

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/189741/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Sun, Shichang, Yang, Haolin, Wang, Haibo, Du, Shun, Liu, Shuang and Lai, Yukun 2026. DACAS: Distortion-Aware Capacity Allocation for Style-augmented domain generalization in computational pathology. Presented at: ICANN 2026, Padua, Italy, 14-17 September 2026. Published in: Pasa, L., Lintas, A., Tetko, I.V., Micheli, A., Navarin, N., Villa, A.E.P., Tortorella, D. and Polato, M. eds. Artificial Neural Networks and Machine Learning – ICANN 2026. Lecture Notes in Computer Science (17092) Springer Nature, pp. 583-594. 10.1007/978-3-032-38407-2_46

Publishers page: https://doi.org/10.1007/978-3-032-38407-2_46

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



DACAS: Distortion-Aware Capacity Allocation for Style-Augmented Domain Generalization in Computational Pathology

Shichang Sun^{1,2}, Haolin Yang¹, Haibo Wang³, Shun Du¹, Shuang Liu^{1*}, and Yukun Lai²

¹ Dalian Minzu University, Dalian, China
shichang2026@outlook.com, abc9562@qq.com, 2298793424@qq.com,
dlnuliushuang@qq.com

² Cardiff University, Cardiff, United Kingdom
laiy4@cardiff.ac.uk

³ Central South University, Changsha, China
wang_haibo2017@163.com

Abstract. Data augmentation improves domain generalization by increasing distributional coverage, but may also introduce structured distortion that degrades semantic fidelity, especially in structure-sensitive tasks such as computational pathology. Existing approaches primarily address this trade-off by optimizing augmentation policies or softening supervision, implicitly treating all augmentation-induced variations as equally important in the learned representation. In this work, we revisit this problem from a representation perspective. We argue that augmentation reshapes the composition of information, requiring the representation to selectively encode these variations. As a result, the key challenge is not only which variations to generate, but which variations should be preferentially encoded and how representational resources should be allocated accordingly. To this end, we propose a distortion-aware capacity allocation framework that operates after augmentation. We introduce an adaptive discrete bottleneck to adapt representation granularity at the sample level, together with a region-aware distortion variable that modulates spatial encoding according to structural sensitivity. In this way, the model redistributes limited representational resources to preserve structure-critical information while absorbing less informative variations. Experiments on colorectal cancer histopathology show that the proposed approach consistently improves generalization under identical augmentation settings. These results highlight post-augmentation representation allocation as a key mechanism for handling heterogeneous distortion in domain generalization.

Keywords: Domain Generalization · Representation Learning · Rate-Distortion Optimization · Vector Quantization · Style Augmentation.

* Corresponding author.

1 Introduction

Style-based data augmentation has been widely adopted to expose models to diverse visual conditions and mitigate domain shift by perturbing low-level appearance while preserving semantic content. By altering color, texture, or staining patterns, these methods aim to encourage models to rely on diagnostically relevant structures rather than dataset-specific appearance cues [3, 13]. However, while augmentation increases distributional coverage, it may also introduce structured distortion that degrades semantic fidelity, especially in structure-sensitive regions. In practice, excessive or uneven perturbations may distort diagnostically critical regions, leading to unstable generalization performance.

A key observation is that the effect of appearance distortion is spatially heterogeneous: different tissue regions tolerate visual perturbations to different degrees. While certain regions can absorb substantial appearance variation without affecting diagnostic cues, others are highly sensitive to distortion and require faithful representation. Treating all perturbations as equally admissible therefore risks inefficient encoding—either overfitting to benign variations or underrepresenting critical structures [17]. This suggests that the limitations of existing approaches arise not only from how variations are generated, but from how they are encoded in the learned representation.

Motivated by this observation, we revisit domain generalization under style perturbations from a representation perspective. Augmentation reshapes the composition of information, producing mixtures of structure-critical and less informative variations. The learned representation must therefore encode these variations selectively rather than uniformly. From this perspective, the central challenge is not only which variations to generate, but which variations should be preferentially encoded and how representational resources should be allocated accordingly.

