ADAPTIVE CURRICULUM LEARNING FOR MULTIMODAL DIFFUSION LANGUAGE MODELS IN BIOLOGY

Anonymous authors
Paper under double-blind review

ABSTRACT

Biological modeling is inherently multimodal, spanning perturbations, cellular phenotypes, biomedical images, spatial structures, and textual descriptions. Multimodal diffusion language models provide a flexible, noncausal framework in these settings. However, multimodal diffusion introduces a training challenge that is largely absent from the usual unimodal view of diffusion: different prediction directions and modality-specific corruption states define distinct training problems. We introduce MaDAC, an adaptive curriculum learning method that allocates training across these multimodal prediction problems. MaDAC uses an outer curriculum to adapt prediction routes and modality-specific corruption states to the learner state, and an inner curriculum to reweight targets, while a reference distribution preserves coverage. Our analysis shows that standard corruption schedules can severely under-sample training states with asymmetric corruption across modalities, and that adaptive allocation can help when their relative utility changes during learning. We evaluate MaDAC on multimodal biological tasks spanning perturbation-phenotype modeling, biomedical imaging, and spatial segmentation. Across settings, MaDAC significantly improves over standard corruption and fixed-allocation curricula, demonstrating the value of adaptive curriculum allocation for multimodal diffusion language models for biological tasks.

1 INTRODUCTION

Biological systems are inherently multimodal. A perturbation can involve chemical structures or genetic interventions, while its effects can be observed through cellular morphology, molecular measurements, or tissue imaging, and further characterized through spatial annotations and textual descriptions. Many real-world biological and drug discovery problems therefore require joint reasoning across these representations rather than modeling each in isolation. For example, Cell Painting data (Chandrasekaran et al., 2023) pair chemical and genetic perturbations with rich cellular morphology profiles, supporting the link of perturbations with desired phenotypic effects (Wu et al., 2026). Biomedical analysis similarly combines visual observations with semantic descriptions and spatially resolved information(Zhao et al., 2025). This multimodal nature is increasingly reflected in biological foundation models, which integrate structural, functional, and textual representations, amongst many others (Liu et al., 2024a; Hayes et al., 2025; Luo et al., 2023).

Autoregressive language models provide a powerful and scalable foundation for generative modeling (OpenAI et al., 2024; Touvron et al., 2023; Nguyen et al., 2023; Brixi et al., 2025; Xiao et al., 2025; Su et al., 2025), but their probability factorization imposes a fixed causal generation order. While this is natural for text generation, many biological generation problems have no predefined sequential ordering. Information may flow in either direction between a perturbation and the measured phenotype, arbitrary regions of a biomedical observation may need to be reconstructed, and multiple partially observed modalities may need to be inferred jointly. By contrast, diffusion language models (DLMs) (Nie et al., 2025; Sahoo et al., 2024; Yang et al., 2025) offer a new and particularly attractive alternative for such settings. DLMs retain the Transformer parameterization and token-level prediction objective familiar from language modeling, while replacing causal next-token prediction with iterative denoising of arbitrary masked subsets. One model can naturally represent marginal, conditional, infilling, and joint generation. Recent multimodal DLMs (Xie et al., 2025; Swerdlow et al., 2025; Yang et al., 2025; AI et al., 2026) further demonstrate that this formulation can unify understanding and generation across heterogeneous modalities such as image and text. These properties make diffusion language modeling an appealing foundation for multimodal biological modeling, preserving the scalable training machinery of LMs, while providing a flexible, noncausal interface to model heterogeneous observed and predicted modalities.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

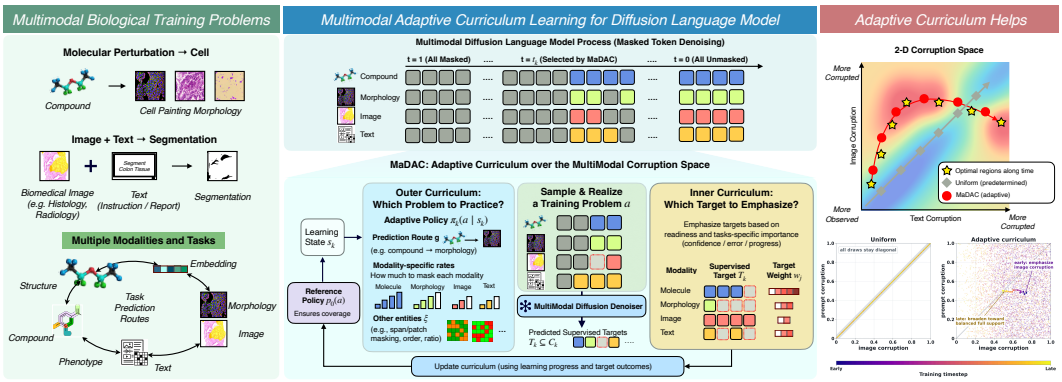

Figure 1: **Overview of MaDAC.** Biological problems couple heterogeneous modalities and tasks. Multimodal diffusion language model learns by reconstructing masked content across modalities. MaDAC treats this corruption process as an adaptive curriculum, allocating training across prediction routes, modality-specific corruption states, and supervised targets. In a multimodal corruption space, the relative utility of training regions evolves as learning progresses. Predetermined schedules follow fixed allocation, while MaDAC adapts training distribution in response to the learner state.

This flexibility, however, introduces a training problem for biology that is largely hidden in conventional diffusion. Each corruption realizes a different prediction problem, determining which modalities needs attention, how much information is removed from each stream, and which targets receive more supervision. The corruption process is thus more than a noise mechanism. It implicitly allocates the finite training budget across a family of denoising problems. In a single modality, a scalar corruption rate already traverses the one-dimensional diffusion noise continuum, making uniform rate sampling a principled and widely-used reference. Multimodality changes this geometry fundamentally. Relevant biological modalities can appear in different observed-target configurations and interact asymmetrically. A conventional uniform policy, explores only the "diagonal" of this higher-dimensional space, while off-diagonal states are necessary. For instance, predicting a heavily corrupted phenotype from a largely visible perturbation representation defines a very different problem from the reverse. Moreover, modality usefulness can change over the course of training.

