MELANOMA DETECTION IN HISTOPATHOLOGICAL IMAGES USING DEEP LEARNING

Salah Alheejawi, Richard Berendt, Naresh Jha and Mrinal Mandal

University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Histopathological images are widely used to diagnose diseases such as skin cancer. As digital histopathological images are typically of very large size, in the order of several billion pixels, automated identification of abnormal cell nuclei would be very helpful for doctors to perform fast diagnosis. In this paper, we propose a technique, using deep learning algorithms, to first segment the cell nuclei in Hematoxylin and Eosin (H&E) stained images and detect the abnormal melanocytes on the histopathological images. The cell segmentation is done by using a novel Convolutional Neural Network (CNN) architecture. The segmented cells are then classified into melanoma and other nuclei using a Support Vector Machine classifier. Experimental results show that the CNN can segment the nuclei with more than 90% accuracy. The proposed technique has a low computational complexity.

KEYWORDS

Histopathological image analysis, Nuclei segmentation, Melanoma Detection, Deep learning.

1. INTRODUCTION

Melanoma is an abnormal growth of melanocytes which mainly occurs on the skin and it can transfer to any part of the body and destroys the tissue. About 7,800 Canadians have been diagnosed with melanoma skin cancer in 2019 and 1,300 of them would be in a fatal stage [1]. The early diagnosis of melanoma is very important as it helps to increase the chances of successful treatment and the survival rate. The Computer-aided diagnosis (CAD) techniques can effectively help doctors to diagnose and detect the melanoma in early stages [2]. The digitized histopathological slides, which are typically obtained by staining and scanning the biopsy slides of the skin tissue, can provide the cell morphological features with a high resolution. The digitized slides are known as Whole Slide Images (WSIs) and with help of CAD techniques that will permit the pathologist for precise diagnosis [3]. Pathologists usually use H&E stained images, because the morphological features of the melanocytes and other cells become vividly clear. In H&E stained image, the cell nuclei contain chromatin and that can be observed in blue shade while the cytoplasm and other connective tissues are observed with varying shades of pink. Fig. 1 shows the section of skin tissue divided into three layers: epidermis, dermis and subcutaneous layer. Due to the large density of cell nuclei in the epidermis layer, the tissue appears in dark purple color. Fig. 1 shows skin tissue image and zoomed patch obtained with H&E stain, where the abnormal melanocytes appear with irregularity in shape and color intensity [4,5].

Several techniques have been proposed to segment the cell nuclei in histopathological images [6-11]. Xu et al. [6] proposed an automated technique (henceforth referred to as the Watershed+Voting technique) to segment the cell nuclei in H&E stained images. The technique detects the nuclei seeds by using voting areas and segments the nuclei cells using marked watershed algorithm. The technique provides a good performance with high computational complexity due the seed detection algorithm. Xu et al. [11] also proposed cell nuclei segmentation technique (henceforth referred to as the gLoG+mRLS technique) using generalized Laplacian of Gaussian (gLoG) filters to detect the seeds nuclei and multiple Radial Lines Scanning (mRLS) algorithm to segment the cells. The mRLS uses high gradient pixel locations and shape information to accurately segment the cell nuclei.
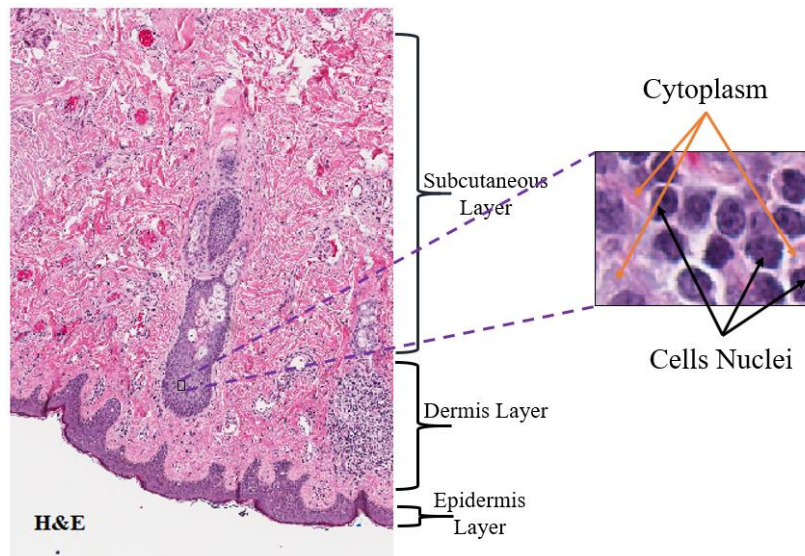

Figure 1. Example of digitized biopsy images of skin tissue image with H&E stain.

