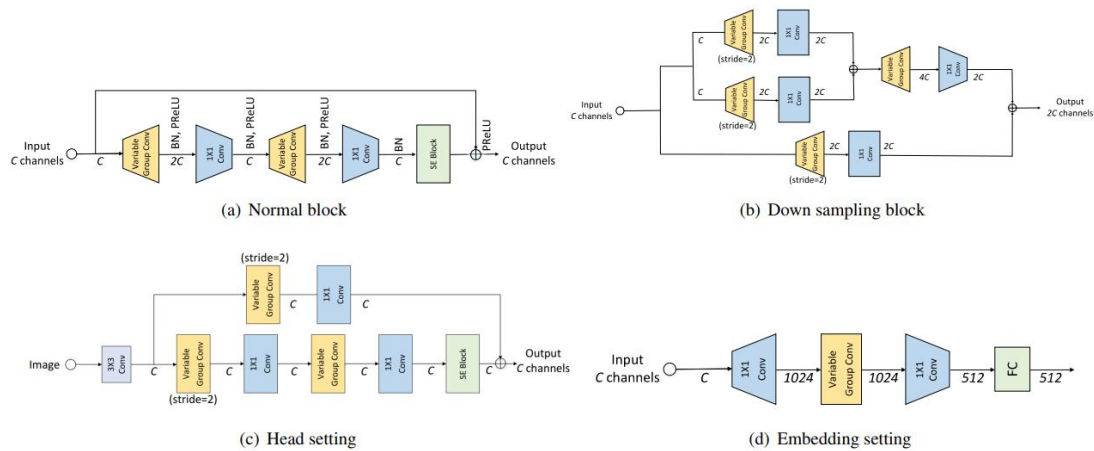


在此次作業中，我採用 VarGFaceNet 的 model，主要使用 Variable group convolution 來減少模型參數，將先介紹模型中的各個模組架構，如下圖所示。



1. Head setting 用來在圖像一開始輸入時，保留必要資訊
2. Down sampling block 用來減半 feature map 的 size
3. Normal block 用來增強判別能力，透過加入 SE block 以及用 PReLU 取代 RELU 來增強判別能力
4. Embedding setting 透過卷積層將維度增加到 1024，以保留必要資訊，再將 feature map 轉變為 1\*1\*512，以減少全連接層中的參數量

下圖為整體模型架構，將圖片先 resize 至(112,112)，透過 Head setting 保留起始必要資訊，且減半 feature map 大小，接下來有 3 個 stage，每個 stage 皆由 Down sampling block 以及 Normal block 組成，feature map 在每個 stage 中會先經過 Down sampling block 將 feature map 的大小減半，再進入 Normal block，Normal block 會維持 feature map 的大小且可重複，其重複次數在每個 stage 中皆不一樣，在 stage 2、3、4 中的重複次數分別為 2、6、3，經過 3 個 stage 後，透過 Embedding setting 將維度增加到 1024，以保留必要資訊，再將 feature map 轉變為 1\*1\*512，以減少全連接層中的參數量，如此即可得到圖片的特徵。

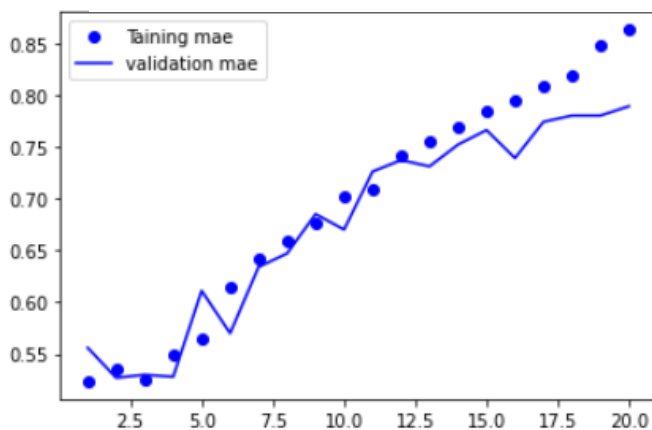
Layer	Output Size	KSize	Stride	Repeat	Output Channels
Image	112x112				3
Conv 1	112x112	3x3	1	1	40
Head Block	56x56		2	1	40
Stage2	28x28		2	1	80
	28x28		1	2	
Stage3	14x14		2	1	160
	14x14		1	6	
Stage4	7x7		2	1	320
	7x7		1	3	
Conv 5	7x7	1x1	1	1	1024
Group Conv	1x1	7x7	1	1	1024
Pointwise Conv	1x1	1x1	1	1	512
FC					512

Table 1. Overall architecture of VarGFaceNet. It only has 1G FLOPs and 5M parameters (memory footprint is 20M saved as float32).

