MICCAI

Semantics-Aware Attention Guidance for Diagnosing Whole Slide Images

Kechun Liu[1,★], Wenjun Wu[1,★], Joann G. Elmore[2], and Linda G. Shapiro[1]

[1] University of Washington, Seattle, WA, 98195, USA
[2] David Geffen School of Medicine, UCLA, Los Angeles, CA 90024, USA
kechun@cs.washington.edu

Abstract. Accurate cancer diagnosis remains a critical challenge in digital pathology, largely due to the gigapixel size and complex spatial relationships present in whole slide images. Traditional multiple instance learning (MIL) methods often struggle with these intricacies, especially in preserving the necessary context for accurate diagnosis. In response, we introduce a novel framework named Semantics-Aware Attention Guidance (`SAG`), which includes 1) a technique for converting diagnostically relevant entities into attention signals, and 2) a flexible attention loss that efficiently integrates various semantically significant information, such as tissue anatomy and cancerous regions. Our experiments on two distinct cancer datasets demonstrate consistent improvements in accuracy, precision, and recall with two state-of-the-art baseline models. Qualitative analysis further reveals that the incorporation of heuristic guidance enables the model to focus on regions critical for diagnosis. `SAG` is not only effective for the models discussed here, but its adaptability extends to any attention-based diagnostic model. This opens up exciting possibilities for further improving the accuracy and efficiency of cancer diagnostics. Our code is publicly accessible at https://github.com/kechunl/SAG.

Keywords: Attention Guidance · Whole Slide Image Diagnosis · Semantic Heuristic · Multiple Instance Learning · Transformers

1 Introduction

In recent years, the landscape of histopathological image analysis has been profoundly reshaped by the advent of deep learning technologies [1,2]. However, learning from gigapixel whole slide images (WSIs) remains a difficult problem, as their size makes end-to-end learning computationally expensive. Thus, WSI classification methods often follow a bag-of-words (BoW) model for learning representations, wherein a large patch of a whole slide image is treated as a bag or set, while smaller image patches inside a bag are treated as words (or instances). Following this BoW model, many studies adopt a multiple instance learning-based (MIL) approach, which involves first extracting word-level feature

---
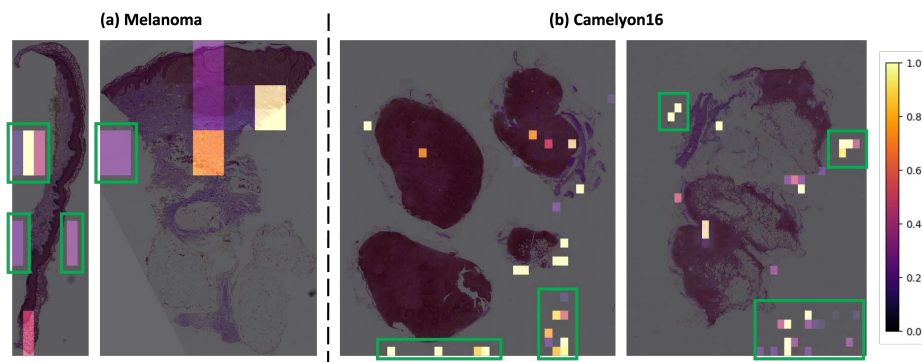
★ contribute equally to this work.

Fig. 1. Visualization of the baseline model's (ScAtNet [13]) attention on (a) skin biopsy WSIs in the melanoma dataset and (b) breast biopsy WSIs in the Camelyon16 dataset. Green boxes show examples of the baseline model mistakenly focusing on background regions. The signal and attention values are normalized for visualization purposes.

representations and then applying global aggregation to bags of word-level representations to obtain WSI-level representations. These approaches [3,4,5,6,7] are good at reducing the computational cost and offer a workaround by segmenting WSIs into smaller, and more manageable patches.

