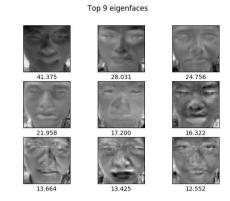
學號:B04902013 系級: 資工二 姓名:鄧逸軒

1.1. Dataset 中前 10 個人的前 10 張照片的平均臉和 PCA 得到的前 9 個 eigenfaces:

答:(左圖平均臉,右圖為 3x3 格狀 eigenfaces, 順序為 左到右再上到下)





1.2. Dataset 中前 10 個人的前 10 張照片的原始圖片和 reconstruct 圖 (用前 5 個 eigenfaces):

答:(左右各為 10x10 格狀的圖, 順序一樣是左到右再上到下)



1.3. Dataset 中前 10 個人的前 10 張照片投影到 top k eigenfaces 時就可以達到 < 1% 的 reconstruction error.

答: (回答 k 是多少)

60

2.1. 使用 word2vec toolkit 的各個參數的值與其意義:

筌:

min_count: 100 default 5 = 出現次數小於這個數值的字會過濾掉

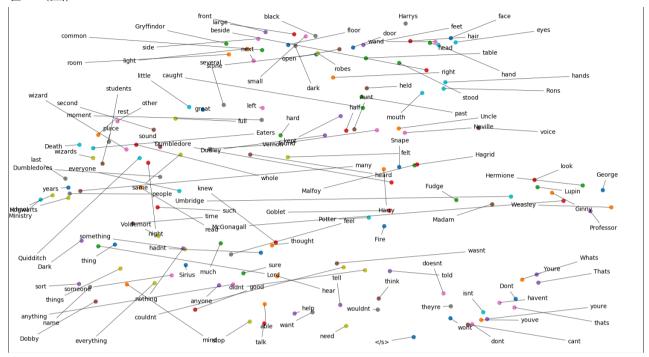
size: 5000 default 100 = 輸出的維度

window: 5 default 5 = 取附近到多遠的詞 negative: 5 default 5 = 可以有多少錯誤的字 alpha: 0.0025 default 0.025 = Learning rate

iter_: 20 default 5 = epochs

2.2. 將 word2vec 的結果投影到 2 維的圖:

答:(圖)



2.3. 從上題視覺化的圖中觀察到了什麼?

答

助動詞都聚集在右下角,右側中間有許多人名,右側上面有許多身體部位,中間上面有許多形容詞,中間下半有許多動詞

3.1. 請詳加解釋你估計原始維度的原理、合理性, 這方法的通用性如何? 答:

在觀察不同維度的資料使用 PCA 的結果之後,發現同維度的資料 Reconstruction error 和 eigenvalues 會很相似,產生出許多 Samples 後用 DNN 判斷 log 後的維度,然後最後把結果 rounding 到 1-60 維。這個方法需要依賴原始資料 Reconstruction error 的絕對數值,在不同 dataset 可能不太通用

3.2. 將你的方法做在 hand rotation sequence datatset 上得到什麼結果?合理嗎?請討論之。

答:

把 480x512 取中間切成 480x480 後,spline 插值成 10x10 的圖片,用上一題的模型得出維度是 6維,真實參數應該會是 6維左右的平移+旋轉,雖然蠻接近的但是我認為是巧合,維度低的主要原因應該是因為圖片間的差異不大,背景又占了大多數的像素,Reconstruction error 不高所以被判定成這個維度,而且上一題的 samples 有 10000 個,這題只有 480 個,所以我覺得這個結果不合理。