Variational Networks for Joint Image Reconstruction and Classification of Tumor Immune Cell Interactions in Melanoma Tissue Sections

Alexander Effland[1], Michael Hölzel[2], Teresa Klatzer[3], Erich Kobler[3], Jennifer Landsberg[4], Leonie Neuhäuser[1], Thomas Pock[3], Martin Rumpf[1]

[1]Institute for Numerical Simulation, University of Bonn
[2]Institute of Clinical Chemistry and Clinical Pharmacology, University of Bonn
[3]Institute of Computer Graphics and Vision, Graz University of Technology
[4]Department of Dermatology and Allergy, University of Bonn
alexander.effland@ins.uni-bonn.de

Abstract. Immunotherapy is currently revolutionizing the treatment of cancer. Detailed analyses of tumor immune cell interaction in the tumor microenvironment will facilitate an accurate prediction of a patient's clinical response. The automatic and reliable pre-screening of histological tissue sections for tumor infiltrating immune cells (TILs) will support the development of TIL-based predictive biomarkers for checkpoint immunotherapy. In this paper, a learning approach for image classification is presented, which allows various pattern inquires for different types of tissue section images. The underlying trainable reaction diffusion model combines classification and denoising. The model is trained using a stochastic generation of training data. The effectiveness of this approach is demonstrated for immunofluorescent and for Hematoxylin and Eosin (H&E) stained melanoma section images. A particular focus is on the classification of TILs in the proximity to melanoma cells in an experimental melanoma mouse model and in human melanoma. This new learning approach for images of melanoma tissue sections will refine the strategy for the practical clinical application of biomarker research.

1 Introduction

The clinical success of immune checkpoint inhibitors has proven the importance of immune surveillance of tumors for the survival of patients with a variety of malignancies, particular with melanoma. However, primary and acquired resistance to immunotherapy limit the therapeutic efficacy for many cancer patients. Thus, there is a growing need to identify predictive biomarkers and to enhance our understanding of the complex interactions between the immune system and tumor cells. Using melanoma tissue sections, it has been shown that tumor infiltrating CD8+ immune cells and their distribution within the tumor microenvironment are promising predictive biomarkers [1,2].

This paper deals with a variational network learning approach for the detection of these biomarkers. There are diverse deep learning approaches for the automatic detection and classification of cell markers in the context of computer-aided cancer diagnosis. E.g. in [3], Sirinukunwattana et al. employed a spatially constrained convolutional neural network for the detection of relevant cell nuclei using a neighboring ensemble predictor to incorporate the spatial structure of the cells. In [4], a classification of epithelial and stromal cells based on a deep convolutional neural network is performed, where the quality of the outcome is compared to handcraft feature extracted datasets. For a recent overview of deep learning methods for classification and segmentation tasks in digital pathology we refer to [5].

In this paper, we take into account the trainable nonlinear reaction diffusion model (TNRD) [6] from image denoising, where the parameters in each time step are degrees of freedom. We minimize an L^2 loss functional over these degrees of freedom. The acquisition of training data from real histological sections is out of reach, since for each section a substantial amount of user interaction of an experienced pathologist would be needed and a robust loss function decay requires large sample sizes of such tissue sections. Hence, we train the proposed model with training data generated via a stochastic cell distribution algorithm, mimicking structure, color distribution and noise of real histological sections of cancer tissues.

2 Materials and Methods

In what follows, we will present the learning approach, the generator of training data and the image acquisition.

2.1 A "deeper" variational network

Assume, we have given a blurry and noisy $N_1 \times N_2$ color image $\mathbf{u}_0^{rgb} \in \mathcal{U}_{rgb} = ([0,1]^3)^{N_1 \times N_2}$ together with an initial segmentation mask $u_0^m \in [0,1]^{N_1 \times N_2}$. Our aim is to compute a combined color- and segmentation image $\mathbf{u} = (\mathbf{u}^{rgb}, u^m) \in \mathcal{U}_{rgbm} = ([0,1]^4)^{N_1 \times N_2}$ by using a "deeper" extension of the TNRD model. The proposed model performs N_t projected gradient steps of the form