We therefore introduce **MaDAC**, **M**ultimodal **D**LM **A**daptive **C**urriculum learning that explicitly allocates training across corruption-defined problems, as illustrated in Figure 1. An outer curriculum adapts the distribution over prediction routes and modality-specific corruption states according to the learner state, while an inner curriculum redistributes supervision across targets within the selected problem using model readiness. A reference component maintains persistent coverage of designated regions of the corruption space. MaDAC leaves the underlying denoising objective unchanged, i.e., the model remains trained through supervised denoising, while the curriculum controls where its learning is spent. Our theoretical analysis explains both why multimodal diffusion benefits from expanding beyond standard corruption and when adaptive allocation can provide additional advantage.

Our main contributions are threefold. First, we formulate multimodal diffusion training as a curriculum over corruption-defined training problems, and theoretically show that standard shared-rate corruption can severely underexpose asymmetric multimodal problems, while an adaptive allocation strategy can be beneficial when the relative utility of corruption states changes during learning. Second, we introduce **MaDAC**, a hierarchical adaptive curriculum that allocates training across prediction routes, modality-specific corruption states, and supervised targets, with explicit coverage preservation. Third, we evaluate MaDAC on biologically distinct multimodal settings, including chemical perturbation with paired chemical perturbations with cellular morphology, compound structures, and bioactivity; biomedical understanding with images, semantic prompts, and spatial segmentation across imaging modalities; and spatial tissue analysis of organ-conditioned spatial segmentation under source shift. Across these settings, MaDAC demonstrates significant improvements over standard and scheduled curricula on both pretrained and from-scratch backbones.

2 Related Work

Multimodal diffusion language models and biological modeling. Diffusion language modeling develops non-autoregressive generation through denoising, including embedding-space, score-based discrete, and masked diffusion (Austin et al., 2021; Li et al., 2022; Lou et al., 2024; Sahoo et al., 2024). LLaDA and DiffuLLaMA (Nie et al., 2025; Gong et al., 2025) demonstrate its applicability to large-scale language modeling. In the multimodal setting, UniDiffuser (Bao et al., 2023)

unifies conditional and joint tasks through modality-specific timesteps, Show-o (Xie et al., 2025) combines autoregressive text modeling with discrete image diffusion, while UniDisc, MMaDA, and LLaDA2.0-Uni (Swerdlow et al., 2025; Yang et al., 2025; AI et al., 2026) develop unified multimodal diffusion formulations for understanding and generation. Biological modeling provides closely related examples. EvoDiff and DPLM (Alamdari et al., 2023; Wang et al., 2024a) apply discrete diffusion to protein sequence generation. ESM3 extends iterative masked generative language modeling across protein sequence, structure, and function modalities (Hayes et al., 2025), and DPLM-2 (Wang et al., 2024b) models sequence and discretized structure with multimodal protein diffusion. Existing work primarily studies model architecture, representation, decoding, or post-training, while MaDAC addresses a different question, i.e., for multimodal diffusion, how should an effective learner distribute its optimization as training progresses?

Curriculum learning for multimodal models. Curriculum learning organizes training experiences by difficulty or usefulness, from predefined easy-to-hard ordering and self-paced selection to competence-based schedules (Bengio et al., 2009; Kumar et al., 2010; Platanios et al., 2019; Graves et al., 2017; Parashar et al., 2026), which improves effectiveness and efficiency for many training perspectives. Multimodal curricula extend these ideas to heterogeneous inputs and supervision, selecting samples according to instance difficulty and model competence (Qian et al., 2025; Saha et al., 2024; Liu et al., 2023; Zhou et al., 2023). However, curricula have been underexplored in diffusion-based multimodal models. Our work fills this gap, framing the multimodal corruption-defined problem and building an adaptive curriculum.

Adaptive optimization of diffusion noising. Diffusion training has long optimized the distribution and weighting of noise levels (Nichol & Dhariwal, 2021; Kingma et al., 2023; Choi et al., 2022; Hang et al., 2024). Related developments extend directly to language and discrete data. Block Diffusion (Arriola et al., 2025) adapts its mask-rate sampling range to reduce training variance, generalized masked diffusion (Shi et al., 2025) learns token-identity-dependent noise schedules, Dream (Ye et al., 2025) reweights masked targets through its noise-rescheduling, and DLM4G (Shahane et al., 2026) uses per-token denoising errors for noising. Our method considers a different perspective over modalities. Moreover, unlike objective-preserving importance sampling, MaDAC reallocates the effective training objective rather than correcting every action by a reference-policy density ratio.

3 Multimodal Adaptive Curriculum for Diffusion-LMs

Masked diffusion training is usually described as choosing a corruption level and reconstructing hidden content. In multimodal models, however, corruption determines which representations are observed and inferred, and where supervision is applied. The same biological record can give rise to different corruption-defined training problems. We introduce MaDAC, which turns the standard corruption distribution into a learner-dependent training policy. The central design is to couple which problem is practiced with how that problem contributes to learning, rather than prescribe a scalar noise schedule or a target-weighting rule independently of the multimodal prediction problem.

