



GNR638 Assignment 1

Bag of Visual Words

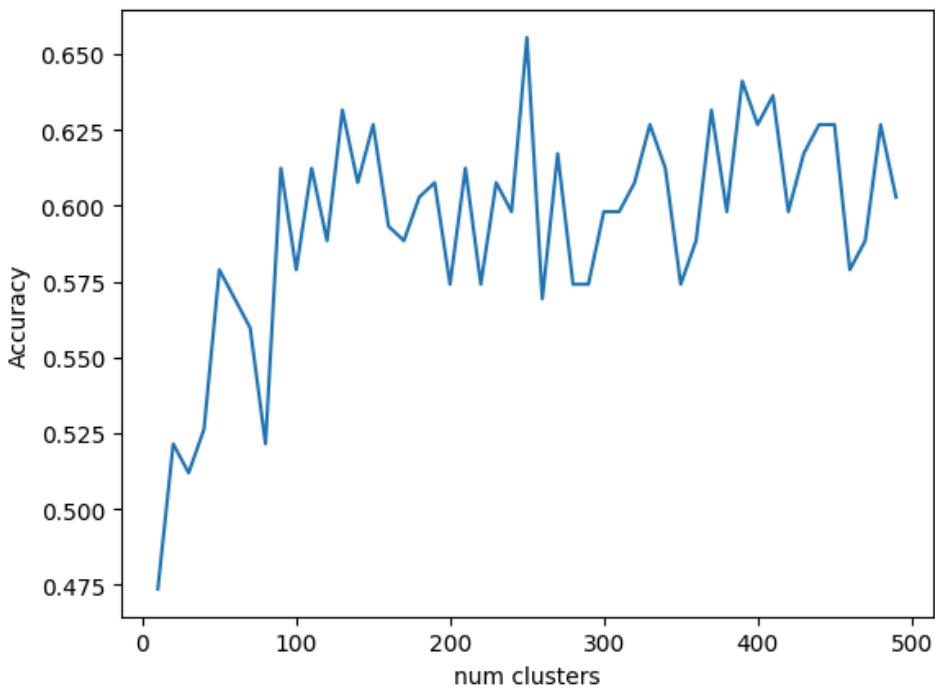
Members:

Desai Utsav Manojkumar 200100054
Ashutosh Mulchandani 200100037

Introduction

In computer vision, Bag of Visual Words(BoVW) is a powerful model for various tasks. It is inspired by bag of words model in Natural language processing (NLP). BoVW provides a systematic framework for image feature extraction and representation, treating images as collections of local visual elements. These local features are quantized into a fixed vocabulary, forming a concise "bag of visual words" that serves as a robust representation. By focusing on local patterns and discarding spatial relationships, BoVW proves effective for various tasks such as object recognition and image retrieval.

