

結構化機器學習模型及其應用

Structural Machine Learning Models and Their Applications

陳奕傑
大數據產學研發班
國立中興大學

A. Self-training with Noisy Student improves ImageNet classification

- 1) train a teacher model on labeled images
- 2) use the teacher to generate pseudo labels on unlabeled images
- 3) train a student model on the combination of labeled images and pseudo labeled images.

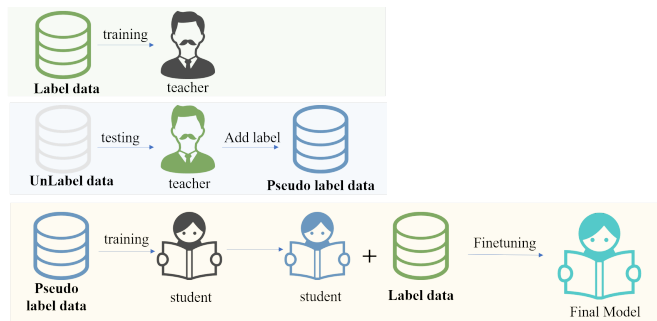


Fig. 2. Billion-scale 流程概念

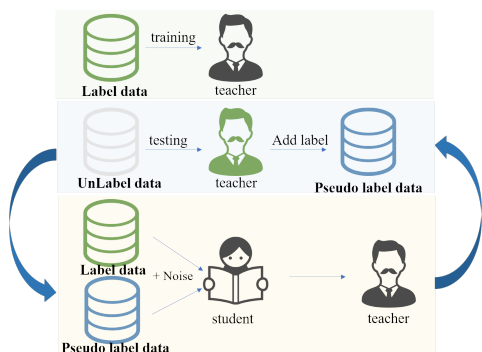


Fig. 1. Noisy student 流程概念

B. Billion-scale semi-supervised learning for image classification

- 1) train a teacher model on labeled images.
- 2) use the teacher to generate pseudo labels on unlabeled images.
- 3) train a student model on the pseudo labeled images. (Select Top K for each classification)
- 4) Finetune student model on the labeled images.

C. Dataset

CIFAR-10

- image size $32 * 32 * 3$
- 這些影像涵蓋了 10 個類別：飛機，汽車，鳥，貓，鹿，狗，青蛙，馬，船以及卡車。資料集總共包含 60000 張圖片，每個類別 6000 張。其中，訓練集包含 50000 張圖片，測試集包含 10000 張圖片

CINIC-10 is an augmented extension of CIFAR-10.

- unlabeled data total numbers: 70,000 (for to be pseudo label data and training)
- Label data total numbers: 500 (for testing)。

	Training	Validation	Test
Teacher	CIFAR 10 Dataset Total: 50,000	CIFAR 10 Dataset Total: 10,000	CINIC Dataset Total: 500
Student 1	CIFAR10 Dataset + Pseudo label Total: 120,000	CIFAR 10 Dataset Total: 10,000	CINIC Dataset Total: 500
Student 2	CIFAR10 Dataset + Pseudo label Total: 120,000	CIFAR 10 Dataset Total: 10,000	CINIC Dataset Total: 500
Student 3	CIFAR10 Dataset + Pseudo label Total: 120,000	CIFAR 10 Dataset Total: 10,000	CINIC Dataset Total: 500

Fig. 3. Noisy student dataset

Noisy student 的部分 Batch size = 64, Image size: $32 * 32 * 3$ 。Pseudo label data 為 CINIC-10 資料集排除跟 CIFAR 10 相同的照片後，進行偽標籤後取 Confidence Score > 0.3 並每類依照 Confidence Score 由高至低排序取前 7000 筆資料，若資料數不夠就重複選取現有的資料。

	Training	Validation
Teacher	CIFAR 10 Dataset Total:50,000	CIFAR 10 Dataset Total:10,000
Student 1 pretrained	Pseudo label Total:40,000 Select K=4,000	CIFAR 10 Dataset Total:10,000
Student 1 finetuning	CIFAR 10 Dataset Total:50,000	CIFAR 10 Dataset Total:10,000

Fig. 4. Billion-scale dataset

Billion-scale dataset 一樣取 CIFAR-10 以及 CINIC-10 進行模型的訓練。

D. 訓練細項

Noisy student:

- Base model: MobileNet v1
- Noise:
- data augmentation with RandAugment
- Stochastic depth (use in residual block) 不使用
- Dropout rate = 0.1