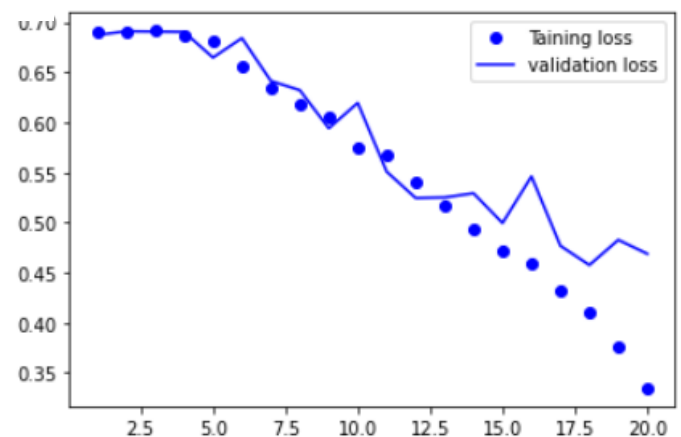
在 face verification 中，使用 LFW 資料集中的 10fold 做為訓練集、test 做為驗證集，10fold 中有 6000 組 pair，而 test 則有 1000 組，我讓同組中兩張人臉圖像分別透過此模型提取  $1 \times 512$  的特徵，將兩組特徵相減取絕對值，再經過一層 64 個神經元的全連接層，最後在 2 個神經元的輸出層透過 softmax 輸出是否為同一人的兩個機率，過程中進行 2 次 dropout(0.5)。

在原始模型訓練設定中，Epoch 設為 20，Optimizer 採用 adam，訓練完後得到的 train loss、train accuracy、validation loss、validation accuracy 分別為 0.335、0.863、0.469、0.789，training、validation accuracy curve 以及 training、validation loss curve 如下圖所示。

training、validation accuracy curve

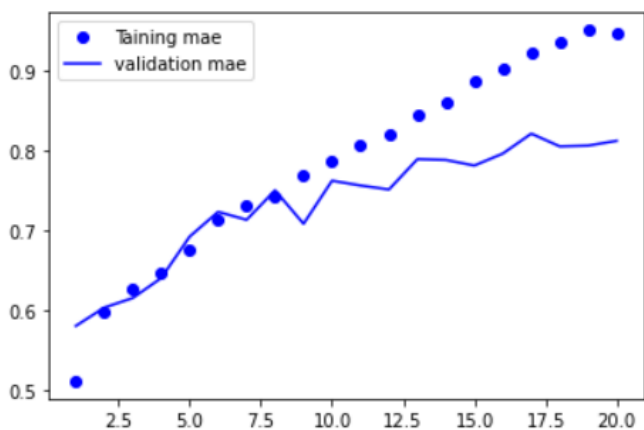


training、validation loss curve

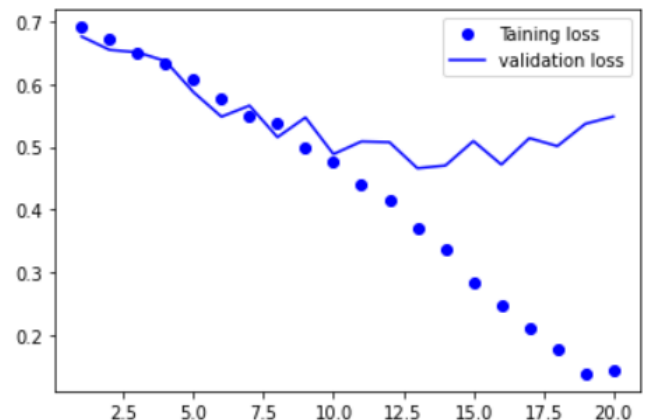


接下來我將模型進行縮減，將 stage3 中原有的 6 個 Normal block 減少至 4 個，以減少計算量及參數，模型訓練設定與原先相同，訓練完後得到的 train loss、train accuracy、validation loss、validation accuracy 分別為 0.144、0.948、0.548、0.813，根據 accuracy curve 以及 loss curve 得知已 overfitting，可推論在模型參數縮減的情況下，不需要訓練太多 epoch，接下來改採訓練 15 個 epoch 即可，本次 training、validation accuracy curve 以及 training、validation loss curve 如下圖所示。