However, how pathologists approach diagnosis is very different from MIL models. Pathologists begin their evaluation by identifying suspicious regions at low magnification to form initial hypotheses. They then switch to high magnification to examine individual cells, mitotic counts, structures like ducts, etc., ultimately reaching a definitive diagnosis [8]. In contrast, by treating image patches independently, MIL models disregard the multi-scale nature of pathology, where zooming in and out is crucial for comprehensive assessment. This limitation in capturing long-range interactions between entities hinders MIL models from effectively capturing the nuanced details critical for accurate diagnosis.

To learn a better global representation, transformer models have been adopted to grasp the interdependencies among patches and formulate comprehensive representations, notably advancing beyond the MIL's limitations [9,10,11,12,13]. A few studies extract features from multiple resolutions and aggregate them hierarchically or concatenate them to predict the diagnosis class [13,14,15]. Specifically, ScAtNet [13] employs a transformer-based end-to-end network that adapts to the information from different input scales through self-attention and predicts the classification label. Results show that ScAtNet outperforms other MIL methods by a large margin in the task of melanoma diagnosis. However, such models often mistakenly focus on non-cancerous regions or just empty spaces, as highlighted by the green boxes in Fig. 1. This problem brings up questions about how well these models can be interpreted, how reliable they are, and if they really match up with the way pathologists diagnose.
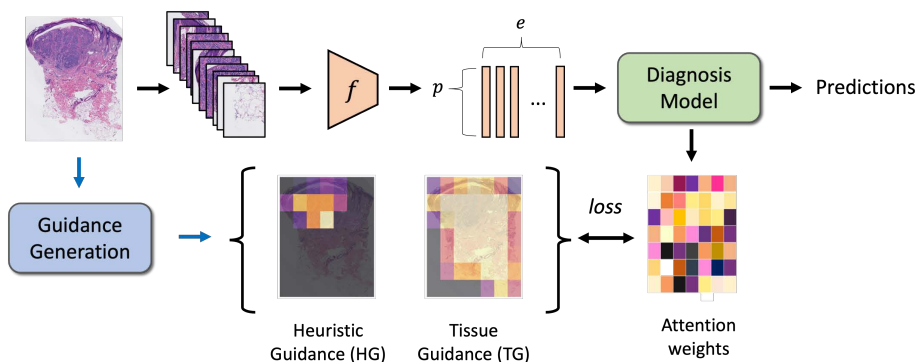
Fig. 2. Overview of the `SAG` approach for improving WSIs diagnosis models. First, a high-resolution histopathological image is divided into p number of non-overlapping patches. Then, patch embeddings are obtained using an off-the-shelf feature extractor f. Subsequently, a diagnostic network utilizes the $p \times e$-dimensional feature map for classification into distinct categories. During training, heuristic guidance (**HG**) and tissue guidance (**TG**) are leveraged to supervise the attention within the diagnosis model, ensuring the focus on diagnostically relevant regions.

In response, integrating additional domain information into diagnostic models has emerged as a promising strategy. Such efforts not only enhance classification accuracy but also improve model performance, especially in scenarios where data is scarce. Miao *et al.* [16] and Tourniaire *et al.* [17] improve attention learning with binary knowledge maps, a step towards integrating prior knowledge into WSI diagnosis. Limited to binary representations, these studies suggest the potential for richer prior knowledge to improve accuracy. Chen *et al.* [10] broadens the scope by leveraging genomics information in addition to WSIs to predict patient outcomes. Yet, their approach lacks adaptability to other modalities and is hard to integrate additional guidance signals.

Recognizing the limitations of current methods, we propose a Semantics-Attention-Guiding framework, `SAG`, whose key contributions are:

– A novel attention guiding module that is applicable to any attention-based multiple instance learning or Transformer models.
– A flexible attention-guiding loss to effectively incorporate varied semantic information, such as tissue and cancerous region masks.
– A heuristic attention-generation method to convert diagnostically relevant entities to heuristic-guidance signals.
– Improving state-of-the-art methods on two datasets of different cancer types.

