

AUTOMATIC DETECTION OF NECROSIS, NORMOXIA AND HYPOXIA IN TUMORS FROM MULTIMODAL CYTOLOGICAL IMAGES

Gustavo Carneiro¹

Tingying Peng²

Christine Bayer³

Nassir Navab² *

¹ Australian Centre for Visual Technologies (ACVT), University of Adelaide, Australia

² Computer Aided Medical Procedure (CAMP), Technical University of Munich, Germany

³ Department of Radiation Oncology, Technical University of Munich, Germany

ABSTRACT

The efficacy of cancer treatments (e.g., radiotherapy, chemotherapy, etc.) has been observed to critically depend on the proportion of hypoxic regions (i.e., a region deprived of adequate oxygen supply) in tumor tissue, so it is important to estimate this proportion from histological samples. Medical imaging data can be used to classify tumor tissue regions into necrotic or vital and then the vital tissue into normoxia (i.e., a region receiving a normal level of oxygen), chronic or acute hypoxia. Currently, this classification is a lengthy manual process performed using (immuno-)fluorescence (IF) and hematoxylin and eosin (HE) stained images of a histological specimen, which requires an expertise that is not widespread in clinical practice. In this paper, we propose a fully automated way to detect and classify tumor tissue regions into necrosis, normoxia, chronic hypoxia and acute hypoxia using IF and HE images from the same histological specimen. Instead of relying on any single classification methodology, we propose a principled combination of the following current state-of-the-art classifiers in the field: Adaboost, support vector machine, random forest and convolutional neural networks. Results show that on average we can successfully detect and classify more than 87% of the tumor tissue regions correctly. This automated system for estimating the proportion of chronic and acute hypoxia could provide clinicians with valuable information on assessing the efficacy of cancer treatments.

Index Terms— Cytological Microscopic Images, Multimodal Classification, Classifier Combination

1. INTRODUCTION

Hypoxic regions (i.e., a region deprived of adequate oxygen supply) are commonly present in human tumors, and they are usually associated with poor clinical prognosis [1]. Specifically, it has been observed that the efficacy of common treatments (such as standard radiotherapy, some O₂-dependent chemotherapy, photodynamic therapy, and immunotherapy) is limited in such hypoxic regions. Hypoxia can be generally

*G.C. thanks the Alexander von Humboldt Foundation (Fellowship for Experienced Researchers). This work was partially supported by the Australian Research Council Centre of Excellence for Robotic Vision (project number CE140100016).

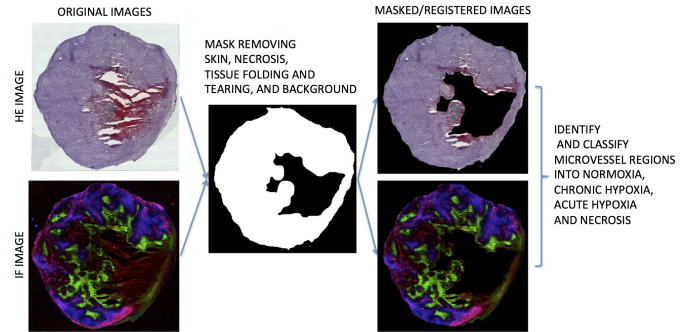


Fig. 1. Manual classification of microvessel regions from HE (top) and IF (bottom) images of the same histological specimen and the rough mask delineating the vital tumor region in both images. The slice represents one whole tumor cryosection, where the pink color channel in HE denotes necrotic region and the three color channels in IF represent three fluorescence stains (red denotes microvessels, green displays hypoxia, and blue shows perfusion).

classified into chronic or acute [1], depending on its causes, duration and consequences [1], where chronic hypoxia is characterized by limitations in oxygen diffusion from tumor microvessels into surrounding tissue, while acute hypoxia is represented by local disturbances in perfusion [1]. The main result of chronic hypoxia is a limitation of tumor growth while acute hypoxia can promote tumor aggressiveness [2]. There is also evidence that fluctuating hypoxia levels with time indicates the development of aggressive survival strategies, such as local invasion, metastasis, and acquired treatment resistance. Therefore, a successful clinical treatment critically depends on the use of medical imaging data for first detecting vital and necrotic tumor tissues, and then for classifying vital tissue regions into normoxia or hypoxia and then to further classify the hypoxia into chronic or acute [3].