- Batch size: 64
- Epoch:50

Billion scale:

- Base model: MobileNet v1
- K=4,000
- Batch size: 64
- Epoch:50

E. 訓練結果

可以明顯發現在 validation 部分，Student model 會比 Teacher model 來的更好。在 testing 的部分也是如此。代表這樣的手法確實有效。

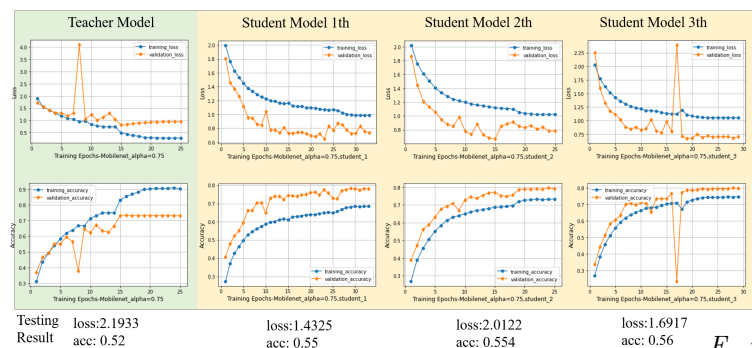


Fig. 5. Noisy student 結果

Fig6. 此部分為比較是否有加 Noise 而有不同的結果，在這裡取第三次疊代的 Student model。可以發現有 image noise 的 model, 其 validation 表現會比 training 來得好，主要是在 training 有加入 noise 導致 training 的圖片在每一次的訓練中有可能是不太一樣的。在單純增加 pseudo data 而沒有加 Noise 時，其結果的表現也不差。

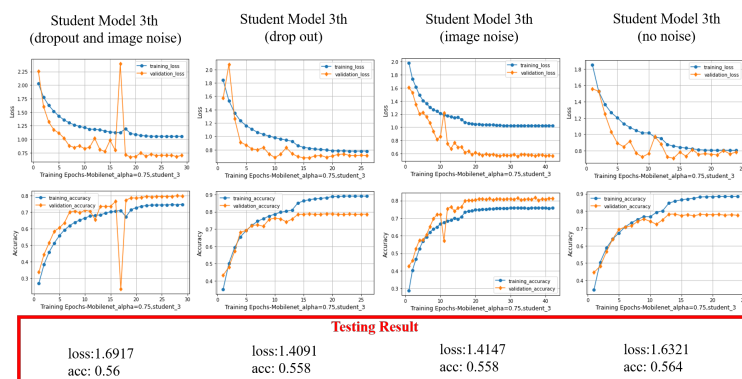


Fig. 6. Noisy 比較

Fig7. Billion-scale semi-supervised learning for image classification 的部分設 K=4,000，因在進行偽標籤時，最少類別的個數為 4000 多。並在此 paper 也無進行照片增量的動作。故只取 K=4000，最後訓練結果也沒有比較好。也可能是因 unlabeled 的數量不太夠。

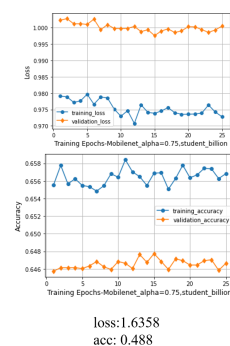


Fig. 7. Billion scale 結果

F. 結論

實驗後的結果，可以發現 Noisy student 的部分是有比較好一點點。Billion 的部分則沒有，覺得可能的問題為 unlabel data 沒有那麼多。在 paper 中 unlabel data 的數量其實是非常多的，以及模型的部分選擇是較輕量的 MobileNet v1，其圖片分類的效果比較 EfficientNet 或是 ResNet 都還是遜色許多。故可能在偽標籤的時候，可能就會有比較大的誤差。導致訓練 data 其實是不太好的，訓練 data 不太好的情況下，其結果可能也比較差強人意。以上為目前覺得實驗後不如預期的兩種原因。

REFERENCES

- [1] BAHULEYAN, Hareesh. Music genre classification using machine learning techniques. arXiv preprint arXiv:1804.01149, 2018.
- [2] TANG, Chun Pui, et al. Music genre classification using a hierarchical long short term memory (LSTM) model. In: Third International Workshop on Pattern Recognition. International Society for Optics and Photonics, 2018. p. 108281B.
- [3] SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.