training、validation accuracy curve

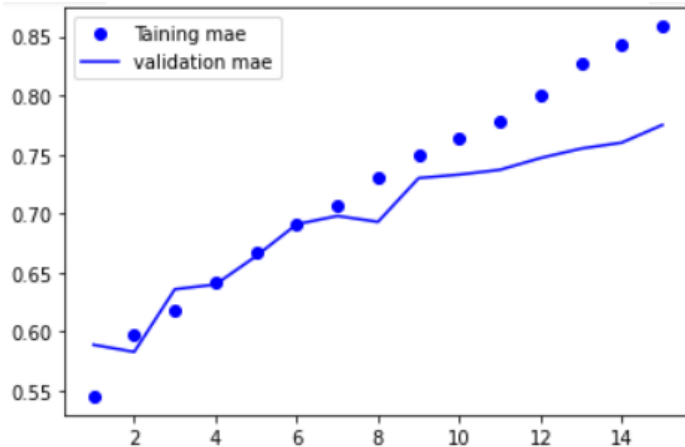


training、validation loss curve

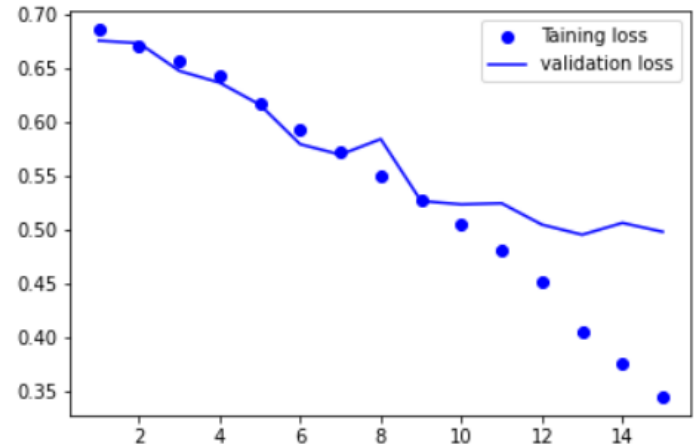


我改採用 15 個 epoch 訓練模型，overfitting 的情況有些許減緩，train loss、train accuracy、validation loss、validation accuracy 分別為 0.346、0.858、0.499、0.775，training、validation accuracy curve 以及 training、validation loss curve 如下圖所示。

training、validation accuracy curve



training、validation loss curve



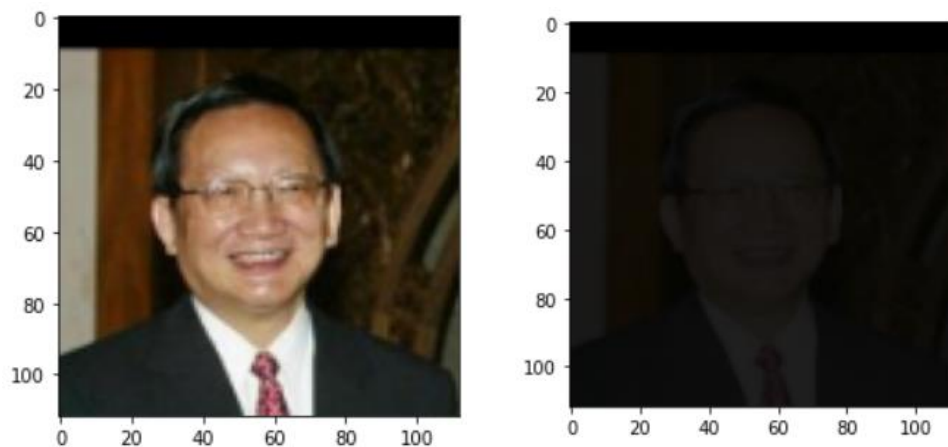
將原模型與縮減後模型的表現進行對比，如下表，可看出縮減模型略遜於原模型，但在 validation 上的表現也不錯。

	Training accuracy	Training loss	Validation accuracy	Validation loss
Original model	0.863	0.335	0.789	0.469
Reduced model	0.858	0.346	0.775	0.499

將原模型以及縮減後模型的參數以及 FLOP 進行比較，如下表所示。

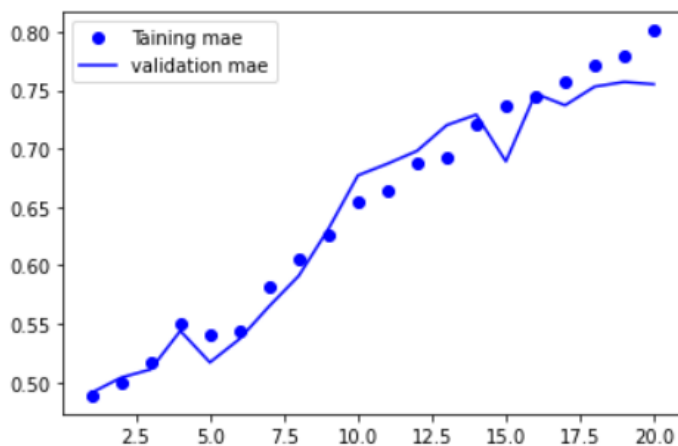
	Total params	Total Flops	Total memory
Original model	5195130	525.62M	55.11MB
Reduced model	4866730	466.26M	50.32MB

