

Looking Under the Hood: Deep Neural Network Visualization to Interpret Whole-Slide Image Analysis Outcomes for Colorectal Polyps

Bruno Korbar, Andrea M. Olofson, Allen P. Miraflor, Catherine M. Nicka, Matthew A. Suriawinata, Lorenzo Torresani, Arief A. Suriawinata, Saeed Hassanpour

Dartmouth College, Hanover, NH

Abstract

Histopathological characterization of colorectal polyps is an important principle for determining the risk of colorectal cancer and future rates of surveillance for patients. The process of characterization is time-intensive and requires years of specialized medical training. In this work, we propose a deep-learning-based image analysis approach that not only can accurately classify different types of polyps in whole-slide images, but also generates major regions and features on the slide through a model visualization approach. We argue that this visualization approach will make sense of the underlying reasons for the classification outcomes, significantly reduce the cognitive burden on clinicians, and improve the diagnostic accuracy for whole-slide image characterization tasks. Our results show the efficacy of this network visualization approach in recovering decisive regions and features for different types of polyps on whole-slide images according to the domain expert pathologists.

1. Introduction

At least half of western adults will have a colorectal polyp in their lifetime and one-tenth of these polyps will progress and transform to cancer [1]. If colorectal polyps are detected early they can be removed before this potential transformation. While there are multiple screening methods to detect colorectal polyps, colonoscopy has become the most common screening test in the United States [2]. Detection and histopathological characterization of colorectal polyps found in the tissue samples after colonoscopy are important parts of colorectal cancer screening, and allow for the distinction between high-risk and low-risk polyps. Specifically, sessile serrated polyps can develop more aggressively into colorectal cancer compared to other colorectal polyps, because of its

associated serrated pathway in tumorigenesis [3]. Differentiation between sessile serrated polyps and benevolent hyperplastic polyps, however, can be a challenging task even for experienced pathologists. This difficulty is likely caused by the lack the dysplastic nuclear changes that are innate to the conventional adenomatous polyps [4-7]. As evidence, a recent study of more than 7,000 patients who underwent colonoscopy in 32 centers in the US, showed that, despite the statistical unlikeliness, no sessile serrated polyp was diagnosed in multiple centers. Therefore, there is still a significant disparity in the performance and education of pathologists with regard to histologic features of colorectal polyps, and their diagnostic consistency and accuracy among pathologists [8]. Considering this challenge, image analysis tools can help clinicians improve the histopathological characterization and diagnosis of colorectal polyps, explain the underlying justification and support of assessment outcomes, and reduce their cognitive burden to enhance their accuracy in this task.

1.1. Related work

Since 2012, deep-learning methods have proven to be the state-of-the-art approach for tasks of classification and segmentation on microscopic whole-slide images, compared to the previous image processing techniques [9-11]. Janowczyk et al. performed a survey of the early applications of deep-learning in pathology, exploring domains such as lymphocyte detection and mitosis detection [11]. Some other applications of deep-learning for analyzing microscopic images include detecting metastatic breast cancer [12], finding mitotically active cells [13], identifying basal-cell carcinoma [14], and grading gliomas [15]. Most models that are mentioned in these applications used standard deep-learning architectures such as AlexNet [16] and VGG [17], which were, at the time, the best performing deep-learning architectures available.

In a relevant application to the domain of our work,

Sirinukunwattana et al. previously presented a deep-learning approach for nucleus detection and classification on microscopic images of colorectal cancer tumors [18]. This work was based on an 8-layer fully convolutional network [19], with an aim of identifying the centers of nuclei and classifying them into four categories as epithelial, inflammatory, fibroblastic, or miscellaneous.

As another related work, our previous work outlines a successful deep-learning-based classification framework for colorectal polyps on whole-slide images [20]. In the ablation study presented in this work, we found that the residual network (ResNet) architecture [21] achieves a superior performance compared to other common deep-learning architectures, such as GoogLeNet [22], VGG [17], and AlexNet [16]. In this paper, we adapt this previously presented ResNet architecture in a deep neural network visualization framework to allow for easier interpretation of whole-slide image analysis outcomes from deep-learning models.

