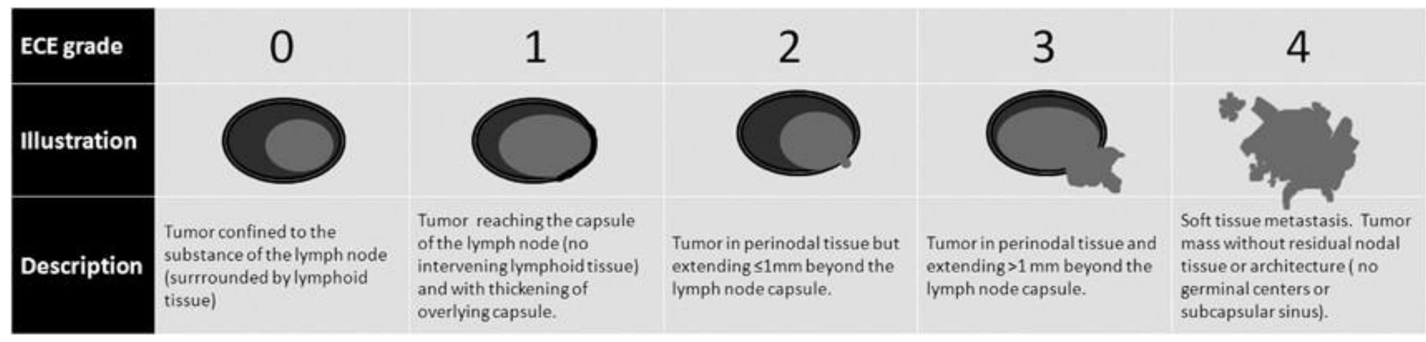
Artificial Intelligence-based Extracapsular Extension Prediction in Head and Neck Cancer Analysis

Yibin Wang[1], Toms V. Thomas[2], William N. Duggar[2], P. Russell Roberts[2], Linkan Bian[1], and Haifeng Wang[1]
[1]Department of Industrial and Systems Engineering, Mississippi State University (MSU), Mississippi State, MS 39762 USA
[2]Department of Radiation Oncology, University of Mississippi Medical Center (UMMC), Jackson, MS 39216 USA

American Association for Cancer Research®
FINDING CURES TOGETHER®

Abstract

Extracapsular extension (ECE) is a decisive indication for treatment planning of patients with head and neck squamous cell carcinoma (HNSCC). It is crucial to identify whether ECE occurs for HNSCC patient treatment.



(Lewis et al., 2011)

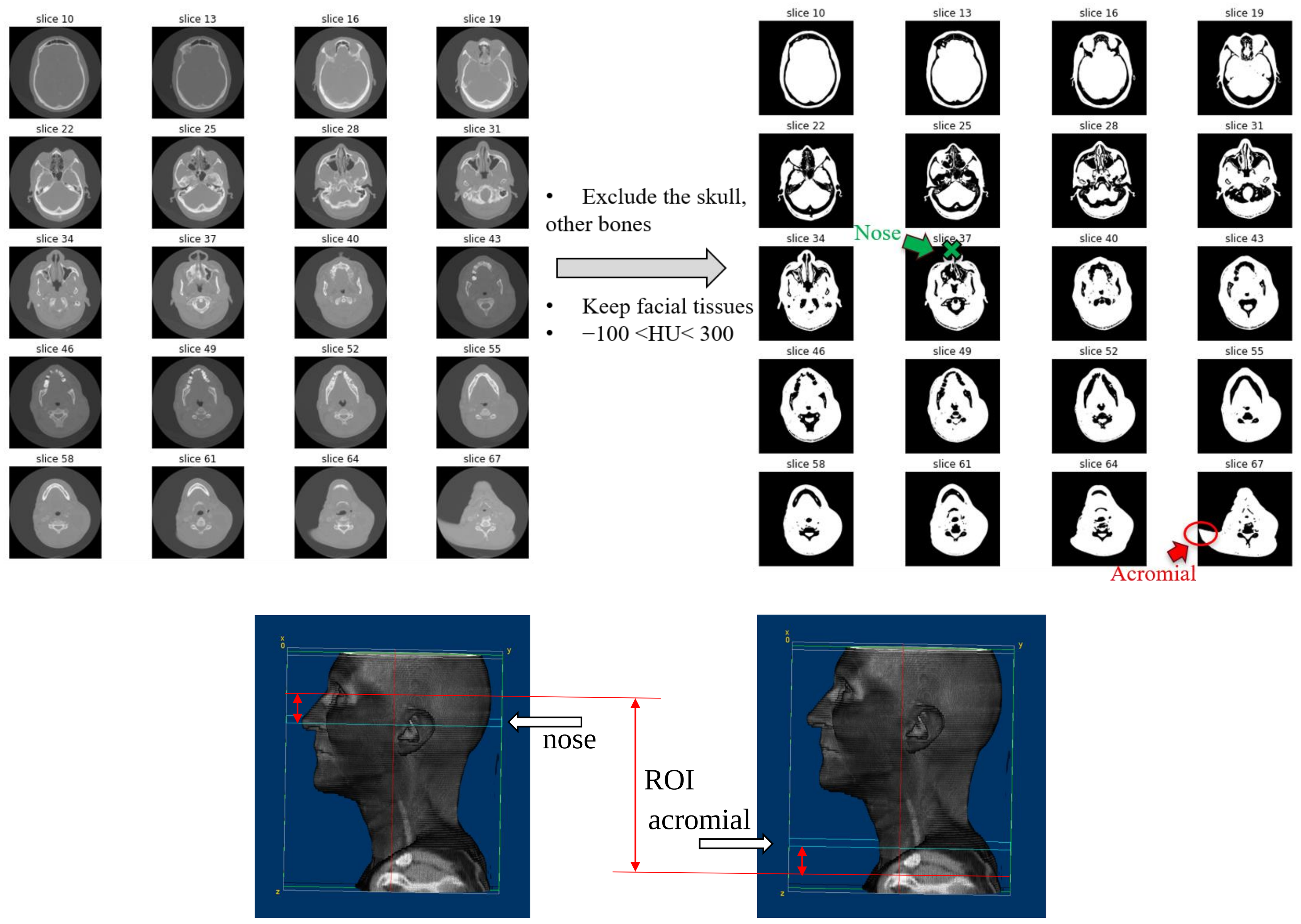
In this research, we propose a systematic machine learning approach to detect and classify ECE from computed tomography (CT) scans. Three machine learning models are implemented. The experimental results have demonstrated that our model is able to classify ECE and non-ECE patients. A test accuracy of 93.64% has been achieved in terms of patch-level ECE classification.