The techniques [6-11] mentioned above are generally based on extracted hand-crafted features that require significant time to process. The deep learning algorithms using CNN have been widely used successfully in medical image analysis. The CNN models can train the feature extraction process to provide high performance with low computational complexity in many different tasks (e.g. classification, detection or segmentation [12-13]). Basrinarayanan et al. [14] proposed the SegNet architecture for object segmentation. The architecture uses a number of sampling and upsampling layers for extracting the features in hierarchical levels. Ronneberger et al. [15] proposed the U-Net architecture for biomedical image segmentation. The U-Net architecture has encoder and decoder sides with number of sampling and upsampling layers, respectively. The upsampling layer outputs are enhanced by concatenating them with features from the encoder side.

In this paper, we proposed an automated technique to segment the cell nuclei and classify them into melanoma and other nuclei using a Support Vector Machine (SVM) classifier. The technique uses a CNN architecture to segment the cell nuclei on H&E stained images. The proposed CNN used several convolutional layers with different size of filters. Experimental results demonstrate high accuracy and low computational complexity of the proposed technique compared to the state-of-the-art techniques.

The organization of the paper is as follows. Section 2 describes the dataset used to train and evaluate the proposed technique. Section 3 describes the proposed technique in detail. Section 4 presents the performance evaluation, followed by the conclusion in Section 5.

2. DATA DESCRIPTION

In this section, the nuclei segmentation and cell classification will be evaluated based on digitized biopsy slides generated at the University of Alberta. The biopsies were collected at the Cross Cancer Institute, Edmonton, Canada in accordance with the protocol for the examination of specimens with skin melanoma. The formalin-fixed paraffin-embedded tissue blocks of these biopsies were cut into thin slices (e.g., 4μm for light microscope). These slices were then mounted to glass slides and stained using H&E stain [5]. The WSIs were obtained by scanning the H&E slides using aperio scanscope scanning system under 40X magnification. The size of a WSI is typically around 40,000×60,000 pixels (in color) and each WSI contains thousands of cell nuclei. The image dataset consists of 64 WSIs for skin tissue and 9 WSIs for lymph node tissue.

3. PROPOSED TECHNIQUE

The schematic of proposed technique is shown in Fig. 2 which consists of two modules: CNN-based nuclei segmentation and nuclei classification. The details of each module are presented in the following.

Figure 2. Schematic of a melanoma detection technique.

3.1. CNN-Based Nuclei Segmentation

In this module, the input H&E stained images are segmented into cell nuclei and background. The nuclei segmentation is done by using the proposed CNN architecture, henceforth referred to as the NS-Net architecture (Nuclei Segmentation Net). The NS-Net architecture, shown in Fig. 3, consists of five convolutional layers (shown in gray color) and one softmax (shown in pink) followed by the pixel classification layer (shown in blue). The convolutional layer in the NS-architecture consists of three operations: convolution, batch normalization [16], and Rectified Linear Unit (ReLU) activation [17]. A brief description of each operation is presented in the following:

(i)   Convolution: In this layer, the input image I is convolved with filters F_j :

$$R_j = I * F_j, \quad j = 1, 2, .., N \tag{1}$$

where N is the number of filters, R_j is the output corresponding to the jth convolution filter, and F_j is the weights of the jth filter.

(ii) Batch normalization: This operation is used to normalize the convolutional layers output to zero mean and unitary variance across a current mini-batch. The normalized $\hat{R}_i$ output will be scaled with σ and shifted by β as follows:

$$y_i = \sigma \hat{R}_i + \beta \qquad (2)$$

(iii) ReLU: It is an activation function that will output the input values y that are greater than zero using the following equation
:

$$f = \max(0, y) \qquad (3)$$

where f is output of the activation function ReLU.

