# Ki67 Hot-Spots Detection on Histopathological Images of Breast Carcinoma using Convolutional Neural Networks

J. Gallego[1], Z. Swiderska-Chadaj[2], O. Deniz[1], G. Bueno[1]

[1] VISILAB Group, Universidad de Castilla-La Mancha, E.T.S.I.I, Ciudad Real, Spain,

[2] Warsaw University of Technology, Faculty of Electrical Engineering, Warsaw, Poland

{Jaime.Gallego, Oscar.Deniz, Gloria.Bueno}@uclm.es

## Abstract

*Hot-spot detection and Ki67 labelling index evaluation in invasive breast cancer have a significant role in routine medical practice. The quantification of cellular proliferation assessed by Ki67 immunohistochemistry is an established prognostic and predictive biomarker which determines the choice of therapeutic protocols. In this paper, we present a deep learning based approach to automatically detect and quantify Ki67 hot-spot areas by means of the Ki67 labelling index. To this end, a dataset composed of 50 Ki67 whole slide images (WSI) belonging to 50 breast cancer cases was used. We propose to apply Convolutional Neural Networks (CNN), to achieve Ki67 regions classification, and create the tumor proliferation map by comparing hot-spot detection to manual diagnosis. The overall accuracy of this novel approach was 96%. Results show the suitability of this CNN-based approach for detecting hot-spots areas of invasive breast cancer in WSI.*

## 1. Introduction

The revolution of deep learning techniques has also reached medical diagnosis and consequently to the study of both the diagnosis and the prognosis of cancer. The same happens in Histopathology with the advances of digital pathology where thanks to the digitization of the microscopic image, image processing and artificial intelligence techniques can already be applied for the diagnosis of cancer [1,2]. Automatic methods can assist pathologists in their task by offering automatic medical diagnosis meanwhile increasing accuracy.

Among the different tools used for studying cell functioning in oncology is the quantification of cellular proliferation. Data obtained from tumor samples may have therapeutic or prognostic impacts. In particular, proliferative activity assessed by Ki67 immuno-histochemistry (IHC), is an established prognostic and predictive biomarker which determines the choice of therapeutic protocols, applied to breast cancer [3,4]. Tumor regions that exhibit high proliferating activity are called hot-spots.

The Ki67 index is evaluated in hot-spot areas, that is, the areas where there is a high immunopositive cell density inside tumor area. Ki67 index is calculated by pathologists as the ratio of immunopositive cells (brown cells) to all cancer cells (immunopositive and immunonegative cells) inside tumor area. The hot-spot may be calculated manually with conventional microscopy reviewing the field of views (FOV) or automatically by review of whole slide images (WSI) using digital image processing.

Some attempts to automatically quantify Ki67 hot-spots have been recently reported in the literature. Lopez *et al.* [5] use agglomerative hierarchical clustering for identifying Ki67 hot-spots in WSI of brain tumor (high-grade gliomas). The clustering is applied on the brown channel after color deconvolution. Niazi *et al.* [6] use minimal graph cuts and multistate difference of Gaussians to detect the individual cells from images of Ki67-stained slides of neuroendocrine tumours.

Lu *et al.* [7], use adaptive filtering to automatically detect Ki67 hot-spot in adrenal cortical carcinomas. The method was tested in 50 WSI and compared to manual scoring. Swiderska *et al.* [8] applied textural and morphological features with support vector machine classifier for automatic hot-spot selection in Ki67 stained meningiomas specimens. Laurinavicius *et al.* [4] and Plancoulaine *et al.* [9] use methods based on stereology to assess Ki67 expression in WSI of 152 surgically removed Ki67-stained breast cancer specimens. They perform a subsampling of the data into a hexagonal tiling (HexT) arrays from which a set of texture and indicators such as entropy may be calculated for the Ki67 variability analysis and hot-spot detection by means of the statistical analysis of those indicators.

Valous *et al.* [10] applied a second order statistical feature, called lacunarity, to deal with the quantification of spatial intratumoral heterogeneity of Ki67 expression.