To address this challenge, we propose **DACAS** (**D**istortion-**A**ware **C**apacity **A**llocation for **S**tyl**e**-**A**ugmented **D**ata), a framework for adaptive representation allocation under augmentation-induced variability. At the sample level, we introduce multiple discretization paths with different capacity budgets, enabling adaptive selection of representation granularity across inputs with varying complexity. At the region level, we incorporate a distortion-aware variable that modulates encoding according to structural sensitivity, ensuring that diagnostically critical regions are encoded with higher fidelity under appearance perturbations. Together, these mechanisms enable the model to redistribute representational resources conditioned on both feature complexity and spatial distortion sensitivity.

We evaluate the proposed framework on colorectal cancer histopathology under diverse style perturbations and domain shifts. Experimental results demonstrate consistent improvements in generalization performance under identical augmentation settings. These findings highlight the importance of post-augmentation representation allocation for robust domain generalization. Our contributions are summarized as follows:

- We show that, beyond augmentation design, domain generalization under style augmentation critically depends on how augmentation-induced variations are encoded.
- We propose a distortion-aware capacity allocation framework that adaptively distributes representation across samples and spatial regions under heterogeneous appearance distortion.
- We demonstrate improved classification robustness on histopathology datasets under both in-distribution and domain-shift settings, supported by ablation and distortion-complexity analyses.

2 Related Work

2.1 Style Augmentation under Domain Shift

Augmentation-based methods have been widely used to improve domain generalization by increasing distributional diversity. Policy learning approaches such as AutoAugment [1] and RandAugment [2] identify effective transformation strategies, while techniques such as MixUp [16] mitigate the adverse effects of augmentation by softening supervision.

Histopathology images suffer from severe stain-induced domain shift across institutions and acquisition protocols, which substantially degrades model generalization. While conventional approaches rely on stain normalization or augmentation [12, 9], recent studies have successfully employed deep networks to predict clinically meaningful MSI (Microsatellite Instability) status, an important biomarker for cancer diagnostics, directly from stained images [4]. Yamashita et al. [14] further validated this task on colorectal cancer through a large-scale diagnostic study, demonstrating its robustness across cohorts. We adopt MSI classification as a representative downstream task for evaluating generalization under style perturbations.

Beyond stain-specific perturbations, domain-agnostic style transfer replaces low-level texture while preserving high-level semantics, providing a more expressive augmentation paradigm [13, 15]. STRAP demonstrates that stylistic perturbations from non-medical sources can outperform traditional stain normalization for domain generalization [13]. SemanticStyleGAN further introduces CLIP-guided region-level manipulation [7]. Building on these advances, CLAP integrates cross-modality style guidance with semantic-region control to generate structure-preserving stain-style variations [8].

Despite the increased diversity and controllability of these methods, they primarily focus on how variations are generated or regularized. They implicitly assume that augmentation-induced variations can be uniformly encoded by the learned representation, without explicitly modeling region-dependent distortion sensitivity.

2.2 Bottlenecked Representation via Vector Quantization

Vector quantization (VQ) provides a discrete representation mechanism closely related to the Information Bottleneck principle [11, 10], enabling robustness

through capacity-constrained latent spaces. Subsequent work explored rate-adaptive and hybrid quantization schemes, primarily targeting efficient compression rather than representation robustness.

In computational pathology, VQ-based bottlenecks have been used to improve robustness under limited supervision and domain shift [6]. However, most existing methods rely on fixed-capacity codebooks, making them ill-suited to handling heterogeneous representational complexity induced by strong style augmentation. Liu et al. [5] proposed Adaptive Discrete Communication Bottlenecks, which dynamically adjust discretization granularity based on input complexity.

Despite their effectiveness, existing bottleneck and adaptive quantization methods typically assume homogeneous distortion and allocate capacity solely based on representation complexity. This becomes suboptimal when distortion is spatially heterogeneous, where different regions require different encoding fidelity. In contrast, our work explicitly models augmentation-induced distortion as region-dependent and incorporates this signal into a conditional rate-distortion formulation, enabling principled, region-aware allocation of representation capacity under structured style perturbations.

3 Method

3.1 Problem Setting

Figure 1 illustrates the adaptive discrete bottleneck that realizes conditional rate-distortion optimization through representation capacity allocation. The allocation policy is instantiated by sample-level path routing and region-level distortion weighting, which are detailed in the following subsections.