2   Methodology

Our `SAG` framework aims to infuse diagnostic models with relevant knowledge, thereby enhancing the diagnostic performance and the interpretability of attention-supervised representations. This versatile framework is compatible with a broad

range of attention-based MIL and transformer methods. Fig. 2 illustrates our SAG pipeline, which includes three main components: 1) generate patchwise embeddings with an off-the-shelf feature extractor, 2) learn diagnostic patterns from these embeddings via a diagnosis network, and 3) utilize an attention-guiding loss that leverages heuristic guidance ($\mathbf{HG}$) and tissue guidance ($\mathbf{TG}$). In the following sections, we give the details of the proposed attention guidance.

2.1   Diagnosis Models

We employ a pre-trained feature extractor f for patch embedding extraction. The implementation detail of f is provided in Sec. 3.2. Moreover, to demonstrate the versatility and model-agnostic nature of our SAG framework, we apply SAG to two state-of-the-art baseline models: a transformer-based model, ScAtNet [13], and an MIL-based model, ABMIL [4].

2.2   Attention Weights

First, we partition an image into p input patches. For transformer-based models, the architecture consists of l layers with h self-attention heads per layer. Given embeddings $q, k, v \in \mathbb{R}^{p \times e_k}$ projected from the inputs, each attention head induces a pairwise similarity from query q and key k to transform the value v. The similarity ($\mathbf{A}$) and the model attention weights ($\mathbf{MA_t}$) of the transformers are computed as follows:

$$\mathbf{A} = \text{softmax}(\frac{qk}{\sqrt{D_h}}) \in \mathbb{R}^{p \times p},$$

$$\mathbf{MA_t} = \frac{1}{p} \sum_{i=1}^{p} A_i \in \mathbb{R}^p. \tag{1}$$

The model attention weights ($\mathbf{MA_m}$) of the MIL methods are formulated as the weighted aggregation of instance embeddings [4]:

$$\mathbf{MA_m} = f(x) \in \mathbb{R}^p, \tag{2}$$

where f denotes the linear layers to learn the attention weights, and $x \in \mathbb{R}^{p \times e}$ denotes the embeddings from p patches.

2.3   Guidance Generation

To regularize the model's attention $\mathbf{MA}$, we induce two types of semantic attention guidance: tissue guidance ($\mathbf{TG}$) and heuristic guidance ($\mathbf{HG}$) (Fig. 3), each represented as a vector $\in \mathbb{R}^p$. The generation of attention guidance is described in two steps: 1) Acquisition of tissue mask and diagnostic heuristics, and 2) Calculation of guidance weights.

To obtain the tissue mask for $\mathbf{TG}$, Otsu's method [18] is used to perform high-quality segmentation of tissue patches. This process transforms the input image shown in Fig. 3a into the binary tissue mask shown in Fig. 3b.
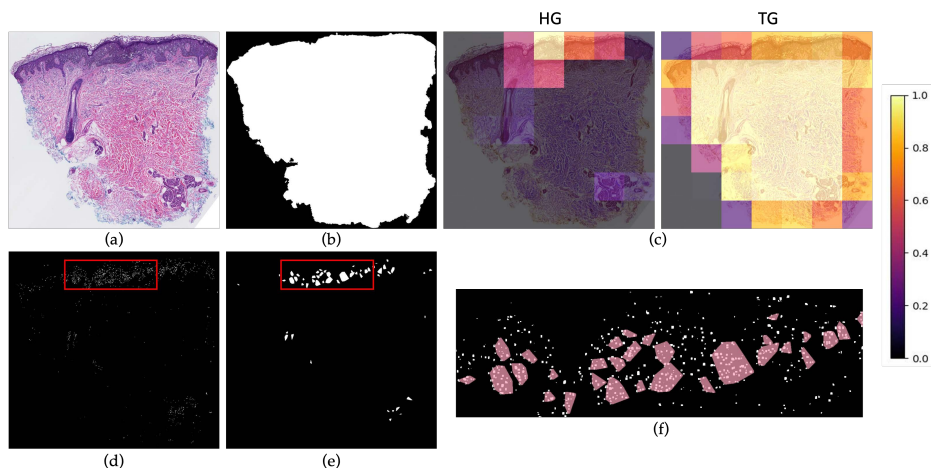
Fig. 3. Generation of attention guidance: (a) H&E sample image. (b) Tissue segmentation mask. (c) **HG** and **TG**. The values are normalized for visualization purpose. (d) Cellular entities detected (**zoom-in for best view**). (e) Convex hull of cellular clusters. (f) A zoomed-in view of the red boxes in (d) and (e). The convex hull is rendered with red color.