$$\mathbf{u}_{t+1} = \mathrm{proj}_{\mathcal{U}_{rgbm}}(\mathbf{u}_t - \nabla E_t(\mathbf{u}_t)) \quad \text{for } t \in \{0, \dots, N_t - 1\}, \tag{1}$$

where the projection operator $\mathrm{proj}_{\mathcal{U}_{rgbm}}$ refers to a simple pointwise truncation to the interval $[0,1]$. Observe that the effective step size of the gradient descent will be defined by means of the parameters of the learned model. Since this model evolves the gradient of a time-dependent variational model E_t, it is also referred to as "variational network" [7]. In this paper we propose to use the following variational model:

$$E_t(\mathbf{u}) = \sum_{\mathcal{F}_2} \phi_t^2(K_t^2 \phi_t^1(K_t^1 \mathbf{u})) + \tfrac{\lambda_t}{2} \|R\mathbf{u} - \mathbf{u}_0^{rgb}\|_2^2. \tag{2}$$

The first term is a smoothness term, which is based on a composition of convolution operators and pointwise non-linear functions. The learned convolution operator $K_t^1 = (K_t^{1,1}, \dots, K_t^{1,N_{f_1}}) : \mathcal{U}_{rgbm} \mapsto \mathcal{F}_1$ implements a set of N_{f_1} 2D convolutions to compute the feature space $\mathcal{F}_1 = \mathbb{R}^{N_1 \times N_2 \times N_{f_1}}$. Next, learned non-linear functions $\phi_t^1 : \mathcal{F}_1 \mapsto \mathcal{F}_1$ are applied in a pointwise manner in order to apply a non-linear transformation to the feature space $\mathcal{F}_1$. This feature space is then again filtered by a second set of convolution operators $K_t^2 = (K_t^{2,1}, \dots, K_t^{2,N_{f_2}}) : \mathcal{F}_1 \mapsto \mathcal{F}_2$ which implement N_{f_2} learned convolution operators. This gives the second feature space $\mathcal{F}_2 = \mathbb{R}^{N_1 \times N_2 \times N_{f_2}}$ which is once more transformed by applying learned non-linear functions $\phi_t^2 : \mathcal{F}_2 \mapsto \mathcal{F}_2$. The final regularization term is then given by taking the sum over all elements of $\mathcal{F}_2$.

The second term is a data fidelity term, which is given by the squared L^2-norm between the current image $\mathbf{u}$ and the initial image $\mathbf{u}_0$. The linear operator $R : \mathcal{U}_{rgbm} \mapsto \mathcal{U}_{rgb}$ is used to restrict the norm to the RGB channels of $\mathbf{u}$.

In our iterative scheme (1), we evolve the gradient of the variational model (2). By virtue of the chain rule, the gradient is given by

$$\nabla E_t(\mathbf{u}) = (K_t^1)^\top \phi_t^{1\prime}(K_t^1 \mathbf{u})(K_t^2)^\top \phi_t^{2\prime}(K_t^2 \phi_t^1(K_t^1 \mathbf{u})) + \lambda_t R^\top (R\mathbf{u} - \mathbf{u}_0^{rgb}) \,, \quad (3)$$

where $\phi_t^{1\prime}$ and $\phi_t^{2\prime}$ denote the derivatives of the pointwise non-linear functions.

Similar to the TNRD model [6], the functions ϕ_t^i are parameterized using Gaussian radial basis functions with weights $w_t^i \in \mathbb{R}^J$ and the convolution operators K_t^1 and K_t^2 are given by small convolution kernels. Thus, the overall parameters for the model that have to be learned are $\theta = (K_t^1, w_t^1, K_t^2, w_t^2, \lambda_t)_{t=1}^{N_t}$. As initial data for the denoising and classification we consider $\mathbf{u}_0$ with $\mathbf{u}_0^{RGB}$ being the actual tissue section image and u_0^m as uniformly distributed noise.