In this architecture, the features are extracted in hierarchical levels by using convolutional filters of different sizes. The change on the convolutional filters can precisely locate the object boundaries that need to be segmented. Most existing CNN architectures include pooling layers. In our experiment, it has been found that the pooling leads to loss of the spatial information that carries important texture and shape features of the nuclei. Therefore, the pooling layer has been omitted in the proposed architecture. Table 1 shows the number and the size of filters in each layer of the NS-Net architecture. The NS-Net architecture is trained and evaluated using a dataset of 24 high resolution H&E stained images (1920×2500 color pixels). Each image is divided into overlapping blocks of 64×64 color pixels to obtain 458 block-images. The block-image dataset is divided into 70% for training, 15% for testing and 15% for validation. The entropy loss function with the stochastic gradient descent with momentum (SGDM) optimizer is used to train the NS-Net architecture [18]. Fig. 4 (a) shows an input H&E stained image and (b) shows the masked nuclei image obtained using the NS-Net architecture.
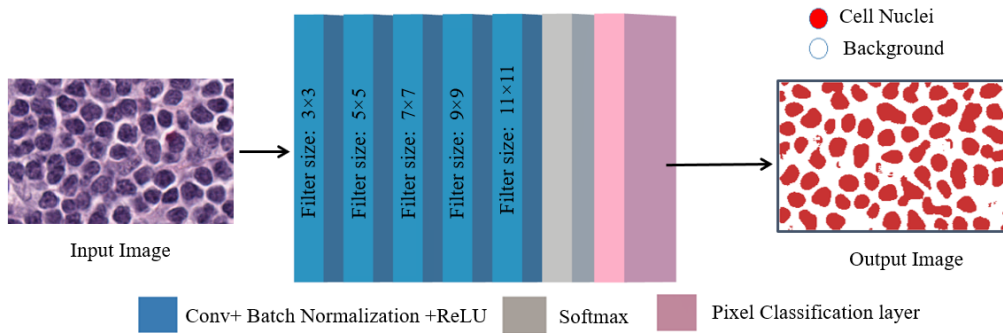

Figure 3. The NS-Net architecture for nuclei segmentation (with 5 convolutional layers).

Table 1. Details of the NS-Net architecture with 5 convolutional layers.
Input image size: M×N pixels (color). Number of classes: C.

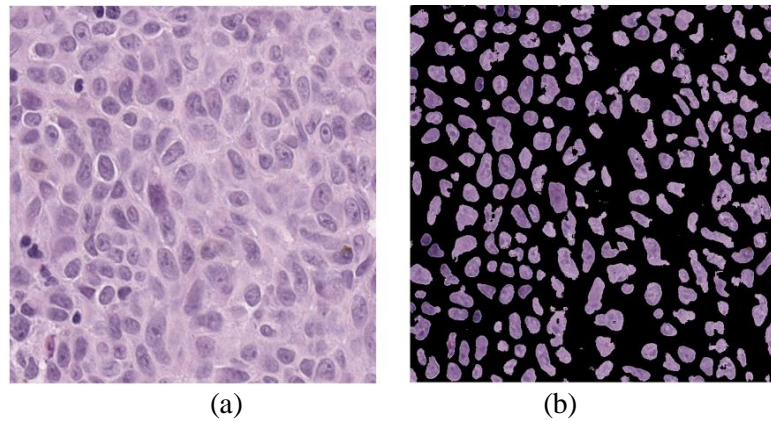|	Number of Filters	Number of Channels	Output Image size	Filter size
Layer-1	64	3	M×N×64	3×3
Layer-2	64	64	M×N×64	5×5
Layer-3	64	64	M×N×64	7×7
Layer-4	64	64	M×N×64	9×9
Layer-5	C	64	M×N× C	11×11
Softmax layer	-	C	M×N×C	

Figure 4. Segmentation results. (a) Input image (b) Segmented image obtained using the NS-Net architecture.

3.2. Nuclei Classification

In this module, the segmented nuclei obtained using the NS-Net architecture is classified into two classes based on hand-crafted features. The feature vector consists of 18 first-order features, 9 Histogram of Oriented Gradient features, 24 Haralick texture features and 3 Morphological features. The features are extracted for each pre-segmented nuclei and described briefly as follows:

(i)   First-order features: It includes histogram-based features such as the mean, standard deviation, third moment, smoothness, entropy, and uniformity for 3-channels (R, G and B) to obtain 18 features (6×3).