We study representation learning under style augmentation with structured distortion. Let $x \in \mathcal{X}$ denote a style-augmented image, and $h = f_{\text{enc}}(x)$, $h \in \mathbb{R}^m$ be its continuous latent representation produced by a feature extractor f_{enc} . While style augmentation increases appearance diversity, it also induces appearance distortion whose impact is spatially heterogeneous. In particular, diagnostically relevant structures are often highly sensitive to distortion, whereas background regions can tolerate substantial appearance variation.

From a representation perspective, the learned features are realized through a finite set of discrete representational units. In this work, we define representation capacity as how information is distributed across such units, including both their number and their assignment across samples and spatial regions. Regardless of model size, information must still be encoded through these units (e.g., tokens, segments, or codebook entries), making selective allocation essential.

We therefore formulate the problem from a rate-distortion perspective, where capacity corresponds to the effective rate of representation. Our objective is not to maximize representation diversity per se, but to allocate capacity in a distortion-aware manner, such that structure-critical information is preserved under augmentation while less informative variations are absorbed.

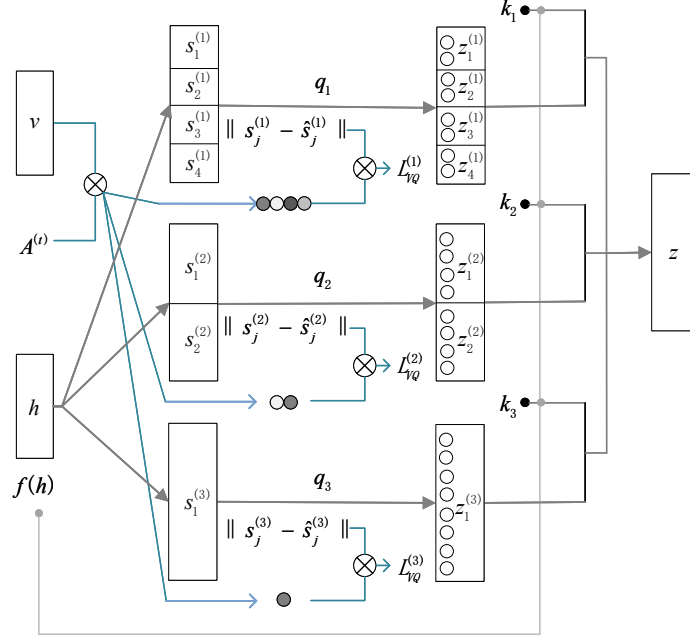


Fig. 1. Distortion-aware capacity allocation via adaptive discrete bottleneck.

Given a latent representation h and a region-aware distortion variable v , we seek a representation allocation policy $\mathcal{A}$ that jointly determines the discretization path and segment-level weighting. The policy minimizes the following Lagrangian:

$$\min_{\mathcal{A}} \mathbb{E}_{(h,v)} [D(h; \mathcal{A}, v) + \beta R(\mathcal{A})], \quad (1)$$

where $D(h; \mathcal{A}, v)$ serves as a task-aligned distortion surrogate that emphasizes errors in distortion-sensitive regions, and $R(\mathcal{A})$ denotes an effective capacity proxy induced by the discretization configuration. This objective is operational rather than strictly information-theoretic: D and R are used to guide practical representation allocation under structured appearance variation, not to estimate exact rate–distortion quantities.

3.2 Region-aware Distortion Variable

The region-aware variable v encodes how sensitive different spatial regions are to appearance distortion. Given an input image x , we extract K semantic region maps $\{r_k\}_{k=1}^K$, where $r_k \in [0, 1]^{H \times W}$ and $\sum_k r_k(p) = 1$ at each spatial location p . Typical regions include nuclei, stroma, and background.