3.1 Multimodal diffusion as corruption-defined training problems

Let $x = (x^{(1)}, \ldots, x^{(D)})$ be a record containing D modalities. For illustration, we consider a biomedical image, a text description, and a molecule. We first choose a prediction route g, which specifies the observed, predicted, and hidden-but-unsupervised modalities (*e.g.,* the image is observed, the molecule is predicted, and the text is hidden without supervision). We then choose corruption rates $\mathbf{r} = (r_1, \ldots, r_D)$, where r_m is the expected fraction hidden in modality m. The route fixes the rates of fully observed and fully hidden streams to zero and one, respectively. The remaining rates are sampled (*e.g.,* $\mathbf{r} = (0, 1, 0.8)$). Together, the route and rates define a corruption configuration $a = (g, \mathbf{r})$. Finally, a fixed masking rule $K(C, T \mid x, a)$ samples the actual hidden token positions C (*e.g.,* all text tokens and 80% of molecule tokens) and supervised targets $T \subseteq C$ (*e.g.,* the hidden molecule tokens). The denoiser receives the corrupted sequence $\widetilde{x}$, obtained by masking positions C in x, and is trained to predict the original tokens at positions T. A training problem is thus defined by $z = (x, a)$ and the masking rule $K(C, T \mid x, a)$. Given a probability distribution $p(z)$ over the tuple $z = (x, a)$, we generate a training problem by first sampling z, and then generate the masking rule according to $K(C, T \mid x, a)$.

For a training problem $z = (x, a)$, the loss at a supervised token position $j \in T$ is $\ell_j(\theta) = -\log p_{\theta,j}(x_j \mid \widetilde{x})$, where $p_{\theta,j}(x_j \mid \widetilde{x})$ is the probability assigned to the original token given the corrupted input. Continuous targets use the corresponding regression loss. We average these losses

within groups T_h, indexed by $h \in \mathcal{H}$. A group typically contains the supervised tokens of one modality. To vary the contribution of individual targets, we assign positive weights w_{hj} and normalize them within each group $\omega_{hj} = \frac{w_{hj}}{\sum_{v \in T_h} w_{hv}}$. Setting all weights equally gives $\omega_{hj} = 1/|T_h|$. The total training loss combines the group averages:

$$\mathcal{L}(\theta; z, C, T, \omega) = \sum_{h \in \mathcal{H}} b_h \sum_{j \in T_h} \omega_{hj} \ell_j(\theta), \tag{1}$$

where the coefficient $b_h \geq 0$ specifies the overall scale of group h. Task-specific loss and normalization conventions are given in Appendix B.1.

3.2 Adaptive curriculum over the multimodal corruption space

At each training step k, the learner state s_k (containing model, optimizer, *etc*) guides two curriculum decisions. An *outer* distribution $q_k(z \mid s_k)$ selects the training problem, while *inner* weights ω_k determine the relative contributions of its supervised targets. Together, these choices define the step-specific training objective:

$$\mathcal{J}_k(\theta) = \mathbb{E}_{q_k, K}[\mathcal{L}(\theta; z, C, T, \omega_k)]. \tag{2}$$

This hierarchy controls both which prediction routes and corruption levels the model practices, and how supervision is distributed among targets within each selected problem.

3.2.1 Outer curriculum: adaptive problem allocation

The outer curriculum selects the training record x, prediction route g, and corruption rates $\mathbf{r}$. Given the learner state s_k, the adaptive policy π_{ϕ_k} parametrized by ϕ_k makes these choices sequentially:

$$\pi_{\phi_k}(z \mid s_k) = \pi_{\text{ex}}(x \mid s_k)\, \pi_{\text{route}}(g \mid x, s_k) \pi_{\text{rate}}(\mathbf{r} \mid g, x, s_k). \tag{3}$$

To preserve exposure to designated training problems, we mix the adaptive policy with a fixed reference distribution p_0:

$$q_k(z \mid s_k) = (1 - \rho)\pi_{\phi_k}(z \mid s_k) + \rho p_0(z), \qquad 0 < \rho \leq 1. \tag{4}$$

Any problem region with positive reference probability thus retains a positive sampling probability. For discrete choices such as prediction routes g, the controller adjusts sampling probabilities using recent learning progress. We periodically evaluate each candidate route on fixed probe data, keeping its corruption protocol and loss normalization unchanged across evaluations. We smooth the resulting losses to reduce measurement noise and define a progress score P_b for route g as the decrease in its smoothed loss since the previous probe. A positive score means that the model is currently improving on that route. At each controller refresh, routes with larger progress scores receive higher sampling probabilities. To encourage exploration, we also add a bonus $d/(1 + n_b)$, where n_b is the number of times route b has been selected and $d \geq 0$ controls the bonus strength.

For continuous choices such as modality-specific corruption rates, the controller similarly shifts probability toward values associated with larger recent improvement. For each active rate factor m, we use a route-specific bounded Beta distribution,

$$v_m \sim \text{Beta}(\alpha_{m,k}, \beta_{m,k}), \qquad r_m = r_m^- + (r_m^+ - r_m^-)v_m. \tag{5}$$

Let ϕ_k denote the parameters of these continuous policy factors. After sampling a training problem, we compute the improvement reward R_k specified for that controller and compare it with the pre-action baseline $\bar{R}_{k-1}$. Rate configurations with above-baseline improvement are reinforced, while those with below-baseline improvement are downweighted. Because the adaptive policy is mixed with the fixed reference distribution in Eq. 4, the score-function update is applied only when the sample came from the adaptive component:

$$\phi_{k+1} = \phi_k + \eta_\phi A_k \big(R_k - \bar{R}_{k-1}\big) \nabla_\phi \log \pi_{\phi_k}(z_k \mid s_k), \tag{6}$$

where $A_k = 1$ for a draw from the adaptive component and $A_k = 0$ otherwise. Thus, the continuous controller reallocates training toward choices that are currently producing greater improvement, while the reference component in Eq. 4 maintains coverage.

3.2.2 Inner curriculum: adaptive target allocation

Once the outer curriculum has selected a training problem, the inner curriculum determines which of its supervised targets should contribute most strongly to the update. The corrupted input and supervised target set are fixed. The inner curriculum only redistributes weight among those targets.

