R05942102 王冠驊

DLCV HW2

Problem 1: Kernel Trick (10%)

$x = [x_1 \; x_2]^T \qquad x' = [x_1' \; x_2']^T$

$$k(x, x') = (x^T x')^2$$

$$= (x_1 x_1' + x_2 x_2')^2$$

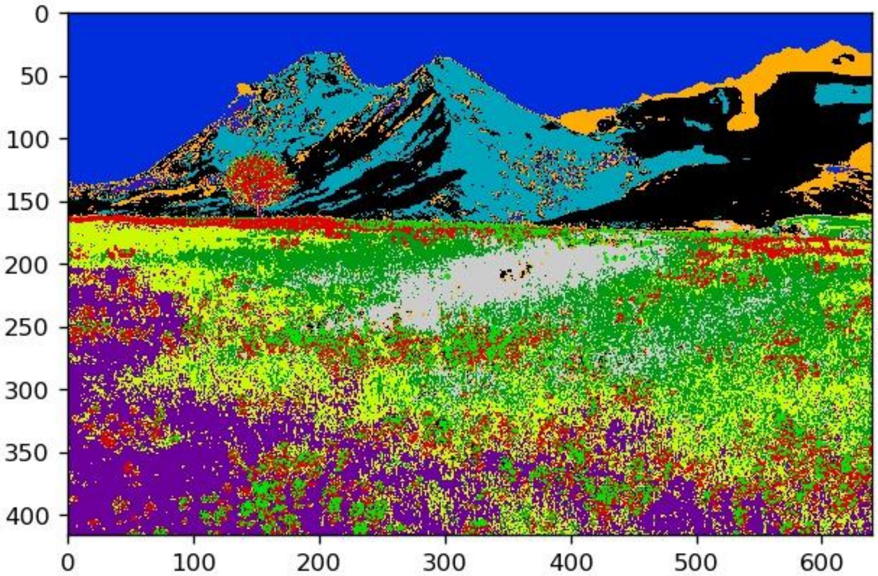
$$= (x_1 x_1')^2 + 2 x_1 x_1' x_2 x_2' + (x_2 x_2')^2 = \Phi(x_1)^T \Phi(x_2)$$

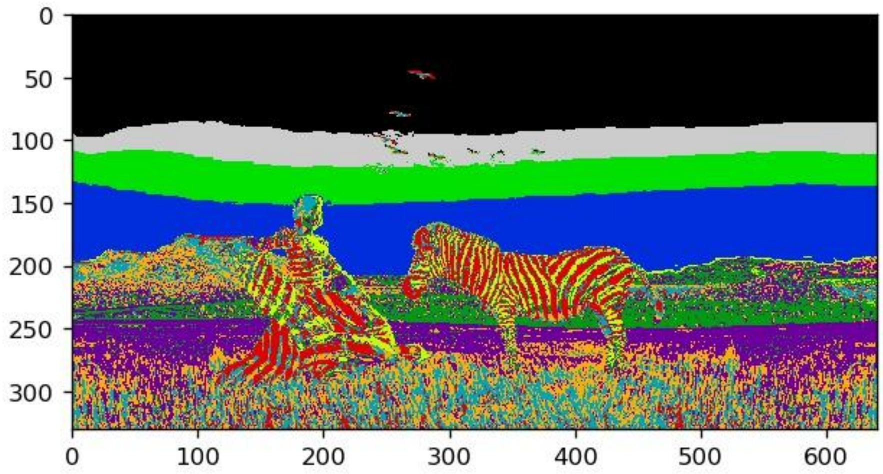
$$\Phi(x) = [x_1^2, \sqrt{2} x_1 x_2, x_2^2]^T$$

Problem 2 : Color and Texture Segmentation (40%)

