



Final Project

組別:名字好難想

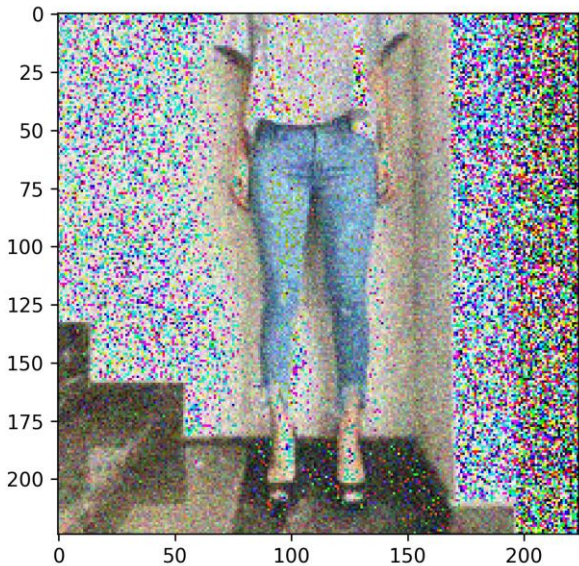


Data Processing

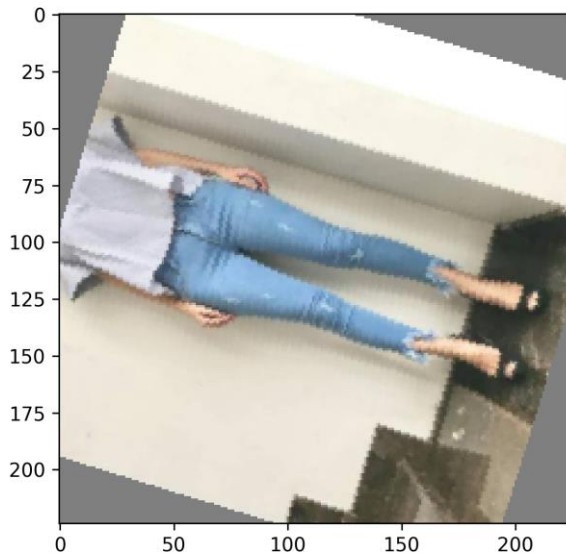
Data augmentation



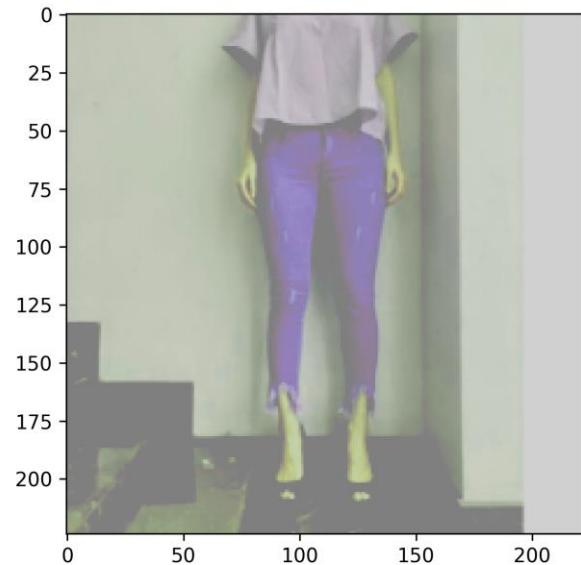
透過資料增強，能夠在有限的資料內，增加訓練樣本的多樣性，並提升模型的Robustness，減少模型過擬合的可能性。