Algorithm 1 MaDAC training

Require: model θ_1, reference p_0, coverage mass ρ, state s_1
for $k = 1, \dots, K$ **do**
  form active factors of $\pi_{\phi_k}(z \mid s_k)$
  draw $A_k \sim \text{Bernoulli}(1-\rho)$; draw $z_k \sim \pi_{\phi_k}$ if $A_k = 1$, else $z_k \sim p_0$; draw $(C_k, T_k) \sim K(\cdot \mid z_k)$
  compute detached target weights
  update θ with $\mathcal{L}(\theta; x, a, C, T, \omega) = \sum_{h \in \mathcal{H}} b_h \sum_{j \in T_h} \omega_{hj} \ell_j(\theta)$
  if a controller update is due **then**
    measure the specified progress or reward signal
    refresh categorical factors $\pi_k(b \mid s_k) \propto \exp\Big(\frac{P_{b,t} + d/(1+n_{b,k})}{\tau_o}\Big)$
    update score-trained factors $\phi_{k+1} = \phi_k + \eta_\phi A_k (R_k - \bar{R}_{k-1}) \nabla_\phi \log \pi_{\phi_k}(z_k \mid s_k)$
  update cached statistics and state s_{k+1}

For a supervised target j, the model prediction provides two signals,

$$c_j = \max_v p_{\theta,j}(v \mid \widetilde{x}), \qquad p_j^\star = p_{\theta,j}(x_j \mid \widetilde{x}), \tag{7}$$

where c_j measures confidence in the model's predicted class and $p_j^\star$ is the probability assigned to the correct target. We define the target error $e_j = 1 - p_j^\star$. For each route, we use confidence / uncertainty as the allocation signal,

$$u_j = c_j \text{ [readiness allocation] or } 1 - c_j \text{ [uncertainty allocation].} \tag{8}$$

Thus, readiness allocation emphasizes targets the current model can already predict confidently, whereas uncertainty allocation emphasizes targets on which the model remains less certain. The choice of direction is fixed by the stated training/calibration protocol. We convert this score into a positive target priority $w_{k,hj} = (1 + \gamma u_j + \kappa e_j)\, \sigma\Big(\frac{u_j - b_0}{\tau_w}\Big)$, where $\gamma, \kappa \geq 0$, $\tau_w > 0$, and σ is the logistic function. The normalized weights $\omega_{k,hj} = \frac{w_{k,hj}}{\sum_{v \in T_h} w_{k,hv}}$ then redistribute supervision within group h without changing its base scale. Continuous outputs follow the same principle at the appropriate loss-unit level. We replace token confidence with a bounded agreement score, such as mapped cosine similarity, and normalize weights across relevant examples. For uninformative agreement signal we use unit weights. Inner-policy statistics are detached from the denoiser update.

The inner curriculum therefore changes how strongly the targets of a selected problem contribute, but not which targets are hidden or supervised, complementing the outer curriculum.

Remark 3.1 (Curriculum over importance sampling). MaDAC intentionally reallocates the reference objective. It does not apply outer density ratios or inverse target-weight corrections. A factor such as $1/r_m$ in the base loss is not the policy ratio needed to undo this change. The resulting step-specific supervised objective need not be an unchanged diffusion likelihood bound.

3.3 Theoretical Analysis

Here we analyze why multimodal adaptive curriculum can help. The analysis follows the method's design, where multimodal corruption selects prediction problems, and their usefulness can change as the model learns, motivating state-dependent allocation. We formalize and study these effects. Proofs and further derivations are in Appendix B.3.

Compared to standard diffusion, MaDAC training avoids the fixed corruption schedule, which is specifically meaningful to multimodal problems. Tying all modalities to the same corruption forces their masked fractions moving together, which causes insufficient learning for some modalities.

Proposition 3.2 (Shared rates underexpose directional multimodal problems)**.** *Consider a fixed route with two modalities of lengths $n_1, n_2 \geq 1$. Let C_m denote the hidden positions of modality m. For thresholds $0 < v < u < 1$, define the directional training region $E_{u,v} = \Big\{ \frac{|C_1|}{n_1} \geq u, \frac{|C_2|}{n_2} \leq v \Big\}$, i.e., modality 1 is substantially more corrupted than modality 2. Let $n_{\text{eff}} = (n_1^{-1} + n_2^{-1})^{-1}$. Under any curriculum that chooses a single shared corruption rate for both modalities and then masks tokens independently conditional on that rate, $\Pr_{\text{shared}}(E_{u,v}) \leq \exp[-2n_{\text{eff}}(u-v)^2]$. In contrast, when the two modality rates are sampled independently and uniformly on $(0,1)$, $\Pr_{\text{independent}}(E_{u,v}) \longrightarrow (1-u)v$ as both stream lengths diverge. Hence, for fixed separation $u - v$, shared-rate curriculum make entire family of asymmetric prediction problems exponentially rare in*

stream length, whereas modality-specific rates retain nonvanishing exposure. The bound also holds when the shared rate is selected adaptively from the training history before fresh Bernoulli masking.

Intuitively, a scalar rate makes modalities easy or hard together, while MaDAC's modality-specific corruption can instead practice "observe one modality, recover the other" problems better. For asymmetric training problems, which are common in real biological tasks, as asymmetry becomes strong, exposure become exponentially rare. The gradient-mixture restriction is stated separately in Prop. B.1. The appendix adds the finite-budget consequence under fixed-route assumptions.