1.2. Our contribution in this paper

Our work presented in this paper emphasizes studying the utility of different visualization methods as an approach to provide greater insight into whole-slide image classification results obtained from deep-learning models. Our proposed approach identifies decisive regions and features on whole-slide microscopic images that shape the classification outcomes for the slides. In this work, we utilize and evaluate several methods of backpropagation projections to achieve this goal. These methods project the predicted classification label for a slide through the non-zero activations in a trained deep-learning classification model to identify features relevant for that class. The manual annotation and computation requirements for our approach proposed in this paper are much simpler than a traditional segmentation approach. This is because our proposed approach relies on image crops as loose regions of interest for polyp types, rather than exact boundaries required by segmentation methods. Also, in contrast to segmentation models, our visualization approach does not require extra training and relies on a single backpropagation pass on a previously trained classification model’s network.

2. Methods

The details of our deep neural network visualization approach are as follows.

2.1. Whole-slide image classification

Our approach to detect indicative features for colorectal polyps in whole-slide images relies on a pre-trained deep-learning model for colorectal polyp classification presented in our previous work [20]. This classification model is based on ResNet architecture, and achieved average accuracy of 91.3% for the classification of 5 colorectal

polyp classes covered in this paper and outperformed other common deep-learning architectures [20]. The standard ResNet architecture in this work consists of 152, 3×3 and 1×1 convolutional layers, with additional mappings or shortcuts that can bypass several convolutional layers.

This model performs classification on small patches from the whole-slide, hematoxylin and eosin (H&E)-stained images. Subsequently, the classification of whole-slide images is done through generating small patches by a sliding window on these images, and performing a maximum-likelihood inference on the patch labels.

2.2. Network visualization and understanding

We argue that the visualization of higher-level features of the model described in the previous section can provide a useful tool in understanding the discriminative features of the colorectal polyps on H&E-stained images. Conceptually, assuming input whole-slide image x , and a neural network $f(x)$ that outputs a probability distribution over a set of colorectal polyp classes, we can use first-order Taylor approximation to define f (which relies on its gradients). Since gradients, by definition, indicate the change of a function f with respect to input x , by visualizing these gradients, we could capture the discriminative features of polyps. In practice, however, visualization of gradients obtained by backpropagation tend to be ineffective due to the interference between positive and negative gradients at a given layer [23-24]. Backpropagation outputs are often modified by approaches such as deconvolution [25] or guided backpropagation [23] in order to reduce the noise in the gradients. While we explored all three of these methods (i.e., backpropagation, guided backpropagation, and deconvolution) in order to gain insight into what the network is learning, we have not found a measurable success to interpret the pixel-space visualizations outputted by these methods.

As the aforementioned network visualization methods alone are hard to interpret to a human observer, we leverage the fact that understanding of a deep-learning model can also be obtained by visualizing only the last layer in these networks. While visualizing a fully connected layer is not feasible, as its activations are scalars, we can take advantage of a fully convolutional ResNet in order to visualize the output tensor in the last layer. This process, also known as class activation mapping (CAM), allows us to visualize a tensor associated with the classification activations of a class label [26]. As a result, this approach can show us “where the network looks” and, in turn, provide the regions of interest for pathologists to examine and verify the classification results of the model.

In addition to CAM, we explore the follow-up work by Selvaraju et al., known as Grad-CAM for network visualization [24]. Grad-CAM allows us to obtain the class-discriminative localization map by computing the gradient

of the score for the predicted class with respect to the feature maps of the convolutional layers. Much like CAM, Grad-CAM can be computed as a rectified weighted-combination of feature maps.