Pilutti *et al.* [11] applied an adaptive thresholding together with a cellularity analysis to automatically identify hot-spot areas over Ki67 stained breast cancer biopsies. 82% of the gold standard hot-spots was successfully recognized by the method. The drawback of these methods is that they are parameter-dependant and cannot be easily tuned for non-homogeneous staining which may variate the hot-spot densities.

The method proposed in this paper makes use of convolutional neural networks (CNN) to detect and characterize Ki67 hot-spots in breast cancer WSI. As far as the authors know this is the first time that the deep learning approach is reported for the detection and

quantification of Ki67 hot-spots. This method solves the problem of feature selection. The average accuracy obtained was 96%. The developed methodology can be used for other tumors where the detection of Ki67 hot-spots and the quantification of a proliferation index is relevant.

The remainder of the paper is organized as follows: Description of the database used is given in Section 2. Section 3 describes the proposed method based on CNNs for the detection of Ki67 hot-spots. Finally, obtained results and conclusions are presented in Section 4 and Section 5 respectively.

## 2. Materials

This study involved digital slides from the AIDPATH breast cancer database (http://aidpath.eu/). It is composed of breast tissue cohorts from four different institutions and pathology labs around Europe, including Spain, Italy and Lithuania. The dataset is formed of samples extracted by expert pathologists from incisional biopsies, with the objective to detect potential cancer regions and its proliferation factor. It comprises 50 breast cancer specimens stained with Ki67/MIB-1 antibody (3,3'-diaminobenzidine tetrahydrochloride, DAB marker, brown color) and counter-stained with hematoxylin (blue).

The slides were prepared serially, using Formalin-Fixed Paraffin-Embedded (FFPE) tissue sections with 4 $\mu m$ of thickness. Ki67 stain have been applied to each sample using the systematic pathological process defined by the highest quality standards.

### 2.1. Slide Digitization

The slides were digitized using Aperio Scanscope CS (Leica Biosystems) whole-slide scanner at 20x and 40x magnification. The area of tissue in a field of view (FOV) at 40x is 0.037mm$^2$, which results a size of 1424x1064 pixels at 40x and 712x532 pixels at 20x.

To reduce the time and effort required to manipulate the digital slides, create the ground truth annotations and the computational processing, we worked at 20x magnification. Therefore, the WSI acquired at 20x magnification have been analysed in full resolution, while slides acquired at 40x magnification have been down-sampled by a factor of 2. Depending on the size of tissue fixed on the slide, digital slides range between 0.5GB to 2GB.

One example of digitized WSI for Ki67 stain is shown in **Figure 1**.

### 2.2. Database creation

Each WSI was reviewed by expert pathologists to generate the ground truth annotations for hot-spot (Ki67) localization areas. Pathologists annotated the hot-spot regions, and, accordingly, evaluated the Ki67 proliferation index of the WSIs. With this information, we created the database automatically by computing the index proliferation of each tile using the automatic cells detector developed by Markiewicz *et al.* [13].

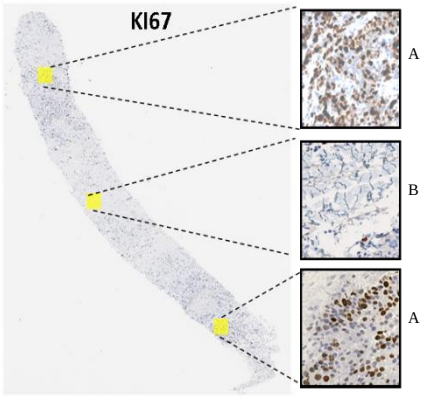



***Figure 1*** *Example of WSI stained Ki67. A- area with hot-spot on the Ki67 image, B- area without tumor. Immunopositive cells are marked by brown and immunonegative cells are marked by blue.*

Pathologists annotations were done with the Aperio ImageScope software.

In order to create the database, the annotated areas are divided into 512x512 non-overlapping pixel image tiles. These Ki67 tiles are shared into 11 classes:

- Non tumor areas.
- Lymphocytes: are added to this analysis to avoid the false positive regions generated by its high proliferation.
- 9 classes with tumor areas: The tumor areas have been shared into 9 groups based on tumor proliferation factor, which has been calculated as ratio between immunopositive cells (brown cells) and all tumor cells (brown and blue cells). The detection of all tumor cells (brown and blue cells) is performed by means of the cell detector developed by Markiewicz *et al.* [13]. **Figure 2** shows a tile sample for each of the classes.