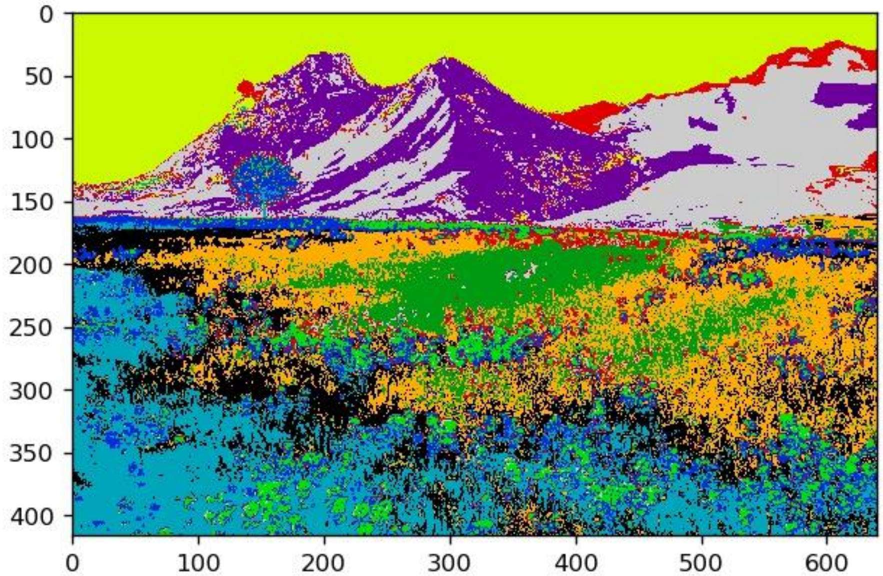
(a) (20%) Color segmentation:

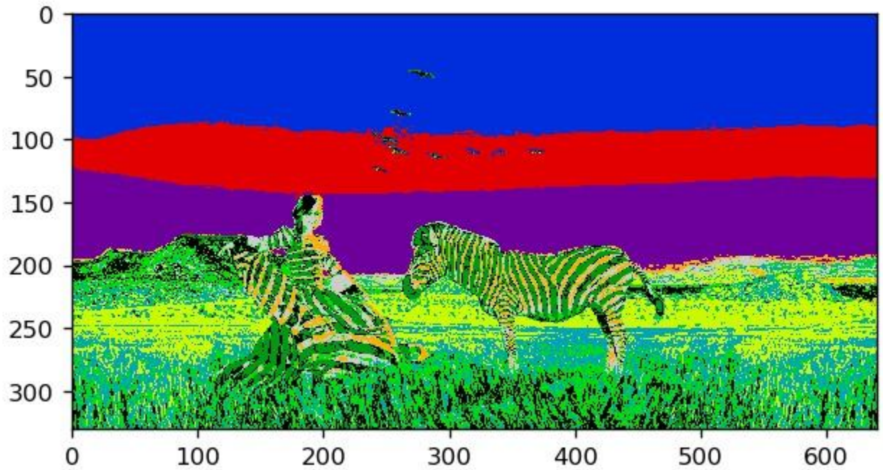
(i) Plot the segmentation results for both images based on your clustering results.





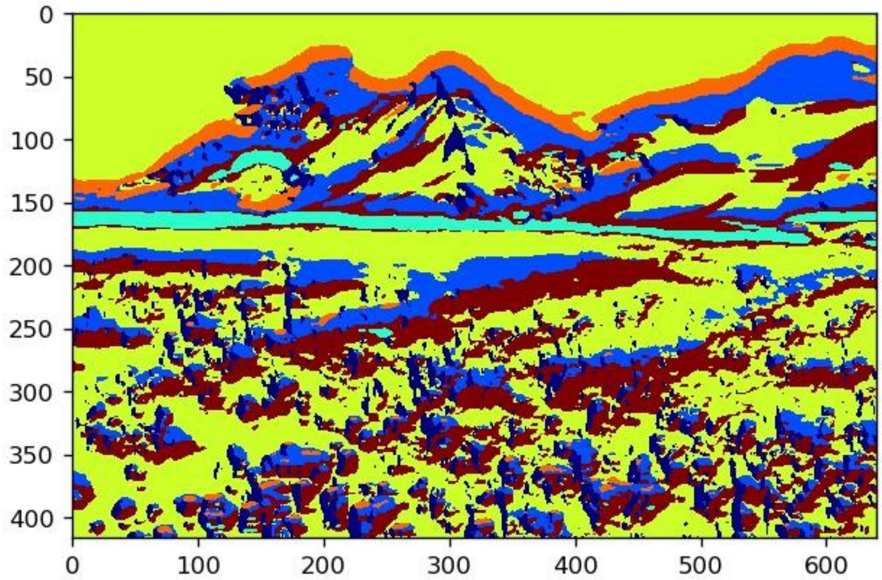
(ii) Convert both RGB images into Lab color space.

