Classifying Tumor Heterogeneity of Human Esophageal Cancer Biopsies by Dynamic Contrast OCT with Deep Learning

Zhen Hua[1], Shadia Jalal, M. D.[3], John Turek[2], and David D. Nolte[1]

[1]Department of Physics and Astronomy, Purdue University, 525 Northwestern Ave, West Lafayette, Indiana 47907, USA; [2]Department of Basic Medical Sciences, Purdue University, 625 Harrison Street, West Lafayette, Indiana 47907, USA; [3]Indiana University School of Medicine, Department of Medicine, Indianapolis, Indiana,USA

ABSTRACT

Tumor heterogeneity is one of the greatest obstacles standing in the way of successful cancer therapy. Cancer in a single patient is not a single disease, but is a host of related diseases, all of which need to respond to a single treatment regimen. We have completed the first human clinical trial in esophageal cancer using dynamic-contrast OCT (DC-OCT) based on full-frame digital holography to assess the spatial heterogeneity of biopsy response to platinum-based chemotherapy. A deep twin neural network successfully identified biopsy sub-phenotypes in the dynamic tissue response that enabled accurate prediction of patient treatment success.

Keywords: optical coherence tomography, machine learning, autoencoder, optical Doppler profiling, intracellular motion, chemotherapy

1. INTRODUCTION

Biodynamic imaging uses *en face* optical coherence tomography with off-axis digital holography, as the coherence gate, to perform tissue dynamics spectroscopy (TDS) of light scattering from living tissues to predict tissue response to applied drugs. This original form of dynamic-contrast OCT[1] has high dynamic range to detect subtle changes in intracellular motions in response to applied chemotherapy drugs. It also penetrates deeper into highly-scattering biopsy tissues than other forms of DC-OCT[2], although it trades off spatial resolution for depth. In clinical trials aimed at measuring patient chemoresistance using TDS, intra-sample variability poses a challenge for the prediction of patient response to therapy. However, the heterogeneous TDS drug spectrograms across subsets of a biopsy display several different dynamic phenotypes that can be used to assess the most representative biopsy subsets that correlate with patient response to chemotherapy.

1.1 Esophageal Cancer Clinical Trial for Chemosensitivity Testing

A clinical trial to assess the efficacy of using TDS to predict patient clinical outcomes to therapy enrolled 28 patients who presented with esophageal adenocarcinoma at the IU School of Medicine Hospital between April 2015 and January 2020. The trial enrollment was halted at the onset of the Covid 19 pandemic. A pinch biopsy was performed for normal diagnostic purposes from which a small volume of between 30 to 70 mm^3 was placed in growth medium to maintain tissue health for up to 8 hours. Samples were transported in chilled medium from the IUSM to Purdue University within 4 hours of the surgery. The biopsy was cut into approximately 1 mm^3 sections to yield between 16 to 32 samples that were immobilized using poly-L lysine in a multi-well plate. Each well received carboplatin+taxol (CT) or cisplatin+5fu (CF) combination therapies, as well as their single-agent components, and the biodynamic response was monitored for up to 18 hours after application of the drug. Of the 28 enrolled patients, 25 received clinical scores.

In the TDS spectrograms of the 28 patients, at least three distinct phenotypes were found in human esophageal cancer among 664 biopsy samples measured in this trial. The characteristic spectrograms are shown in Figure 1a), showing blue-shift, mid-frequency enhancement, and red-shift, respectively. The similarity matrix for the samples is shown in Figure 1b) after unsupervised clustering, with the associated dendrogram in Figure 1c). The similarity matrix is approximately block-diagonal, corresponding to the three most-dominant spectrogram phenotypes. These average spectral shifts represent the background upon which specific drug-response shifts occur when the sample wells are

Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXVII, edited by Joseph A. Izatt, James G. Fujimoto, Proc. of SPIE Vol. 12367, 123670L · © 2023 SPIE · 1605-7422 · doi: 10.1117/12.2647602

treated with the chemotherapy agents. The background behavior can represent the overall metabolic state of the biopsy sample, which may affect how the sample responds to the chemotherapy. This generates a well-to-well heterogeneity that contains more information than spectrograms averaged over many wells. Therefore, the goal of this work is to tap this well-by-well heterogeneity information using deep learning in neural networks.

