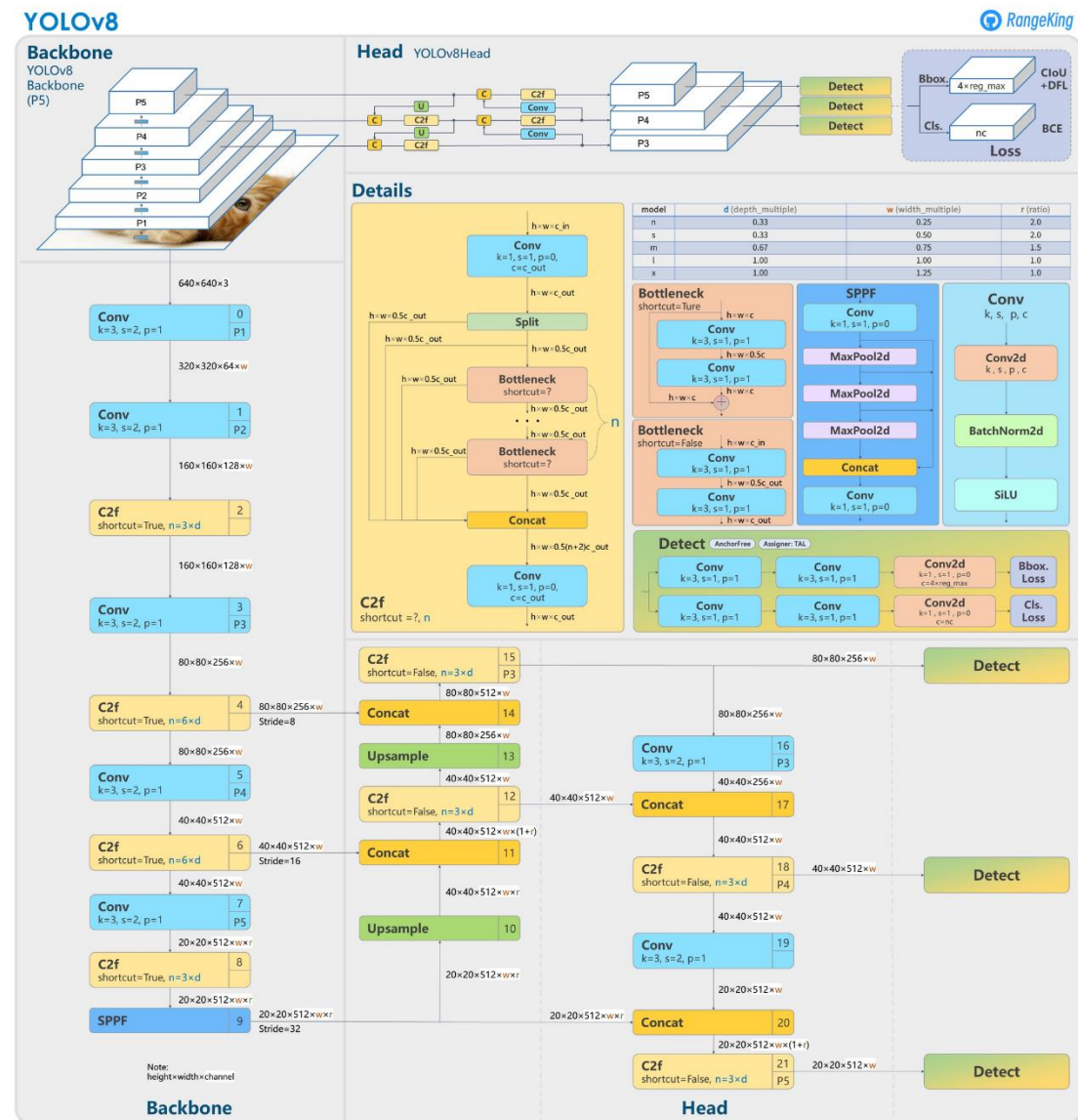


Course. Computer Vision Practice with Deep Learning

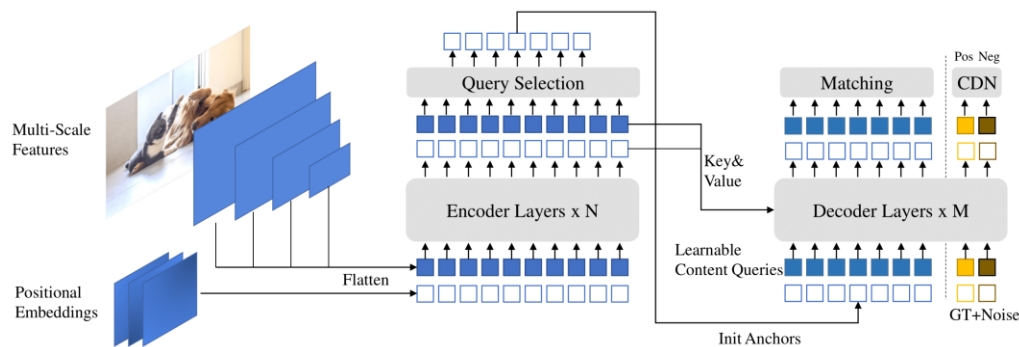
No.61147050S. Department. NTNU CSIE. Name.林詠閔

1. (5%) Draw the architectures for both CNN-based and Transformer-based methods

CNN-based Architecture:



Transformer based architecture:DINO



2. (10%) Report and compare the performance of two methods on validation set

a. at least with mAP@[50:5:95], mAP@50, mAP@75

b. use table to organize the results

	mAP@[50:5:95]	mAP@50	mAP@75	map_small	map_medium	map_large
yolo v8m	0.4805	0.7578	0.5058	0.1557	0.3618	0.5859
DETR	0.4475	0.7022	0.4656	0.1483	0.3392	0.5789

3. (10%) Report the implementation details of both methods

a. Ex: augmentation, loss function, cross validation method, ...etc.

DINO:

1. 在 DINO 模型中，decoder 的輸入是 learnable content queries 和 dynamic 4D anchor box formulation of queries。
2. 透過 mixed query selection 方法來改善 decoder positional queries，進一步提升模型性能。其中，只使用 top-K encoder features 來增強 positional queries，並保持 content queries 為可學習的狀態。