Add noise



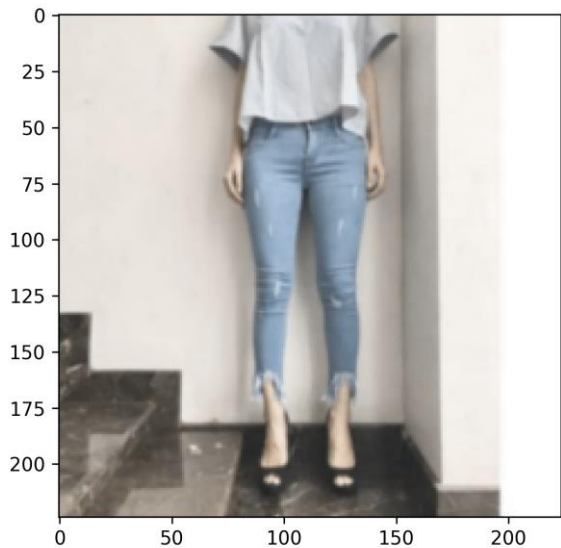
Rotate



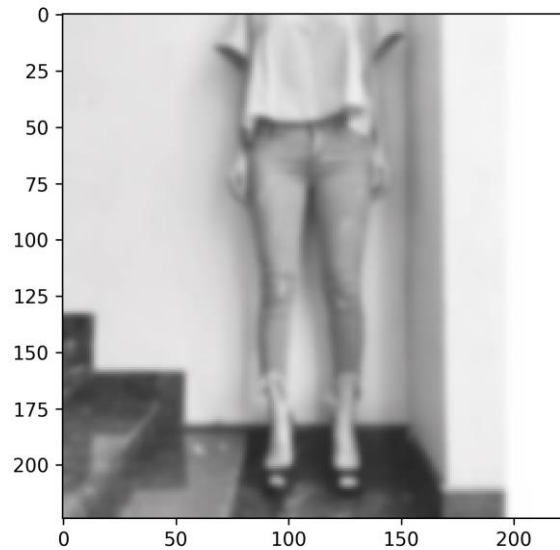
Change Color



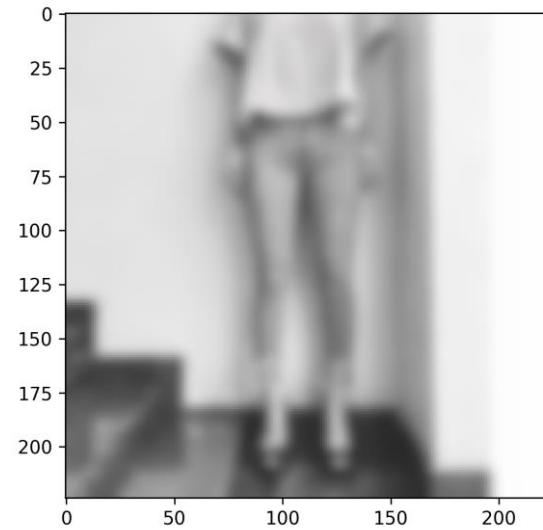
Data augmentation



gaussian_filter, sigma = 1



gaussian_filter, sigma = 3.5



gaussian_filter, sigma = 7



Image Normalize

mean=[0.485, 0.456, 0.406],

std=[0.229, 0.224, 0.225]

These values are calculated based on millions of images.

<https://pytorch.org/vision/stable/models.html>



透過將特徵標準化，除了能夠有效提升模型收斂速度，還能有效提高模型精準度。

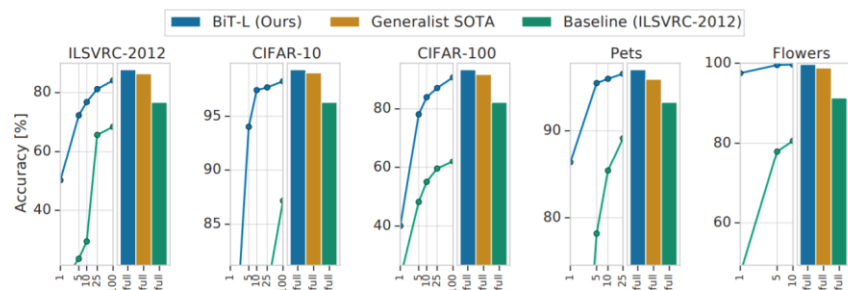


Model

Big Transfer (BiT): General Visual Representation Learning

- BiT是Google最新開源的視覺任務pre-trained模型，而模型架構是使用ResNet152。
- BiT使用了三種不同的pre-training dataset：ILSVRC-2012 (1.3M images)、ImageNet-21k(14M images)、JFT(300M images)。分別以BiT-S、BiT-M、BiT-L對應。
- 目前似乎只公開BiT-S、BiT-M的pre-trained weight

在這次的競賽中使用的model為BiT-M的pre-trained model.

