Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy

Lili Zhang[1,2#], Yongzheng Wu[1#], Bin Zheng[3#*], Lizhong Su[3], Yuan Chen[4], Shuang Ma[4], Qinqin Hu[4], Xiang Zou[5], Lie Yao[5], Yinlong Yang[6], Liang Chen[5], Ying Mao[5], Yan Chen[1*], Minbiao Ji[1,2*]

[1] State Key Laboratory of Surface Physics and Department of Physics, Fudan University, Shanghai 200433, China

[2] Human Phenome Institute, Multiscale Research Institute of Complex Systems, Key Laboratory of Micro and Nano Photonic Structures (Ministry of Education), Fudan University, Shanghai 200433, China

[3] Department of Otolaryngology, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014, China

[4] Department of Pathology, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014, China

[5] Department of Neurosurgery, Department of Pancreatic Surgery, Huashan Hospital, Fudan University, Shanghai 200040, China

[6] Department of Breast Surgery, Fudan University Shanghai Cancer Center; Department of Oncology, Shanghai Medical College; Fudan University, Shanghai 200040, China

[#] These authors contributed equally.

***CORRESPONDENCE:** zhengbin@hmc.edu.cn, yanchen99@fudan.edu.cn, minbiaoj@fudan.edu.cn

Supplementary materials

Table S1: Descriptive statistics of the case series used in the deep learning model, which includes 45 cases for training and 33 patients for testing. Note the cases with grey shading were also used in the survey comparing SRS to H&E histology with frozen sections.

Patient	Admission Number	Sex	Age	Diagnosis	Date
Train-01	20328749	M	46	Neoplastic	20160426
Train-02	20353752	M	64	Neoplastic	20160613
Train-03	14063532	M	39	Neoplastic	20161112
Train-04	20389461	M	62	Neoplastic	20161117
Train-05	20553852	F	44	Neoplastic	20161208
Train-06	20649021	F	56	Neoplastic	20161208
Train-07	20349210	M	53	Neoplastic	20161211
Train-08	15208439	M	65	Neoplastic	20161211
Train-09	13481307	M	58	Neoplastic	20170315
Train-10	15066877	M	54	Neoplastic	20170315
Train-11	15106000	M	61	Neoplastic	20170315
Train-12	17165301	F	50	Neoplastic	20170316
Train-13	20225629	M	43	Neoplastic	20170317
Train-14	20558765	F	63	Neoplastic	20170317
Train-15	20279153	M	56	Neoplastic	20170322
Train-16	13789538	M	73	Neoplastic	20170322
Train-17	16269310	M	60	Neoplastic	20170324
Train-18	18743201	F	55	Neoplastic	20170324
Train-19	19745521	F	72	Neoplastic	20170324
Train-20	20592246	M	48	Neoplastic	20170426
Train-21	20592211	M	56	Neoplastic	20170426
Train-22	20594436	M	42	Neoplastic	20170426
Train-23	14049400	M	60	Neoplastic	20170920
Train-24	20548457	M	54	Neoplastic	20181201
Train-25	15146972	F	53	Normal	20170920
Train-26	12347527	M	68	Normal	20160816
Train-27	20367211	M	58	Normal	20160613
Train-28	12570137	F	59	Normal	20160816
Train-29	13742068	M	47	Normal	20160816
Train-30	20448209	M	58	Normal	20161117
Train-31	20305120	F	37	Normal	20161215
Train-32	20578776	M	68	Normal	20161215
Train-33	20573974	M	51	Normal	20170209
Train-34	20576405	F	26	Normal	20170209
Train-35	20189042	F	46	Normal	20170209
Train-36	12006970	M	42	Normal	20170209
Train-37	18342589	M	62	Normal	20170316
Train-38	19723649	M	54	Normal	20170316
Train-39	20625992	M	45	Normal	20170330