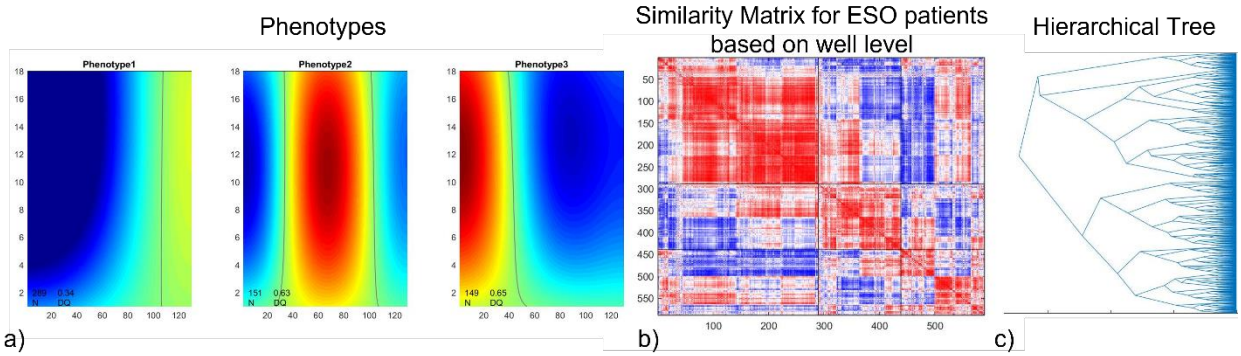

Figure 1 a) Three main phenotypes of spectrogram from esophageal biopsies. b) Spectrogram clustering result of all esophageal biopsies. c) Hierarchical tree of all esophageal biopsies based on their feature vectors.

1.2 Deep Learning

To address the problem of intra-biopsy heterogeneity in the context of chemosensitivity testing[3] for cancer patients, we introduce a Deep-Twin Neural Network[4] methodology. Deep learning in neural networks favors high input diversity for two reasons. First, highly-variable data are difficult to analyze using conventional multivariate analysis. Second, the action of multiple hidden layers in a neural network learns nonlinear relationships among the input variables, especially when the inputs are conditionally dependent on other inputs. The highly variable character of OCT tissue dynamics spectra among the biopsy samples for a single patient presents an ideal opportunity for deep learning. Our previous work on tissue dynamics of cancer biopsies[5] used averaging approaches which suppressed the conditional dependences. Here, we retain the full well-to-well variability.

The Deep-Twin Neural Network is shown in Figure 2. It is composed of two identical neural networks that are presented with a pair of data inputs in the form of feature vectors that have been extracted from the drug-response spectrograms. The loss function is minimized through contrastive loss in which the transformed coordinates in latent space are nearby when the inputs are from the same class, or else are greater than a set margin when they are from different classes. The latent-space is followed by a decoder that reconstructs the input space from the low-dimensional latent space representation with a reconstruction loss. There is in addition an L1 regularization term in the loss function that minimizes the input weights in the first layer of neurons to help identify the input channels that carry the most information. The network is trained using feature vectors from individual wells that are formed through bootstrap aggregation using sampling with replacement from the full well-based set of feature vectors.