Style conditions (e.g., text prompts or reference styles) are encoded into a style embedding u . A lightweight modulation network maps u to distortion sensitivity weights defined over K components:

$$\omega_k = g_{\text{mod},k}(u), \quad \omega = [\omega_1, \omega_2, \dots, \omega_K]^\top, \quad (2)$$

where a larger ω_k indicates higher distortion sensitivity for component k . The spatial distortion-sensitivity map is computed as

$$V(p) = \sum_{k=1}^K \omega_k r_k(p), \quad (3)$$

where $V(p)$ denotes pixel-level sensitivity. We then pool the region-specific responses to obtain a compact region-aware vector $v \in \mathbb{R}^K$, which is used by the bottleneck.

Importantly, v does not directly modify the latent representation h . Instead, it specifies region-level distortion sensitivity that modulates how representation capacity is allocated in the adaptive discrete bottleneck.

3.3 Adaptive Discrete Bottleneck

Given $h \in \mathbb{R}^m$, we define a set of T discretization paths $\mathcal{Q} = \{q^{(t)}\}_{t=1}^T$. Each path t defines a vector quantization operator parameterized by a segment number G_t and a codebook size L_t , which maps continuous features to discrete latent representations.

Specifically, path t partitions h into G_t segments, which serve as discrete representational units for capacity allocation:

$$h = [s_1^{(t)}, \dots, s_{G_t}^{(t)}], \quad s_j^{(t)} \in \mathbb{R}^{m/G_t}, \quad (4)$$

and quantizes each segment using a path-specific codebook of size L_t . The discretized representation is denoted as $q^{(t)}(h)$.

To select among discretization paths, we define a conditional path distribution $\pi(t | h)$. Specifically, h is projected via a lightweight function $f(\cdot)$. Each path t is associated with a learnable prototype vector k_t . The unnormalized score for path t is $a_t = k_t^\top f(h)$, and the conditional distribution is obtained by softmax normalization:

$$\pi(t | h) = \frac{\exp(a_t)}{\sum_{j=1}^T \exp(a_j)}. \quad (5)$$

The effective representation rate is approximated as

$$R(\mathcal{A}) = \sum_{t=1}^T \pi(t | h) R_t = \sum_{t=1}^T \pi(t | h) G_t \log L_t, \quad (6)$$

reflecting the relative granularity of discretization paths rather than an exact information-theoretic rate.

3.4 Distortion-modulated Vector Quantization

Vector quantization loss is computed over representation segments. To align region-level distortion sensitivity with latent segments, we project v to segment space using a lightweight learnable linear mapping $A^{(t)} \in \mathbb{R}^{G_t \times K}$ followed by a non-negative activation:

$$v^{(t)} = A^{(t)}v, \quad v_j^{(t)} \geq 0. \quad (7)$$

For a given path t , let $E^{(t)} = \{e_l^{(t)}\}_{l=1}^{L_t}$ denote the path-specific codebook. Each segment $s_j^{(t)}$ is assigned to its nearest codeword: $o_j^{(t)} = \arg \min_{l \in \{1, \dots, L_t\}} \|s_j^{(t)} - e_l^{(t)}\|_2^2$. The discrete representation is $z_j^{(t)} = e_{o_j^{(t)}}^{(t)}$, and the reconstructed segment is $\hat{s}_j^{(t)} = \text{sg}(z_j^{(t)})$, where $\text{sg}(\cdot)$ denotes the stop-gradient operator.

We define a segment-conditioned distortion surrogate for path t as

$$d_j^{(t)}(s_j^{(t)}, \hat{s}_j^{(t)} | v) = \|\sqrt{1 + \rho v_j^{(t)}} (s_j^{(t)} - \hat{s}_j^{(t)})\|_2^2, \quad (8)$$

which assigns higher cost to reconstruction errors in distortion-sensitive segments. The coefficient ρ controls the strength of distortion sensitivity in the quantization loss. The distortion is then given by

$$D(h; \mathcal{A}, v) = \sum_{t=1}^T \pi(t | h) \sum_{j=1}^{G_t} d_j^{(t)}(s_j^{(t)}, \hat{s}_j^{(t)} | v). \quad (9)$$

The distortion-modulated VQ loss for path t is derived as

$$\begin{aligned} L_{\text{VQ}}^{(t)}(h, v) = \sum_{j=1}^{G_t} & \left(\|\sqrt{1 + \rho v_j^{(t)}} (s_j^{(t)} - \text{sg}(e_{o_j^{(t)}}^{(t)}))\|_2^2 \right. \\ & \left. + \|\text{sg}(s_j^{(t)}) - e_{o_j^{(t)}}^{(t)}\|_2^2 \right). \end{aligned} \quad (10)$$

During training, we obtain a soft discrete representation by aggregating path-specific outputs under the routing distribution:

$$z = \sum_{t=1}^T \pi(t | h) q^{(t)}(h). \quad (11)$$

At inference time, the allocation becomes deterministic by selecting $\hat{t} = \arg \max_t \pi(t | h)$ and setting $z = q^{(\hat{t})}(h)$.