To obtain **HG**, we exploit dataset- and disease-specific prior knowledge, such as structures, tissues, and cells. In the example shown in Fig. 3, we first perform cell segmentation for a specific cell type (Fig. 3d). Then, groups of cells are aggregated via the density-based spatial clustering algorithm DBSCAN [19]. Next, the convex hull [20] is generated for each cluster (Fig. 3e) and utilized as the semantic signal for attention supervision (Fig. 3f).

To calculate the guidance weight $W \in \mathbb{R}^p$, we leverage Eqn. 3 to transform the heuristic signals (**HG**) and the tissue masks (**TG**) into the attention supervision (Fig. 3c):

$$W_i^k = \frac{M_i^k}{\sum_{j=1}^p M_j^k}, \quad k \in \{\mathbf{TG}, \mathbf{HG}\}, \tag{3}$$

where W_i^k denotes the guidance weight of patch i, and M_i^k is the mask area ratio of patch i.

2.4   Loss Functions

Since heuristic guidance (**HG**) reflects the relevance to the diagnosis, we employ the mean squared error (MSE) loss, L_{mse}, to regularize **MA**:

$$L_{mse} = \frac{1}{p} \sum_{i=1}^p (W_i^{\mathbf{HG}} - \mathbf{MA}_i)^2. \tag{4}$$

On the other hand, tissue guidance (**TG**) is useful in guiding the model to focus on tissue patches and ignore the background and artifact patches. Thus, we

employ a less constrained loss, $L_{in\&out}$, which sums the attention weights outside of the tissue and the negative attention weights inside the tissue, as defined in Eqn. 5 below:

$$L_{in\&out} = \frac{1}{p}(- \sum_{i,W_i^{\mathbf{TG}}>0}^{p} \mathbf{MA}_i + \sum_{i,W_i^{\mathbf{TG}}=0}^{p} \mathbf{MA}_i).$$ (5)

For joint learning, we leverage uncertainty weighting, $\mathcal{UW}$ [21], which weighs multiple loss functions by considering the homoscedastic uncertainty of each task. The overall loss function is defined as:

$$L = \mathcal{UW} \otimes \{L_{cls}, L_{mse}, L_{in\&out}\},$$ (6)

where L_{cls} is the cross entropy loss for the classification task.

3 Experiments and Results

3.1 Datasets

Melanoma. The melanoma diagnosis dataset used in this study was acquired as a part of the MPATH study (R01CA151306) [22] and consists of 222 H&E stained WSIs. There are four classes in this dataset: 1) mild and moderate dysplastic nevi, 2) melanoma in situ, 3) invasive melanoma stage pT1a, and 4) invasive mealnoma stage $\geq$ pT1b. In our study, we use a random split of 89/22/111 samples for training, validation and testing. We follow the preprocessing steps in ScAtNet[13] which crops the slice into 25, 49, and 81 patches in 7.5x, 10x, and 12.5x magnifications.

Camelyon16. Camelyon16 [23] is a public dataset comprising 400 H&E stained WSIs from breast cancer. The WSIs are diagnosed into two classes: normal and tumor. We use the official split of 271/129 slides for training and testing. To train ABMIL, we follow DSMIL [6], which crops the WSI into 224x224 sized non-overlapping patches in 20x magnification, and excludes background patches, leaving around 15K patches per bag on average. To train ScAtNet on breast biopsies, we adapted the original skin biopsy patch size while adjusting the number of patches per WSI (10x magnification) to maintain similar content per patch. The result is 35×35, or $1,225$ number of crops. This ensures consistent representation and preserves model architecture.