Methodology

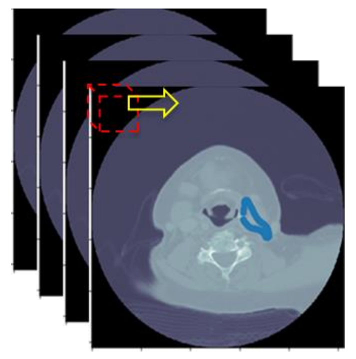
Sliding cube to extract small 3D patches → 3D Haralick texture features collection for ECE and non-ECE patches → Machine learning models with feature extraction and selection approaches → Classify the patient based on patch-level classification result

A. Data Prepossessing

- The HNSCC dataset collected by the UMMC. 82 patients were retrospectively reviewed with the diagnosis of HNSCC between 2008 and 2014. Built and trained machine learning models based on three selected patient data.
- For data prepossessing, we first narrowed the entire CT scans down to a few particular regions where lymph nodes locate, so as to remove irrelevant background.
- Hounsfield unit (HU) is widely used in CT scanning to express values in a standardized and convenient form. After applying HU threshold of -100 to 300, the bones are excluded and facial tissues are retained. Then the slices of nose and acromial can be identified in the CT scans. We selected ROI based on the slice of nose 3 centimeters upward and the slice of acromial 3 centimeters downward.



- A sliding cube approach to extract relatively small samples for training and testing within the CT scan volume.
- The size of the cubes is 20×20×20 pixels with 50% overlap.
- 15000 small patches are collected from ROIs in total with 7500 ECE samples and 7500 non-ECE samples.
- Then, 3D texture features are collected for each cube. Classification task are performed to differentiate ECE vs. non-ECE samples.



B. Classification Methods

- Machine learning: 1) gradient boosting; 2) random forest; 3) support vector machine
- Feature selection: 1) low variance threshold; 2) linear support vector classification
- Feature extraction: 1) principle component analysis; 2) feature agglomeration; 3) fast independent component analysis

Results

Five-fold cross validation applied. A set of results including accuracy, F-1 score, precision, recall has been collected. The best test accuracy of 93.64% has been achieved with the f1-score of 93.64%, precision of 93.66%, recall of 93.65% among the training scenarios.

Accuracy	FE	FS	Model	avgf1-score	avgprecision	avgrecall	w-f1-score	w-precision
0.9364	**None**	**None**	**RF**	**0.9364**	**0.9366**	**0.9363**	**0.9365**	**0.9366**
0.9284	None	None	GradBoost	0.9283	0.9289	0.9279	0.9284	0.9286
0.9182	PCA	None	RF	0.9181	0.9180	0.9183	0.9182	0.9183
0.9145	PCA	None	GradBoost	0.9144	0.9147	0.9143	0.9145	0.9146
0.9118	None	None	SVM	0.9114	0.9139	0.9108	0.9116	0.9133

avgf1-score: average f1-score; avgprecision: average precision; avgrecall: average recall; w-f1-score: weighted f1-score; w-precision: weighted precision

Conclusions

This research studies different machine learning models for ECE detection and classification in head and neck cancer. The experimental results show that random forest algorithm without any feature extraction and selection methods outperforms other methods with regard to model performance measurements. The implemented machine learning techniques are competitive among the automated diagnosis methods. Further study will focus on advanced deep learning models on ECE detection task.

Acknowledgement

Department of Industrial and Systems Engineering, Mississippi State University
Department of Radiation Oncology, University of Mississippi Medical Center

Reference

Lewis, J. S., Carpenter, D. H., Thorstad, W. L., Zhang, Q., & Haughey, B. H. (2011). Extracapsular extension is a poor predictor of disease recurrence in surgically treated oropharyngeal squamous cell carcinoma. Modern Pathology, 24(11), 1413-1420.

THE UNIVERSITY OF MISSISSIPPI MEDICAL CENTER

MISSISSIPPI STATE UNIVERSITY™ JAMES WORTH BAGLEY COLLEGE OF ENGINEERING