Abstract # 1393

NantOmics

Automated adeno/squamous-cell NSCLC classification from diagnostic slide images: A deep-learning framework utilizing cell-density maps

CONTRIBUTING RESEARCHERS

Mustafa I. Jaber[1], Liudmila Beziaeva[2], Christopher W. Szeto[3], John Yahya Elshimali[2], Shahrooz Rabizadeh[1], and Bing Song[1]

[1] NantOmics LLC., Culver City, CA; [2] NantHealth, Culver City, CA; [3] NantOmics LLC., Santa Cruz, CA

BACKGROUND

The most common form of lung cancer, **non-small cell lung cancer** (**NSCLC**), is further classified into two major histopathological subtypes: ~40% **Adenocarcinoma (LUAD)**, and ~30% **squamous cell carcinoma (LUSC)**.

Classifying tumors accurately is important for prognosis and therapy decisions, but requires costly pathologist review.

Here we present an automated **algorithm to differentiate LUAD and LUSC subtypes using diagnostic whole slide images** (WSIs).

METHODS

- 488 subtyped NSCLC high-resolution diagnostic WSIs were obtained from TCGA sources. Samples were randomly split into 338 (70%) training and 150 (30%) testing sets.
- All 100 micron 2D color patches were transformed into 1D descriptive vectors using the inception v3 deep learning framework.
- We trained an expert system (ResNet-34 convolutional neural network) to identify tumor patches from adjacent-normal tissue, and such regions were analyzed separately. See Figures 1 and 2.
- We also trained an algorithm to count cells in each 2D color patch. The cell count system has modules for Color Deconvolution, Local Drain, and Watershed Segmentation. Sample cell maps are shown in Figure 3.
- The generated tumor mask and cell density map (in addition to 1D descriptive vectors of 100 micron 2D color patches in target WSI) were used as inputs into the adeno/squamous-cell NSCLC classifier in Figure 4.
- 1D descriptive vectors were placed into 10 discrete bins based on their cell-density (i.e. 20-30 cells per patch, 30-40, etc. up to >110 cells per patch).
- Ten LUAD/LUSC linear SVM classifiers (one for each cell-density bin) were trained on such transformed data.
- Subtype prediction in the held-out 30% of unseen testing samples was achieved by averaging subtype predictions from the 10 subsequent models

RESULTS

Figure 1. **Expert guided tumor/normal masking.** A browser-based tool was developed to capture expert opinion on tumor or normal tissue points. These points were used to generate whole-slide masks. The mask was then iteratively refined by selection of more tumor and normal points (i.e. human-in-the-loop training)

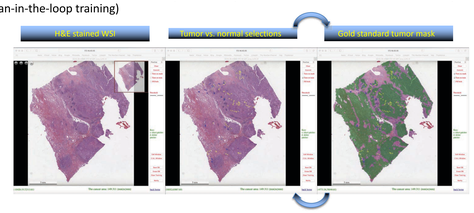

Figure 2. **Assessment of trained tumor/normal masks**. Examples of deep-learning generated tumor mask, based on the expert-guided system in Figure 1.

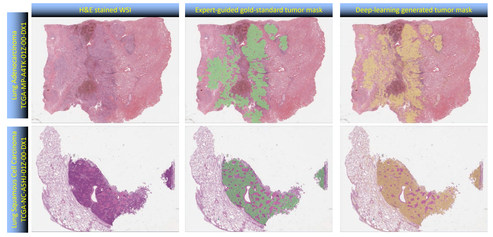

Figure 3. **Cell Density Maps** generated by Color Deconvolution, Local Drain, and Watershed Segmentation modules developed by Nant.

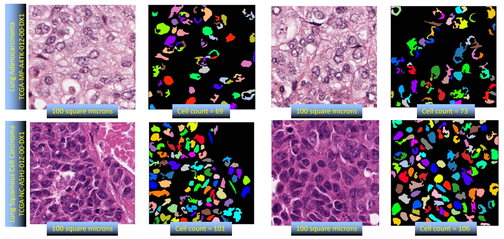

Figure 4. **Block diagram of the automated adeno/squamous-cell NSCLC classifier**. The classifier is based on tumor patches (represented by 1D vector in logits layer of Inception-v3 Neural Net) of pre-specified ranges of cell count used in tower of SVMs.