3.2 Implementation Details

Feature Extraction and Attention Guidance. For the melanoma dataset, an ImageNet pre-trained MobileNetV2 [24] extracts a 1280-dimensional feature vector for each patch, as described in Sec. 2.1. Since melanocytes are believed to be highly informative about melanoma diangosis, an open-sourced off-the-shelf melanocyte detection model [25] is employed to generate the cellular entity

map that eventually transforms to **HG**, as described in Sec. 2.3. To cluster the cell entities, DBSCAN in the scikit-learn package [26] is used with `eps=20` and `min_samples=5`. **TG** is generated using Otsu thresholding [18].

For Camelyon16 [23], a SimCLR pretrained by DSMIL [6] extracts a 512-dimensional feature vector for each patch. Moreover, the metastasis mask and tissue mask in the dataset are utilized for **HG** and **TG**.

Diagnosis Models and Training Details. `SAG` is applied to two models: a transformer model, ScAtNet [13], and a MIL model, ABMIL [4]. For ScAtNet, we impose **TG** across all attention heads and impose **HG** on half of the attention heads. This maintains the model's adaptability and accommodates potential noise in **HG**. For ABMIL, we apply both **HG** and **TG** on the melanoma dataset, while we only apply **HG** to Camelyon16 as the dataset already exclude background patches. We use ABMIL's [4] and ScAtNet's [13] public codebase for implementation and train models under their experimental settings.

3.3   Results

Table 1 compares the overall performance of `SAG` on different datasets and backbone models, demonstrating its consistent ability to enhance diagnostic performance in histopathological image analysis. For each setting, we conduct 15 runs of experiments with randomly sampled seeds for model initialization and report the average.

Table 1. Experimental Results of SAG across single-scale (SC) and multi-scale (MC) configurations for Melanoma and Camelyon16 datasets. Baseline methods are indicated with a †. Performance metrics include Accuracy (Acc), Precision (P), Recall (R), and Area Under the Curve (AUC).

	SAG		Melanoma				Camelyon16			
Methods	HG	TG	Acc	P	R	AUC	Acc	P	R	AUC
ScAtNet (SC)†[13]			55.03	57.17	55.36	77.38	67.79	58.17	57.51	70.28
ScAtNet (SC)	✓		57.14	59.57	57.31	78.75	68.71	58.50	64.01	72.39
ScAtNet (SC)	✓	✓	56.67	60.27	56.66	79.72	**71.60**	**64.45**	61.22	71.87
ScAtNet (MC)†			58.16	61.54	58.21	79.54	66.82	55.98	61.22	69.45
ScAtNet (MC)	✓		59.95	64.77	60.13	81.58	67.91	57.28	**66.39**	72.26
ScAtNet (MC)	✓	✓	**62.71**	**65.23**	**63.34**	**82.03**	70.13	60.53	62.58	**73.13**
Best Improvement △			+4.55	+3.69	+5.13	+2.49	+3.81	+6.28	+6.50	+3.68
ABMIL†[4]			45.55	48.23	46.42	68.07	93.02	92.47	92.79	97.52
ABMIL	✓		51.59	**57.42**	51.02	**74.68**	**94.73**	**94.61**	**94.17**	**97.80**
ABMIL	✓	✓	**52.01**	56.25	**51.84**	74.35		*Not Applicable*		
Best Improvement △			+6.46	+9.19	+5.42	+6.28	+1.71	+2.14	+1.38	+0.28