3. 提出了一種新的盒子預測方法，稱為 look forward twice。該方法使用兩層的損失函數來影響每一層的參數更新，從而優化預測盒子的精度。
4. 透過將 b_i 和 Δb_{i+1} 相加，來獲得預測的下一層盒子，從而提高模型的精度。
5. $\text{Update}(\cdot, \cdot)$ 函數用於通過預測的盒子偏移量 Δb_i 來修正初始盒子 b_{i-1} 。在此，採用了 Deformable DETR 中相同的方法來進行盒子更新。
6. 在執行時有多個 backbone 可使用，因我的 gpu 僅能負荷 resnet50，並用 batch=1 做訓練，資料集不多的情況下，使用 pretrain model 進行調整。

Yolo v8

1. Augmentation：YOLOv8 使用了多種數據擴增方法來擴增訓練數據，這些方法包括縮放、旋轉、水平翻轉、裁剪等等。這些方法可以增加數據

集的多樣性，提高模型的泛化能力，也提供了 yam1 file 給我們選擇要使用的 augmentation。

2. Loss Function：YOLOv8 使用了一個稱為 YOLOv5 Loss 的損失函數，它由多個項組成，包括位置損失 L1 loss、置信度損失 Binary cross-entropy loss 和類別損失 Cross-entropy loss。這個損失函數的設計旨在最小化目標檢測中的誤差。
3. Cross Validation Method：YOLOv8 使用了 k 折交叉驗證方法，通常設置為 5 或 10，這個方法可以將數據集切分為 k 個子集，每次使用其中一個子集作為驗證集，其它 k-1 個子集作為訓練集。這樣可以有效地利用數據集進行模型的驗證和訓練。
4. 在 detect 時也有多個 argument 可做篩選，可在官方文檔中 <https://docs.ultralytics.com/usage/cfg/> 中查詢。
5. 資料集不多的情況下，使用 pretrain model 進行調整。

4. (5%) Visualization: draw the bounding boxes of two methods on this test image.

a. IMG_2574_jpeg.jpg.rf.ca0c3ad32384309a61e92d9a8bef87b9

b. Result should be something like this

YOLOV8



DINO