Selvaraju et al. also showed that by combining the Grad-CAM approach with guided backpropagation, one can obtain even finer and more localized fine-grained gradient visualizations in a pixel space [24]. Therefore, we adapt the same approach, however, before presenting the projections on a pixel space, we normalize the projected weights to obtain an even more discriminative heat-map for the input image, which we later process to obtain candidate regions of interest.

3. Experimentation

The following section details our experimentation with different visualization methods on whole-slide, H&E-stained images that contain the five most common colorectal polyps.

3.1. Dataset

The whole-slide images required to test and evaluate our methods were collected from patients who underwent colorectal cancer screening at Dartmouth-Hitchcock Medical Center, an academic quaternary care center, since January 2010. All these slides were scanned using high-throughput Leica Aperio whole-slide microscopes, at 200x magnification. Our dataset includes H&E-stained, whole-slide images for five types of colorectal polyps: hyperplastic polyp, sessile serrated polyp, traditional serrated adenoma, tubular adenoma, and tubulovillous/villous adenoma. A total of 176 whole-slide images have been collected in this study, and the distribution of which can be seen in Table 1. The use of data for this project was approved by our Institutional Review Board (IRB).

Table 1. The distribution of colorectal polyp types in H&E-stained, whole-slide images used for evaluation in this work.

Colorectal polyp type	Acronym	# Cases
Hyperplastic polyp	HP	37
Sessile serrated polyp	SSP	42
Traditional serrated adenoma	TSA	34
Tubular adenoma	TA	31
Tubulovillous/villous adenoma	TVA/V	32
Total	—	176

3.2. Image annotation

Our domain expert pathologist collaborators annotated the five different types of colorectal polyps in the H&E-stained, whole-slide images in our dataset. We used these annotations as reference standards for to evaluate our deep-learning visualization methods and to provide further insight into colorectal polyp classification on whole-slide images. Whole-slide, high-resolution images consist mostly of normal tissue, with only a small segment being related to the colorectal polyps. In this image annotation, two collaborator pathologists independently reviewed the whole-slide images in order to characterize the type of polyp present in it. Furthermore, these pathologists outline the regions in which the colorectal polyp is present by generating crops that are focused on colorectal polyps in whole-slide images. Following our previously described procedure [20], we extracted smaller crops at the same magnification of the whole-slide images, as this proved to yield superior results in previous histopathology applications [27]. Any disagreement in either classification or cropping was resolved through consultation with a third, senior gastrointestinal pathologist. When the agreement could not be reached on a polyp type or the correct cropping, the image was discarded.

3.3. Architecture and framework

Following the results from the ablation study in our previous work [20], we adopted the ResNet architecture, first introduced by Microsoft [21], with slight modifications. First, we replaced the last fully connected layer with a convolutional layer with the same input as the number of features, and one feature map output for each class. Then, we reduced the number of input features from the second-to-last convolutional layer from 2,048 to 1,024. Finally, we reduced the depth of the network to 101 layers, down from the original 152 layers. This is because we did not find any noticeable performance loss with the smaller network, while the smaller network gained marginal computational speed.

3.4. Classification Model Training

Our deep neural network visualization approach relies on a pre-trained, deep-learning classification model for colorectal polyps. To obtain this classification model, we trained a deep-learning classification model on a dataset consisting of the labeled patches of H&E-stained whole-slide images. These patches were collected as part of our previous colorectal polyp classification study [20]. A total of 2,074 H&E-stained crops were available in this dataset, with a balanced distribution among the five types of colorectal polyps discussed in this paper. We used 80% of