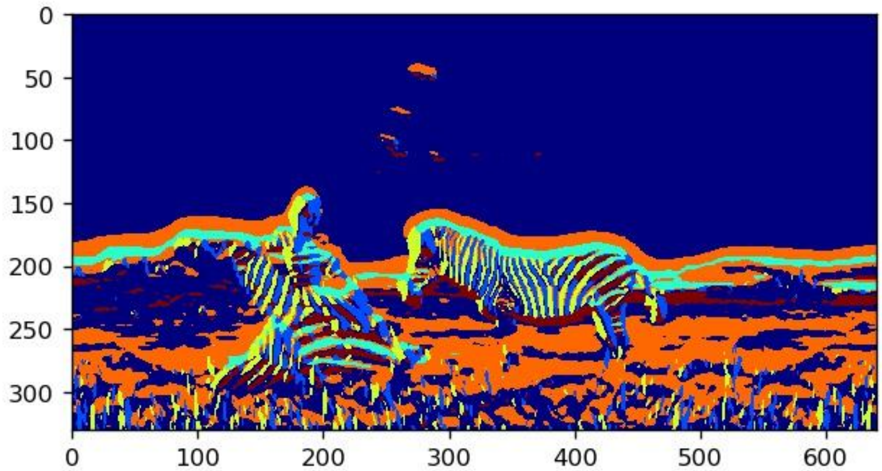




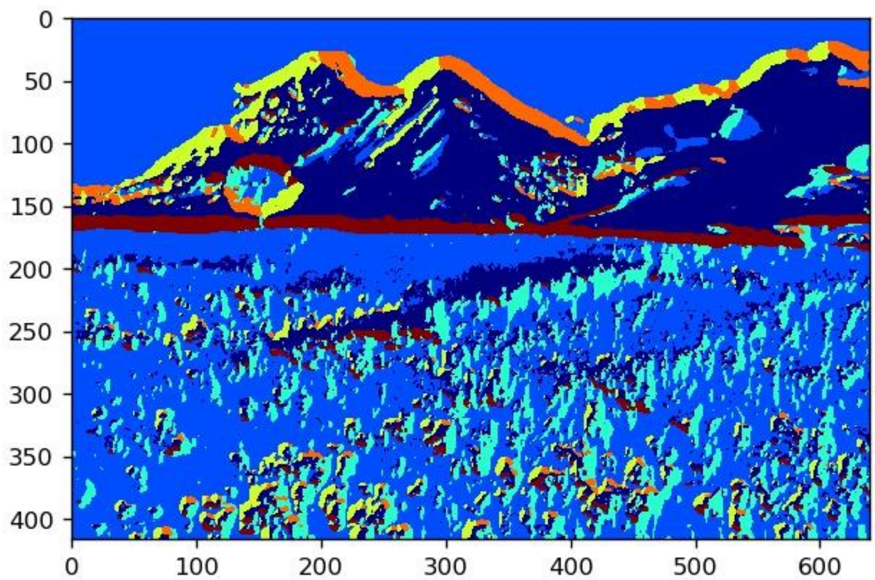
(b) (20%) Texture segmentation:

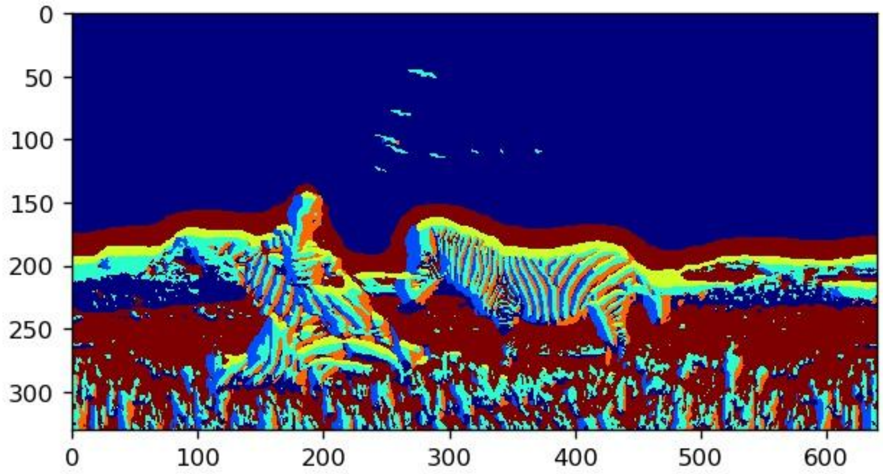
(i) Please plot the texture segmentation results for both images,





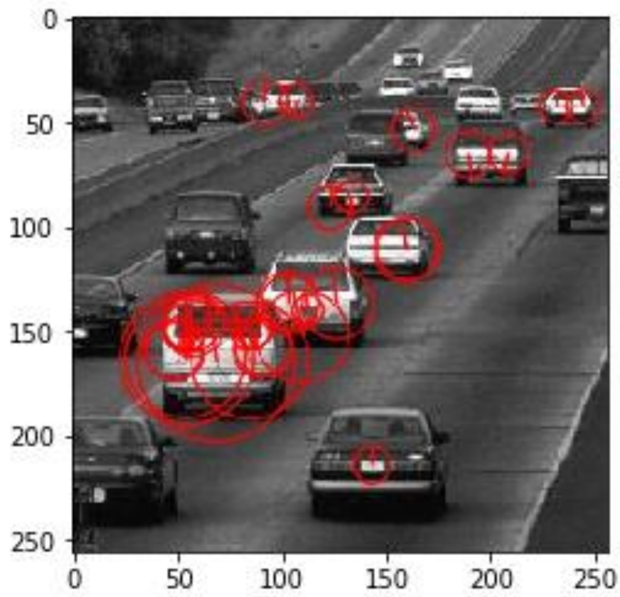
(ii) Combine both color and texture features for image segmentation.



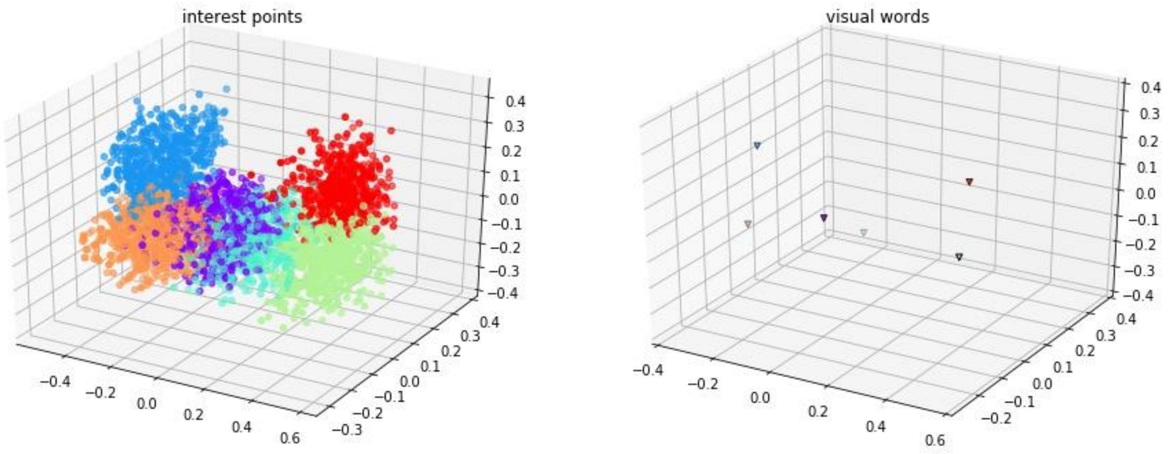


Problem 3 : Recognition with Bag of Visual Words (60%)