Train-40	20586327	F	59	Normal	20170330
Train-41	20563147	F	41	Normal	20170330
Train-42	20614821	F	65	Normal	20170330
Train-43	90169106	F	73	Normal	20181201
Train-44	90169497	M	59	Normal	20181201
Train-45	90169322	F	55	Normal	20181201
Test-01	12481038	M	61	Neoplastic	20170315
Test-02	15184278	M	85	Neoplastic	20170316
Test-03	20632246	M	78	Neoplastic	20170317
Test-04	18992682	F	62	Neoplastic	20170322
Test-05	17652803	F	55	Neoplastic	20170322
Test-06	20618306	M	66	Neoplastic	20170324
Test-07	16334096	F	43	Neoplastic	20170324
Test-08	14063512	M	57	Neoplastic	20170426
Test-09	14163385	M	38	Neoplastic	20170920
Test-10	14903762	F	64	Neoplastic	20170920
Test-11	20329524	M	62	Neoplastic	20160426
Test-12	20340756	M	66	Neoplastic	20160613
Test-13	12759295	M	59	Neoplastic	20160816
Test-14	13953812	F	49	Neoplastic	20161112
Test-15	20395032	F	53	Neoplastic	20161112
Test-16	20148209	M	63	Neoplastic	20161117
Test-17	20548552	F	57	Neoplastic	20161208
Test-18	20541661	M	61	Neoplastic	20161208
Test-19	20450800	M	66	Neoplastic	20161211
Test-20	14482136	M	50	Neoplastic	20161211
Test-21	20530000	M	47	Normal	20161117
Test-22	15018669	F	30	Normal	20161211
Test-23	20551749	F	52	Normal	20161211
Test-24	14178163	M	34	Normal	20161215
Test-25	15018825	F	54	Normal	20161215
Test-26	20550717	M	66	Normal	20161215
Test-27	16034000	F	59	Normal	20170209
Test-28	12015470	F	62	Normal	20170209
Test-29	15001928	M	31	Normal	20170209
Test-30	20579775	M	39	Normal	20170209
Test-31	20585571	M	47	Normal	20170209
Test-32	20548781	F	51	Normal	20170209
Test-33	20594576	M	48	Normal	20170324

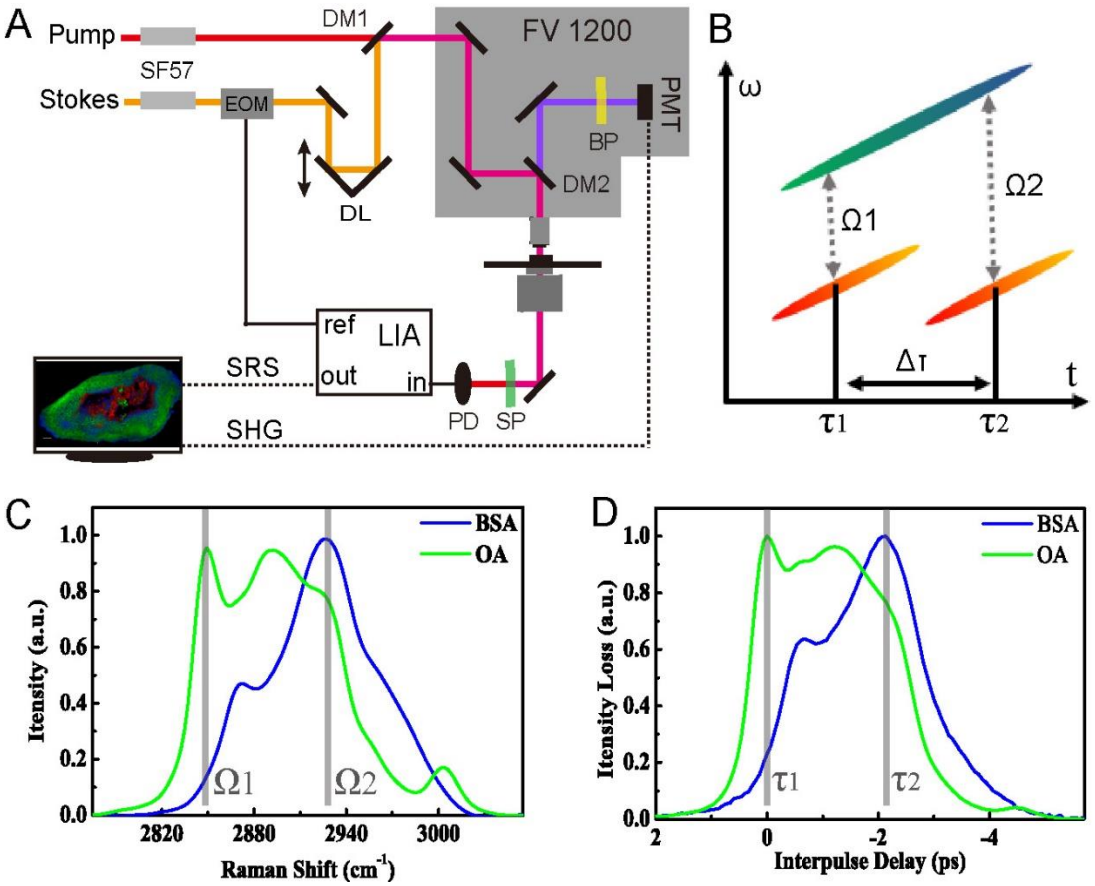

Figure S1. A, Optical layout of SRS/SHG microscope. B, Sketch for the spectral focusing of stimulated Raman scattering (SRS). C-D, Spontaneous Raman (C) and SRS spectra (D) of standard lipid (Olic Acid, OA) and protein (bovine serum albumin, BSA) samples, the marked grey line represent the Raman shifts and interpulse delays of lipid and protein. EOM: electro-optical modulator; LIA: lock-in amplifier; DM: dichroic mirror; PMT: photomultiplier tube; BP: band-pass filter; SP: short-pass filter; PD: photodiode.