Next we study the condition when adaptive selection improves over standard. Other than richer multimodal support, MaDAC adds adaptation. The route and corruption configuration that is most useful can change as cross-modal representations mature. To isolate this effect, we compare policies at a common distribution μ of learner states. Let $b \in \mathcal{B}$ denote a candidate MaDAC training problem. We define its one-step intervention utility as $Q(s,b) = J(\theta_s) - \mathbb{E}[J(\theta_s - \eta\widehat{d}_b) \mid s, b]$, where J is a common validation objective and $\widehat{d}_b$ is the supervised update generated by practicing problem b. Thus $Q(s,b)$ measures how useful that problem is at the current learner state. Let $V_\mu(q) = \mathbb{E}_{S\sim\mu, b\sim q(\cdot|S)} Q(S,b)$ and define $\Gamma_\mu = \mathbb{E}_\mu \max_b Q(S,b) - \max_b \mathbb{E}_\mu Q(S,b)$. The quantity Γ_μ is the oracle adaptivity gap. It is zero when one fixed problem is optimal at almost every state, and positive when the identity of the most useful problem changes over learning.

Theorem 3.3 (Adaptive Curriculum Advantage)**.** *Assume a finite candidate set $\mathcal{B}$ and finite displayed expectations. Suppose the curriculum estimates state-dependent utility by $\widehat{Q}$ with worst-case estimation error $\max_b |\widehat{Q}(s,b) - Q(s,b)| \le \varepsilon(s)$, and define its selection gap $\zeta(s) = \max_b \widehat{Q}(s,b) - \mathbb{E}_{b\sim\pi(\cdot|s)}\widehat{Q}(s,b)$, which measures how far the adaptive policy π is from choosing the highest-estimated-utility problem. For the coverage-anchored MaDAC policy $q = (1-\rho)\pi + \rho p_0$, let V^ρ_{stat} denote the best value achievable by any state-independent curriculum with the same coverage component, $(1-\rho)\bar{\pi} + \rho p_0$. Then*

$$V_\mu(q) - V^\rho_{\mathrm{stat}} \ge (1-\rho)\big[\Gamma_\mu - 2\mathbb{E}_\mu\varepsilon(S) - \mathbb{E}_\mu\zeta(S)\big]. \tag{9}$$

Consequently, whenever $\Gamma_\mu > 2\mathbb{E}_\mu\varepsilon(S) + \mathbb{E}_\mu\zeta(S)$, the state-dependent allocation has strictly larger expected utility than every fixed curriculum with the same coverage mass. For the softmax allocation $\pi(b \mid s) \propto \exp(\widehat{Q}(s,b)/\tau_o)$, $\zeta(s) \le \tau_o \log|\mathcal{B}|$.

Thus adaptive curriculum is useful when the preferred point in that space moves with the learner state. The gain available from tracking this moving optimum is Γ_μ, while noisy learning-progress estimates and imperfect selection consume that gain. Continuous-action and multi-step interpretations are in the appendix. We also study the design of the two levels and how they interact.

Proposition 3.4 (Coupling of problem and target allocation)**.** *Given a differentiable objective J, and common base scales. Let $d_\omega(z), d_0(z)$ denote conditional expected gradients under adaptive and uniform inner weights. Define $\mathcal{A}_0(z) = \langle \nabla J, d_0(z)\rangle$ and $\Delta\mathcal{A}(z) = \langle \nabla J, d_\omega(z) - d_0(z)\rangle$. For positive detached weights and finite expectations,*

$$\Delta\mathcal{A}(z) = \mathbb{E}_{K(\cdot|z)} \sum_h b_h \frac{\mathrm{Cov}_h(w,\chi)}{\bar{w}_h}, \tag{10}$$

where $\chi_{hj} = \langle \nabla J, \nabla \ell_j\rangle$, $\bar{w}_h = |T_h|^{-1}\sum_j w_{hj}$, and covariance uses the uniform target law within group h. The composed alignment gain is

$$\mathbb{E}_q[\mathcal{A}_0 + \Delta\mathcal{A}] - \mathbb{E}_{p_0}\mathcal{A}_0 = \underbrace{(\mathbb{E}_q - \mathbb{E}_{p_0})\mathcal{A}_0}_{\textit{problem allocation}} + \underbrace{\mathbb{E}_{p_0}\Delta\mathcal{A}}_{\textit{target allocation}} + \underbrace{(\mathbb{E}_q - \mathbb{E}_{p_0})\Delta\mathcal{A}}_{\textit{curriculum interaction}}. \tag{11}$$

The two curriculum levels reinforce one another when MaDAC visits problems for which its target reweighting also points the update more strongly toward validation improvement. This explains why the levels are coupled rather than automatically additive. Under Lipschitz gradients and finite second moments, the appendix gives a finite-step improvement bound for their composition. Finally, we discuss what remains protected while the curriculum concentrates.

Proposition 3.5 (Coverage and positive target allocation)**.** *Under equation 4, let A be a fixed measurable problem region with $p_0(A) > 0$ and $N_K(A) = \sum_{k=1}^K \mathbf{1}\{z_k \in A\}$. Then*

$$\Pr[N_K(A) = 0] \le (1 - \rho p_0(A))^K \le e^{-K\rho p_0(A)}. \tag{12}$$

Table 1: JUMP multimodal retrieval and prediction. Retrieval tasks report 100-way Hits@10 (H@10) for SMILES→morphology, structure→morphology, and morphology→compound retrieval; structure-based morphological bioactivity prediction and morphology-guided MoA virtual screening report standardized pAUROC@0.1 (pA). We report mean±SD.