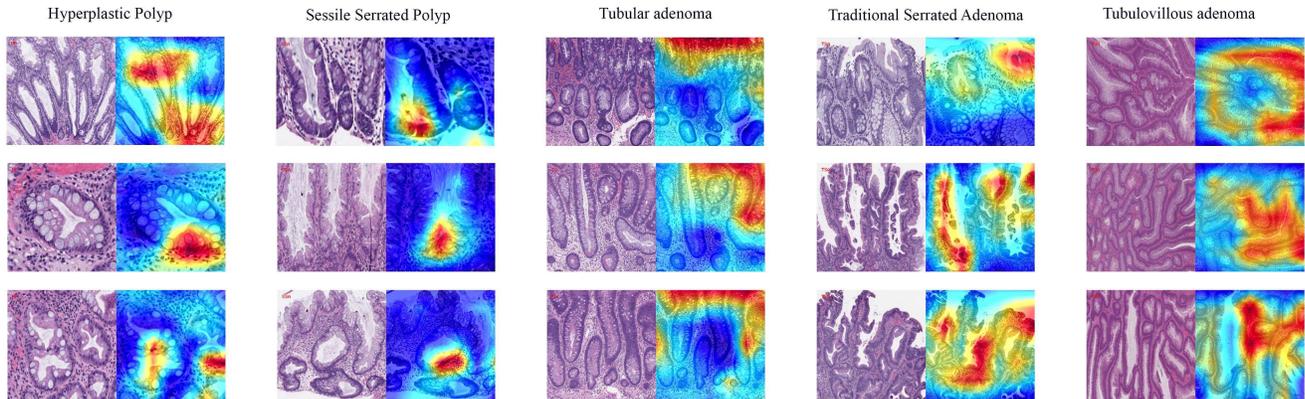


Figure 1. Projected activations for five different predicted types of colorectal polyps on portions of whole-slide images.

the crops as a training set, 10% as a validation set, and reserved approximately 10% as a test set. The use of data for this project is approved by an appropriate IRB. Each crop is processed as follows.

First, a patch is re-scaled to conform to the median of the respective axis that is computed on the training dataset. If the image size was below the median, we used zero-padding to conform it to the aspect ratio. We then subtracted the mean in order to compensate for deviation in staining concentration, and performed color jittering. Finally, we rotated the image and flipped it horizontally to make them invariant with respect to the orientation. We trained the optimal model for 100 epochs on the augmented training set, with an initial learning rate of 0.1, decreasing it 0.1 times each 50 epochs, and with 0.9 momentum. The overall training time was 26 hours on two NVIDIA K40c GPUs. The model reached optimal test error at the 86th epoch. The overall classification accuracy on the test set was 94.1%, which was slightly better than the overall classification results previously reported on the same dataset [20].

3.5. Projecting Class Activations

As discussed in the Methods section above, the visualization of indicative features and regions for colorectal polyp classification on whole-slide images relies on a gradient-based projection of the classification scores to the input pixels. To acquire this projection in our approach, we classified the whole-slide image and obtained the classification scores for different types of colorectal polyps through a forward pass on the pre-trained classification model. Subsequently, for each polyp class with a classification score greater than 0.7, we set the classification score to one, while setting the scores of the other classes to zero in the last layer of the classification network. This threshold for classification scores was proposed in our previous work through a cross-validation for binarizing the classification results [20]. Subsequently, we tried all discussed gradient-based backpropagation methods discussed earlier in the Methods section to project the scores of the predicted classes from the last layer to the

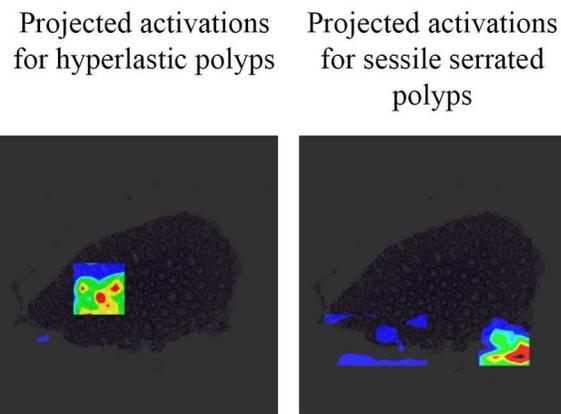


Figure 2. Projected activations for two types of colorectal polyps on a single H&E-stained, whole-slide image.