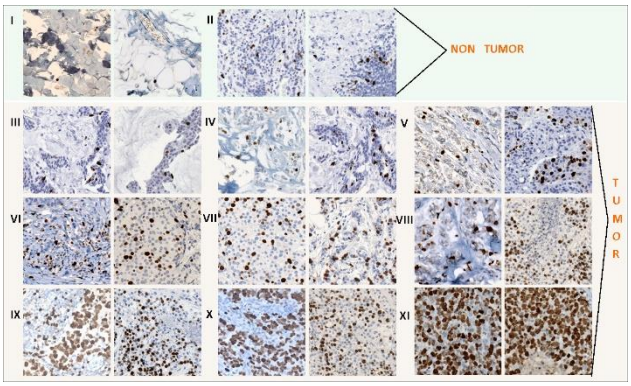



***Figure 2*** *Example of tiles for each class. I, no tumor class. II, lymphocytes. III-IX, tumor.*

Augmentation procedures are applied on the training data to increase the number of samples and the data variety for each class, which is known to enhance the network classification accuracy. The following methods were applied: color modification, 90º rotation, flip up and Gaussian blurring. Color modification has been performed with Reinhard's method [14]. For each staining, we use four color patterns corresponding to the different sites or institutions from where the databases come from. In order to achieve a similar number of tiles

in each class, for classes with much less data (such as lymphocytes), we applied additional data augmentation, which includes: 45º rotation, flip up and Gaussian blurring. Finally, data extracted from WSIs and augmented results in 210,312 training tiles. For testing we have used 25,264 tiles.

231 training tiles. For testing, we have used 25,264 tiles.

## 3. Method

We propose to obtain tumor proliferation map for each WSI under analysis by using Deep Learning CNN. The proposal follows a three-step work-flow: i) Tissue area segmentation, ii) Hot-spot detection and iii) Resultant tumor proliferation map. The workflow of this method is depicted in **Figure 3**.

### 3.1. Tissue area segmentation

Tissue segmentation is the first step in WSI analysis. It allows to limit the region under analysis, thus reducing the computational time. A basic background segmentation based on color histogram analysis is done to detect the WSI tissue regions.

### 3.2. Ki67 CNN Hot-spot detection

Hot-spot detection is central to achieve the correct tumor proliferation index. We use Ki67 data to train a CNN to detect hot-spot areas on the WSI. There are two main processes involved in this step: CNN model training and patch-based WSI classification.

#### 3.2.1 CNN model training

The CNNs are set up using the AlexNet model [15]. This CNN has eight layers of neurons: five convolutional and three fully-connected. We apply the full training and fine-tuning (FT) methods to the training process. FT is one of the transfer learning techniques based on transferring the features and the parameters of the network from a broad domain to the specific one. The re-training operation allows for optimization of the network to minimize the error in the other, more specific domain. In a case of FT, the CNN model was fine-tuned by modifying the network weights and by trimming the number of outputs in the last layer.

The full training process has been performed with the MatConvNet environment. The CNN model returns a vector with class scores representing the affinity of the image to each of the predefined classes. Classification of image tiles by the CNN is computed by finding the maximum class score for each tile. Finally, classification results from all tiles were saved as an image that had the same size as the full resolution WSI.

#### 3.2.2 Patch-based WSI classification

The network classifies each tile into one of the 11 classes: no tissue, lymphocytes and 9 class of tumor ranging from low to high proliferation index. High value of tumor class implies high tumor proliferation index, and this defines the hot-spot areas, by selecting the regions classified as high tumor density.