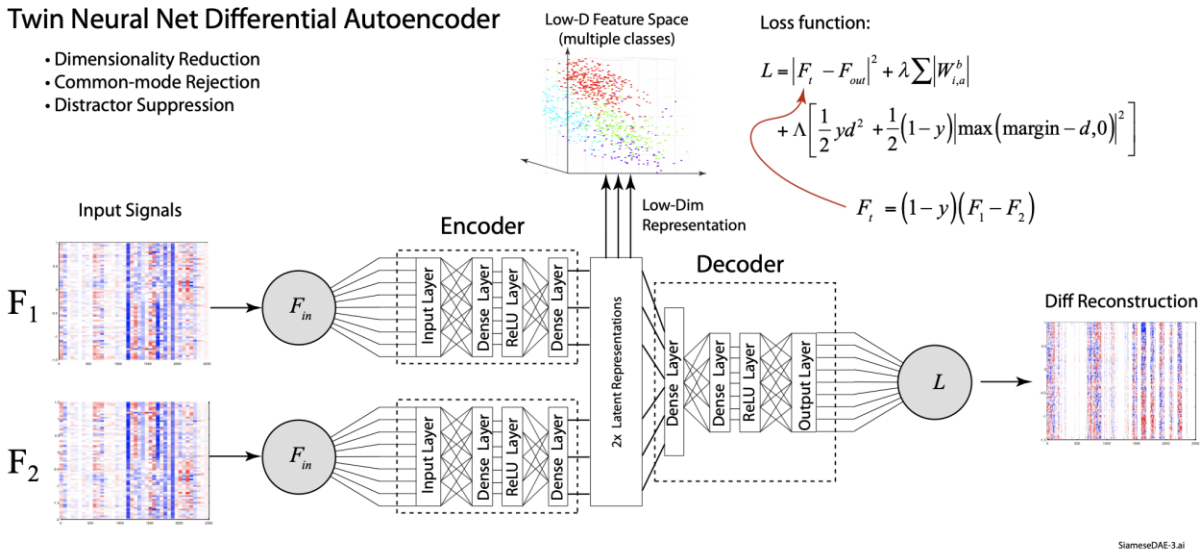

Figure 2 The "twin" network for dimensionality reduction consists of a network pair. Data structures are applied in pairs, then the contrastive loss is used to update the parameters. The network in this figure has two layers of ReLUs that are fully connected to the output neurons. The latent-space neurons are the input to a decoder that reconstructs the pair differences. The loss function optimizes locality-sensitive hashing, decoder reconstruction, and L1 regularized input weights.

2. CANCER PATIENT CHEMOSENSITIVITY CLASSIFICATION

The well-based phenotyping that is illustrated in Figure 1 enables the exclusion of background behavior that correlates most weakly with patient drug sensitivity. This down-selection enriches the relevant drug responses that are input to the neural network, improving overall accuracy in patient chemosensitivity classification.

2.1 Post-selection of Well-based Phenotypes

The post-selection of well-based phenotypes that have enhanced high frequencies (blue shifts), and the exclusion of the wells that had low-frequency enhancements (red shifts), improved overall accuracy of the chemosensitivity assay from approximately 85% to 95%. Because tissue-dynamics spectroscopy is based on Doppler light scattering, there is an immediate interpretation of the red shift of the non-representative samples. The red shift signifies that the average intracellular activity is decreasing throughout the duration of the 12 to 16-hour assays. Lower activity is indicative of reduced metabolic efficiency that can be interpreted as an overall reduction in sample health. Conversely, blue shifts indicate a tissue that is increasing its metabolic health through the duration of the assay. In simple terms, this means that the excluded wells with red-shifted samples contained samples whose overall health was declining. The response of these red-shifted tissues to the exogenous challenge of the chemotherapy drugs is not well correlated with the patient outcome. Therefore, this well-based phenotyping approach highlights the dual advantages of tissue-dynamics spectroscopy to assess the overall health of biopsy samples as well to assess the likelihood of patient response to chemotherapy.

2.2 Patient Response to Chemotherapy

The Twin Neural Network performs dimensionality reduction, taking the high-dimensional feature vectors of the TDS spectrograms and transforming the data into a low-dimensional latent space. This process is illustrated in Figure 3a) and b). The patients are clustered by the contrastive loss algorithm into three classes: chemoresistant, partial response, and chemo-sensitive. The hold-out validation is shown in Figure 3c) as a network structure and in Figure 3d) as relative probabilities for a patient to be classified into one of the three classes. The network was trained on 22 of the patients

with clinical outcomes. An additional 3 patients had non-representative phenotypes, and 3 more patients did not have clinical outcomes (eso19, eso4 and eso26).