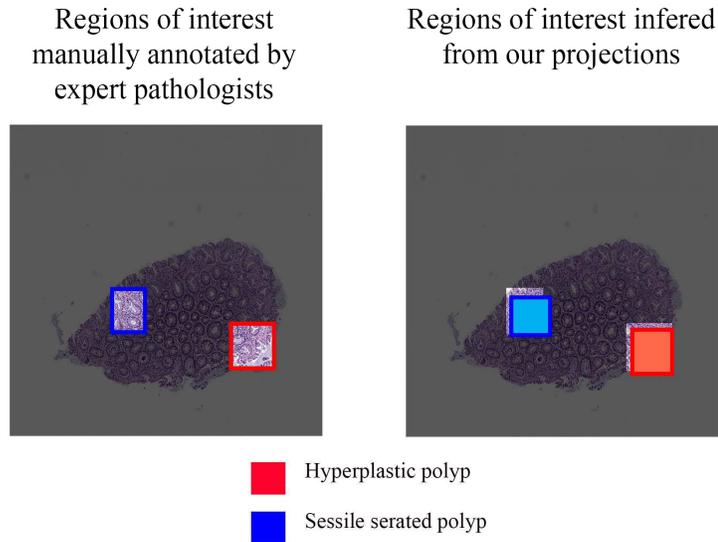


Figure 3. Manually annotated regions of interest by the domain expert pathologists and predicted regions of interest by our deep neural network visualization approach (guided Grad-CAM with boxes) for two types of colorectal polyps on a whole-slide image.

first layer of the network. These projections produce activation maps for the predicted classes on the the input pixel space. Figure 1 shows the projected activations for the five different predicted classes on parts of whole-slide images.

3.6. Generating Regions of Interest

We assign a pixel to the regions of interest of a certain class if the projected activation value of a pixel for that class is larger than a certain threshold. This threshold is learned through cross-validation on one-third of the images in our dataset. This threshold is 0.95 for the Grad-CAM projection method, and it is 0.85 for guided Grad-CAM method. Imposing thresholds on the projected activation values will produce a binary mask on the whole-slide image, which constructs the regions of interest in our application. To generate the regions of interest in guided Grad-CAM with boxes visualization method, we compute and utilize a surrounding rectangular bounding box for each region of interest on the slide. Figure 2 shows the projected activations for two types of polyps by guided Grad-CAM on a whole-slide image. Figure 3 shows the reference standard and the predicted regions of interest from these activations on the same slides.

4. Evaluation

In our evaluation, we consider a pixel that is predicted to be in the region of interest for a colorectal polyp type as a true positive if it is also recognized in the standard reference crops generated by our domain expert pathologist annotators for that polyp type. If the pixel in the predicted region of interest is not in a reference standard region, it is

considered to be a false positive. For this evaluation, we measured the standard object detection evaluation metric of intersection over union (IOU) between the predicted regions of interest and the reference standard crops for a whole-slide image and a type of colorectal polyp.

5. Results

We applied the proposed methods in this paper on our dataset of 176 H&E-stained, whole-slide images to identify discriminative regions and features for detecting colorectal polyps. Table 2 shows average IOU results for different visualization methods on whole-slide images with different types of colorectal polyps. Of note, the results from backpropagation, guided backpropagation, deconvolution, and CAM visualization methods were not included in this table because their corresponding average IOU measures were low and less than 0.2.

6. Discussion

A major shortcoming of current deep-learning models for microscopic image analysis and classification is their “black box” approach to outcomes. These methods are mostly focused on the efficacy of final results and rarely provide sufficient evidence and details on factors that contribute to their outcomes. In this paper, we provide a complimentary method to assist pathologists in the diagnosis of colorectal polyps and decision-making on their characterization, in addition to creating new training opportunities. Therefore, we leverage deep neural network visualization methods to provide insight into the colorectal polyp characterization results of a pre-trained ResNet classification model.

Table 2. Average IOU results across different colorectal polyp types in our test set for different deep neural network visualization methods.