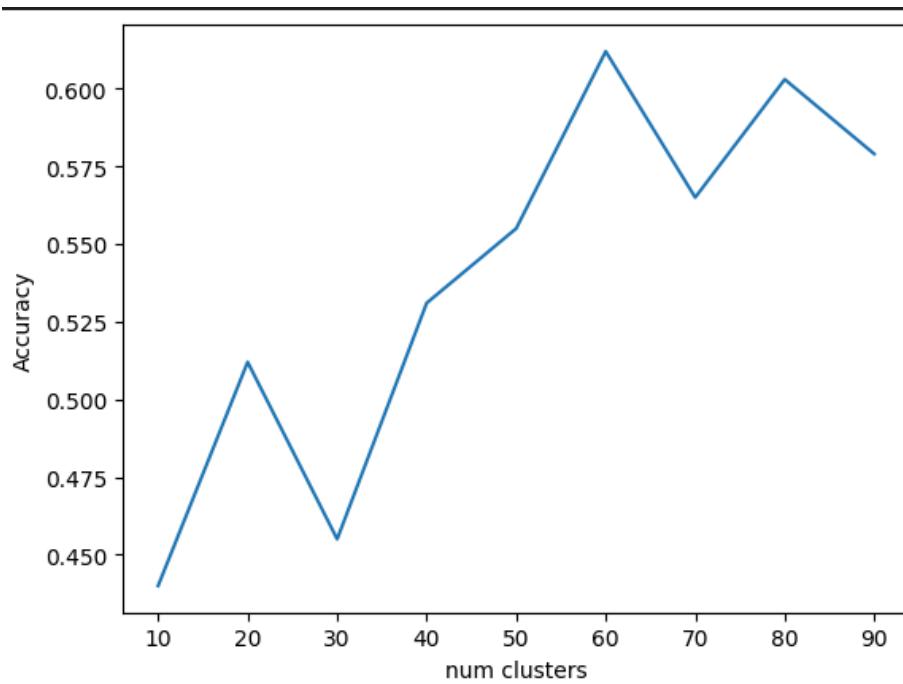
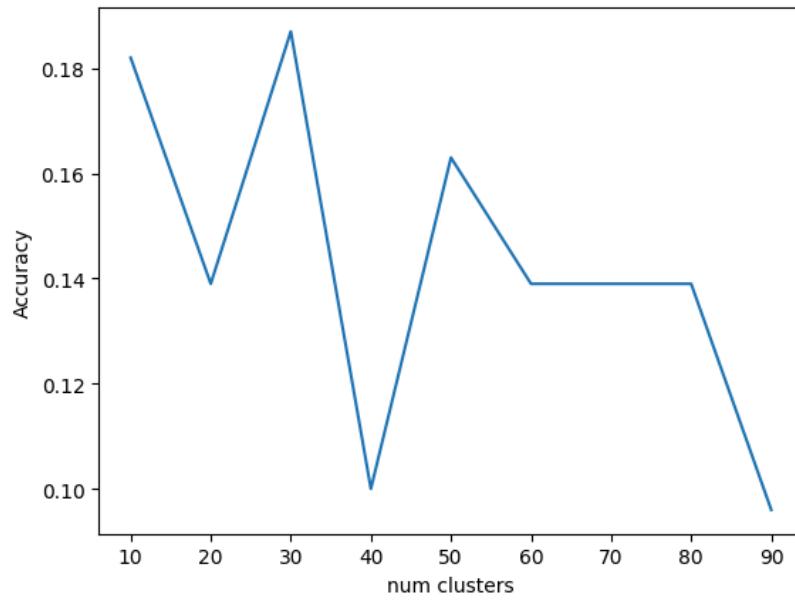
- Sensitivity analysis on K



From the above graph we can see that, the accuracy of the model increases until the number of clusters are couple hundreds, but after that, there is no significant increase in the accuracy, rather it highly fluctuates as the k changes. So, without glcm features, the accuracy almost reaches close to 60%. Since the increase in number of clusters also increases the computation time, $k \sim 150$ is the optimal number of clusters for our dataset.

- GLCM Features Extraction

- We were not getting good accuracy with GLCM features while using L2 normalization.



In the context of Bag of Visual Words (BoVW), Visual Polysemy and Synonyms refer to certain challenges faced while representing the image information using bag of words concept.

Visual Polysemy:

Polysemy, in general, refers to the phenomenon where a single word has multiple meanings. Similarly, in the context of BoVW, Visual Polysemy arises when a visual word in the vocabulary represents different objects in different contexts or images.

For example, a visual word representing a patch of blue sky may also be present in images with blue water. This ambiguity can lead to confusion in the classification process, as the same visual word might correspond to different objects or scenes.



Synonyms:

In the context of BoVW, synonyms refer to different visual words that represent similar or related visual concepts. This can happen due to variations in lighting conditions, viewpoints, or other factors affecting the appearance of an object.

For instance, two visual words might represent slightly different textures of the same object, such as a textured wall. In the BoVW framework, treating these visually similar words as synonyms can lead to better generalization in image classification tasks. In the below images, we can see that both the boxes have different textures, but both represent a wall.