Method	SMILES→Morphology H@10 (↑)	Structure→Morphology H@10 (↑)	Morphology→Compound H@10 (↑)	Morphological bioactivity pA (↑)	Morphology-guided MoA pA (↑)
GNEprop + kNN	–	.221 ± .003	.232 ± .005	.550 ± .002	.520 ± .010
Morgan + kNN	–	.285 ± .002	.300 ± .002	.557 ± .002	.507 ± .008
GNEprop + linear	.373 ± .006	.301 ± .010	.301 ± .018	.612 ± .001	.567 ± .019
Morgan + linear	.346 ± .006	.316 ± .002	.340 ± .002	.614 ± .001	.559 ± .009
Ridge	.367 ± .006	.367 ± .006	.395 ± .006	–	–
PhenoCompass	–	.549 ± .012	.556 ± .012	.677 ± .003	.739 ± .015
Frozen (LLaDA2)	.254 ± .005	.608 ± .006	.631 ± .006	.562 ± .005	.548 ± .015
Uniform (LLaDA2)	.535 ± .005	.596 ± .005	.637 ± .006	.676 ± .007	.746 ± .024
MaDAC (LLaDA2)	.529 ± .006	.615 ± .006	.636 ± .006	.676 ± .007	.747 ± .024
Uniform (UniDisc)	.486 ± .006	.606 ± .006	.632 ± .006	.683 ± .007	.755 ± .023
MaDAC (UniDisc)	**.564 ± .006**	**.623 ± .006**	**.642 ± .006**	**.684 ± .007**	**.757 ± .024**

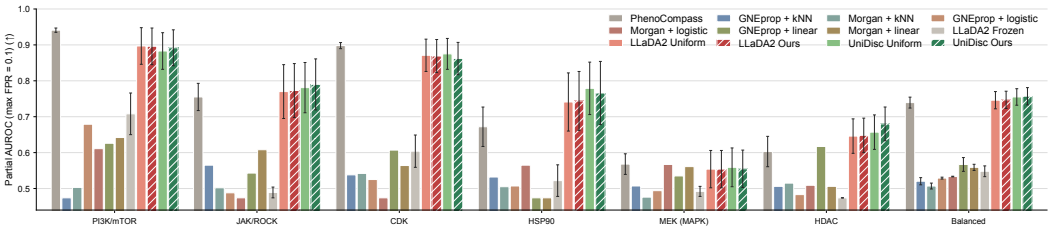

Figure 2: JUMP morphology-guided MoA virtual screening. Bars show pAUROC@0.1 for each mechanism and the balanced mean across six MoAs. Further results and comparisons are in Appendix 6.

No independence between successive adaptive decisions is required. Within a group, $0 < w_- \leq w_{hj} \leq w_+$ *also implies* $\omega_{hj} \geq \frac{w_-}{|T_h| w_+}$.

The reference mixture preserves exposure and positive weights preserve target contributions. Together, the results support the design of MaDAC. Multimodality creates a richer corruption action space, changing utility across learning states creates a reason to allocate that space adaptively, and explicit base-policy mixing ensures that adaptation does not silently remove required capabilities.

4 Experiments

We evaluate MaDAC on three biologically distinct multimodal settings: **JUMP** for molecular perturbation-phenotype modeling, including prediction, retrieval, and bioactivity; **BioMed** for biomedical image segmentation conditioned on semantic prompts, and **HuBMAP** for organ-conditioned spatial tissue segmentation under source shift. Together, the benchmarks span molecule↔phenotype, image+text→mask, and image+organ-prompt→segmentation settings, with both discrete and continuous targets and multiple directions of biological understanding. **Backbone models.** We instantiate MaDAC with two unified multimodal DLM backbones. LLaDA2.0-Uni (AI et al., 2026) is a 16B pretrained model with semantic visual tokenizer, mixture-of-experts diffusion-language backbone over text and visual tokens. UniDisc (Swerdlow et al., 2025) formulates text-image understanding and generation as joint discrete denoising, and we train its 147M transformer backbone from scratch on each benchmark. Task-specific protocols are given in Appendix C. Additional results and analyses are in Appendix D.

4.1 JUMP: Multimodal Perturbation-Phenotype Modeling

Data and tasks. We use JUMP Cell Painting data (Chandrasekaran et al., 2023) following the tasks and metrics of PhenoCompass (Wu et al., 2026). Each chemical perturbation links molecular structure, SMILES (Weininger, 1988), and corresponding embeddings, with a morphology profile, and bioactivity labels. We evaluate SMILES→morphology retrieval, structure→morphology retrieval, morphology→compound retrieval, structure-based morphological bioactivity prediction, and morphology-guided MoA virtual screening. Details are in Appx. C.1.1.

Setting and baselines. Our primary comparisons follow PhenoCompass (Wu et al., 2026) with MaDAC, fixed-policy controls, and PhenoCompass operating in the same aligned target space. Baselines comprise PhenoCompass, GNEprop embeddings (Scalia et al., 2026)/Morgan fingerprint (Rogers & Hahn, 2010) kNN and linear predictors, ridge regression, and matched frozen or uniform-curriculum controls.

Results. Table 1 compares the three retrieval tasks, structure-based morphological bioactivity prediction, and morphology-guided MoA virtual screening. UniDisc-based MaDAC training achieves the best results on all five endpoints and improves over UniDisc-based uniform training as well as

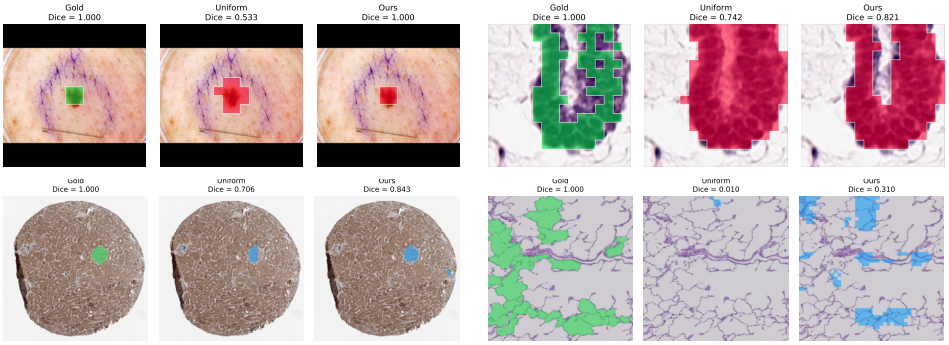

Figure 3: Visualization examples. The top row shows BioMed IID ISIC (Codella et al., 2019) and PanNuke (Gamper et al., 2019); the bottom row shows HuBMAP kidney and lung cases. Each triplet shows Gold, LLaDA2-based uniform, and LLaDA2-based MaDAC results. Green denotes ground-truth masks, while red (BioMed) and blue (HuBMAP) denote predictions. More visualization results are in Appendix D.9.