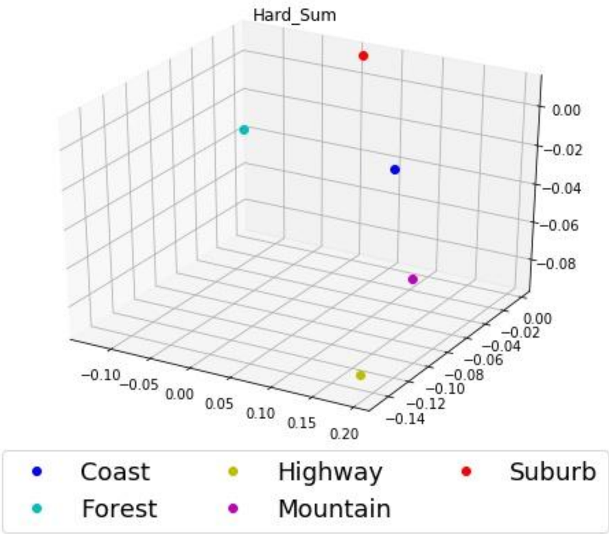
(a) (5%) Detect interest points and calculate their descriptors for this image using SURF.

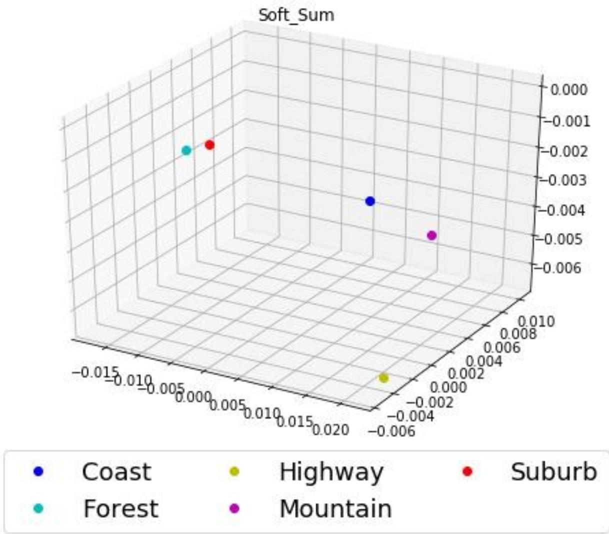


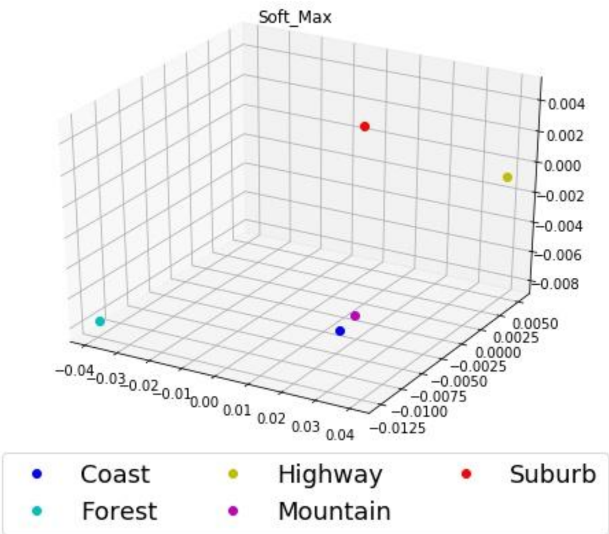
(b) (10%) Plot the visual words and the associated interest points in this PCA subspace.



(c) (20%) Choose one image from each category and plot their Hard-Sum, Soft-Sum, and Soft-Max, respectively.







