

Machine Learning Final Report

Small Data Training for Medical Images

NTU_b05901009_易昀競選總部

b05901009 高瑋聰, b05901022 許睿洋, b05901034 劉奎元

Abstract

醫療資料充足及各種電腦視覺技術蓬勃發展的今日，衍生出運用機器學習技術輔助醫療辨識的各式方向。其中在胸腔疾病的判別方面，CheXNet(2017)曾經做出能達成相當成功的結果。然而其訓練運用了大量且充分的醫療資料，在資料匱乏的狀況下可能有所缺失。本實驗以CheXNet為基礎進行變化，運用ImageNet pretrained model以及Auto-Encoder一同訓練的方式，最終達成AUROC=0.79998的實驗成果，大幅縮減少量與巨量資料進行訓練的差距。

1. Introduction & Motivation

本次Project為利用 NIH Chest X-ray dataset進行疾病預測。在資料方面，有約10000筆labeled data和約70000筆unlabeled data。此dataset的input為胸腔的X光片，期望輸出該X光片中代表患有14種疾病之中的哪幾種，同時也限制只能用約整個dataset中1/10之一(10000筆)labeled data做訓練。當初會選擇這個題目的原因有三個，第一是因為想要挑戰當訓練資料很有限時該如何做出一個好的model，第二是覺得這個題目可能可以使用的技巧很多變化也大，希望可以從中找出一個最有用的技巧。第三是因為比較少有機會接觸到醫療相關的dataset，想藉著這個機會嘗試一下醫療相關的data以及應用。

2. Related Works

在過去電腦視覺的領域，已經有許多研究提出不同的模型，特別是從2012年AlexNet在ImageNet Large Scale Visual Recognition Competition (ILSVRC)的良好表現開始，基於神經網路的電腦視覺方法越來越受到關注。隨著GPU運算速度的提升，神經網路的深度也能隨著增加，例如VGG19等，但隨著深度越來越深，model的訓練也慢慢出現許多問題，例如梯度消失 (gradient vanishing)等等，所幸有了ResNet的出現，訓練加深後的神經網路的方法獲得了解答。其後也有許多模型繼承ResNet的想法，發展出表現更為提升的模型，例如本次final project使用的DenseNet等等。

此外，針對NIH Chest X-ray dataset，也有過去的研究實作過深度學習的方法，本次final project也將參考其想法，利用DenseNet來進行本次的supervised learning。

在unsupervised learning的領域，Deep Auto-Encoder是一個典型的方法。透過數量較多的unlabeled data，經過encode-decode的兩個階段，期望model能夠還原回原本的輸入，並在其中維度較低的一層中抽取壓縮後的特徵。Auto-encoder原來並非容易訓練的模型，可能需要一些pretrain來初始化(例如Restricted Boltzmann Machine, RBM)，但由於Batch Normalization的新發現，在此次final project中，我們沒有使用RBM initialization仍能夠將auto-encoder訓練起來。

3. Data Preprocessing

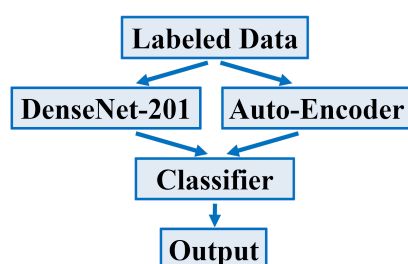
由於原本的圖片大小長寬皆為1024，對於我們使用的model來說非常大，且若要使用ImageNet pretrained model，輸入必須長寬皆為224。我們最後採取直接將輸入的圖片隨機擷取一個長寬皆為224的區塊，並做隨機的水平和翻轉來當作Data augmentation以避免overfitting。不同的輸入圖片大小將在Experiment中討論。

4. Model Description

本次使用的model主要分為兩部分，supervised部分採用DenseNet結構，即model由多個dense block組成，一個dense block內有數層CNN，但較接近輸入圖片的CNN層的輸出會再加到下一層的輸入去，由於層與層之間的連結較ResNet緊密，因此稱為DenseNet。同時為了避免channel增加太快，DenseNet內部還有使用1x1 convolution來控制channel的數量，同時也做channel間的混合。使用ImageNet pretrain的model可以使我們的model在較少的時間內達到較高的準確率。

此外根據此次提供的unlabeled data，我們用來訓練一個Auto-Encoder (架構參照附錄)，訓練(至MSE loss約為0.16)後將其中的encoder加入到model中，因此輸入圖片會分別進入DenseNet與encoder，最後將兩個特徵合併放入最後的classifier得到結果。