<i>Colorectal polyp type</i>	<i>Grad-CAM</i>	<i>Guided Grad-CAM</i>	<i>Guided Grad-CAM w/ boxes</i>
HP (N=37)	0.38	0.50	0.59
SSP (N=42)	0.28	0.52	0.61
TSA (N=34)	0.18	0.42	0.49
TA (N=31)	0.10	0.42	0.47
TVA/V (N=32)	0.22	0.45	0.51
Total (N=176)	0.24	0.47	0.55

The proposed deep-learning understanding methods project the detected class activations back to the input pixel space to reveal which parts of the input image are important for classification. This projection approach provides a visualization that will identify regions and features of interest of the detected colorectal polyp on the whole-slide image. The identified regions and features are responsible for strong activations at higher layers in the colorectal polyp detection model. All the model understanding methods utilized in this work rely on gradient-based visualization approaches that provide this identification with minimal computational costs and easy interpretability. Given an image, a predicted class for the image, and a deep-learning classification model, a gradient-based approach assigns a score to the pixels of the image based on their influence on the predicted class using a single backpropagation pass through the model’s network.

In a gradient-based approach, after applying the model to an image and detecting a specific colorectal polyp, we set the score of the predicted class to maximum (i.e., 1.0) in the last layer of our classification network, while setting the scores of other classes to zero. Subsequently, through backpropagation of gradients in the network, we map the detected class score back to the input pixel space and compute an influence score for each pixel of the input image. This influence score indicates the contribution of each pixel to the classification results of the whole-slide image.

While in the domain of colorectal polyps, we found little to no use for pixel-space visualizations, regions of interest candidates proved to be a promising step for providing a clinical decision support mechanism for the characterization of colorectal polyps. Due to the nature of guided Grad-CAM, which discards negative gradients at each step, we can obtain highly-specific regions of interest

for each type of colorectal polyps. Therefore, our proposed method not only allows the clinician to correctly classify the predominant class of polyp (which we previously showed to be feasible [20]), but also to localize multiple instances and multiple types of polyps on the same whole-slide image.

Through our proposed approach, the influential regions of a whole-slide image for histopathological characterization are identified in the form of a mask, which highlights the features that contribute to the results of our deep-learning characterization model for classification of colorectal polyps on that image. Of note, our method for detecting influential features of the detected polyps uses a pre-trained classification model. Therefore, no additional annotations, segmentations, or training is required to utilize our visualization methods. In addition, the detection of important regions and features of a whole-slide image is fast because our method only requires a single backpropagation pass on the classification model’s network. As future work, we plan to apply and evaluate the deep-learning visualization framework presented in this paper to other microscopic image analysis applications, such as histopathological characterization of lung cancer, breast cancer, and melanoma tumors on whole-slide images, to assist clinicians with interpreting deep-learning image analysis outcomes.

Acknowledgements

This research was supported in part by a research grant from Dartmouth College Neukom Institute for Computational Science. The authors would like to thank the Visual Learning Group at Dartmouth College, Naofumi Tomita, and Lamar Moss for their help with this manuscript.

References

- [1] N. A. C. S. Wong, L. P. Hunt, M. R. Novelli, N. A. Shepherd, and B. F. Warren, “Observer agreement in the diagnosis of serrated polyps of the large bowel,” *Histopathology*, vol. 55, no. 1, pp. 63–66, 2009.
- [2] D. A. Lieberman, D. K. Rex, S. J. Winawer, F. M. Giardiello, D. A. Johnson, and T. R. Levin, “Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer,” *Gastroenterology*, vol. 143, no. 3, pp. 844–857, 2012.
- [3] B. Leggett and V. Whitehall, “Role of the serrated pathway in colorectal cancer pathogenesis,” *Gastroenterology*, vol. 138, no. 6, pp. 2088–2100, 2010.
- [4] H. T. Vu, R. Lopez, A. Bennett, and C. A. Burke, “Individuals with sessile serrated polyps express an aggressive colorectal phenotype,” *Dis. Colon Rectum*, vol. 54, no. 10, pp. 1216–1223, 2011.