Currently, this classification is performed manually using (immuno-)fluorescence (IF) and hematoxylin and eosin (HE) stained images of a histological specimen (see Fig. 1). These two image modalities need to be registered [4] to allow the delineation of vital tumor tissue in both images (see mask in the middle of Fig. 1), but note that this delineation

is usually rough, and patches of necrotic tissue can still be found within the mask. The vital tumor tissue regions are then classified based on the categorization of the oxygenation and perfusion status of individual microvessel regions (see Fig. 2) [3]. Hence, each microvessel region must firstly be detected from the IF image and then classified into necrotic or vital using the HE and IF images, and then the vital tissue is classified into normoxia, chronic and acute hypoxia using the three color channels of the IF image, where red denotes microvessel, green represents hypoxia and blue indicates perfusion. Currently, this is a lengthy manual process that requires expertise, which is not widespread in clinical practice.

We propose a new methodology that takes the masked and registered IF and HE images of the same histological specimen and automatically detects and classifies the microvessel regions into: a) necrotic, b) normoxic, c) chronic hypoxic or d) acute hypoxic. The proposed solution uses as input multimodal data, i.e., the IF and HE images. Furthermore, instead of proposing a single classifier for this problem, we propose a principled combination of the following state-of-the-art classifiers: 1) Adaboost [5], 2) linear support vector machine (linear SVM) [6], 3) random forest [7], and 4) deep convolutional neural networks [8]. The combination of the classification results is done in two different ways: 1) using a joint classification from the results of the four classifiers; and 2) using a conditional random field (CRF) model [9] with four unary potentials (representing the results of the four classifiers) and a binary potential that encodes contrast dependent labeling homogeneity. Using a dataset with 16 pairs of IF and HE stained images from different specimens (but each pair of IF and HE images from the same specimen), which have been partially annotated (i.e., only a subset of the microvessel regions is annotated), we run a 4-fold cross validation that produces 87% classification accuracy, which shows that this system is potentially useful in clinical practice. In particular, this system will reduce significantly the time needed to annotate new IF and HE stained images of a histological specimen and will alleviate the need of experts that are usually absent in clinical sites.

2. LITERATURE REVIEW

In this paper we are concerned with multimodal cytology image analysis using a combined automated classification method that detects and classifies sparse regions into four different classes. We are not aware of other approaches that perform a similar task, so we will discuss briefly other papers that use a subset of the contributions of our work. The use of machine learning in medical image analysis is now widespread, with the use of boosting [10], random forests [11], and convolutional deep neural nets [12]. The combination of classifiers in cytology is proposed by Daskalakis et al. [13], but their work does not use multimodal input data. When considering the application requirements mentioned above, we notice some similarities with other problems, particularly in medical imaging. For instance, the detection and segmentation of brain structures [14] or the detection and segmentation of