(ii)   Histogram of Oriented Gradient features: It measures the gradient of (9) orientations in localized portions of the segmented nuclei image.

(iii)   Haralick texture features: It is calculated from a Gray Level Co-occurrence Matrix, (GLCM). It includes six GLCM features (correlation, energy, homogeneity, contrast, entropy, and dissimilarity) in 4 directions (i.e., $0°$, $45°$, $90°$ and $135°$) to obtain 24 features (6×4).

(iv)   Morphological features: It includes the eccentricity, solidity, and the ratio of major and minor axes of the cell nuclei to obtain (3) features.

The extracted feature vector (size 54) of each cell nuclei are then classified into normal and melanoma using SVM classifier [19-20]. The SVM is a very efficient supervised classifier that can handle even a non-linearly separable features and create hyperplane to separate melanocyte from other cells. In the proposed technique, the SVM model is trained and tested on 800 cell nuclei (70% for training and 30% for testing). In this technique, the SVM classifier is applied with different kernels and it shown that the Gaussian kernel provides the best classification results (see Table 4 in section 4).

After the nuclei in the lymph node are classified in the lymph node, the staging of melanoma is done based on the number and size of melanocytic clusters present in the lymph node image.

4. RESULTS AND DISCUSSIONS

In this section, we present the performance of the proposed technique using the dataset mentioned in Section 2.

The segmentation performance is evaluated and compared with handcrafted feature-based algorithms such as gLoG+mRLS and Voting+Watershed techniques and with trained CNN features techniques such as SegNet and U-Net architecture. The segmentation performance is evaluated using Accuracy, Precision, Recall and BF-score [21] measures defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

$$BF\ Score = \frac{2TP * TN}{TP * TN + 2TN * FP + TN * FN} \times 100\%$$

where *TP*, *TN*, *FN* and *FP* denote the number of true positives, true negatives, false negatives and false positives, respectively. Table 2 shows the segmentation performance of different techniques. It is observed that the deep learning algorithms provide excellent performance compared to the classical feature-based algorithms. This is because the classical features are less sensitive to the diversity of the cell nuclei in the skin tissue. For example, the abnormal melanocytes tend to have light and inhomogeneous color (see Fig. 5) and that causes miss detection of the melanocytes in the gLoG+mRLS and Voting+Watershed techniques.

Table 2. Segmentation Performance of the deep learning algorithms and the classical feature-based algorithms.

Technique:	Accuracy	Precision	Recall	BF-Score	Execution time (in s)
Voting+Watershed [6]	83.64	78.24	84.64	81.31	143.71
gLoG+mRLS [11]	76.67	79.27	60.25	68.46	128.57
SegNet [14]	87.84	84.16	87.53	85.81	15.37
U-Net [15]	78.79	87.41	57.87	69.63	20.82
NS-Net	90.21	87.20	89.90	88.52	14.27

In this work, the NS-Net, SegNet and U-Net architectures are trained with the same number of training images. The NS-Net architecture is also evaluated with CNN architecture in terms of the required parameters need to be train as shown in Table 3.

Figs. 6 (b)-(f) shows the subjective segmentation performance of SegNet [14], U-Net [15], gLoG+mRLS [11], Voting+Watershed [6] and the NS-Net architecture, respectively. It is observed that the NS-Net architecture provides excellent nuclei segmentation, whereas gLoG+mRLS, Voting+Watershed techniques miss a few cell nuclei due to the inhomogeneity in the cell nuclei color. It is also observed that the U-Net architecture does not perform well

compared to the other techniques because the overfitting due the large number of the filters that are used in the cell nuclei segmentation.

Table 3. Properties of CNN architectures used in performance evaluation.

