

Abstract and Specific Aims

The main goal of this proposal is to design a prototype Computer Aided Recognition (CAR) system for recognition of lung tumors utilizing the proposed hierarchical deep fusion-learning scheme. The PI will employ the proposed prototype system to obtain extensive preliminary results to be nationally competitive as well as utilize the fund to build a Biomedical Informatics Research Lab in the department of Computer Science.

In this proposal, the PI aims to explore a hierarchical deep fusion network that combines multi-channel and multi-perspective decisions. In addition, the PI will investigate performance of the proposed network integrated with association rule mining based decision making at the fully connected layer.

The idea proposed in this research is innovative such that the model learns how to make fusion of different predictions in contrast to using a voting scheme or simple descriptive approach. The idea of utilizing association rules proposed in this research is also a novel approach in object recognition.

As professional development objectives, the PI aims to build a Biomedical Informatics Research Lab where he will advise undergraduate and graduate students at Southeastern Louisiana University (Southeastern). As an adjunct faculty and a former non-tenure track faculty at Louisiana State University (LSU), the PI has a strong relationship with LSU in terms of academic activities such as advising graduate students, teaching graduate level courses, and collaboration in research. Through this proposal, the PI plans to extend his biomedical research collaboration at LSU and in other local institutions such as Pennington Biomedical Research Center and Mary Bird Perkins Cancer Center. Funding of this proposal will greatly support cross-university and cross-disciplinary research of the PI.

Background and Preliminary Results

Rationale:

The PI Dr. Soysal is an Assistant Professor in the Department of Computer Science at Southeastern Louisiana University as well as an adjunct faculty in the School of Electrical Engineering and Computer Science at Louisiana State University. He has advised 14 graduate level theses and supervised several projects in his former employment at LSU while working as a non-tenure track Research Assistant Professor. He has managed the development of several software applications. In 2009, he proposed a novel hierarchical fusion-learning scheme for recognition of lung tumors; he developed a prototype Computed Aided Recognition application using MATLAB. After appointment as a Research Assistant Professor, he has extended his research on hierarchical learning and lung tumor detection. Currently, his focus is integrating deep learning into his hierarchical learning scheme for concept mining of 3D/4D spatio/temporal data.

The PI has been working on building a Biomedical Informatics Research Lab in the Department of Computer Science. In this lab, he will conduct research in biomedical informatics working closely with undergraduate and graduate students to prepare them for the work force. In this regard, the PI will develop a curriculum that teaches students how to conduct scientific research and translate the outcome of the research into technology.

A student research assistant who has good programming skills will be hired to assist in research, implementation of the proposed algorithms, and preparation of data for experiments. This project will greatly facilitate the PI's ability as an independent researcher who can manage a research project successfully. As a result, the PI will be more competitive nationally.

The PI foresees three barriers to be competitive at the national level:

- 1) Publication in the field of lung cancer recognition: It is important to collect more publications in the field to be a strong candidate at the national level.
- 2) Extensive preliminary results: The PI needs to conduct more experiments using a larger set of data to strengthen his proposed approach for lung cancer detection.
- 3) Independent researcher: The PI has supervised several projects in his previous employment and currently working on his startup project. If funded, this proposal will make the PI more competitive at nationwide funding opportunities.

Significance and Potential Impact: Although lung cancer is the second most commonly diagnosed cancer in both men and women, it is the leading cancer type that causes mortality in both men and women [1]. Lung nodule detection is a very challenging task. The research team in [2] explored the effect of the low-dose CT scans in cancer mortality. Utilizing either low-dose CT or chest radiography, they screened around 53K high lung cancer risk patients three times a year between August 2002 and April 2004. The results of their study show that there is a 20% reduction in mortality of the patients who were screened by low-dose CT scan. Even though the CT scan helps to reduce the mortality rate, the radiologists' decision may differ significantly in identification of the lung nodules from the CT scans. As an example, [3] shared the results of two radiologists' examinations over 25 CT scans; the results show that one of the radiologists detected 20 nodules, whereas the other radiologist detected 63 nodules from the same CT scans.

A Computer Aided Recognition (CAR) system increases the performance of the nodule detection substantially. The study conducted by [4] showed that the CAR system significantly reduced the number of false positives. The research in [5] that studied the effect of a CAR system in detection of small nodules shared the results of six radiologists' examinations over 52 CT scans with/without a CAR system. The results show that the CAR system improves a radiologist's performance considerably. In [6], the performance of the commercial software Lung-CAD VB10A and Siemens AG Healthcare were compared with the performance of two independent readers for detecting the pulmonary nodules in NELSON dataset. The study showed that sensitivity of CAR was 96.7% with a 3.7 FPs/scan and sensitivity of double reader was 78.3% with 0.5 FPs/scan. Therefore, the CAR system with a higher nodule detection rate can be a highly useful tool for radiologists to decrease the number of missed nodules, particularly, the small nodules in their early stages.

Innovation: The outcome of this proposed research will contribute to the increase of our knowledge about the most recent learning approach, deep learning, and its application to lung nodule detection. In object recognition, the idea of learning-based fusion is a new approach in contrast to using a voting scheme. The proposed idea is also novel in applying the approach for the lung tumor recognition. The proposed learning scheme hierarchically improves the decision about the class of an input 3D volume. In addition, association rule mining (ARM) based decision making will be integrated to the proposed hierarchical deep fusion network to improve recognition

performance. The proposed fusion approach and integration of ARM is a novel approach in the field.

Preliminary Results:

The PI's earlier preliminary study [7] shows that the performance of a hierarchical fusion-learning decision engine improves at each successive layer as seen in Figure 1. In this proposal, the PI aims to extend his earlier idea utilizing a deep learning scheme. The PI recently proposed a deep learning method that integrates a well-known data mining (DM) technique, Association Rule Mining, into a deep convolutional neural network (DCNN) at the final fully connected layer [8]. The PI used this DCNN+DM learning machine for prediction of action units (AU), which is a very challenging multi-label classification problem. The preliminary results show that this new approach is promising in recognition of facial expressions. Table 1 summarizes performance of DCNN integrated data mining techniques for action unit classification. In this proposal, the PI will explore the performance of such content representation for lung nodule detection within the deep fusion-learning schema.

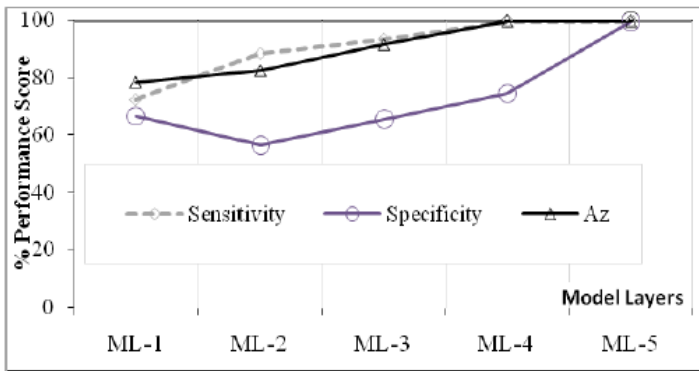


Figure 1 Classification performance through hierarchical fusion-learning

Table 1 Classification performance for AUs

| Method | Accuracy | Specificity | Sensitivity |
|----------------------|----------|-------------|-------------|
| DCNN | 63.77% | 49.53% | 47.53% |
| DCNN+Causality Max | 49.53% | 52.09% | 45.00% |
| DCNN+Causality Count | 47.53% | 49.21% | 47.67% |
| Decision Tree | 53.61% | 47.31% | 58.22% |
| DCNN+CBA | 69.70% | 47.36% | 64.88% |