Mammographic images detection

For the detection and diagnosis of cancer from microscopic biopsy images, the histopathologists normally look at the specific features in the cells and tissue structures. The various common features used for the detection and diagnosis of cancer from the microscopic biopsy images include shape and size of cells, shape and size of cell nuclei, and distribution of the cells. The brief descriptions of these features are given as follows.

(A) **Shape and Size of the Cells**. It has been observed that the overall shape and size of cells in the tissues are mostly normal. The cellular structures of the cancerous cells might be either larger or shorter than normal cells. The normal cells have even shapes and functionality. Cancer cells usually do not function in a useful way and their shapes are often not even.

(B) **Size and Shape of the Cell’s Nucleus**. The shape and size of the nucleus of a cancer cell are often not normal. The nucleus is decentralized in the cancer cells. The image of the cell looks like an omelet, in which the central yolk is the nucleus and the surrounding white is the cytoplasm. The nuclei of cancer cells are larger than the normal cells and deviated from the centre of the mass. The nucleus of cancer cell is darker. The segmentation step mainly focuses on separation of regions of interests (cells) from background tissues as well as separation of nuclei from cytoplasm.

(C) **Distribution of the Cells in Tissue**. The function of each tissue depends on the distribution and arrangements of the normal cells. The numbers of healthy cells per unit area are less in the cancerous tissues. These adjectives of microscopic biopsy images have been included in shape and morphology based features, texture features, color based features, Color Gray Level Cooccurrence Matrix (GLCM), Law’s Texture Energy (LTE), Tamura’s features, and wavelet features which are more biologically interpretable and clinically significant.

Steps in image processing:

enhancement of microscopic images

contrast limited adaptive histogram equalization

segmentation of background cells

k-means segmentation algorithm

features extraction

biologically interpretable and clinically significant shape and morphology based features are extracted from the segmented images which include gray level texture features, color based features, color gray level texture features, Law’s Texture Energy (LTE) based features, Tamura’s features, and wavelet features

finally the classification

K-nearest neighborhood (KNN), fuzzy KNN, and support vector machine (SVM) based classifiers are examined for classifying the normal and cancerous biopsy images

enables detection of early signs of breast cancer such as masses, calcifications, architectural distortion and bilateral asymmetry.

On the other hand, a malignant mass usually has a spiculated, rough and blurry boundary

These approaches are tested on four fundamental tissues (connective, epithelial, muscular, and nervous) of randomly selected 1000 microscopic biopsy images.

There are seven sets of features used for computing the feature vector of microscopic biopsy images explained as follows.

(i) Texture Features (F1–F22). [[32](https://www.hindawi.com/journals/jme/2015/457906/#B12)–[34](https://www.hindawi.com/journals/jme/2015/457906/#B53)] Autocorrelation, contrast, correlation, cluster prominence, cluster shade, difference variance, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference (INV), inverse difference normalized (INN), and inverse difference moment normalized are major texture features which can be calculated using equations of the texture features.

(ii) Morphology and Shape Feature (F23–F32). In papers [[35](https://www.hindawi.com/journals/jme/2015/457906/#B10), [36](https://www.hindawi.com/journals/jme/2015/457906/#B11)] authors describe the shape and morphology features. The considered shape and morphological features in this paper are area, perimeter, major axis length, minor axis length, equivalent diameter, orientation, convex area, filled area, solidity, and eccentricity.

(iii) Histogram of Oriented Gradient (HOG) (F33–F68). Histogram of oriented gradient is one of the good features set to deify the objects [[32](https://www.hindawi.com/journals/jme/2015/457906/#B12)]. In our observation it will be included for better and accurate identification of objects present in microscopic biopsy images.

(iv) Wavelet Features (F69–100). Wavelets are small wave which is used to transform the signals for effective processing [[3](https://www.hindawi.com/journals/jme/2015/457906/#B2)]. The wavelets are useful in multiresolution analysis of microscopic biopsy images because they are fast and give better compression as compared to other transforms. The Fourier transform converts a signal into a continuous series of sine waves, but the wavelet precedes it in both time and frequency. This accounts for the efficiency of wavelet transforms [[37](https://www.hindawi.com/journals/jme/2015/457906/#B14)]. Daubechies wavelets have been used because they have fractal structures and they are useful in the case of microscopic biopsy images. In this paper mean, entropy, energy, contrast homogeneity, and sum of wavelet coefficients are taken into consideration.

(v) Color Features (F101–F106). The components of these models are hue, saturation, and value (HSV) [[34](https://www.hindawi.com/journals/jme/2015/457906/#B53)]. This is represented by the six sided pyramids, the vertical axis behaves as brightness, the horizontal distance from the axis represents the saturation, and the angle represents the hue. Here mean and standard deviation of HSV components are taken as features.

(vi) Tamura’s Features (F107–F109). Tamura’s features are computed on the basis of three fundamental texture features: contrast, coarseness, and directionality [[3](https://www.hindawi.com/journals/jme/2015/457906/#B2)]. Contrast is the measure of variety of the texture pattern. Therefore, the larger blocks that make up the image have a larger contrast. It is affected by the use of varying black and white intensities [[32](https://www.hindawi.com/journals/jme/2015/457906/#B12)]. Coarseness is the measure of granularity of an image [[32](https://www.hindawi.com/journals/jme/2015/457906/#B12)]; thus coarseness can be represented using average size of regions that have the same intensity [[38](https://www.hindawi.com/journals/jme/2015/457906/#B13)]. Directionality is the measure of directions of the grey values within the image [[32](https://www.hindawi.com/journals/jme/2015/457906/#B12)].

(vii) Law’s Texture Energy (LTE) (F110–F115). These features are texture description features which mainly used average gray level, edges, spots, ripples, and wave to generate vectors of the masks. Law’s mask is represented by the features of an image without using frequency domain [[39](https://www.hindawi.com/journals/jme/2015/457906/#B31)]. Laws significantly determined that several masks of appropriate sizes were very instructive for discriminating between different kinds of texture features present in the microscopic biopsy images. Thus its classified samples are based on expected values of variance-like square measures of these convolutions, called texture energy measures. The LTE mask method is based on texture energy transforms applied to the image classification used to estimate the energy within the pass region of filters [[40](https://www.hindawi.com/journals/jme/2015/457906/#B55)].