CNN Architecture	Convolutional layers	No. of Trained parameters	Filter size	No. of Filters
SegNet [14]	8	225,542	3×3	64
U-Net [15]	11	905,472	3×3	(64, 128, 256)
NS-Net	5	150,336	(3×3)- (11×11)	64

(a) An H&E stained image patch   (b) SegNet [14]   (c) gLoG+mRLS [11]

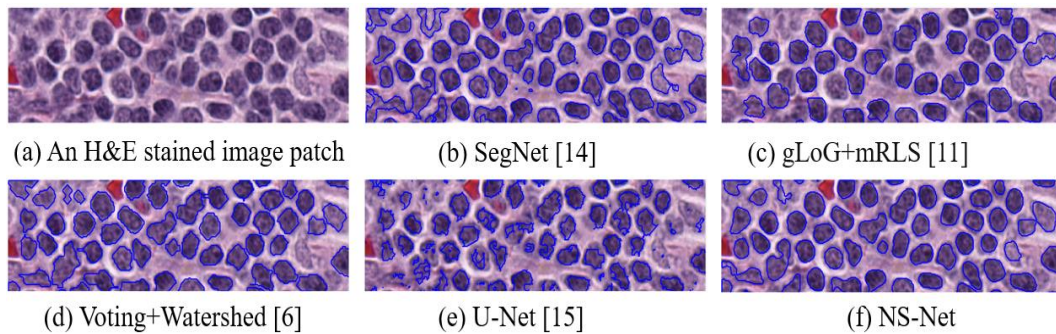(d) Voting+Watershed [6]   (e) U-Net [15]   (f) NS-Net

Figure 5. Subjective comparison of cell nuclei segmentation results (contoured in blue color) (a) original test image, (b)-(f) Segmentation results for SegNet [14], gLoG+mRLS [11], Voting+Watershed [6], U-Net [15] and NS-Net techniques, respectively.

The classification performance is evaluated in terms of the Accuracy, Precision, Recall and BF Score measures. The SVM classifier has been evaluated with different kernels (such as Gaussian, linear and polynomial kernels [18]). The results are shown in Table 4. It is observed that the SVM classifier with Gaussian kernel provides the best performance for the dataset. Fig. 6(a)-(b) shows the nuclei classification results obtained using Gaussian kernel, where the melanoma and other nuclei are contoured in red and blue colors, respectively.

Table 4. Performance of the nuclei classification using different SVM kernels.

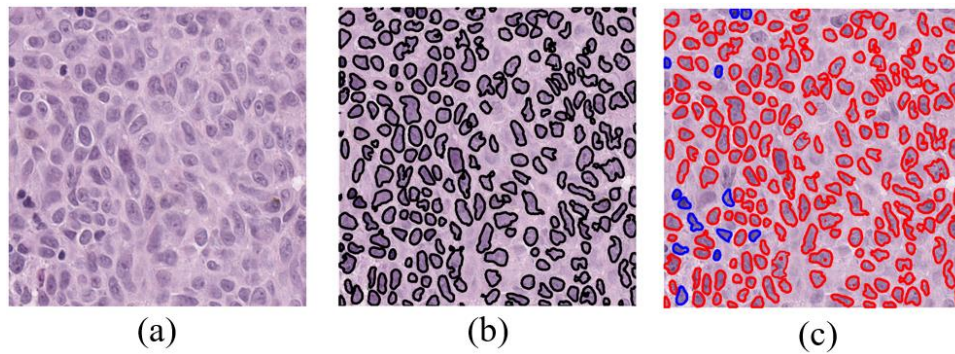| Evaluation Measures	SVM Kernel		
	Linear	Polynomial	Gaussian
Accuracy	80.52	57.28	85.72
Precision	77.43	57.28	80.04
Recall	93.14	96.65	97.42
BF Score	84.56	71.93	87.87

Figure 6. Example of classification results. (a) NS-Net input image (b) NS-Net segmented output image (c) Classified image obtained using SVM, where melanoma and other cell nuclei are contoured with red and blue color, respectively.

5. CONCLUSION

This paper proposes an automated technique for melanoma detection on the skin/ lymph node histopathological images. The technique segments the cell nuclei in H&E stained lymph node image using the NS-Net architecture. The NS-Net architecture segments the image into background and cell nuclei regions. The proposed CNN architecture provides an excellent segmentation performance with a low computational complexity. The segmented cell nuclei are then classified into melanocytic and other cell nuclei using an SVM classifier with Gaussian kernel. After the nuclei classification, the staging of melanoma can be done by the doctors based on the number and size of melanocytic clusters present in the image.

REFERENCES

[1].   Rebecca L Siegel, Kimberly D Miller, & Ahmedin Jemal (2019), "Cancer Statistics, 2019," CA: A Cancer Journal for Clinicians, Vol. 69, No. 1, pp.7–34.