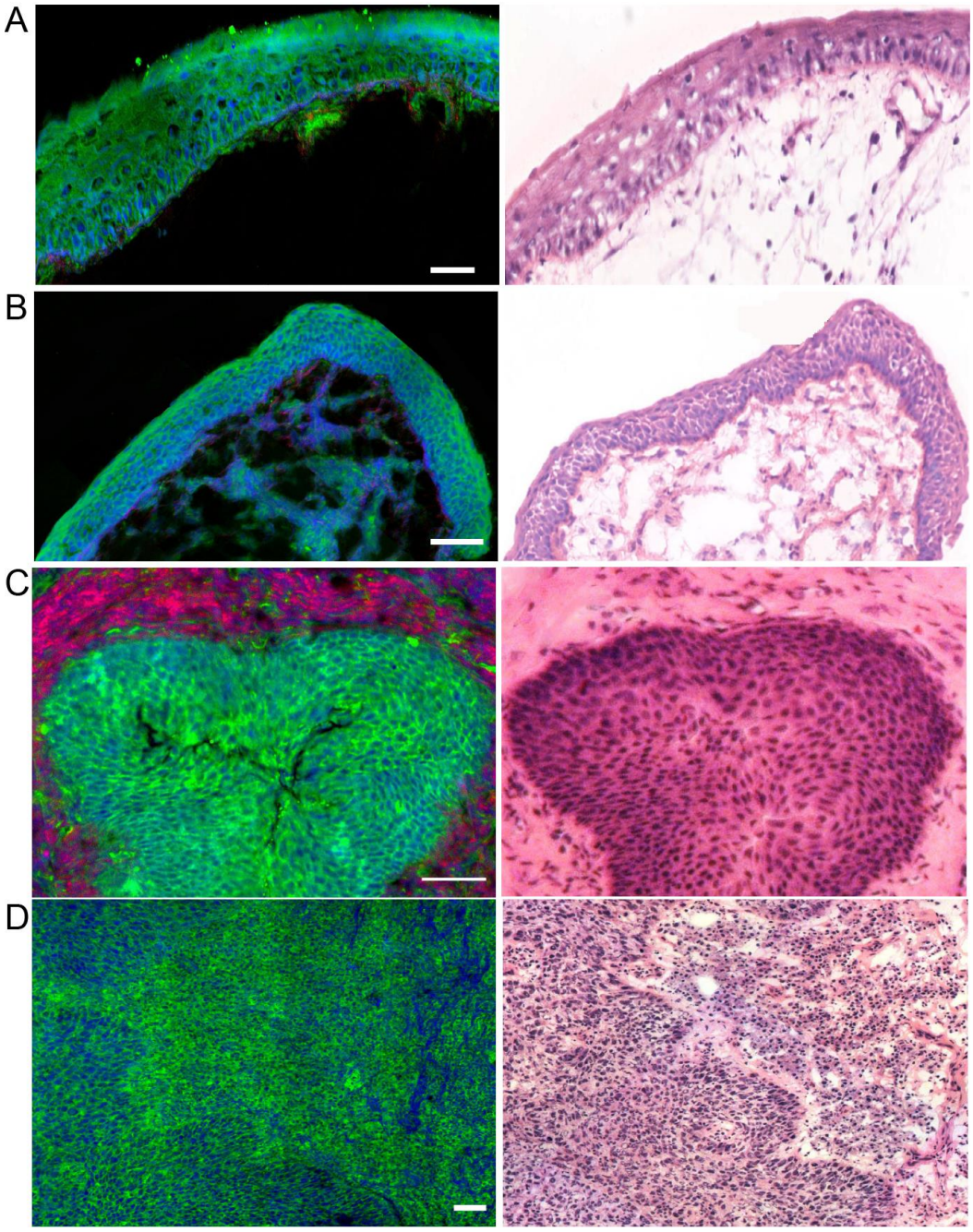

Figure S2. Typical normal (A-B) and neoplastic (C-D) SRS images used in the quantitative survey. Scale bars: 50 μm.