3.5 Training Objective

The overall learning objective follows an operational rate–distortion-inspired Lagrangian:

$$\mathcal{L} = \mathbb{E}_{(h,v)} \left[\mathcal{L}_{\text{task}}(z) + \alpha \sum_{t=1}^T \pi(t | h) L_{\text{VQ}}^{(t)}(h, v) + \beta \sum_{t=1}^T \pi(t | h) R_t \right]. \quad (12)$$

Here, $\mathcal{L}_{\text{task}}$ denotes the downstream task loss applied to the aggregated discrete representation z , $R_t = G_t \log L_t$ penalizes the effective representation rate of path t , and α and β control the trade-off between distortion-modulated quantization and representation rate. This formulation follows Eq. (1): the first two terms instantiate the task-aligned distortion surrogate, and the last term implements the capacity proxy.

4 Experiments

4.1 Experimental Settings

We conduct experiments to evaluate how DACAS allocates representation capacity under heterogeneous and spatially uneven appearance distortions.

Datasets We evaluate the proposed method on publicly available colorectal histopathology datasets commonly used for studying appearance and staining variability. CRC-DX consists of H&E-stained colorectal whole-slide images with MSI/MSS labels. Stanford-CRC is used as an external training source in the out-of-distribution setting. PanNuke provides nucleus-level annotations across multiple tissue types; in this work, it is used only to define anatomical region categories for region-aware perturbation and sensitivity estimation.

Implementation Details All models use the same preprocessing, backbone, and optimization protocol. Images are scaled to $[0, 1]$, normalized by dataset statistics, and evaluated without augmentation. The encoder uses stride-2 convolutions, residual blocks, and a 1×1 projection to a 64-dimensional latent space. DACAS uses three VQ paths with $\{G_t\} = \{4, 8, 16\}$ and $\{L_t\} = \{32, 64, 128\}$. CLAP text embeddings are mapped to the distortion weights ω by a small MLP. Unless stated otherwise, $\rho = 1.0$, $\alpha = 1.0$, $\beta = 0.25$, the Adam learning rate is 1×10^{-4} , and the batch size is 64.

Augmentation Protocols In this work, augmentation is not treated as a contribution in itself, but as a mechanism for inducing diverse distortion patterns to evaluate distortion-aware capacity allocation. We adopt a style-based appearance perturbation scheme that modifies color, texture, and staining patterns while preserving tissue structure. Perturbations are guided by text descriptions of appearance attributes, such as staining tone, contrast, saturation, and texture granularity, and are applied in a region-aware manner to induce spatially heterogeneous distortion. The prompts describe appearance only and do not include tissue types or diagnostic labels.

For the main experiments, each source image is augmented with a fixed set of text prompts and region configurations, so the number of augmented samples is controlled across methods. We do not tune augmentation strength with

scalar parameters; instead, text-guided style variation and region-aware application jointly produce the distortion patterns used to test whether representation capacity is adaptively allocated across samples and regions. In addition to stain-based baselines (SA, MSIfromHE, MSINet) and style augmentation baselines (STRAP, CLAP), we include the bottleneck-based ADCB [5] with the same backbone. All methods use identical source images, prompt sets, region definitions, and augmentation scale, ensuring that performance differences reflect representation allocation rather than augmentation design. Additional dataset statistics and baseline descriptions are provided in the Supplementary Material.