[2].   Lieve Brochez, Evelien Verhaeghe, Edouard Grosshans, Eckhart Haneke, Gérald Piérard, Dirk Ruiter, & Jean-Marie Naeyaert (2002), "Inter-observer variation in the histopathological diagnosis of clinically suspicious pigmented skin lesions". J. Pathol. 196(4), pp.459–466.

[3].   Cheng Lu, Muhammad Mahmood, Naresh Jha, & Mrinal Mandal, (2013), "Automated segmentation of the melanocytes in skin histopathological images". IEEE J. Biomed. Health Inf. 17(2), pp. 284–296.

[4].   Marcial García Rojo, Ana Morillo Castro, & Luis Gonçalves, (2011), "COST action "eurotelepath": digital pathology integration in electronic health record, including primary care centres," Diagnostic Pathology 6, S6 (2011), pp. S1–6.

[5].   Ronald S Weinstein, Anna R Graham, Lynne C Richter, Gail P Barker, Elizabeth A Krupinski, Ana Maria Lopez, Kristine A Erps, Achyut K Bhattacharyya, Yukako Yagi, & John R Gilbertson, (2009) "Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for thefuture," Human pathology, vol. 40, no. 8, pp. 1057–1069.

[6].   Hongming Xu, Cheng Lu, & Mrinal Mandal (2013), "An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm". IEEE J. Biomed. Health Inf. 18(5), pp. 1729–1741.

[7].   Olcay Sertel, (2012), "Image analysis for computer-aided histopathology," Ph.D. dissertation, The Ohio State University.

[8].   Juliana M Haggerty, Xiao N Wang, Anne Dickinson, Chris J O'Malley & Elaine B Martin, (2014), "Segmentation of epidermal tissue with histopathological damage in images of haematoxylin and eosin stained human skin". BMC Med. Imaging. 14(1), pp. 7.

[9].   Jierong Cheng, Jagath C. Rajapakse, (2009), "Segmentation of clustered nuclei with shape markers and marking function," IEEE Transactions on Biomedical Engineering, vol. 56, no. 3, pp. 741–748.

[10]. Bahram Parvin, Qing Yang, Ju Han, Hang Chang, Bjorn Rydberg, Mary Helen & Barcellos-Hoff, (2007), "Iterative voting for inference of structural saliency and characterization of subcellular events," IEEE Transactions on Image Processing, vol. 16, no. 3, pp. 615–623.

[11]. Hongming Xu, Cheng Lu, Richard Berendt, Naresh Jha & Mrinal Mandal, (2017), "Automatic nuclei detection based on generalized Laplacian of Gaussian filters," IEEE Journal of Biomedical and Health Informatics (JBHI), vol.21, no.3, pp.826-837.

[12]. Jonathan Long, Evan Shelhamer, & Trevor Darrell, (2015), "Fully convolutional networks for semantic segmentation," in CVPR, pp. 3431–3440.

[13]. David Eigen, & Rob Fergus, (2015) "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture", in Proceedings of the International Conference on Computer Vision, pp. 2650–2658.

[14]. Vijay Badrinarayanan, Alex Kendall, & Roberto Cipolla (2015). "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." *arXiv*. Preprint arXiv: 1511.0051.

[15]. Olaf Ronneberger, Philipp Fischer, & Thomas Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 9351, pp. 234–241.

[16]. Sergey Ioffe, & Christian Szegedy, (2015), "Batch normalization: accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167.

[17]. Vinod Nair, & Geoffrey E. Hinton (2010), "Rectified linear units improve restricted Boltzmann machines," in Proceedings of the 27th international conference on machine learning, pp. 807–814.

[18]. Christian Robert, (2014) "Machine learning, a probabilistic perspective," CHANCE, vol. 27, no. 2, pp. 62–63.

[19]. Corinna Cortes & Vladimir Vapnik, (1995), "Support-vector networks". Machine Learning. 20 (3), pp. 273–297.

[20]. Asa Ben-Hur, David Horn, Hava Tova Siegelmann, & Vladimir N Vapnik, (2001) "Support vector clustering," Journal of Machine Learning Research, vol. 2, pp. 125-137.

[21]. Gabriela Csurka, Diane Larlus, & Florent Perronnin (2013), "What is a good evaluation measure for semantic segmentation?" Proc. of the British Machine Vision Conference, pp. 32.1-32.11.