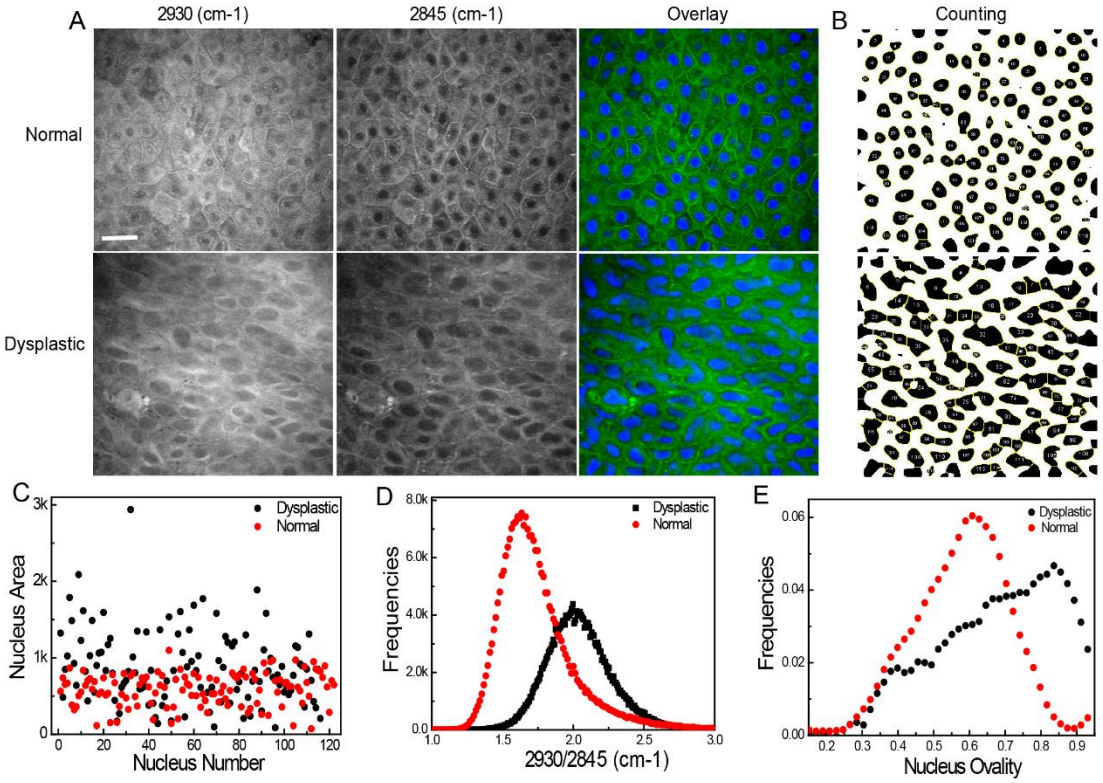

Figure S3. Quantitative analysis the SRS images from normal and neoplastic laryngeal tissues. A. SRS images of normal and neoplastic laryngeal tissues. B. Cellular nucleus counting, showing 122 and 116 nucleus for normal and neoplastic SRS images in A. C. Distribution of cellular nucleus areas, showing an increased area for neoplastic tissue compared with normal tissue. D. Histogram of the intensity ratio of the SRS image, indicating increased protein content for neoplastic tissue compared with normal tissue. E. Histogram of the nucleus ovality, showing an increased eccentricity for neoplastic tissue compared with normal tissue. Scale bars: 30 μm.