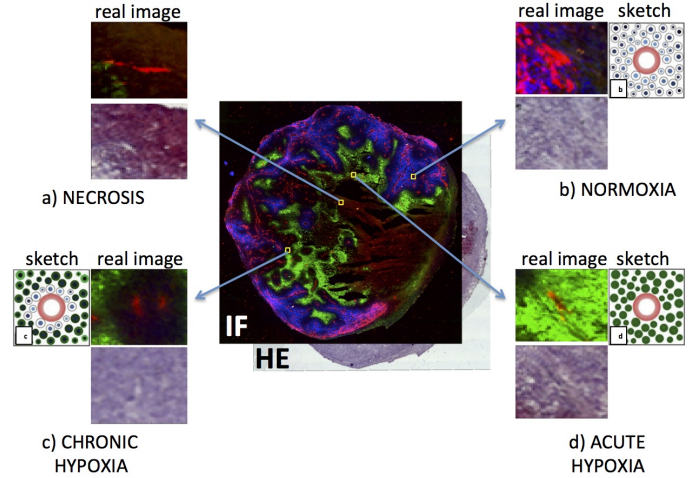


Fig. 2. Samples of the four classes of microvessel regions with the red region in the center (in the IF image) representing a microvessel, and the square region around it denoting the region. The sketches of the microvessel region appearance for 3 classes (normoxia, chronic hypoxia and acute hypoxia) are from [3].

organs from full body scans [15] are somewhat similar given that they also deal with a sparse multi-class problem, but the structures being detected have a consistent shape and can be found in specific places in the data, as opposed to the microvessel regions that do not have consistent shapes and can be found in relatively random places. The detection and segmentation of tumors from several organs [16] also deals with sparse detection and classification from random locations and uses less consistent shape models, but it is in general a binary problem, as opposed to the multi-class problem being dealt here. Therefore, we are dealing with a relatively new problem, and for this reason we propose a new methodology for the detection and classification of microvessel regions from IF and HE images into necrosis, normoxia, chronic hypoxia and acute hypoxia.

3. METHODOLOGY

We assume that the HE image is represented by $\mathbf{x}_{HE} : \Omega \rightarrow \mathbb{R}^3$ ($\Omega \in \mathbb{R}^2$ represents the image lattice), the IF image is denoted by $\mathbf{x}_{IF} : \Omega \rightarrow \mathbb{R}^3$, and the classification is represented by $y : \Omega \rightarrow \mathcal{Y}$, with $\mathcal{Y} = \{1, 2, 3, 4\}$, with 1, 2, 3, 4 indicating necrosis, normoxia, chronic hypoxia and acute hypoxia, respectively (see Fig. 2 for samples of each class in both image modalities). We formulate the classification problem as:

$$\mathbf{y}^* = \arg \max_{\mathbf{y}} P(\mathbf{y} | \mathbf{v} = 1, \Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF}), \theta), \quad (1)$$

where \mathbf{y} denotes a vector with all classification results in Ω , \mathbf{v} denotes a vector with elements $v : \Omega \rightarrow \{0, 1\}$ indicating the presence of a microvessel (this means that in positions without a microvessel, this classification is undefined), $P(\cdot)$ is a

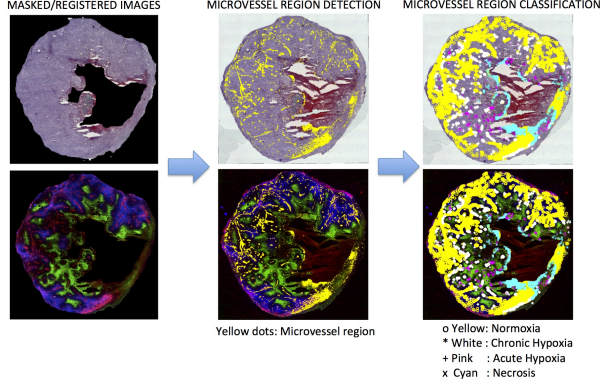


Fig. 3. The two steps of our proposed method: microvessel region detection and classification.