圖1.model架構圖



5. Experiment and Discussion

以CheXNet論文的實驗結果為基礎，我們進行了幾種實驗方式以獲得較好的結果:不同的pretrained model、不同的data augmentation、不同的圖片大小、不同的optimizer與regularization、加入unsupervised model一同訓練、使用ensemble。

5.1 Pretrained model

針對少量資料的訓練，我們嘗試過幾種較小的ImageNet pretrained model例如:VGG-16、VGG-19來進行訓練，然而這些小型的模型皆無法成功訓練。因此，我們改採CheXNet論文中使用的DenseNet-121進行訓練，結果可以成功訓練。在進行optimizer與regularization的微調後(參照5.4)，可以達成更好的成果。在搭配unsupervised model(參照5.5)的狀況下，使用DenseNet-201則可以獲得最佳的單一模型訓練結果。

表格1. 使用不同Pretrained model (單一model)的AUROC變化。

Pretrained Model	AUROC
VGG-16	0.5
VGG-19	0.5
DenseNet-121	0.667
DenseNet-121 (with better optimizer and regularization)	0.740
DenseNet-201 (with Auto-Encoder)	0.771

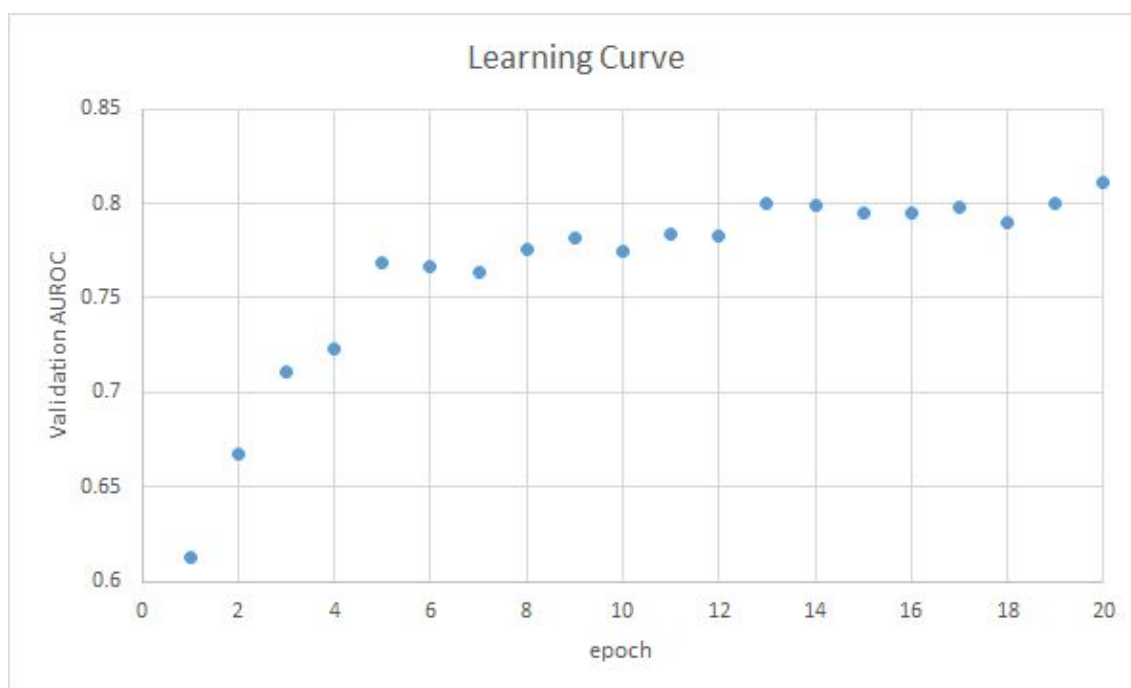


圖2. 使用ImageNet pretrained的DenseNet-201搭配本文中的Optimizer及Regularization及Auto-Encoder一同訓練時，單一次訓練的學習曲線。

5.2 Data Augmentation

我們嘗試過幾種傳統的data augmentation方式:rotation、shear、colorjitter、RandomResizeCrop、RandomHorizontalFlip。使用不同程度的rotation、shear、colorjitter皆無法造成更佳的表现結果，因此在最後的模型中並未使用。RandomResizeCrop與RandomHorizontalFlip則皆為CheXNet論文中使用的data augmentation方式，前者根據輸入圖片大小的改變(參照5.3)而選擇使用；後者則是與DenseNet-121搭配使用，可以微幅提升訓練效果。

