

* Local Binary Patterns (LBP)

- Is a type of visual descriptor used for classification.
- It's known to be a powerful feature for texture classification, also, further when combined with HOG descriptors, it improves the detection performance considerably ~~can~~ on some datasets.

• Concept

LBP feature vector is created in the following manner:

- Divide the examined window into cells (e.g. 16×16 pixels for each cell)
- For each pixel in a cell, compare the pixel to each of its 8 neighbours (on its left-top, left-middle, left-bottom, right-top, etc). Follow the pixels along a circle, i.e. clockwise or counter-clockwise.
- Where the center pixel's value is greater than the neighbour's value, write "0". Otherwise, write "1". This gives an 8-digit binary number (which is usually converted to decimal for convenience).
- Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center). The histogram can be seen as 256-dimensional feature vector.
- Optionally normalize the histogram.

→ concatenate (normalized) histogram of all cells. This gives a feature vector for the entire window.

This feature vector can be used for ~~any~~ classification using any machine learning algorithms. These are used for face recognition or texture analysis.

Bag of Visual Words

- * commonly used in image classification
- * adapted from information retrieval and NLP's bag of words.
- * In bag of words (BOW), we count the number of each word appears in a document, use the frequency of each word to know the keywords of the document, and make a frequency of each word to histogram from it. We treat a document as bag of words.
- * Putting the same ~~over~~ concept, we use image features as the "words".

* Bag of visual words (BOVW) is to represent an image as a set of features. Features consist of keypoints and descriptor.

Keypoints \Rightarrow "stand out" or unique points in an image, regardless of rotation, shrink or expansion, it stays the same.

Descriptor \Rightarrow description of the keypoint.

* We use the keypoints and descriptors to construct vocabularies and represent each image as a frequency histogram of features that are in the image. From this, we find another similar images or predict the category of the image.

* Building a BOVW:

\rightarrow Detect features, extract descriptors from each image in the dataset, and build a visual dictionary.

→ Next, we make a cluster from the descriptors (K-means, DBSCAN, etc). The center of ~~each~~ each cluster will be used as the visual dictionary's vocabularies.

→ Finally, for each image we make a frequency histogram from the vocabularies and the frequency of the vocabularies in the image. These histograms are our bag of visual words.

* By using BOVW representation, we can compute this image's nearest neighbours. We can do it using nearest neighbours algo. or any other algo.