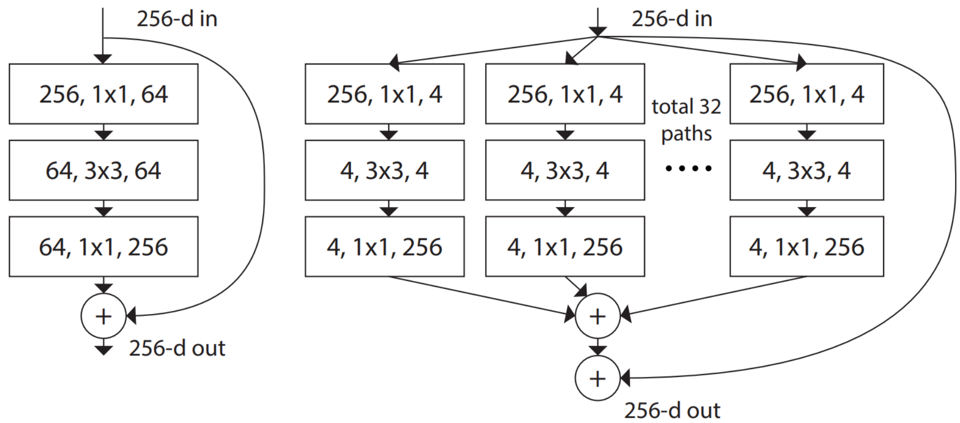
- Pneumonia Introduction:
  肺炎是一種感染，會導致一個或兩個肺部的氣囊發炎。氣囊可能充滿液體或膿液（化膿性物質），引起帶痰或膿液的咳嗽、發燒、寒戰和呼吸困難。包括細菌、病毒和真菌在內的多種生物體可引起肺炎。
  肺炎的嚴重程度可以從輕微到危及生命。對於嬰幼兒、65歲以上的人以及有健康問題或免疫系統較弱的人來說，這種情況最為嚴重。

- Introduction to Pneumonia Dataset:
  他的數據集被組織成 3 個文件夾（train、test、val）並包含每個圖像類別（Pneumonia/Normal）的子文件夾。有 5,863 張 X-ray 圖像和 2 個類別（肺炎/正常）。

2. Experiment setups

2.a Model architechure

- resnext50
  ResNeXt的架構類似於 ResNet，但不一樣的地方在於 ResNeXt 把輸入分成許多相同寬度的卷積層，之後再合起來。如下圖所示，左邊為一般 ResNet 的 bottleneck block，右邊是 ResNeXt 的 block 架構。分支的數量稱為 cardinality。



  下表為 ResNeXt50 與 ResNet50 的模型結構，可以看到兩者模型複雜度相近。(其中 C 為 cardinality)

stage	output	ResNet-50	ResNeXt-50 (32×4d)
conv1	112×112	7×7, 64, stride 2	7×7, 64, stride 2
conv2	56×56	3×3 max pool, stride 2	3×3 max pool, stride 2
		$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128, C{=}32 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3	28×28	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256, C{=}32 \\ 1\times1, 512 \end{bmatrix}\times4$
conv4	14×14	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512, C{=}32 \\ 1\times1, 1024 \end{bmatrix}\times6$
conv5	7×7	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 1024 \\ 3\times3, 1024, C{=}32 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	global average pool 1000-d fc, softmax	global average pool 1000-d fc, softmax
# params.		**25.5**$\times10^6$	**25.0**$\times10^6$
FLOPs		**4.1**$\times10^9$	**4.2**$\times10^9$

Table 1. (**Left**) ResNet-50. (**Right**) ResNeXt-50 with a 32×4d template (using the reformulation in Fig. 3(c)). Inside the brackets are the shape of a residual block, and outside the brackets is the number of stacked blocks on a stage. "C=32" suggests grouped convolutions [24] with 32 groups. *The numbers of parameters and FLOPs are similar between these two models.*

可以看出在 ImageNet-1K 資料集 ResNeXt 的效果 比 ResNet 好。

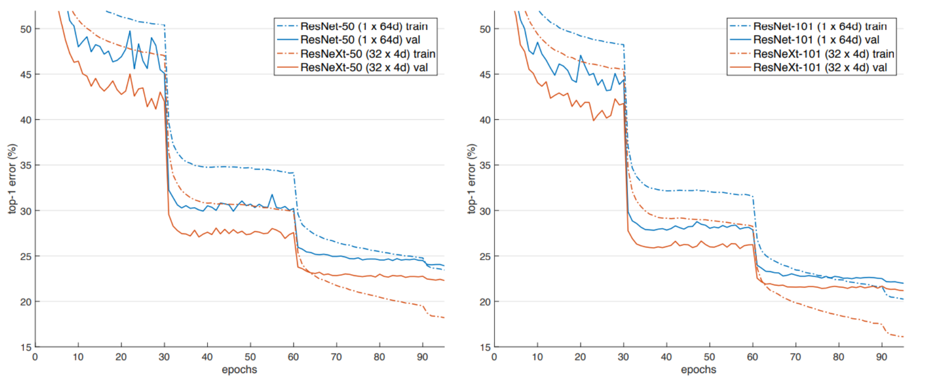


Figure 5. Training curves on ImageNet-1K. (**Left**): ResNet/ResNeXt-50 with preserved complexity (~4.1 billion FLOPs, ~25 million parameters); (**Right**): ResNet/ResNeXt-101 with preserved complexity (~7.8 billion FLOPs, ~44 million parameters).