4.2 MSI Classification Performance

Based on the experimental protocol described above, we evaluate DACAS on MSI/MSS classification under both in-distribution and out-of-distribution settings. In the in-distribution setting, models are trained on CRC-DX-TRAIN and evaluated on CRC-DX-TEST. In the OOD setting, models are trained on Stanford-CRC and tested on CRC-DX-TEST to assess robustness to cross-center staining and acquisition shifts. Bootstrapping with the percentile method and 3,000 resamples is used to estimate 95% confidence intervals (CI). The results are summarized in Table 1.

As shown in Table 1, augmentation-only baselines improve generalization compared to standard training, while fixed-capacity bottlenecks provide limited additional benefit. The ADCB baseline further improves performance by adaptively selecting representation granularity across samples. Our full method (DACAS) consistently achieves the highest AUROC (Area Under the Receiver Operating Characteristic Curve), indicating that incorporating region-aware distortion sensitivity enables more effective allocation of representation capacity under heterogeneous appearance variation.

Table 1. Averaged in-distribution generalization performance in genetic subtype classification.

Model	AUROC	Confidence Interval
SA[9]	0.815	[0.709, 0.915]
MSIfromHE[4]	0.758	[0.633, 0.871]
MSINet[14]	0.814	[0.716, 0.917]
STRAP[13]	0.847	[0.738, 0.931]
CLAP[8]	0.871	[0.721, 0.923]
ADCB + backbone[5]	0.880	[0.760, 0.925]
DACAS (Ours)	0.905	[0.813, 0.942]

We further evaluate generalization under an out-of-distribution setting, where models are trained on the single-domain Stanford-CRC dataset and tested on the multi-source CRC-DX-TEST dataset. To mitigate selection bias due to

the limited size of Stanford-CRC, we employ four-fold cross-validation. The averaged AUROC, standard deviation, and paired t -test results are reported in Table 2, with DACAS serving as the reference method.

As shown in Table 2, adaptive bottleneck models consistently outperform augmentation-only baselines under domain shift. Notably, DACAS achieves the highest AUROC and statistically significant improvements over all competing methods. These results suggest that treating style-induced variability as distortion and explicitly allocating representation capacity—rather than relying on augmentation alone—leads to more robust and generalizable representations for structure-sensitive tasks such as MSI classification.

Table 2. Averaged out-of-distribution generalization performance in genetic subtype classification.

Model	AUROC	Std.	P-value
SA[9]	0.814	0.021	0.001*
MSIfromHE[4]	0.741	0.012	0.001*
MSINet[14]	0.785	0.021	0.007*
STRAP[13]	0.876	0.015	0.009*
CLAP[8]	0.893	0.017	0.009*
ADCB + backbone[5]	0.905	0.016	0.010*
DACAS (Ours)	0.939	0.024	REF

4.3 Robustness under Increasing Distortion Complexity

To test robustness under increasing distortion complexity, we vary the number of sampled text prompts $N_p \in \{1, 2, 3, 4\}$ while keeping the region configuration and the total number of augmented samples fixed. The same augmented sets are reused for all methods.

Figure 2 shows that performance decreases as N_p increases, but DACAS degrades slowest and remains more stable than the fixed bottleneck and DACAS without v . This suggests that adaptive granularity and spatial capacity allocation improve robustness under complex distortions.

4.4 Ablation Study

We conduct ablations with the same backbone, classifier, and training protocol, comparing a single-path fixed VQ bottleneck, DACAS without v , and full DACAS.

Table 3 shows that adaptive routing improves over a fixed bottleneck, while adding region-aware distortion sensitivity provides a further gain. This indicates that DACAS benefits from distortion-aware allocation rather than merely from increased capacity or stronger augmentation.

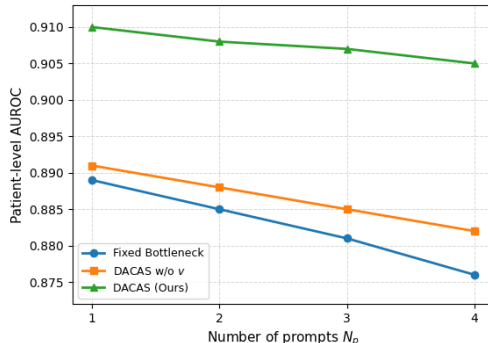


Fig. 2. Performance under increasing distortion pattern complexity. The complexity level is controlled by the number of sampled prompts N_p , while the region configuration and total augmentation protocol are kept fixed. DACAS exhibits a slower degradation rate as distortion patterns become more complex.