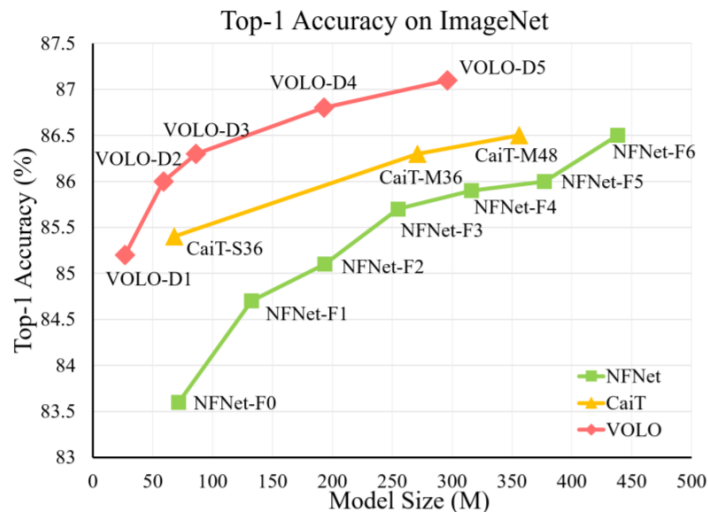


Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., & Hounsby, N. (2020). Big transfer (bit): General visual representation learning. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16* (pp. 491-507). Springer International Publishing.

VOLO: Vision Outlooker for Visual Recognition

來自新加坡 Sea 集團旗下、顏水成教授領導的 Sea AI Lab (SAIL) 團隊提出了一種新的深度學習網絡模型結構: Vision Outlooker (VOLO)，用於高性能視覺識別任務。它是一個簡單且通用的結構，在不使用任何額外數據的情況下，實現了在 ImageNet 上圖像分類任務 87.1% 的精度目標

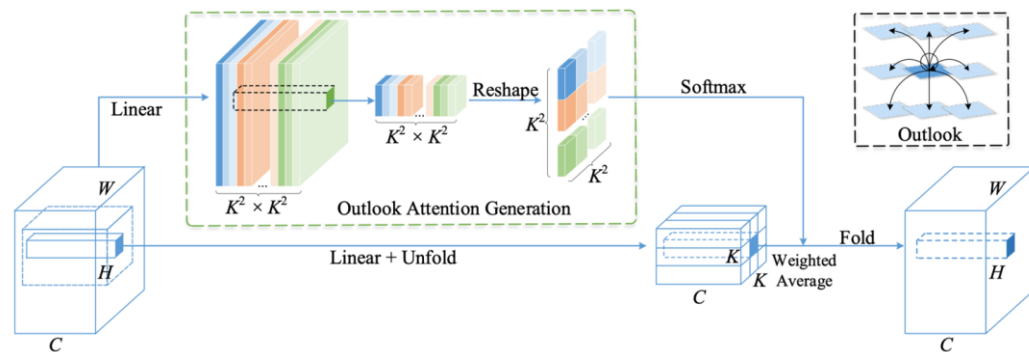
在這次的競賽中使用的model為VOLO-D5



Yuan, L., Hou, Q., Jiang, Z., Feng, J., & Yan, S. (2021). Volo: Vision outlooker for visual recognition. *arXiv preprint arXiv:2106.13112*.

VOLO: Vision Outlooker for Visual Recognition

- 由Outlooker組成，針對圖像細緻特徵編碼
- Outlooker內含outlook attention機制，記錄空間上的特徵與附近點的關係



Yuan, L., Hou, Q., Jiang, Z., Feng, J., & Yan, S. (2021). Volo: Vision outlooker for visual recognition. *arXiv preprint arXiv:2106.13112*.



Model Ensemble



利用具有一定預測能力的model，將無label的圖片進行psudo label，再放入模型進行訓練，能在label 資料有限的狀況下提升模型表現。

前面的兩個model預測的結果均為0.855多，為了使結果更進一步的提高，利用Ensemble的方式，成功的把結果提升到0.86多。