Notably, incorporating `SAG` into single- and multi-scale ScAtNet models on the melanoma dataset yields significant improvements, particularly with multi-scale inputs achieving a 4.55% accuracy increase (Table 1). Similar trends are observed on Camelyon16, where `SAG` boosts accuracy across ScAtNet configurations (3.81% for multi-scale) and increases ABMIL's accuracy by 1.71% (Ta-
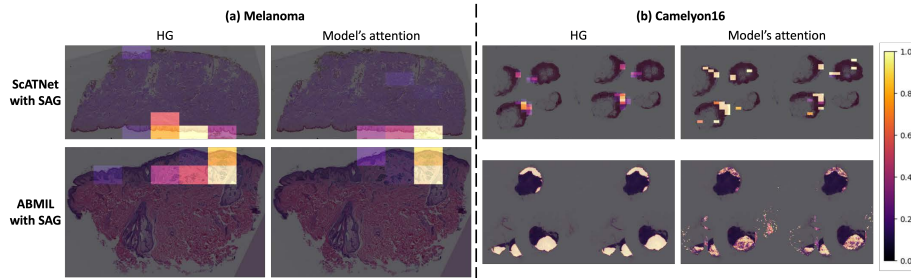
Fig. 4. Comparative visualizations of **HG** and the models' attention under `SAG`'s training on the melanoma and Camelyon16 datasets. The images are sampled from test set. The **HG** and attention values are normalized for visualization purpose.

ble 1). These improvements highlight `SAG`'s effectiveness in refining focus and enhancing the models' diagnostic performance.

In our analysis, we observe that ABMIL exhibits superior diagnostic performance on the Camelyon16 dataset (94.73% vs. 71.60%), whereas ScAtNet is more effective on the melanoma dataset (62.71% vs 45.52%). This distinction in model efficacy can be attributed to the intrinsic characteristics of these datasets and the models' specific designs. Notably, our melanoma dataset, presenting a four-class classification problem, requires a comprehensive understanding of the entire image at multiple scales and holistic levels. This aligns well with ScAtNet's transformer-based architecture, which excels at capturing long-range dependencies and aggregating multi-scale information through attention mechanisms [13]. In contrast, the Camelyon16 dataset, being a binary classification problem, prioritizes local feature identification for diagnosis, which aligns with ABMIL's MIL-based approach, suggesting why ABMIL outperforms in this context. On the other hand, ScAtNet's complexity and multi-scale inputs may not offer significant benefits here due to overfitting risks. This highlights the importance of choosing an appropriate method based on the specific data characteristics.

To further illustrate, Fig. 4 visualizes the attention patterns of ScAtNet and ABMIL on both datasets compared to **HG**. We notice that `SAG` encourages the model to focus on diagnostically relevant regions. These visualizations effectively demonstrate `SAG`'s capacity to guide attention and improve interpretability. Additional visualizations are available in the appendix for further exploration.

4   Conclusion

Motivated by our observation of misplaced attention on irrelevant regions in previous approaches, we propose a novel framework called Semantics-Aware Attention Guidance (`SAG`). `SAG` integrates tissue and heuristic attention guidance to better emulate the diagnostic process of pathologists, focusing on meaningful interconnections within WSIs. This targeted approach enables `SAG` to enhance model performance across various datasets with limited size and potentially noisy

annotations, highlighting its contribution to improving the precision and reliability of computational diagnostics.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Echle, A., Rindtorff, N.T., Brinker, T.J., Luedde, T., Pearson, A.T., Kather, J.N.: Deep learning in cancer pathology: a new generation of clinical biomarkers. British journal of cancer 124(4), 686–696 (2021)
2. Chen, X., Wang, X., Zhang, K., Fung, K.M., Thai, T.C., Moore, K., Mannel, R.S., Liu, H., Zheng, B., Qiu, Y.: Recent advances and clinical applications of deep learning in medical image analysis. Medical Image Analysis 79, 102444 (2022)
3. Hou, L., Samaras, D., Kurc, T.M., Gao, Y., Davis, J.E., Saltz, J.H.: Patch-based convolutional neural network for whole slide tissue image classification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2424–2433 (2016)
4. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: International conference on machine learning. pp. 2127–2136. PMLR (2018)
5. Mercan, C., Aygunes, B., Aksoy, S., Mercan, E., Shapiro, L.G., Weaver, D.L., Elmore, J.G.: Deep feature representations for variable-sized regions of interest in breast histopathology. IEEE journal of biomedical and health informatics 25(6), 2041–2049 (2020)
6. Li, B., Li, Y., Eliceiri, K.W.: Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 14318–14328 (2021)
7. Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F.: Data-efficient and weakly supervised computational pathology on whole-slide images. Nature biomedical engineering 5(6), 555–570 (2021)
8. Mello-Thoms, C., Mello, C.A., Medvedeva, O., Castine, M., Legowski, E., Gardner, G., Tseytlin, E., Crowley, R.: Perceptual analysis of the reading of dermatopathology virtual slides by pathology residents. Archives of pathology & laboratory medicine 136(5), 551–562 (2012)
9. Myronenko, A., Xu, Z., Yang, D., Roth, H.R., Xu, D.: Accounting for dependencies in deep learning based multiple instance learning for whole slide imaging. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 329–338. Springer (2021)