Let us now briefly describe our learning procedure. Given a set of sample pairs $((\mathbf{u}_0^{rgb}, u_0^m)_s, (\mathbf{g}^{rgb}, g^m)_s)_{s=1}^S$, the training problem is defined as

$$\min_{\theta \in \mathcal{T}} \sum_{s=1}^{S} \frac{1}{6} \|(\mathbf{u}_T^{rgb})_s - (\mathbf{g}^{rgb})_s\|_2^2 + \frac{1}{2} \|(u_T^m)_s - (g^m)_s\|_2^2 \,, \quad (4)$$

where $\mathbf{g}^{rgb}$ and g^m define the target images and masks, respectively. The set $\mathcal{T} = \{(K_t^1, w_t^1, K_t^2, w_t^2, \lambda_t) : \|K_t^{1,i}\|_2 \leq 1,\ \|K_t^{2,j}\|_2 \leq 1,\ \lambda_t \geq 0,\ i = 1...N_{f_1},\ j = 1...N_{f_2},\ t = 1...N_t\}$ is the set of admissible model parameters. Observe that these constraints are essential to avoid a scaling problems between the filter and the corresponding non-linear functions. The coefficients in (4) ensure a balance of the loss between image reconstruction and cell detection. To solve this variational problem, the Adam optimizer [8] is employed, where a projection of θ onto $\mathcal{T}$ is performed after each gradient step.

2.2 Stochastic generator of training data

Our training data for the two prototypic scenarios considered here is based on a comparably simply structural description: ellipsoidal shapes for the different cell

Table 1. Cell specific data of all scenarios.

	category	mean color	semi axes (pixels)	number/placement
(1)	tumor cell	green	$20-35$	60 cells, random placement
	tumor nucleus	light blue	$8-12$	only 60% visible
	immune cell	red	$8-20$	40 cells, random placement, no overlapping
	immune cell nucleus	blue	$7-8$	only 60% visible
	stroma/ dediff. melanoma	blue	$5-15$	40 cells, random placement, no overlapping
(2)	tumor cell	light purple	$15-25$	5 of 7 Voronoi regions filled, no overlapping
	tumor nucleus	purple	$3-10$	only 60% visible
	immune cell	violet	$4-8$	2 of 7 Voronoi regions filled, no overlapping

components, a stochastic placement of these primitives, and an additive noise generation for the cells and background color. We consider two different tissue section image classes:

The aim of scenario 1 is the detection of the direct contact of immune cells with melanoma cells in immunofluorescent stained melanoma sections. Scenario 2 analyzes the pattern/invasion of tumor infiltrating immune cells in H&E stained melanoma sections.

The essential guidelines for the data generation are (for details see Table 1):

- All cells and cell nuclei are modeled as ellipses, where both the length of the semi axes and the inclination are drawn from a uniform distribution.
- The cells are placed randomly or in cells of a precomputed random Voronoi tessellation.
- Overlapping of cells or the location of cell midpoints inside other cells can be excluded for certain cell types.
- The colors of cells/nuclei are drawn from a multivariate normal distribution with expectation/variance extracted from a small sample of real images.
- Realistic noise is added in the composition of the geometric primitives, with covariance matrices estimated via a covariance analysis of the above sample.
- Finally, background noise is added with average color and variance again extracted from this sample.

Examples of the training data for the two problem classes are shown in Fig. 2.

Our method classifies certain types of immune cells, depending on a fluorescence marking and their local cell environment. Depending on the scenario, immune cells are marked as classified if they are located in a proximity of tumor

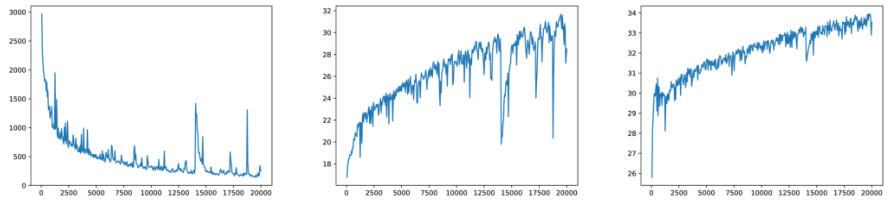

Fig. 1. Loss function (first column), PSNR for RGB channel (second column) and masking channel (third column) for scenario 1.

cells (scenario 1) or if the concentration of immune cells in a neighborhood is sufficiently large (scenario 2). Local cell concentrations are computed based on the volumetric measure of cells in circular neighborhoods. A classified immune cell is indicated with a value 1.0 in the marking channel at all underlying pixels on a 0 background, i.e. for a circular cell concentration beyond a threshold 0.4 (scenario 1) or 0.2 (scenario 2). Since this measure is prone to small perturbations of the cell structures, we could improve the stability of the classification in scenario 1 by assigning to all pixels in the ground truth mask the value 0.5 if the cell concentration is in the range $[0.2, 0.4]$.