Table 3. Ablation study on distortion-aware capacity allocation for MSI classification.

Model Variant	Adaptive Region-aware Routing	v	AUROC
Fixed Bottleneck	✗	✗	0.876
DACAS without v	✓	✗	0.882
DACAS (Ours)	✓	✓	0.905

5 Conclusion

We presented DACAS, a distortion-aware capacity allocation framework for style-augmented domain generalization in computational pathology. Instead of further tuning augmentation policies, DACAS allocates representation capacity after augmentation through sample-level adaptive granularity and region-level distortion sensitivity. Experiments on colorectal MSI classification show that this allocation improves both in-distribution and out-of-distribution robustness under heterogeneous staining variations. The current study focuses on colorectal cancer, but the same capacity-allocation principle can be extended to other structure-sensitive medical imaging tasks.

References

1. Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V., Le, Q.V.: Autoaugment: Learning augmentation policies from data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
2. Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V.: Randaugment: Practical automated data augmentation with a reduced search space. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2020)

3. Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F.A., Brendel, W.: Imagenet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. In: International Conference on Learning Representations (ICLR), Oral (2019)
4. Kather, J.N., Pearson, A.T., Halama, N., Jäger, D., Krause, J., Loosen, S.H., Marx, A., Boor, P., Tacke, F., Neumann, U.P., Grabsch, H.I., et al.: Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature Medicine* **25**(7), 1054–1056 (2019)
5. Liu, D., Lamb, A., Xu, J., Tikeng Notsawo, P., Mozer, M., Bengio, Y., Kawaguchi, K.: Adaptive discrete communication bottlenecks with dynamic vector quantization for heterogeneous representational coarseness. *Proceedings of the AAAI Conference on Artificial Intelligence* **37**(7), 8825–8833 (2023)
6. Santhirasekaram, A., Kori, A., Winkler, M., Rockall, A., Glocker, B.: Vector quantisation for robust segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI) (2022)
7. Shi, Y., Yang, X., Wan, Y., Shen, X.: Semanticstylegan: Learning compositional generative priors for controllable image synthesis and editing. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR). pp. 11254–11264 (2022)
8. Shichang, S., Zhihao, Y., Hongfei, L., Jiana, M., Jincheng, Z., Rui, W.: Efficient style transfer for computational pathology with cross-modality local manipulation. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 1680–1683 (2024)
9. Tellez, D., Litjens, G., Bándi, P., Bulten, W., Bokhorst, J.M., Ciompi, F., van der Laak, J.: Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Medical Image Analysis* **58**, 101544 (2019)
10. Tishby, N., Pereira, F.C., Bialek, W.: The information bottleneck method. *Proceedings of the 37th Annual Allerton Conference* (2000)
11. Van Den Oord, A., Vinyals, O., et al.: Neural discrete representation learning. *Advances in neural information processing systems* **30** (2017)
12. Wang, Y., Chen, H., Zhang, J., Rueckert, D.: Adaptive style perturbation for robust medical image generalization. *Medical Image Analysis* **92**, 103102 (2025)
13. Yamashita, R., Long, J., Banda, S., Shen, J., Rubin, D.L.: Learning domain-agnostic visual representation for computational pathology using medically-irrelevant style transfer augmentation. *IEEE Transactions on Medical Imaging* **40**(12), 3945–3954 (2021)
14. Yamashita, R., Long, J., Longacre, T., Peng, L., Berry, G., Martin, B., Higgins, J., Rubin, D.L., Shen, J.: Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *The Lancet Oncology* **22**(1), 132–141 (2021)
15. Yoon, J.S., Oh, K., Shin, Y., Mazurowski, M.A., Suk, H.I.: Domain generalization for medical image analysis: A review. *Proceedings of the IEEE* (2024)
16. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond empirical risk minimization. In: International Conference on Learning Representations (ICLR) (2018)
17. Zhang, Z., Yang, S., Dang, Q., Wu, Z., Gu, L.: Split-and-combine: Enhancing style augmentation for single domain generalization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 15616–15625 (2025)