SYLVESTER COMPREHENSIVE CANCER CENTER
UNIVERSITY OF MIAMI HEALTH SYSTEM
UNIVERSITY OF MIAMI DEPARTMENT OF BIOMEDICAL ENGINEERING

Cancer Treatment Selection using Artificial Intelligence on Electronic Medical Records and Radiomics

T. Bejarano[1], M. Samuels[2], and I. B. Mihaylov[2]
1University of Virginia, Charlottesville, VA
2University of Miami, Miami, FL

JULY 12–16 2020 VIRTUAL
JOINT AAPM | COMP MEETING
EASTERN TIME [GMT-4]

INTRODUCTION

HPV-associated oropharyngeal squamous cell carcinoma (OPSCC) can be treated with definitive radiotherapy (RT), combined chemoradiotherapy (CRT), and/or transoral robotic surgery (TORS). TORS is a novel approach that provides a comparable overall survival (OS) rate (~80%) to radiotherapy with the potential for less toxicity and lower cost.[1-3] However, approximately 33% of TORS patients require adjuvant chemoradiation (CRT), with the highest toxicity profile due to extranodal extension (ENE) or positive surgical margins on the final pathology.[4-7] In retrospect, the decision for these patients to undergo TORS was unfavorable.

AIM

This study describes an artificial intelligence (AI) algorithm, designed to identify TORS suitable candidates with low risk of ENE or positive margins, based on EMR data and diagnostic imaging.

METHOD

- Electronic medical records (EMR) for **84** subjects were reviewed retrospectively.
- Patient data was obtained from the EMR, including: **demographics, history, vital signs, and labs.**
- In a second approach**, radiomics from diagnostic CT images** were also utilized.
  - After statistical pre-processing and remapping to equivalent numerical representations, three statistically significant textures were added to the analyses.
- Extracted data were used as an input into a **deep neural network (DNN**), with a probabilistic value for TORS being the post-surgical pathology.
- The DNN was implemented using Google's Tensor flow library and consisted of 40+ input nodes, 3 hidden layers, and 2 output nodes.
- A supervised learning technique was used to train the DNN. The treatment strategy was used as ground truth.
- A **5-fold cross-validation** method was used to evaluate the performance of the algorithm.

RESULTS

A DNN was successfully implemented, with variable parameters, using Google's TensorFlow library. The machine learning algorithm was trained, with EMR data and CT textures, to classify subjects into one of two groups: 1) TORS or TORS+RT and 2) TORS+CRT.

Group 1 patients were recommended for TORS, while Group 2 patients were contraindicated for TORS with a high risk for adjuvant CRT.

The DNN was trained in **< 1 min** and reproduced patient outcomes with an accuracy of ~70% (only EMR data) and ~80% (EMR + CT data).

EMR Data

The first prediction algorithm used only EMR data as input. The iterative change in training cost is displayed in **Figure 1**. Cost of the DNN quantifies the error between predicted and actual classifications; a lower cost indicates a better ability to fit the data.

The performance of the model is summarized using an ROC curve in **Figure 2**. The true positive rate (y-axis) and the false positive rate (x-axis) quantify the algorithm's ability to distinguishing between groups with an average AUC of 0.725.

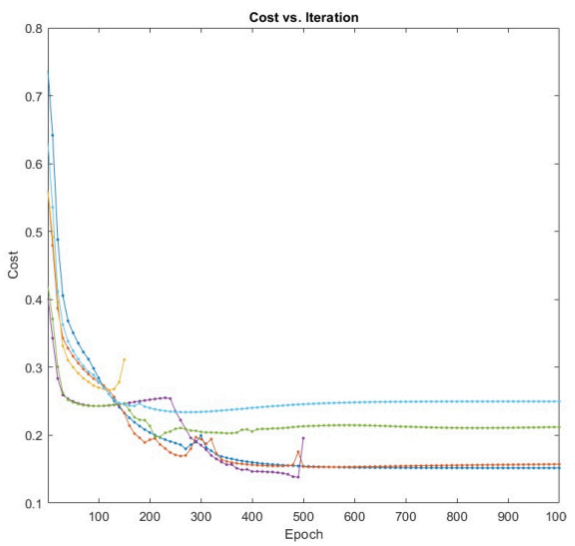

Figure 1. Cost vs. Iteration of the DNN using only EMR Data. The calculated cost for each k-fold cycle is displayed as a function epoch (training cycle). If the change in cost is greater than the previous two training cycles, the algorithm ends the training process. A shorter training process is for the purple and gold traces.

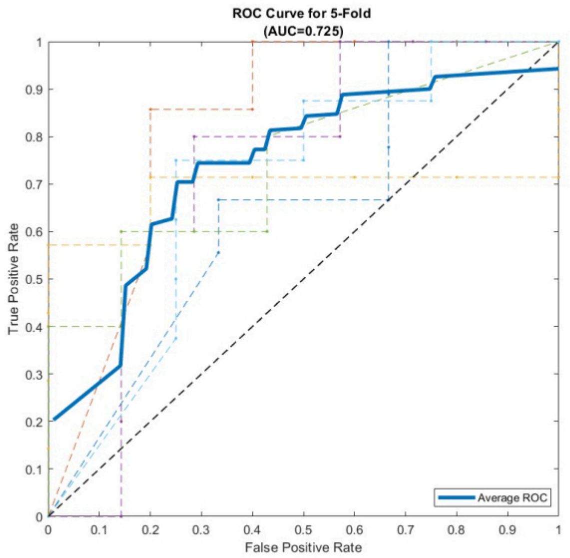

Figure 2. ROC (receiver operating characteristic) plot and AUC (area under the curve) for the DNN using only EMR data. The y-axis denotes the true positive rate, while the x-axis denotes the false positive rate. An AUC of 0.725 represents a model that makes the "correct" prediction 72.5% of the time. Dotted colored lines show the ROC curve for individual testing training runs of the k-fold validation. The dotted black line shows the reference to of an AUC of 0.5.