probability function parameterized by θ , and $\Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF})$ represents the following set of unary and pairwise potential functions: $\{\phi_A : \mathcal{Y} \times \mathcal{X} \rightarrow \mathbb{R}, \phi_S : \mathcal{Y} \times \mathcal{X} \rightarrow \mathbb{R}, \phi_R : \mathcal{Y} \times \mathcal{X} \rightarrow \mathbb{R}, \phi_C : \mathcal{Y} \times \mathcal{X} \rightarrow \mathbb{R}, \phi_P : \mathcal{Y} \times \mathcal{Y} \times \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}\}$, with \mathcal{X} and \mathcal{X} denoting the space of features extracted from the IF and HE images. The unary potential functions above are defined as follows: 1) Adaboost $\phi_A(y, \mathbf{x})$ [5], 2) linear SVM $\phi_S(y, \mathbf{x})$ [6], 3) random forests $\phi_R(y, \mathbf{x})$ [7], and 4) convolutional neural network (CNN) $\phi_C(y, \tilde{\mathbf{x}})$ [8, 17], all encoding the confidence of classifying the feature \mathbf{x} (or $\tilde{\mathbf{x}}$ for $\phi_C(\cdot)$) with class $y \in \mathcal{Y}$. The binary potential function encodes the distance dependent labeling homogeneity, defined by $\phi_P(y_i, y_j, \mathbf{p}_i, \mathbf{p}_j) = (1 - \delta(y_i - y_j)) \left(\frac{1}{\|\mathbf{p}_i - \mathbf{p}_j\|_2} \right)$, where $\mathbf{p}_i, \mathbf{p}_j \in \mathbb{R}^2$ are the 2-D positions of lattice indexes i and j , and $\delta(\cdot)$ is the Dirac delta function. Also note that in (1), the detection of a microvessel presence is automatically done by verifying the red channel of the IF image, which should have a value above a certain threshold (specifically, the lattice positions where $\mathbf{v} = 1$ has the red channel of the IF image with a value bigger than 25 (from the original range $[0, 255]$)¹.

The probability $P(\cdot)$ in (1) is defined in two ways. The first way assumes that the pairwise potential function is inactive, which essentially means that

$$P(\mathbf{y}|\mathbf{v} = 1, \Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF}), \theta) = \prod_{i \in \Omega} P(y_i|v_i = 1, \Psi_i(\Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF})), \theta). \quad (2)$$

The model parameters θ in (2) can be estimated with the training methods of the same classifiers listed above (i.e., Adaboost, linear SVM, random forest and CNN) using, as input features, the results of the individual classifiers for all 4 classes: $\Psi_i(\Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF})) = \{\phi_A(y_j, \mathbf{x}_i), \phi_S(y_j, \mathbf{x}_i), \phi_R(y_j, \mathbf{x}_i), \phi_C(y_j, \tilde{\mathbf{x}}_i)\}_{y_j=1}^4$, and the inference is simply based on selecting the class that locally produces the highest confidence. The second way of defining $P(\cdot)$ in (1) is as

¹In Sec. 4.1, we explain in more detail each of the IF channels.

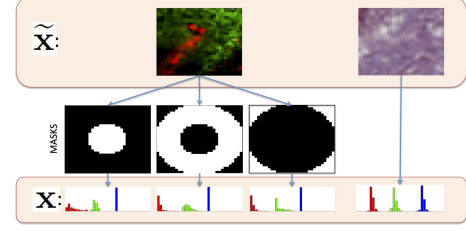


Fig. 4. Feature $\tilde{\mathbf{x}}$ represented by the microvessel region pixels (top) and \mathbf{x} denoted by the histograms of each of the depicted regions from the IF image and the histogram of the HE image (histograms are represent by the amount of red, green and blue pixels in the regions, shown with the respective color bars).

follows [9]:

$$P(\mathbf{y}|\mathbf{v} = 1, \Phi(\mathbf{y}, \mathbf{x}_{HE}, \mathbf{x}_{IF}), \theta) = \frac{1}{Z} \exp\{-\sum E(\Phi, \theta)\}, \quad (3)$$

with Z denoting the partition function and

$$E(\Phi, \theta) = \sum_{\{i \in \Omega | v_i = 1\}} (w_A \phi_A(y_i, \mathbf{x}_i) + w_S \phi_S(y_i, \mathbf{x}_i) + w_R \phi_R(y_i, \mathbf{x}_i) + w_C \phi_C(y_i, \tilde{\mathbf{x}}_i)) + \sum_{(i,j) \in \mathcal{E}} w_P \phi_P(y_i, y_j, \mathbf{p}_i, \mathbf{p}_j), \quad (4)$$