接下來在低光照情況下，測試模型的性能，下圖為原圖片以及低光照圖片的對比。

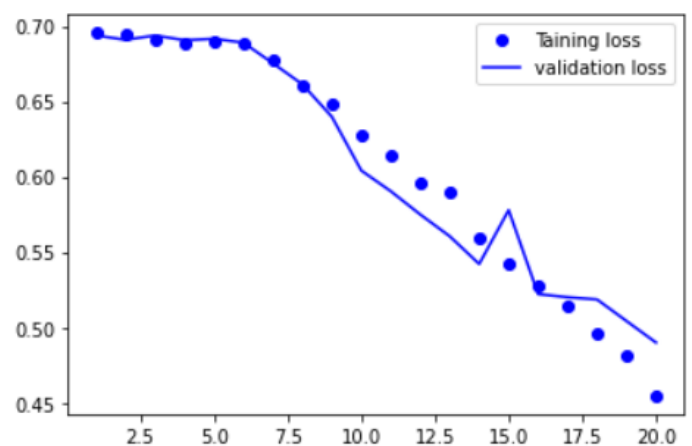


首先，在原模型訓練低光照圖片，epoch 為 20，train loss、train accuracy、validation loss、validation accuracy 分別為 0.455、0.801、0.49、0.755，training、validation accuracy curve 以及 training、validation loss curve 如下圖所示。

**training、validation accuracy curve**

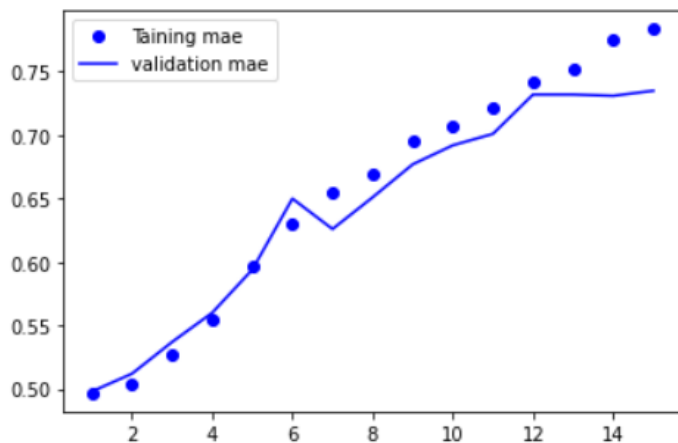


**training、validation loss curve**

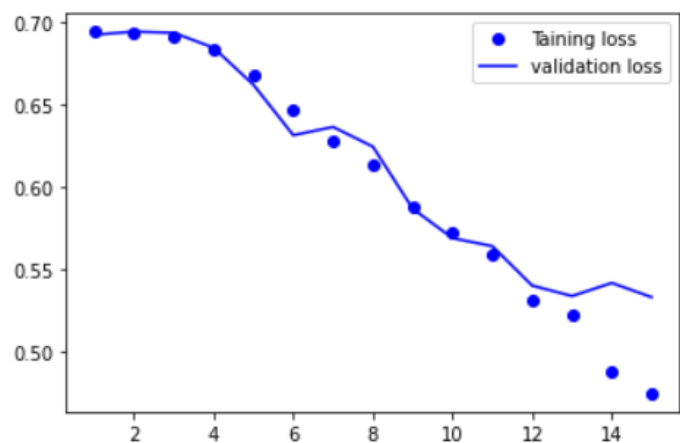


接下來在縮減模型上訓練低光照圖片，epoch 為 15，train loss、train accuracy、validation loss、validation accuracy 分別為 0.474、0.784、0.533、0.735，training、validation accuracy curve 以及 training、validation loss curve 如下圖所示。

training、validation accuracy curve



training、validation loss curve



將原模型與縮減後模型的表現進行對比，如下表，縮減後模型也是略遜於原模型，但表現上並沒有差太多。

	Training accuracy	Training loss	Validation accuracy	Validation loss
Original model	0.801	0.455	0.755	0.49
Reduced model	0.784	0.474	0.735	0.533