EMR + CT Data

The second approach utilized both EMR data and radiomics from CT images to train the ML algorithm. A visualization of the training cost for each 5-fold run is presented in **Figure 3**.

The overall performance of the model is summarized using an ROC curve in **Figure 4**. An average AUC of 0.844 was calculated. The higher mean AUC is indicative of a better ability to predict optimal treatment strategy.

The difference in model performance also underscores the sensitivity of the final output to input training parameters.

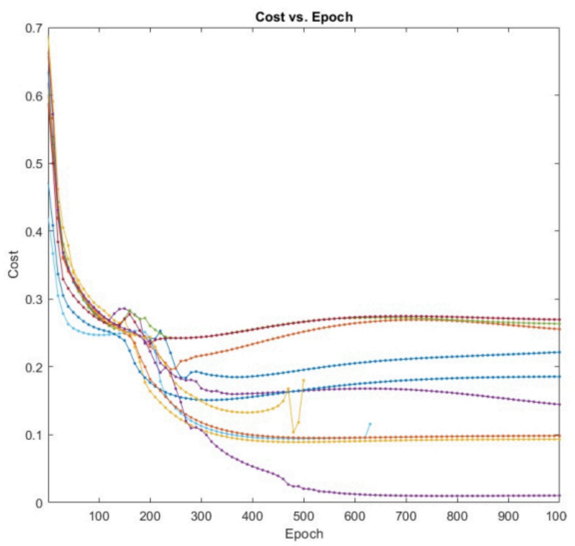

Figure 3. Cost vs. Iteration of the DNN using EMR Data and significant imaging Data. The calculated cost for each k-fold cycle is displayed as a function epoch (training cycle). If the change in cost is greater than the previous two training cycles, the algorithm ends the training process. A shorter training process is for the purple and gold traces.

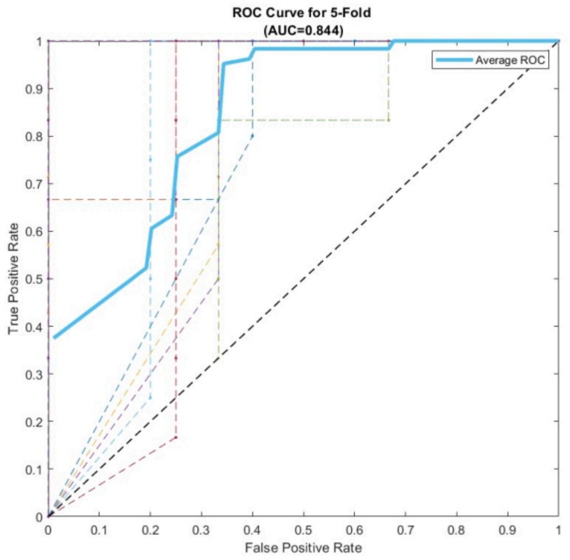

Figure 4. ROC (receiver operating characteristic) plot and AUC (area under the curve) for the DNN. The y-axis denotes the true positive rate, while the x-axis denotes the false positive rate. An AUC of 0.844 represents a model that can moderately distinguish between patient groups.

CONCLUSIONS

The presented findings indicate that EMR and CT textures can aid HPV-associated OPSCC patient treatment selection process. ML techniques can help in discriminating those who would ultimately require tri-modality therapy with the highest toxicity profile. With this knowledge, these patients would not be recommended for TORS.

It should be noted that the CTs were very heterogenous - acquired on different CT scanners with different protocols. EMR data in CNN require consistency across patients to be successfully implemented.

Future directions for research include expanding the database to include more patient information, enhancements on DNN architecture for better results, and determining specific variables in DNN decision making.

This algorithm can be easily extended to be used with other disease sites and treatment types.

ACKNOWLEDGEMENTS

We would like to thank University of Miami clinical and research teams who helped with patient treatments and data summary.

REFERENCES

1. **Chen MM**, Roman SA, Kraus DH, Sosa JA, Judson BL. Transoral robotic surgery: a population-level analysis. Otolaryngol Head Neck Surg. 2014;150(6):968–75.
2. **Moore EJ**, Hinni ML. Critical review: transoral laser microsurgery and robotic assisted surgery for oropharynx cancer including human papillomavirus related cancer. Int J Radiat Oncol Biol Phys. 2013;85(5):1163–7.
3. **Kelly K**, Johnson-Obaseki S, Lumingu J, Corsten M. Oncologic, functional and surgical outcomes of primary Transoral Robotic Surgery for early squamous cell cancer of the oropharynx: a systematic review. Oral Oncol. 2014;50(8):696–703.
4. **Hutcheson KA**, Holsinger FC, Kupferman ME, Lewin JS. Functional outcomes after TORS for oropharyngeal cancer: a systematic review. Eur Arch Otorhinolaryngol. 2015;272(2):463-71.
5. **Cooper JS**, Pajak TF, Forastiere AA, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. N Engl J Med. 2004;350(19):1937–44.
6. **Bernier J**, Domenge C, Ozsahin M, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. N Engl J Med. 2004;350(19):1945–52.
7. **Bernier J**, Cooper JS, Pajak TF, et al. Defining risk levels in locally advanced head and neck cancers: a comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (#9501). Head Neck. 2005;27(10):843–50.

CONTACT INFORMATION

Tatiana Bejarano, txb287@miami.edu.
Ivaylo Mihaylov, i.mihaylov@med.miami.edu.