where \mathcal{E} denotes the set of neighboring nodes that are considered for the computation of the pairwise potentials. The estimation of the model parameters $\theta = [w_A, w_S, w_R, w_C, w_P]$ in (3) is based on the margin-rescaled structured SVM training [6] that uses the decomposable label loss function (between the annotations \mathbf{y} and $\hat{\mathbf{y}}$): $\Delta(\mathbf{y}, \hat{\mathbf{y}}) = \sum_{i \in \Omega | v_i = 1} 1 - \delta(y_i - \hat{y}_i)$. The inference is efficiently performed with graph cuts [18]. Fig. 3 displays the two steps of the proposed methodology, i.e., the detection of microvessel regions followed by their classification into the aforementioned four classes.

The feature used in the potentials $\phi_A(y, \mathbf{x})$, $\phi_S(y, \mathbf{x})$ and $\phi_R(y, \mathbf{x})$ is extracted from a microvessel region (here defined to be a square region of size $200\mu\text{m}$ around a microvessel [1] - see Fig. 2), following the general description provided by Bayer et al. [1] (see Fig. 2). In general, necrotic regions are represented by a pink color in the HE image, normoxia is represented by a region containing a red center (i.e., a microvessel) and blue pixels around it (indicating perfusion) in the IF image, chronic hypoxic regions contain a red center with blue pixels around it and green pixels towards the borders (also from the IF image), and acute hypoxia has a red center, with green pixels around it (again, from the IF image). This description motivates the feature \mathbf{x} based on the histograms of three separate regions of the IF image (from the center, ring around the center, and borders), as represented in Fig. 4, and the histogram of the whole HE image. For the CNN potential

feature \tilde{x} , we use the original IF and HE microvessel region pixels [8] (see Fig. 4), where each modality is fed to independent networks that are joined only at their last stage [17]. Finally, the graph structure represented by the set \mathcal{E} in (4) is estimated with a Delaunay triangulation of the points in the lattice that are selected as being part of a microvessel (i.e., the set $\{i \in \Omega | v_i = 1\}$).

4. EXPERIMENTS

4.1. Materials and Methods

We use the material available from the work by Maftai et al. [3, 2], consisting of five xenografted human squamous cell carcinoma lines of the head and neck (FaDu), which were transplanted subcutaneously into the right hind leg of nude mice that were whole body irradiated with 4 Gy (200 kV X-rays, 0.5 mm Cu filter) 2-5 days before transplantation. Each whole tumor cryosection was scanned and photographed using the AxioVision 4.7 and the multidimensional and mosaix modules. The IF images for the tumor cryosection were prepared with three separate stainings. Pimonidazole was used for hypoxia stain (green regions), CD31 was used for vessel stain (red regions), and Hoechst 33342 was used for perfusion stain (blue). Then, the cover slip was removed to stain the same slice with HE in order to detect the necrotic regions. This procedure can cause severe tearing and folding in HE images, as clearly shown in Fig. 1. After staining, the whole tumor cryosections were scanned at the same pixel size and photographed with the same settings as the IF images. In total, 16 pairs of IF and HE images from five different tumors were produced using the procedure above. These pairs of IF and HE images were then downsampled to 1000×1000 pixels (of resolution $10\mu\text{m}$ per pixel), registered [4] and shading corrected [19]. Finally, a manually delineated mask was also used in order to remove major necrotic regions, skin, background and tissue folding and tearing (see Fig 1).