5.3 Input size

我們嘗試過將輸入的圖片大小維持原本大小(1024x1024)、縮小成512x512、256x256或值皆能放入ImageNet pretrained model的224x224，原先預期若一口氣將圖片縮到太小，可能會丢失許多資訊，但若不是224x224的圖片

則須在model前增加幾層convolution以讓輸入符合大小。就實驗的結果而言，這四個尺寸的輸入對最後model的表现影響不大，但是輸入越大張運算的時間會大幅提升，model收斂的時間也會變久，因此最後決定採用直接縮小到224x224的做法。

5.4 Optimizer and Regularization

我們以Adam及其pytorch預設參數為基礎，進行learning rate及regularization的調整。Learning rate的部分，我們從CheXNet論文提供的 $\alpha=1e-3$ 調整為 $\alpha=1e-4$ ，以及L2 regularization($\lambda=1e-5$)，可以獲得相當巨大的進步幅度(參照表格1)。由這些參數所進行的實驗結果，可以獲得相當平穩的訓練過程。

5.5 Unsupervised model

在這項實驗中，我們運用MSE loss訓練至0.16的Auto-Encoder與supervised model一同進行訓練。在配合Auto-Encoder的情況下，DenseNet-121的訓練能將AUROC提升0.007，與ensemble方法(參照5.6)一同使用後能夠獲得相當大幅的進步。但使用較大的DenseNet-201時，Auto-Encoder的提升效果則不顯著。

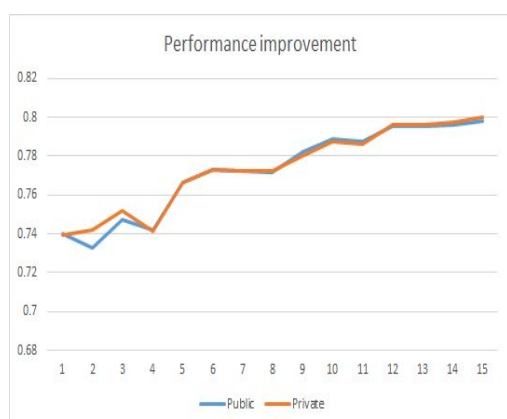
5.6 Ensemble

我們使用bagging取平均的方式來進行ensemble的實踐。在少量model的狀態下，可以獲得巨大的進步。在堆疊到十筆以上的model後，雖然進步幅度逐漸收斂，仍然能夠在每增加一筆model的狀況下AUROC進步0.001。

在堆疊六筆model後，也曾經嘗試進行將單一成果最差的model去除以進行觀察，其結果不如沒去除的狀況；嘗試手動增加單一成果最佳的model之權重，其結果亦是下降。

透過ensemble的方式，我們獲得了最終最佳的結果。

圖3. Ensemble後的進步曲線



6. Conclusion

本次final project利用醫療相關的資料，並同時模擬了當擁有的labeled

data量不多的情況。我們同時運用了supervised learning與unsupervised learning的技術，並運用了transfer learning與fine tuning的技巧來加速訓練。結合以上的方法與參數調整，使我們在public與private的kaggle排名皆為本次的第一名。

7. Future Work

本次提供的data有約10000筆labeled data且開放使用ImageNet pretrained model，因此在supervised的部份已經有相當程度的發揮，但在unsupervised learning這一部份，礙於輸入的圖片size太大以至於記憶體不足，使得我們沒能完成如smoothness assumption中的graph based approach，也礙於開發時間有限與training難度，我們來不及完成以Generative Adversarial Network (GAN)為基礎的Data Augmentation方法，例如:Data Augmented GAN來增加比傳統方法更合適的augmentation，或是利用Augmented CycleGAN來訓練一組GAN，輸入任意的label以產生對應的圖片。若在此次final project中新增更多利用unlabeled data的方法，應可更進一步提昇準確率。

在未來的應用上，我們可以運用此次訓練model的技巧，套用到其他醫學相關疾病的判斷，或者其他領域，尤其是針對資料不易取得或不易用人工進行label的領域如智慧製造的品質檢測等等。對於本次project的成果，也能更進一步從domain knowledge出發，分析利用深度學習方法與傳統人工方法的異同，藉此輔助與加快肺部相關疾病的診斷速度與準確率。

7. Reference

Dietterich, Thomas G. "Ensemble methods in machine learning." *International workshop on multiple classifier systems*. Springer, Berlin, Heidelberg, 2000.

Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. Ieee, 2009.

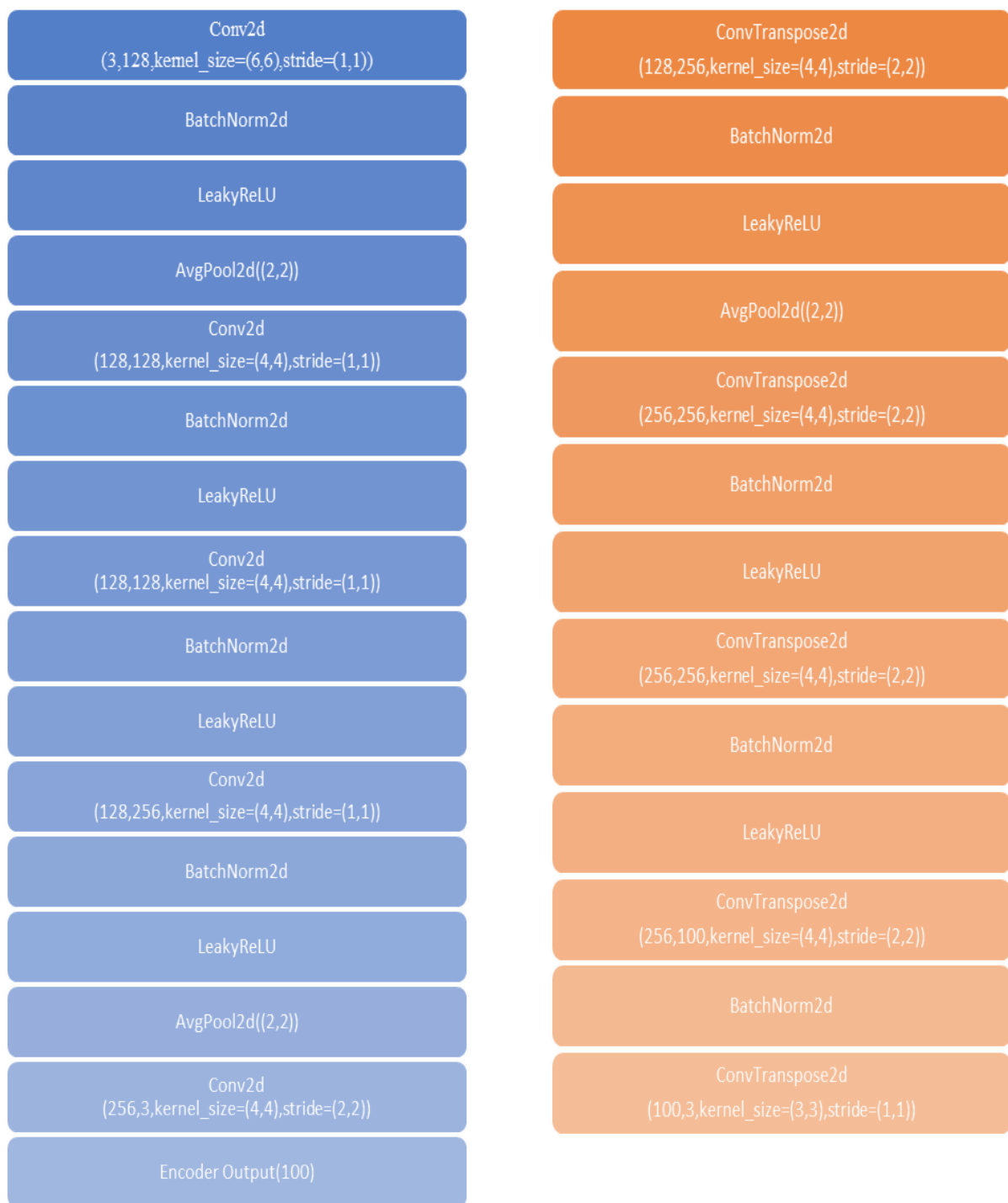
He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *arXiv preprint arXiv:1502.03167* (2015).

Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).

Rajpurkar, Pranav, et al. "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning." *arXiv preprint arXiv:1711.05225* (2017).



附錄1. Auto-Encoder架構
(左:Encoder 右:Decoder)