2.3 Acquisition of histological sections

To analyze immune cell and melanoma cell interactions, we used ten representative immunofluorescent stains for the immune cell marker CD45 in red, the melanocytic marker gp100 in green and the nuclei DAPI in blue of murine melanomas (staining protocol has been described previously in [9]) and ten H&E stained human melanoma metastases of the Skin Cancer Center Bonn of the University Hospital in Bonn. H&E stains were performed according to standard protocols. Stained sections were examined with a Leica DMBL immunofluorescence microscope, all images were acquired with a JVC digital camera KY-75FU.

3 Results

We consider $N_t = 10$ steps in the TNRD model and 200 random 500×500 training images. In each step, $N_{f_1} = 24$ different filters of size $11 \times 11 \times 4$ and $N_{f_2} = 24$ filters of size $5 \times 5 \times 24$, and for the activation function $J = 31$ Gaussian radial basis functions in the interval $[-1.2, 1.2]$ were used. Following [8], we performed 20000 iterations of the Adam optimizer with step size 0.001 using a batch size of 2 and the exponential decay rate parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The plots of the associated loss function and the peak signal-to-noise ratio (PSNR) for the RGB and the masking channel are shown in Fig. 1.

Fig. 2 depicts pairs of RGB images and associated masking channels for training input data (first pair), computed pair of denoised image and estimated

classification (second pair) and ground truth image and classification (third pair) for scenario 1 (first row) and scenario 2 (fourth row). In the remaining rows, the input data (first column), the denoised image (second column), and the corresponding classification (third column) are shown for scenario 1 (second/third row) and scenario 2 (fifth/sixth row). As an outcome, nearly all immune cells to be classified are actually detected.

4 Discussion

In this work, we investigated a deep learning approach using variational networks for joint image reconstruction and segmentation on melanoma tissue sections to detect direct interactions of immune cells with melanoma cells and patterns of tumor infiltrating immune cells. We were able to provide spatial localization and distribution of immune cells within the tumor microenvironment. Obvious limitations of the method are the restriction to ellipsoidal shaped cells and cell nuclei and the currently small number of cell or nuclei types, e.g. in the first image in Fig. 2 associated with scenario 2 small regions of extracellular matrix structures are erroneously classified as immune cells, probably because these structures are not treated explicitly as cell types in the data generation process. Future analyses should consider the complex heterogeneity of tumor cells.

References

1. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJM, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature. 2014;515:568–571.
2. Jacquelot N, Roberti MP, Enot DP, Rusakiewicz S, Ternès N, Jegou S, et al. Predictors of responses to immune checkpoint blockade in advanced melanoma. Nat Commun. 2017;8.
3. Sirinukunwattana K, Raza SEA, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Trans Med Imaging. 2016;35(5):1196–1206.
4. Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A Deep Convolutional Neural Network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing. 2016;191:214–223.
5. Janowczyk A, Madabhushi A. Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases. J Pathol Inform. 2016;7.
6. Chen Y, Pock T. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1256–1272.
7. Kobler E, Klatzer T, Hammernik K, Pock T. Variational Networks: Connecting Variational Methods and Deep Learning. In: Ger Pattern Recognit Conf; 2017. p. 281–293.
8. Kingma DP, Ba JL. Adam: A Method for Stochastic Optimization. In: International Conference on Learning Representations; 2015. .
9. Landsberg J, Kohlmeyer J, Renn M, Bald T, Rogava M, Cron M, et al. Melanomas resist T-cell therapy through inflammation-induced reversible dedifferentiation. Nature. 2012;490:412–416.