The manual labeling of the microvessel regions is performed using an active learning scheme. Specifically, one image is arbitrarily chosen (from the set of 16 images described above) to be initially annotated, from which 500 microvessel regions are manually labeled (these 500 regions to be annotated are randomly selected from the currently detected microvessel regions). Then the four classifiers described in Sec. 3 are trained and applied to additional 500 regions in the same image (also randomly selected), and when they disagree on the classification of any of the additional regions, the user is requested to label the region. This process is repeated for the remaining 15 images, which means that we have 1000 microvessel regions annotated per image. The 16 images are then divided in 4 different ways into 8 images for training, 4 for validation and 4 for testing, which allows for a 4-fold cross validation testing of our methodology.

The assessment is based on the computation of the error $\Delta(\mathbf{y}, \mathbf{y}^*) = \sum_{i \in \{\Omega | v_i = 1\}} 1 - \delta(y_i - y_i^*)$ (defined above in Eq. 4), assuming that \mathbf{y} represents the manual annotation and \mathbf{y}^* denotes the automated annotation produced by (1). Note

Table 1. Mean and standard deviation of the errors produced by the proposed methods in the 4-fold cross validation test.

Method	Training	Testing
Joint (2)	0.047 ± 0.013	0.125 ± 0.040
CRF (3)	0.080 ± 0.024	0.130 ± 0.041
Adaboost	0.132 ± 0.042	0.151 ± 0.053
RandForests	0.080 ± 0.024	0.130 ± 0.041
linear SVM	0.183 ± 0.058	0.210 ± 0.071
CNN	0.047 ± 0.016	0.195 ± 0.055

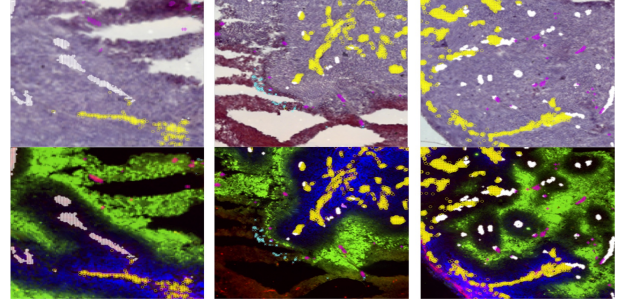


Fig. 5. Zoomed in results on the test image of Fig. 3 using the joint model (2) with the IF and HE images. The labels are organized as follows: 'o yellow': normoxia, '* white': chronic hypoxia, '+ pink': acute hypoxia, 'x cyan': necrotic.

that for the joint classifier in (2), we try all four classifiers studied in this paper (i.e., Adaboost, linear SVM, CNN and random forest), but the best result in the validation set is obtained with Adaboost, so in the results, please assume that the joint classifier in (2) is produced by the Adaboost classifier.

4.2. Results

In Table 1 we show the mean and standard deviation of the training and testing errors in the 4-fold cross validation test using the classifiers in Eq. 2 (row 1) and Eq. 3 (row 2) that combine the results of the Adaboost, random forests, linear SVM and CNN classifiers (we also show the errors produced by each of these classifiers for comparison in rows 3-6). Fig. 5 shows the results of (2) on a test image in some challenging regions of the HE and IF images (registered).

5. DISCUSSION AND CONCLUSIONS

The results from Sec. 4.2 show that the combined classifier in (2) produces the best classification result on the test set, while the results produced by the CRF model in (3) does not improve over the best unary potential (random forests). These results show that the random forests classifier is particularly important in this framework. Although the CRF model does not seem to be competitive enough, it is important to show it as a baseline result. Furthermore, these results indicate that this methodology is potentially useful in clinical practice [1]. Finally, we plan to make this dataset available in the near future, and the results presented here can then be used as a baseline for future research.