Can you expect which BoW strategy results in better classification results and why?

從此小題的視覺化結果，我們觀察到使用 Hard-Sum 時可以使 5 類的圖片的樣本點分的最開。而當使用 Soft-Sum 以及 Soft-Max 時則無法完全分開所有不同種類的圖片。故根據此小題的結果，我們推測使用 Hard-Sum 應該會得到最好的分類效果。

(d) (25%)

(i) Use Train-10 as the training data and Test-100 for testing. Report the classification accuracy using Hard-Sum, Soft-Sum, and Soft-Max.

在此小題的實驗中，我們測是不同的參數 c=[50, 100, 200], k=[1, 3, 5, 7, 9]。以下我們列出分別使用 Hard-Sum, Soft-Sum 與 Soft-Max 所得到的最好的分類結果以及個別的 c 與 k 的值:

	accuracy	c	k
Hard-Sum	**63.4%**	100	5
Soft-Sum	55.4%	200	1
Soft-Max	59.2%	200	5

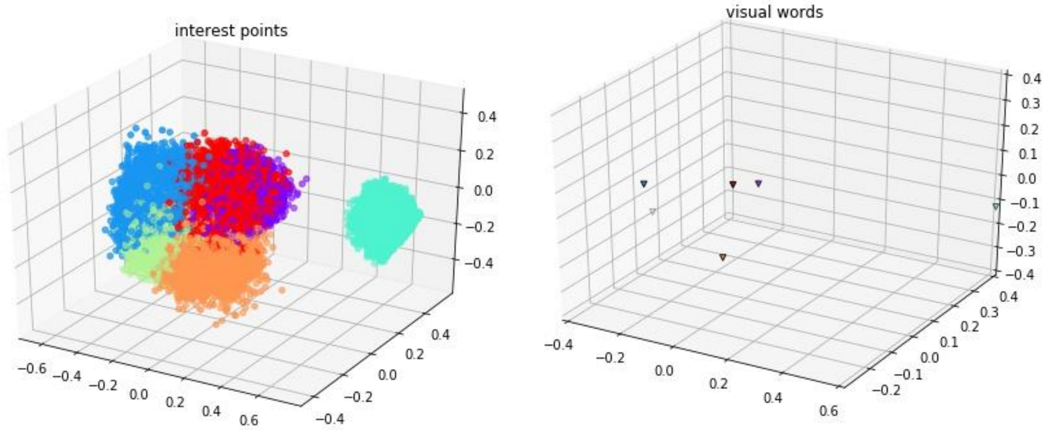
Are the results as expected (based on your observation on different BoW features in (c))? If not, why?

