**機器學習Hw4**

1. **想法**

本次作業我有將transformer替換成conformer，另外再寫一個Self-Attention Pooling，分別如下:

一張含有 文字, 螢幕擷取畫面, 字型, 軟體 的圖片

自動產生的描述

一張含有 文字, 螢幕擷取畫面, 軟體 的圖片

自動產生的描述

另外，我使用ensemble的方法訓練3個模型，最後讓這三個模型推論的結果平均，結果如下:



至於，超參數的設定，就是try and error的方式。