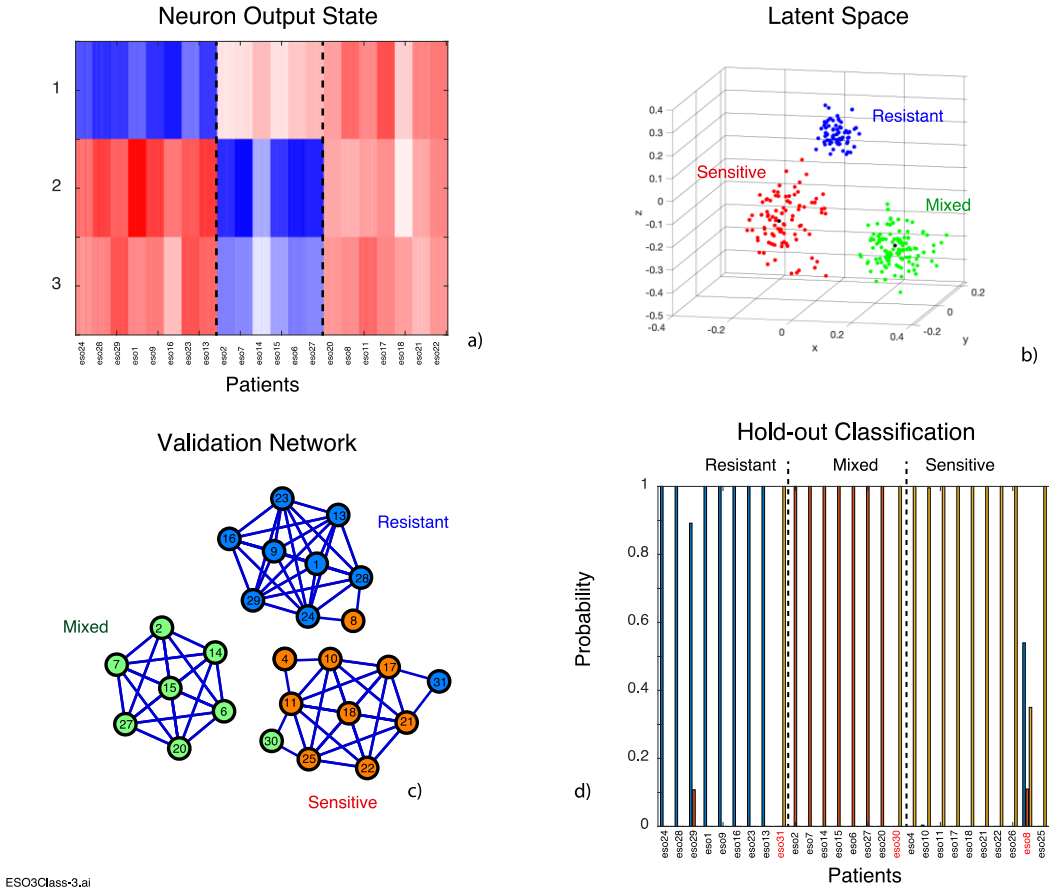

Figure 3 Neural network transformation to a low-dimensionality latent space. a) The output states of the Twin neural network for three esophageal cancer patient classes (sensitive, mixed, resistant) and three neurons. b) Three-dimensional representation of a batch of Patient Charts showing strong clustering into 3 classes. c) Network graph generated from the latent-space representation. d) One-hold-out patient classification for 22 training patients and 6 non-training patients.

REFERENCES

[1] D.D. Nolte, R. An, J. Turek, and K. Jeong, "Tissue dynamics spectroscopy for phenotypic profiling of drug effects in three-dimensional culture", Biomed. Opt. Express 3, 2825-2841 (2012).

[2] C. Apelian, F. Harms, O. Thouvenin, A. C. Boccara, Dynamic full field optical coherence tomography: subcellular metabolic contrast revealed in tissues by interferometric signals temporal analysis. *Biomedical Optics Express* **7**, 1511-1524 (2016).

[3] H. Choi, Z. Li, H. Sun, D. Merrill, J. Turek, M. Childress, and D. Nolte, "Biodynamic digital holography of chemoresistance in a pre-clinical trial of canine B-cell lymphoma," Biomed. Opt. Express 9, 2214-2228 (2018).

[4] D. D. Nolte and L. J. Pyrak-Nolte, "Monitoring Fracture Saturation with Internal Seismic Sources and Twin Neural Networks", *Journal of Geophysical Research: Solid Earth* **127** e2021JB023005 (2022).

[5] Z. Li, R. An, W. M. Swetzig, M. Kanis, N. Nwani, J. Turek, D. Matei, and D. Nolte, "Intracellular optical doppler phenotypes of chemosensitivity in human epithelial ovarian cancer," Scientific Reports, vol. 10, no. 1, p. 17354, 2020/10/15 (2020)