2.b Dataloader

- ImageFolder: 用來讀取資料夾內的data與label，默認數據集已經自覺按照要分配的類型分成了不同的文件夾，一種類型的文件夾下面只存放一種類型的圖片
- Transform (training phase): 每次取出data時，先將圖片分別做
  - resize to 128 * 128
  - RandomEqualize
  - RandomRotation: (-25, 20)

    - CenterCrop: 64 * 64
    - Normalize: ([0.485, 0.456, 0.406],[0.229, 0.224, 0.225])
- Transform (testing phase): 每次取出data時，先將圖片分別做
    - resize to 128 * 128
    - Normalize: ([0.485, 0.456, 0.406],[0.229, 0.224, 0.225])

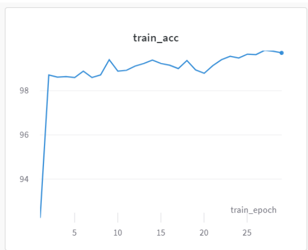
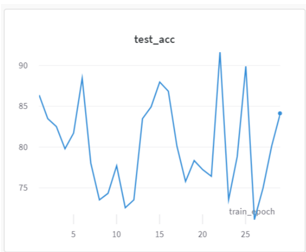
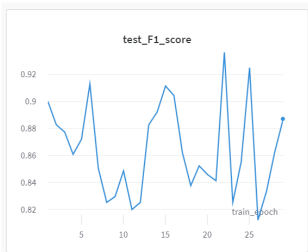
3. Experiment result

3.a

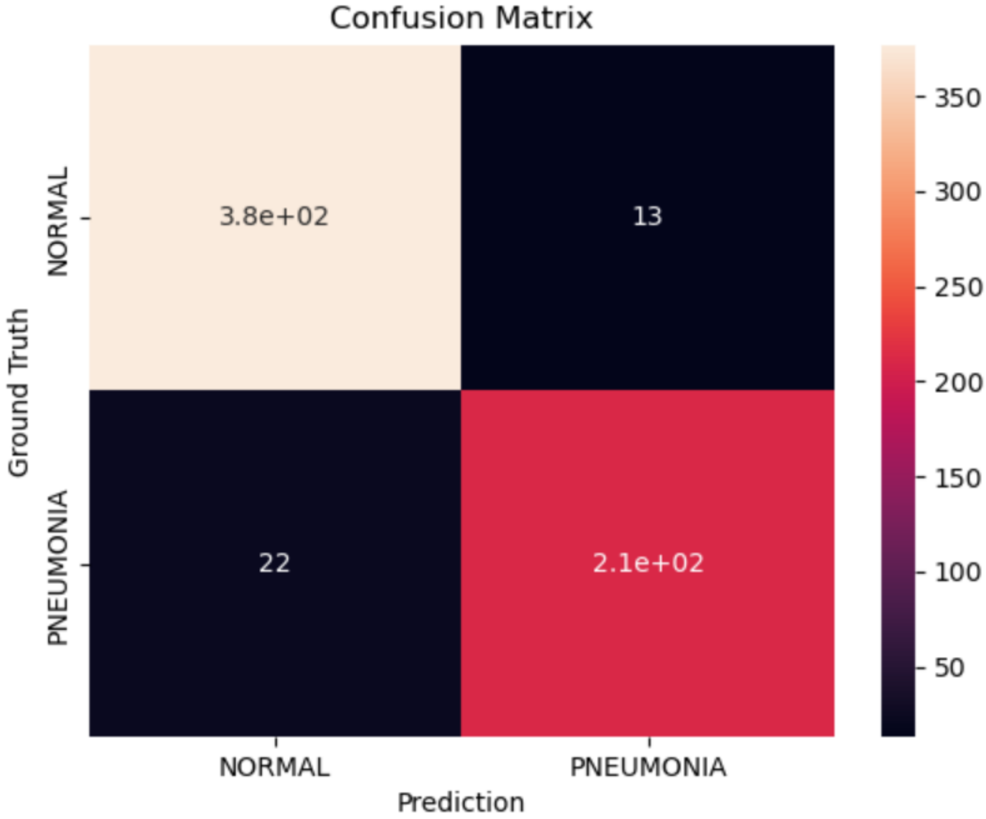
- Accuracy: 94.39%
- F1-score: 0.9556

3.b

We can see that training accuracy gets higher with respect to epoch. The value of testing accuracy and testing F1-score is quite similar. Therefore, their plots are also similary.

Training accuracy	Testing accuracy	Testing F1-score
		

- Test Confusion matrix:



3.c Results of different models

- We also try different models, using same hyperparameter. resnext50 has the best performance, and efficientnet-b3 is the worst.

Model	Accuracy
resnext50	94.39%
densenet161	91.67%
resnet18	87.34%
resnet50	83.65%
efficientnet-b3	79.33%

Discussion

Comparison of different model

We find that resnext50 performs best. On the contrary, efficientnet-b3 performs worst. However, efficientnet-b3 actually performs better than resnext50 on imagenet dataset. It is quite interesting, so we find a paper related to this problem. From Alexander[1], they proposed that there is no relationship between ImageNet performance and CheXpert performance. They suggests that the choice of model family influences performance more than size within a family for medical imaging tasks. They also find that ImageNet pretraining models yields a statistically significant boost in performance. In our experiments, the results are also same as the arguments they made. Model selection drastically influence the performance of this task.

Github link

https://github.com/eritup45/AIMI_Lab1_pneumonia_classification.git (https://github.com/eritup45/AIMI_Lab1_pneumonia_classification.git)

Refernence

- **[1] CheXtransfer: Performance and Parameter Efficiency of ImageNet Models for Chest X-Ray Interpretation** (https://arxiv.org/pdf/2101.06871.pdf)