Table 2: BioMed results on IID and source-held-out tests. We report Dice and AUROC metrics. Further results including broader BiomedParse FM are reported in Appx. D.

Method	IID		Source-held-out	
	Dice (↑)	AUROC (↑)	Dice (↑)	AUROC (↑)
BiomedParse	0.724 ± 0.117	0.951 ± 0.001	0.353 ± 0.073	0.825 ± 0.009
SAM ViT-B	0.521 ± 0.002	0.806 ± 0.002	0.178 ± 0.002	0.631 ± 0.004
MedSAM ViT-B	0.664 ± 0.002	0.903 ± 0.001	0.313 ± 0.004	0.799 ± 0.004
Frozen LLaDA2	0.282 ± 0.001	0.512 ± 0.000	0.104 ± 0.002	0.513 ± 0.002
Uniform (LLaDA2)	0.775 ± 0.006	0.959 ± 0.002	0.469 ± 0.019	0.926 ± 0.003
Easy-to-Hard (LLaDA2)	0.799 ± 0.008	0.968 ± 0.003	0.471 ± 0.051	0.921 ± 0.018
MaDAC (LLaDA2)	$\mathbf{0.860 \pm 0.002}$	$\mathbf{0.985 \pm 0.001}$	$\mathbf{0.616 \pm 0.025}$	$\mathbf{0.968 \pm 0.003}$
Uniform (UniDisc)	0.353 ± 0.004	0.652 ± 0.021	0.198 ± 0.043	0.745 ± 0.029
MaDAC (UniDisc)	0.371 ± 0.021	0.662 ± 0.037	0.210 ± 0.020	0.763 ± 0.112

Table 3: HuBMAP segmentation results. Dice is higher-is-better and empty-image rate is lower-is-better. We report performance on held-out test.

Method	Image Dice(↑)	Macro Dice(↑)	Empty (%; ↓)
SegFormer MiT-B4	0.415 ± 0.050	0.418 ± 0.037	21.540 ± 1.420
DeepLabV3 ResNet-50	0.151 ± 0.007	0.152 ± 0.008	24.550 ± 1.310
U-Net ResNet-34	0.235 ± 0.007	0.251 ± 0.008	15.400 ± 1.070
Frozen (LLaDA2)	0.396 ± 0.003	0.390 ± 0.000	20.310 ± 6.940
Uniform (LLaDA2)	0.387 ± 0.004	0.382 ± 0.012	8.370 ± 3.000
MaDAC (LLaDA2)	$\mathbf{0.431 \pm 0.039}$	$\mathbf{0.418 \pm 0.042}$	$\mathbf{0.450 \pm 0.630}$
Frozen (UniDisc)	0.173 ± 0.083	0.213 ± 0.077	54.350 ± 15.310
Uniform (UniDisc)	0.337 ± 0.005	0.368 ± 0.006	14.060 ± 6.000
MaDAC (UniDisc)	0.345 ± 0.002	0.375 ± 0.003	11.500 ± 1.100

PhenoCompass on the three retrieval directions. LLaDA2-based MaDAC training improves retrieval over its uniform counterpart and retains morphological bioactivity and morphology-guided MoA performance. Figure 2 reports morphology-guided MoA scores for the six mechanisms. UniDisc-based MaDAC training leads on HSP90, JAK/ROCK, HDAC, and overall performance, while PhenoCompass leads on PI3K/mTOR, and CDK. Appendix D.1 reports further results with morphology/structure co-embedding space and raw-feature space, and Appendix D.5 reports policy ablations.

4.2 BioMed: multimodal biomedical segmentation

Dataset and task. We construct BioMed from the public BiomedParse Data collection (Zhao et al., 2025), a multi-source collection spanning pathology, radiography, dermoscopy, endoscopy, fundus imaging, mammography, and ultrasound. Each example pairs biomedical image and harmonized target prompt with a binary segmentation mask. The primary task is to predict the mask from the image and prompt. Source composition, split sizes, and mask construction are in Appendix C.1.2.

Setting and baselines. We compare MaDAC with matched frozen, uniform, and easy-to-hard curriculum (Bengio et al., 2009) controls and with BiomedParse, SAM (Kirillov et al., 2023), and MedSAM (Ma et al., 2024) specialists trained on the same BioMed data. Grounding DINO (Liu et al., 2024b) converts the shared image-prompt input into the box consumed by SAM and MedSAM.

Results. Table 2 reports aggregate IID and source-held-out segmentation performance. LLaDA2-based MaDAC training improves over LLaDA2-based uniform training while also achieving the highest performance under the same data. UniDisc-based MaDAC training improves over its uniform counterpart on all four metrics. Table 4 resolves performance by biomedical modality, and LLaDA2-based MaDAC training is strongest among the matched training variants in all 11 reported columns, indicating that the improvement extends across the trained modalities. Figure 3 visualizes representative results, showing that LLaDA2-based MaDAC predictions are close to gold.

4.3 HuBMAP: organ-conditioned spatial segmentation

Dataset and task. We use the labeled post-competition release of the 2022 HuBMAP+HPA Hacking the Human Body benchmark (Howard et al., 2022). It pairs histology images from five organs with functional tissue unit (FTU) masks. The task is image+organ-prompt→mask prediction and evaluate on donor-disjoint HPA images and HuBMAP images. Task details are in Appendix C.1.3.