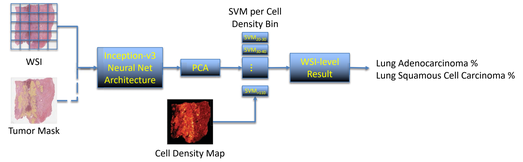

Figure 5. **Performance evaluation of the classifier.** SVMs in the proposed system were trained using tumor patches only, however area under the ROC curves of three runs of 150 test WSIs were used in the the evaluation: **(a)** all valid patches of test WSIs were used. i.e. the entire WSI was considered tumor, **(b)** tumor patches of test WSIs were used (based on the deep-learning framework in Figure 2), and **(C)** adjacent normal patches alone which may provide additional insights into tumorigenesis/invasion mechanisms upon further analysis.

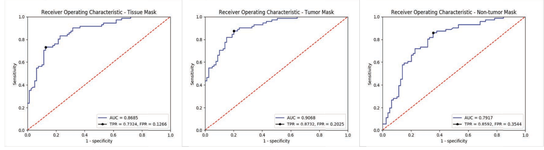

Table 1. **Performance of the Nant classifier relative to state-of-the-art methods.** Our algorithm showed comparative or better performance while maintaining higher spatial resolution of tissues used to give overall adeno/squamous call in test images.

Test Set		Patch size in square microns	Average number of test patches per WSI	Area under the ROC Curve	Accuracy
Yu, K.-H. et al. (2016) Train and test on TCGA diagnostic & frozen tissue WSIs		250	10	0.7500	-
Coudray, N. et al. (2018) Train on TCGA frozen tissue WSIs and test on diagnostic WSIs		256	-	0.8825	-
		1024	-	0.9180	-
Graham, S. et al. (2018) – 64 WSIs from 2017 Computational Precision Medicine Challenge		256	2,056	-	81.00%
Nant Classifier Train and test on TCGA diagnostic WSIs	Test using all tissues	100	14,906	0.8685	80.67%
	Test using tumor mask	100	6,722	**0.9068**	**83.33%**
	Test using adjacent normal mask	100	8,184	0.7917	74.67%

KEY FINDINGS

- **An automated NSCLC subtype classifier based on cell-count based tumor patched was developed by training on an expert system and utilizing a novel method of cell density mapping.**
- **The proposed system achieved an area under the ROC Curve of 0.9068 in test samples, corresponding to a classification accuracy of 83.33%.**
- **The (heretofore excluded) adjacent normal regions were classified correctly and almost as accurately as tumor regions (74.7%).**

CONCLUSIONS

This fully-automated histopathology-based subtyping classifier generates maps of regions-of-interest within WSIs, providing novel spatial information on tumor organization. For example, our results on test data show tumor patches of 100 square microns in size with 60 to 100 cells distinguish LUAD from LUSC better than other cell-density ranges. Moreover, this classifier reveals that adjacent normal tissue may provide additional insights into tumorigenesis/invasion mechanisms. This deep-learning system outperforms similar efforts using CellProfiler features (Yu *et al.* 2016), and provides additional explanatory information beyond systems with similar performance (Courdray *et al.* 2018).

REFERENCES

1. Yu, K.-H. et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nature Communications, Vol. 7, Article number: 12474 (2016).
2. Coudray, N. et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning, Nature Medicine, Vol. 24, pages 1559–1567 (2018).
3. Graham, S. et al. Classification of lung cancer histology images using patch-level summary statistics. Medical Imaging 2018: Digital Pathology conference, International Society for Optics and Photonics, Vol. 1058119 (2018).

ACKNOWLEDGEMENT

Thank you to Patricia Spilman for editing help.