Abstract

Breast cancer is the most common cancer and is the primary cause of cancer related deaths in women worldwide. Early diagnosis of breast cancer could be helpful in the reduction of mortality rate, cost of the treatment and the anxiety of the patients. Breast cancer can be diagnosed in the non-palpable stage with the help of mammography. Microcalcification cluster is one of the early indicator of breast cancer found in mammograms. The microcalcifications appear as small high intensity spots with the size ranging from 0.05 mm to 1 mm. Small microcalcifications clusters are very difficult to interpret and are often missed by the radiologists. There is massive workload on the radiologists to conduct screening mammography for huge population which may lead to decrease in detection accuracy of the breast cancer. In the screening program, majority of the mammograms are normal or do not have any cancer signs. Thus, there is a need of screening tool which can reliably discard normal mammograms and thereby reduce the burden on radiologists.

Since microcalcifications are associated with high frequency components of an image, multiscale unsharp masking is proposed to highlight microcalcifications of variable sizes. A simple iterative thresholding is proposed to get potential microcalcification regions. Various features such as shape-based, texture-based, gradient-based features are extracted from foreground regions to discriminate true microcalcifications and falsely detected regions. The proposed multiscale unsharp masking based method achieved higher sensitivity and lower false positive compared to simulated results of existing techniques.

Two multiscale approaches based on non linear energy operator (NEO) technique are proposed viz. max multiscale 2-D NEO and mean multiscale 2-D NEO to emphasize microcalcifications of variable sizes. A new majority class data reduction technique based on data distribution is proposed to counter the data imbalance in the training set. Since some of the false positives occur due to the detection of vascular calcifications, a principle component analysis based approach is proposed to remove vascular calcifications. The proposed max multiscale 2-D NEO achieved more than 97% sensitivity at lower false positive rate compared to existing techniques. The proposed mean multiscale 2-D NEO is able to achieve 100% sensitivity and least false positive compared to existing techniques. Thus, the proposed mean multiscale 2-D NEO can act as second reader to assist radiologists in detecting microcalcification clusters. Since the proposed mean multiscale 2-D NEO is able to detect all of the microcalcification clusters, it can identify mammograms without microcalcification clusters. Therefore, it can act as a first read and discard significant percentage of mammograms without microcalcification clusters and reduce the workload of radiologists.

Keywords: Computer-aided detection; classification; feature extraction; feature selection; breast cancer; microcalcification clusters; vascular calcifications; multilevel thresholding; data imbalance; shape-, gradient-, texture-based features; screening tool.