6. REFERENCES

- [1] Christine Bayer, Kuangyu Shi, Sabrina T Astner, Constantin-Alin Maftai, and Peter Vaupel, "Acute versus chronic hypoxia: why a simplified classification is simply not enough," *International Journal of Radiation Oncology* Biology* Physics*, vol. 80, no. 4, pp. 965–968, 2011.
- [2] Constantin-Alin Maftai, Christine Bayer, Kuangyu Shi, Sabrina T Astner, and Peter Vaupel, "Quantitative assessment of hypoxia subtypes in microcirculatory supply units of malignant tumors using (immuno-) fluorescence techniques," *Strahlentherapie und Onkologie*, vol. 187, no. 4, pp. 260–266, 2011.
- [3] Constantin-Alin Maftai, Christine Bayer, Kuangyu Shi, Sabrina T Astner, and Peter Vaupel, "Changes in the fraction of total hypoxia and hypoxia subtypes in human squamous cell carcinomas upon fractionated irradiation: evaluation using pattern recognition in microcirculatory supply units," *Radiotherapy and Oncology*, vol. 101, no. 1, pp. 209–216, 2011.
- [4] Tingying Peng, Mehmet Yigitsoy, Abouzar Eslami, Christine Bayer, and Nassir Navab, "Deformable registration of multi-modal microscopic images using a pyramidal interactive registration-learning methodology," in *Biomedical Image Registration*, pp. 144–153. Springer, 2014.
- [5] Ji Zhu, Hui Zou, Saharon Rosset, and Trevor Hastie, "Multi-class adaboost," *Statistics and Its*, 2009.
- [6] Ioannis Tsochantaridis, Thomas Hofmann, Thorsten Joachims, and Yasemin Altun, "Support vector machine learning for interdependent and structured output spaces," in *Proceedings of the twenty-first international conference on Machine learning*. ACM, 2004, p. 104.
- [7] Leo Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [9] Martin Szummer, Pushmeet Kohli, and Derek Hoiem, "Learning crfs using graph cuts," in *Computer Vision—ECCV 2008*, pp. 582–595. Springer, 2008.
- [10] Gustavo Carneiro, Bogdan Georgescu, Sara Good, and Dorin Comaniciu, "Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 9, pp. 1342–1355, 2008.
- [11] Antonio Criminisi, Jamie Shotton, and Ender Konukoglu, "Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning," *Foundations and Trends® in Computer Graphics and Vision*, vol. 7, no. 2–3, pp. 81–227, 2012.
- [12] Dan C Cireşan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*, pp. 411–418. Springer, 2013.
- [13] Antonis Daskalakis, Spiros Kostopoulos, Panagiota Spyridonos, Dimitris Glotsos, Panagiota Ravazoula, Maria Kardari, Ioannis Kalatzis, Dionisis Cavouras, and George Nikiforidis, "Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely h&e-stained cytological images," *Computers in biology and medicine*, vol. 38, no. 2, pp. 196–203, 2008.
- [14] Michael Wels, Yefeng Zheng, Martin Huber, Joachim Hornegger, and Dorin Comaniciu, "A discriminative model-constrained em approach to 3d mri brain tissue classification and intensity non-uniformity correction," *Physics in medicine and biology*, vol. 56, no. 11, pp. 3269, 2011.
- [15] Antonio Criminisi, Jamie Shotton, Duncan Robertson, and Ender Konukoglu, "Regression forests for efficient anatomy detection and localization in ct studies," in *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, pp. 106–117. Springer, 2011.
- [16] Chao Lu, Sudhakar Chelikani, David A Jaffray, Michael F Milosevic, Lawrence H Staib, and James S Duncan, "Simultaneous nonrigid registration, segmentation, and tumor detection in mri guided cervical cancer radiation therapy," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 6, pp. 1213–1227, 2012.
- [17] Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Ng, "Multimodal deep learning," in *International Conference on Machine Learning*, 2011.
- [18] Yuri Boykov, Olga Veksler, and Ramin Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [19] Tingying Peng, Lichao Wang, Christine Bayer, Sailesh Conjeti, Maximilian Baust, and Nassir Navab, "Shading correction for whole slide image using low rank and sparse decomposition," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2014*, pp. 33–40. Springer, 2014.