- [5] E. Aptoula, N. Courty, and S. Lefèvre, “Mitosis detection in breast cancer histological images with mathematical morphology,” in *Signal Processing and Communications Applications Conference (SIU), 2013 21st*, 2013, pp. 1–4.
- [6] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, “Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential,” *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 97–114, 2014.
- [7] M. Veta, P. J. Van Diest, S. M. Willems, H. Wang, A. Madabhushi, A. Cruz-Roa, F. Gonzalez, A. B. L. Larsen, J. S. Vestergaard, A. B. Dahl, and others, “Assessment of algorithms for mitosis detection in breast cancer histopathology images,” *Med. Image Anal.*, vol. 20, no. 1, pp. 237–248, 2015.
- [8] K. Abdeljawad, K. C. Vemulapalli, C. J. Kahi, O. W. Cummings, D. C. Snover, and D. K. Rex, “Sessile serrated polyp prevalence determined by a colonoscopist with a high lesion detection rate and an experienced pathologist,” *Gastrointest. Endosc.*, vol. 81, no. 3, pp. 517–524, 2015.
- [9] Y. Xie, X. Kong, F. Xing, F. Liu, H. Su, and L. Yang, “Deep voting: A robust approach toward nucleus localization in microscopy images,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 374–382.
- [10] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. J. Snead, I. A. Cree, and N. M. Rajpoot, “Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1196–1206, 2016.
- [11] A. Janowczyk and A. Madabhushi, “Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases,” *J. Pathol. Inform.*, vol. 7, no. 1, p. 29, 2016.
- [12] A. A. Cruz-Roa, J. E. A. Ovalle, A. Madabhushi, and F. A. G. Osorio, “A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013, pp. 403–410.
- [13] M. G. Ertosun and D. L. Rubin, “Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks,” in *AMIA Annual Symposium Proceedings*, 2015, vol. 2015, p. 1899.
- [14] C. D. Malon, E. Cosatto, and others, “Classification of mitotic figures with convolutional neural networks and seeded blob features,” *J. Pathol. Inform.*, vol. 4, no. 1, p. 9, 2013.
- [15] H. Wang, A. Cruz-Roa, A. Basavanthally, H. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, “Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection,” in *SPIE Medical Imaging*, 2014, p. 90410B--90410B.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [17] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep inside convolutional networks: Visualising image classification models and saliency maps,” *arXiv Prepr. arXiv1312.6034*, 2013.
- [18] K. Sirinukunwattana, D. R. J. Snead, and N. M. Rajpoot, “A stochastic polygons model for glandular structures in colon histology images,” *IEEE Trans. Med. Imaging*, vol. 34, no. 11, pp. 2366–2378, 2015.
- [19] B. B. Le Cun, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Handwritten digit recognition with a back-propagation network,” in *Advances in neural information processing systems*, 1990.
- [20] B. Korbar, A. M. Olofson, A. P. Miraflor, K. M. Nicka, M. A. Suriawinata, L. Torresani, A. A. Suriawinata, and S. Hassanpour, “Deep-Learning for Classification of Colorectal Polyps on Whole-Slide Images,” Mar. 2017.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *arXiv Prepr. arXiv1512.03385*, 2015.
- [22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [23] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for simplicity: The all convolutional net,” *arXiv Prepr. arXiv1412.6806*, 2014.
- [24] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,” Oct. 2016.
- [25] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *European Conference on Computer Vision*, 2014, pp. 818–833.
- [26] B. Zhou, A. Khosla, L. A., A. Oliva, and A. Torralba, “Learning Deep Features for Discriminative Localization,” *CVPR*, 2016.
- [27] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, “Deep learning for identifying metastatic breast cancer,” *arXiv Prepr. arXiv1606.05718*, 2016.