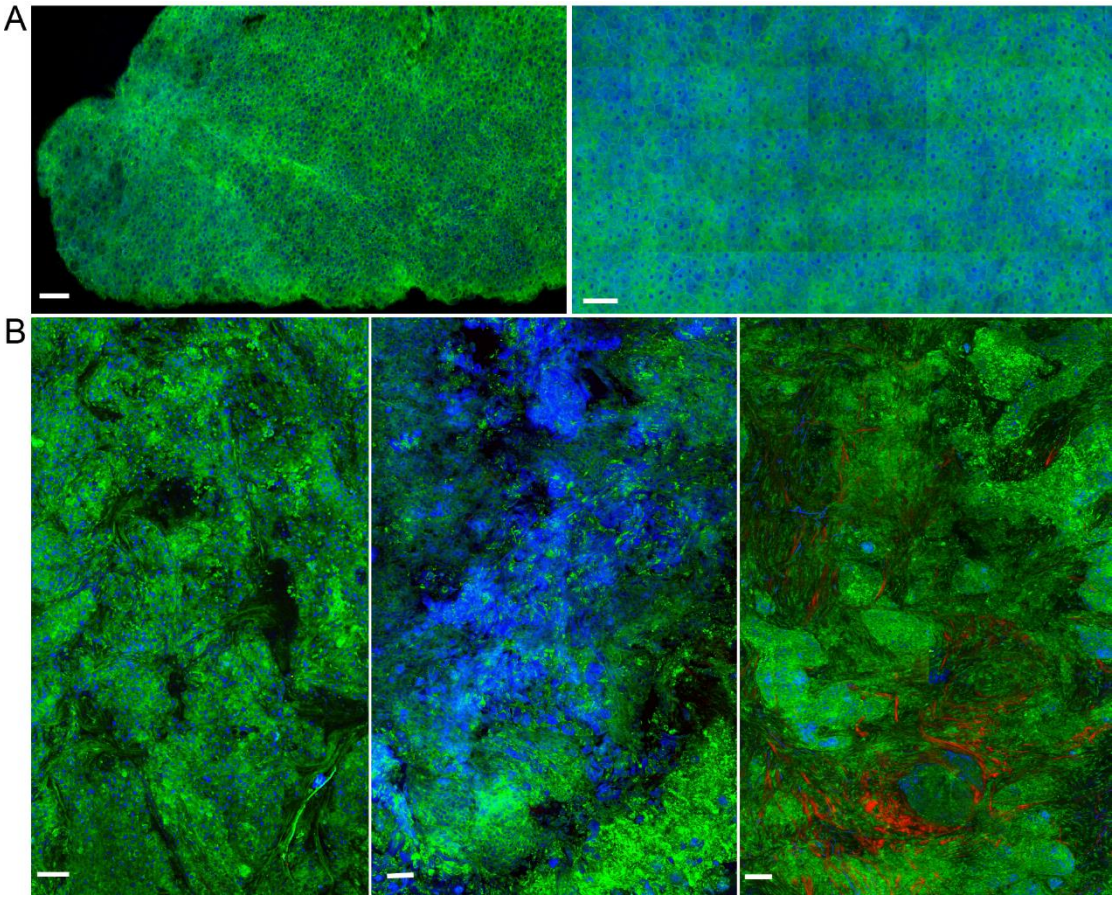

Figure S4. Typical images used in deep learning algorisms. A. Normal tissues. B. Neoplastic tissue. Scale bars: 100 μm.

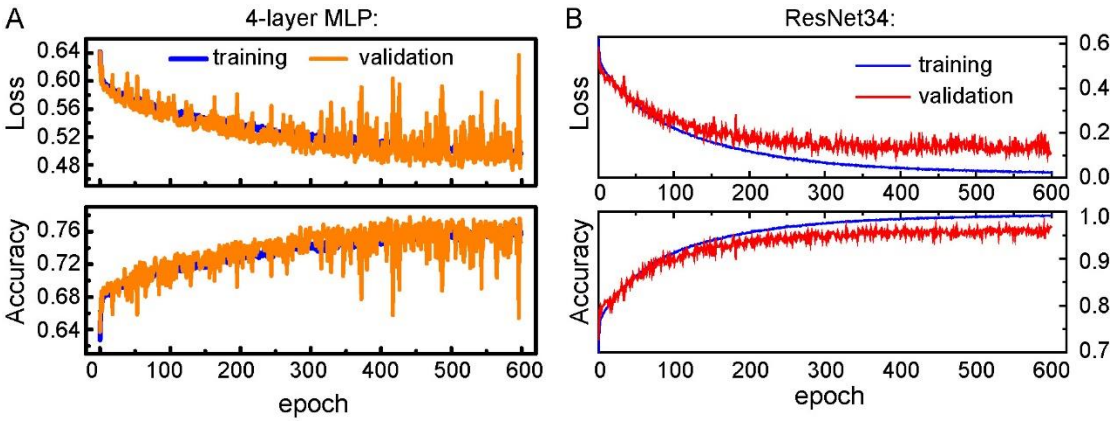

Figure S5. Comparison between ResNet34 and 4 layer MLP models. Plotted are the training and validation loss and accuracy of the two models. Comparing the number of weight parameter, our 34-layer ResNet contains 21,285,698 weight parameters, for image tiles of 200*200 pixels, and 3 color channels; whereas the 4-layer MLP model has: 200*200*3*1024=122,880,000 (assuming 1024 neurons in the first layer). Although the number of weight parameters in MLP is larger than ResNet34, the performance of ResNet34 is much better than MLP as can be seen from the loss function and accuracy of the two models when training the same datasets.