10. Chen, R.J., Lu, M.Y., Weng, W.H., Chen, T.Y., Williamson, D.F., Manz, T., Shady, M., Mahmood, F.: Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4015–4025 (2021)

11. Chen, R.J., Chen, C., Li, Y., Chen, T.Y., Trister, A.D., Krishnan, R.G., Mahmood, F.: Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16144–16155 (2022)

12. Zheng, Y., Gindra, R.H., Green, E.J., Burks, E.J., Betke, M., Beane, J.E., Kolachalama, V.B.: A graph-transformer for whole slide image classification. IEEE transactions on medical imaging 41(11), 3003–3015 (2022)

13. Wu, W., Mehta, S., Nofallah, S., Knezevich, S., May, C.J., Chang, O.H., Elmore, J.G., Shapiro, L.G.: Scale-aware transformers for diagnosing melanocytic lesions. IEEE Access 9, 163526–163541 (2021)

14. Guo, Z., Zhao, W., Wang, S., Yu, L.: Higt: Hierarchical interaction graph-transformer for whole slide image analysis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 755–764. Springer (2023)

15. Shi, J., Tang, L., Li, Y., Zhang, X., Gao, Z., Zheng, Y., Wang, C., Gong, T., Li, C.: A structure-aware hierarchical graph-based multiple instance learning framework for pt staging in histopathological image. IEEE Transactions on Medical Imaging (2023)

16. Miao, K., Gokul, A., Singh, R., Petryk, S., Gonzalez, J., Keutzer, K., Darrell, T.: Prior knowledge-guided attention in self-supervised vision transformers. arXiv preprint arXiv:2209.03745 (2022)

17. Tourniaire, P., Ilie, M., Hofman, P., Ayache, N., Delingette, H.: Ms-clam: Mixed supervision for the classification and localization of tumors in whole slide images. Medical Image Analysis 85, 102763 (2023)

18. Zhang, J., Hu, J.: Image segmentation based on 2d otsu method with histogram analysis. In: 2008 international conference on computer science and software engineering. vol. 6, pp. 105–108. IEEE (2008)

19. Ester, M., Kriegel, H.P., Sander, J., Xu, X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: kdd. vol. 96, pp. 226–231 (1996)

20. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17, 261–272 (2020)

21. Kendall, A., Gal, Y., Cipolla, R.: Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7482–7491 (2018)

22. Elmore, J.G., Barnhill, R.L., Elder, D.E., Longton, G.M., Pepe, M.S., Reisch, L.M., Carney, P.A., Titus, L.J., Nelson, H.D., Onega, T., et al.: Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. bmj 357 (2017)

23. Bejnordi, B.E., Veta, M., Van Diest, P.J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J.A., Hermsen, M., Manson, Q.F., Balkenhol, M., et al.: Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama 318(22), 2199–2210 (2017)
24. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)
25. Liu, K., Li, B., Wu, W., May, C., Chang, O., Knezevich, S., Reisch, L., Elmore, J., Shapiro, L.: Vsgd-net: Virtual staining guided melanocyte detection on histopathological images. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 1918–1927 (2023)
26. Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., Varoquaux, G.: API design for machine learning software: experiences from the scikit-learn project. In: ECML PKDD Workshop: Languages for Data Mining and Machine Learning. pp. 108–122 (2013)