|  |  |
| --- | --- |
| Neural Networks | |
| HW3 Computer Project2 CNN | |
| 授課老師 | 王振興 |
| Due date  12/16 | |
| 學號 | **N26094883** |
| 姓名 | **杜冠勳** |
| 系所 | **電機所** |

|  |  |
| --- | --- |
| **derivative process** | |
| **handcraft** | **CNN** |
| **Model** | **RBF** |
| **CNN** | **參考論文:** REF.A gentle explanation of Backpropagation in Convolutional Neural Network Son Nguyen February 27, 2020 |
| **Gradient & BP 推導** |
| **首先先從 softmax 層往回推**    **因為 是採用 softmax 故使用 crossentrophy 當作loss 依據**    **藉由上式第二行可推**    **而因為**    **可得**    **利用chain rule**    **而因為**    **將T 用代數帶換**      **若是 label hits 的case 令**    **C 為一常數**  **可得**        而當label 不相同時 亦是此方式得到其gradient  Softmax 階段 小結        **接下來再往回推 要更新FC layer 的weight**      **回推到maxpooling & convolution層 要更新**  **該bias**    **Maxpooling 前 為 convolution 3D output shape 為 (8 8 12)**  **故 bias 更新數量應為12**  **若今天更新回去的路徑上 是maxpooling 最大值的元素則**    **令fk=Pmxy**    **By chain rule 從output 端口回推**      **今 convolution 為1D arrary 則**    **今 convolution 為3D arrary 則**    **若今天 要對pooling不是最大值的 nuron 更新 可以發現 其不需要更新 因為其梯度造成的影響甚小**  **小結** |
| **RBFN** | **.**  **N為 hidden layer 的neuron 個數 Xi 代表的是中心點位置 w 為 權重 y 為 該neuron 的輸出值**  **淨向基函數將 input 與一定點去做距離比較**  **距離的計算方式有許多種 本次作業採用了**[**高斯函數**](https://zh.wikipedia.org/wiki/%E9%AB%98%E6%96%AF%E5%87%BD%E6%95%B0)**逆多二次函數** |

|  |  |  |
| --- | --- | --- |
| **Result** | | |
| **Model\_name** | **lr** | **PICTURE** |
| **CNN** | **0.05** |  |
| **CNN** | **0.01** |  |

|  |  |  |
| --- | --- | --- |
| **CNN+RBF** | **0.05** |  |
| **CNN+RBF** | **0.005** |  |
| **RBF** | **0.05** |  |

|  |
| --- |
| **RBF&& NN 比較與討論** |
| * **從結果上來看純粹用 Dense + RBF 的網路 因為沒有經過特徵篩選 故會相較於CNN based RBF 網路在相同的learning rate 下較慢 提升 accuracy** * **選用CNN 取代 dense 除了獲取到較好的特徵外 也減少了 weight 的數量 對硬體上設計也是較好的** * **相同的 learning rate 下 CNN 相較於 CNN +RBF 可以接受較大的learning rate 來找到 更好的performance** * **同樣的網路架構 learning rate 調過大可能導致無法收斂 CNN +RBF 需用較小的learning rate tune 出model 但也不排除是因為網路較為複雜的緣故。** |