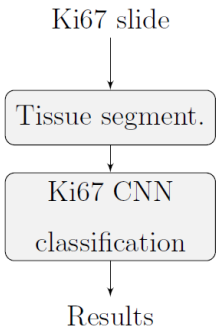
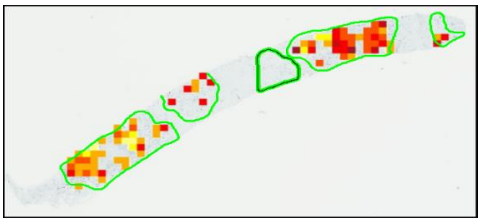


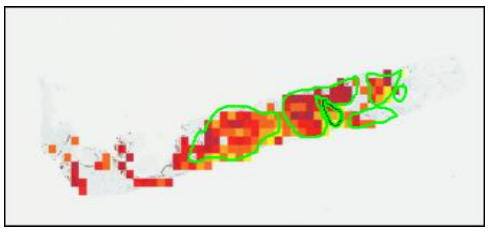

***Figure 3*** *Workflow for hot-spot detection using Ki67 data with CNN. Ki67*

## 4. Results

We performed a complete evaluation of the proposed study for automatic WSI Ki67 hot-spot detection. The goal of medical doctors is to select only a few hot-spot areas in WSI, while the automatic evaluation can detect a significantly higher number of these regions. The proposed methods allow to create a proliferation index map. **Figure 4**, shows results of automatic hot-spot detection in two specific samples. High tumor proliferation areas have been marked by red (dark red), whereas area with low tumor proliferation is marked by yellow.



a)



b)

***Figure 4*** *Automatic tumor proliferation map examples. In Green, manual annotations made by medical doctors. In yellow-red scale, automatic regions classification.*

We can observe how detected hot-spot regions match with manual pathologists' annotations. In Figure 4a, false negative detections appear in a region with too low density of cancer cells. In Figure 4b, false positive regions appear in a region where cancer cells are not infiltrating, but in-situ. These regions are not considered by pathologists in the specimen analysis.

Quantitative analysis has been performed by computing the classification accuracy. To perform the testing, we used 25,264 tiles not used in the training process. The average accuracy results are shown in **Table 1**. In this

table, we consider a binary detection between Non-tumor (classes I and II) and Tumor (classes III to IX).

| Accuracy | Sensitivity | Specificity |
|---|---|---|
| 96% | 47% | 98% |

***Table 1*** *Average values of statistical results of tiles classification (N=25,264).*

The results confirm correct tile detection, with 96% of accuracy, but confused classification between adjacent classes results in a decrease of sensitivity to a 47%.

**Table 2** shows confusion matrix for the classification within different tumor groups (low, medium and high tumor proliferation). The analysis of the confusion matrix shows that non-tumor regions are correctly classified, meanwhile, the rest of the classes present some confusion with other adjacent tumor proliferation classes. This is due to the similarity between adjacent classes, which is important but not decisive for medical diagnosis. The one that can be considered as central for pathologists is that lymphocytes are not correctly classified. This is important because the presence of lymphocytes means that the patient is reacting against tumor, and must not be wrongly classified as a false positive hot-spot. The main reason is the similarity between lymphocytes and cancer cells, which leads us to consider that Ki67 data could not be enough to achieve this differentiation.

| no tumor | | tumor proliferation [%] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| no | lympho | low | | medium | | hight | | | | |
| tumor | cytes | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 | >80 |
| 1,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 0,00 | 0,00 | 0,04 | 0,00 | 0,16 | 0,00 | 0,00 | 0,00 | 0,00 | 0,60 | 0,20 |
| 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| 0,02 | 0,00 | 0,03 | 0,34 | 0,54 | 0,06 | 0,01 | 0,00 | 0,00 | 0,00 | 0,00 |
| 0,02 | 0,00 | 0,01 | 0,05 | 0,55 | 0,20 | 0,10 | 0,01 | 0,01 | 0,01 | 0,03 |
| 0,02 | 0,00 | 0,00 | 0,00 | 0,14 | 0,39 | 0,17 | 0,18 | 0,06 | 0,01 | 0,03 |
| 0,01 | 0,00 | 0,00 | 0,00 | 0,05 | 0,11 | 0,33 | 0,34 | 0,06 | 0,05 | 0,05 |
| 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,02 | 0,13 | 0,51 | 0,20 | 0,07 | 0,05 |
| 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,00 | 0,03 | 0,16 | 0,32 | 0,35 | 0,13 |
| 0,01 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,02 | 0,10 | 0,34 | 0,51 | 0,01 |
| 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,01 | 0,01 | 0,10 | 0,88 |

***Table 2 Confusion matrix for tiles classification (N=25,264)***

## 5. Conclusions

In this paper, we have presented a method for breast cancer hot-spot detection on Ki67 WSI. It is based on CNN regions classification between 11 classes: non-tumor, lymphocytes and 9 tumor classes ranging according to tumor proliferation.