這樣的結果與欲期的結果相同。(c)小題中，使用Hard-Sum可以使不同類的圖片分的最開，在這裡我們也可以看到使用Hard-Sum可以得到最高的分類準確率。

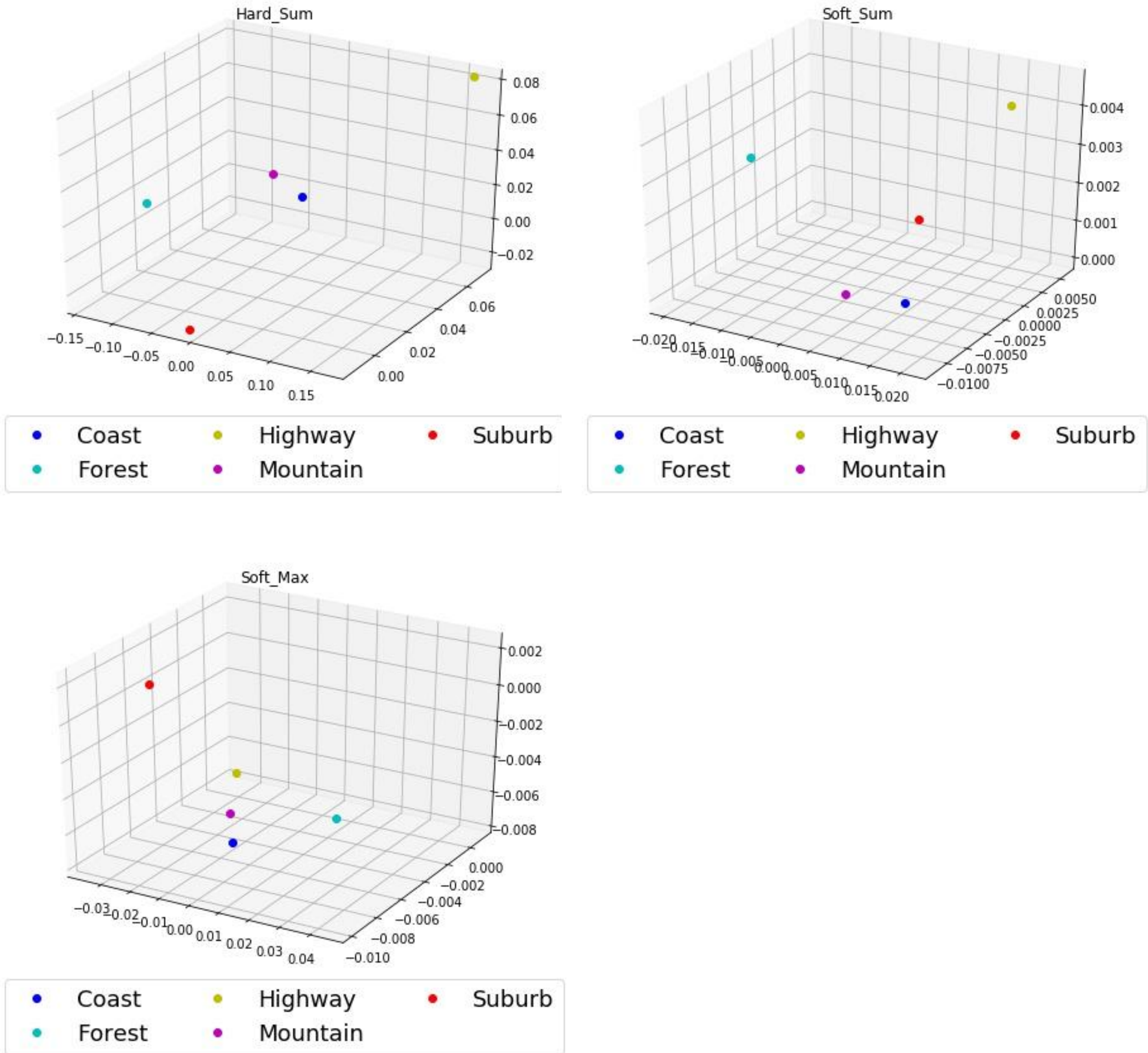
(ii) Repeat (a) to (c) using Train-100 as the training data.

(a)

(b)


interest points
visual words

(c)


Hard_Sum
Soft_Sum
Soft_Max
Coast
Highway
Suburb
Forest
Mountain

在 Train-100 當做 training data 可以發現使用 Soft-Sum 時不同種類的圖片彼此會分的最開，故我們推測使用 Soft-Sum 時會有最好的分類效果。

(d)

在此小題的實驗中，我們測是不同的參數 c=[50, 100, 200], k=[1, 3, 5, 7, 9, 11, 13, 15, 17, 19]。以下我們列出分別使用 Hard-Sum, Soft-Sum 與 Soft-Max 所得到的最好的分類結果以及個別的 c 與 k 的值:

	accuracy	c	k
Hard-Sum	72.4%	50	7
Soft-Sum	**72.8%**	200	7
Soft-Max	70.8%	200	5

Do you observe improved classification results? Please report and explain your results.

隨著 training data 的增加，分的準確率也有顯著的提升。其中參數 k 也有變大的趨勢，代表我們可以參考更多相近的圖片的標籤提升分類的準確率。此外，最佳的分類準確率發生在我們使用 Soft-Sum，此結果也與我們使用 Train-10 當做 training data 時不同。