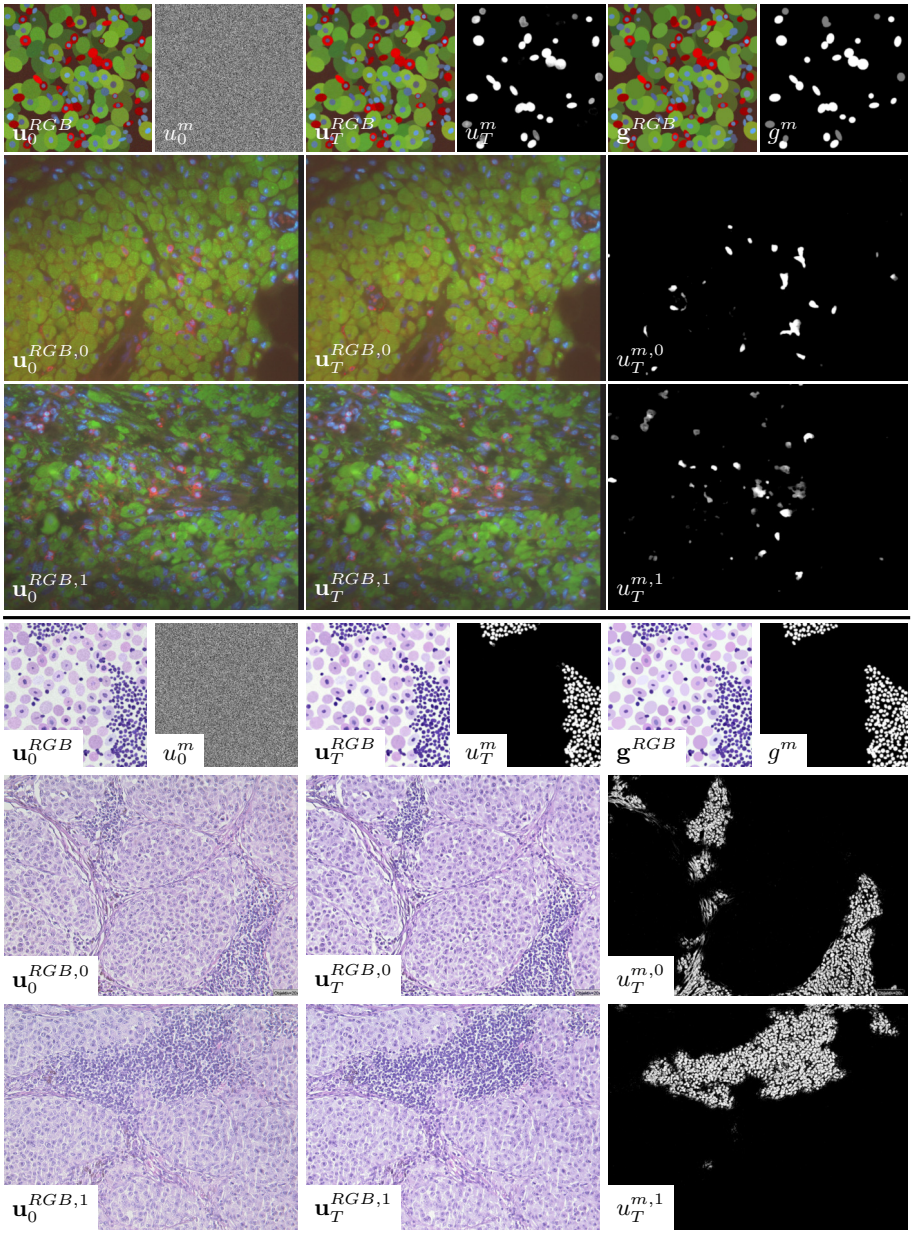

Fig. 2. Training data (pairs of image and mask of input, output and ground truth) for scenario 1 (first row) and scenario 2 (fourth row). $\mathbf{u}_0^{RGB,i}$, $\mathbf{u}_T^{RGB,i}$ and $u_T^{m,i}$ $(i = 0, 1)$ for histological sections in scenario 1 (second/third row) with two representative images of immunofluorescent stains of melanoma: with blue (DAPI, cell nuclei), red (CD45, immune cell marker), green (gp100, melanocyte marker), and scenario 2 (fifth/sixth row) H&E stains of melanoma.