

AbstractID: 4739 Title: Information-Theoretic CAD System in Mammography: Investigation of an Entropy-Based Indexing Scheme for Improved Computational Efficiency and Robust Performance

Purpose: We previously presented a Knowledge-Based Computer Assisted Detection (KB-CAD) system for the detection of mammographic masses based on information theoretic similarity measures. The purpose of this study is to integrate an entropy-based indexing strategy that improves the speed of analysis without compromising the diagnostic performance of the system.

Materials and Methods: The KB-CAD system is designed to compare a query mammographic region with mammographic templates of known ground truth stored in the knowledge databank. The system utilizes normalized mutual information (NMI) as the similarity measure. Previous studies showed that the system's diagnostic performance consistently improves as the knowledge databank increases. However, comparing each query to every template stored in the databank becomes computationally impractical. To address this concern, we implemented an entropy-based indexing scheme. Instead of comparing the query with the whole knowledge databank, we restrict the comparisons to those stored templates that are most relevant or informative. Only those templates are retrieved for more detailed analysis using the NMI similarity measure. Image entropy was used to determine the level of information or relevance for each available template.

Results: We used a database of 2,318 mammographic regions (905 masses, 1,413 normal) extracted from the Digital Database of Screening Mammography (DDSM). Based on a leave-one-out sampling scheme, different clustering implementations of the entropy-based indexing scheme were investigated. Using entropy-based indexing, the KB-CAD system achieved the same performance (ROC area index $A_z=0.91\pm 0.01$ for malignant masses) using only a subset of 600 templates from the knowledge databank (compared to the full set of 2,317 available templates).

Conclusion: This study showed that to sustain detection performance while maintaining computational efficiency, an optimized subset of relevant templates could be determined using a sophisticated entropy-based indexing strategy.

Conflict of Interest: This work was supported in part by the National Cancer Institute under grant R01 CA101911.