Our study shows that hot-spot maps are correctly computed close to the medical doctors' analysis, achieving 96% of accuracy. Due to the similarity between tumoral adjacent classes, the sensitivity of this proposal is 47%. This confusion can be difficult to correct in samples located at the boundaries between classes, but can be improved in the case of lymphocytes, where more data involving other stains could be necessary to correctly detect them.

Therefore, in this paper we have shown that automatic diagnosis based on CNN allows to develop an efficient tool to support Ki67 evaluation in breast cancer specimens.

## Acknowledgements

This work is partially founded by the EU FP7 Program, AIDPATH project, grant number 612471.

## References

[1] G. Bueno, M. Fernandez-Carrobles, O. Deniz, M. Garcia Rojo, New trends of emerging technologies in digital pathology, Pathobiology 83 (2016) 61-69

[2] G. Litjens, C. Sanchez, N. Timofeeva, et al., Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis, Scientific reports 6 (2016) 26286.

[3] T. Khoury, G. Zirpoli, S. M. Cohen, et al., Ki-67 expression in breast cancer tissue microarraysassessing tumor heterogeneity, concordance with full section, and scoring methods, American Journal of Clinical 474 Pathology 148 (2) (2017) 108-118

[4] A. Laurinavicius, B. Plancoulaine, A. Rasmusson, et al., Bimodality of intratumor ki67 expression is an independent prognostic factor of overall survival in patients with invasive breast carcinoma, Virchows Archiv 468 (4) (2016) 493-502.

[5] X. Lopez, O. Debeir, C. Maris, et al., Clustering methods applied in the detection of ki67 hot-spots in whole tumor slide images: An efficient way to characterize heterogeneous tissue-based biomarkers, Cytometry Part A 81A (9) (2012) 765-775.

[6] M. Khan Niazi, M. Yearsley, X. Zhou, F. WL., G. MN., Perceptual clustering for automatic hotspot detection from ki-67-stained neuroendocrine tumour images, Journal of Microscopy 256 (3) (2014) 213-225.

[7] H. Lu, T. G. Papathomas, D. van Zessen, et al., Automated selection of hotspots (ash): enhanced automated segmentation and adaptive step finding for ki67 hotspot detection in adrenal cortical cancer, Diagnostic Pathology 9 (1) (2014) 216.

[8] Z. Swiderska, A. Korzynska, T. Markiewicz, et al., Comparison of the manual, semiautomatic, and automatic selection and leveling of hot spots in whole slide images for ki-67 quantification in meningiomas, Analytical Cellular Pathology 2015 (498746) (2015) 15.

[9] N. Elie, B. Plancoulaine, J. Signolle, P. Herlin, A simple way of quantifying immunostained cell nuclei on the whole histologic section, Cytometry Part A 56A (1) (2003) 37-45.

[10] N. A. Valous, B. Lahrmann, N. Halama, et al., Spatial intratumoral heterogeneity of proliferation in immunohistochemical images of solid tumors, Medical Physics 43 (6 Part1) (2016) 2936-2947.

[11] R. Paulik, T. Micsik, G. Kiszler, et al., An optimized image analysis algorithm for detecting nuclear signals in digital whole slides for histopathology, Cytometry Part A 91 (6) (2017) 595-608.

[12] D. Pilutti, V. Della Mea, E. Pegolo, F. La Marra, F. Antoniazzi, C. Di Loreto, An adaptive positivity thresholding method for automated ki67 hotspot detection (akhod) in breast cancer biopsies, Computerized Medical Imaging and Graphics.

[13] T. Markiewicz, S. Osowski, J. Patera, W. Kozlowski, Image processing for accurate cell recognition and count on histologic slides, Analytical and quantitative cytology and histology 28 (5) (2006) 281-291.

[14] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Computer graphics and applications 21 (5) (2001) 34-41.

[15] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks (2012) 1097-1105.