Setting and baselines. The fixed image+organ-prompt→mask route uses capped loss-aware sampling (Nichol & Dhariwal, 2021) at the outer level and readiness weighting over example-level losses at the inner level. We compare against matched frozen and uniform curriculums and with SegFormer MiT-B4 (Xie et al., 2021), DeepLabV3/ResNet50 (Chen et al., 2017; He et al., 2016),

Table 4: BioMed AUROC by biomedical modalities. IID and source-held-out results are shown for each method; the held-out split contains only endoscopy, fundus, pathology, and ultrasound. Comparisons within each family use a common evaluation population. The LLaDA2-based Frozen, Uniform, Easy-to-Hard, and MaDAC variants form a matched curriculum comparison; bold marks the best-performing among them.

	IID test							Source-held-out test			
Method	Derm.	Endo.	Fundus	Mammo.	Path.	US	X-Ray	Endo.	Fundus	Path.	US
BiomedParse	.883 ± .007	.728 ± .013	.873 ± .007	.870 ± .014	.912 ± .003	.789 ± .011	.962 ± .001	.733 ± .015	.965 ± .003	.739 ± .018	.850 ± .010
SAM	.900 ± .009	.728 ± .008	.630 ± .007	.702 ± .009	.858 ± .002	.702 ± .015	.833 ± .002	.804 ± .005	.874 ± .002	.823 ± .009	.665 ± .006
MedSAM	.905 ± .009	.819 ± .006	.896 ± .003	.888 ± .009	.885 ± .002	.865 ± .008	.939 ± .001	.836 ± .005	.966 ± .001	.850 ± .011	.745 ± .007
Frozen (LLaDA2)	.481 ± .014	.525 ± .021	.507 ± .008	.505 ± .007	.499 ± .003	.461 ± .025	.504 ± .003	.519 ± .010	.497 ± .003	.506 ± .002	.501 ± .008
Uniform (LLaDA2)	.955 ± .010	.936 ± .011	.991 ± .001	.826 ± .044	.779 ± .011	.926 ± .019	.978 ± .001	.893 ± .008	.989 ± .001	.779 ± .019	.881 ± .011
Easy-to-Hard (LLaDA2)	.951 ± .010	.924 ± .012	.993 ± .001	.848 ± .037	.779 ± .014	.910 ± .026	.984 ± .001	.900 ± .006	.990 ± .001	.798 ± .013	.862 ± .011
MaDAC (LLaDA2)	**.976 ± .007**	**.954 ± .011**	**.997 ± .000**	**.861 ± .027**	**.883 ± .010**	**.962 ± .016**	**.993 ± .001**	**.940 ± .005**	**.996 ± .000**	**.850 ± .013**	**.928 ± .007**

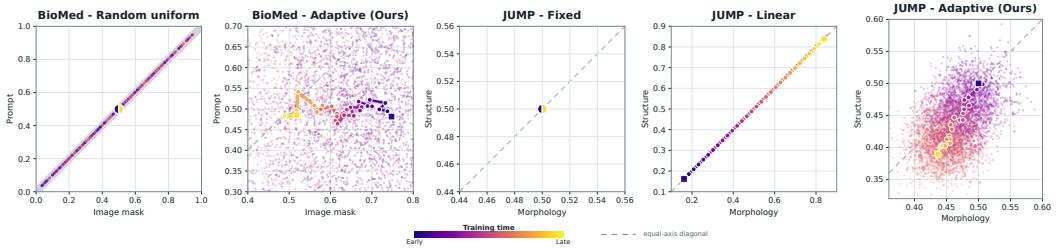

Figure 4: **Curriculum geometry across modalities.** Faint points show realized training actions, and connected markers show means over successive 200-step windows from early (purple) to late (yellow) training. The dashed line denotes equal control values. Fixed, uniform, and linear schedules are restricted to a point or the diagonal, whereas MaDAC explores off-diagonal regions and adapts its allocation over training.

and U-Net/ResNet-34 (Ronneberger et al., 2015; He et al., 2016). Backbone-specific spatial decoders and Further details are in Appendix C.

Results. Table 3 reports image-balanced Dice, macro-organ Dice, and empty-image rate on the source-held-out test. LLaDA2-based MaDAC training improves image Dice and macro-organ Dice over LLaDA2-based uniform training, while reducing the empty-image rate significantly to 0.45%. It also exceeds the dense specialists in image Dice and matches SegFormer's rounded macro-organ Dice. UniDisc-based MaDAC training improves over its uniform counterpart on all three measures. Figure 3 shows representative HuBMAP kidney and lung masks in its bottom row.

4.4 Supporting the Theory: Multimodal Curriculum Geometry

Figure 4 visualizes adaptive curriculum actions in two-dimensional control spaces. The baseline policies illustrate increasingly flexible but still constrained allocations. Random Uniform samples a scalar corruption level jointly so its actions remain on the diagonal, while Linear follows a predetermined trajectory. These cases correspond directly to the shared-rate setting analyzed in Proposition 3.2, not exposing off-diagonal states where the two modalities or targets are treated differently. The adaptive policies exhibit a qualitatively different geometry, illustrating our theory on expanded multimodal support and adaptive allocation within that support. In BioMed, MaDAC with modality-specific control produces off-diagonal states across mask and prompt modalities. In JUMP, the policy changes the relative allocation between morphology and structure. Moreover, the adapted temporal centers move through these regions as training progresses. This behavior is consistent with the state-dependent allocation in Theorem 3.3, that effective curriculums adjust which part of the multimodal problem it emphasizes as the learner state evolves. Additional details on the visualization are provided in Appendix D.4.

Ablation, scaling, and training dynamics studies are in Appendix D. Appendix D.5 identifies effects of outer-, inner-, and route-allocation, while Appendix D.6 extends analysis to scales and dynamics.

5 Conclusion and Discussion

We introduced MaDAC, an adaptive curriculum framework for multimodal diffusion language models that allocates training across corruption-defined prediction problems while preserving coverage. Across JUMP, BioMed, and HuBMAP, MaDAC improves primary task metrics over matched standard training, across both pretrained and from-scratch backbones. These results span perturbation-phenotype modeling, biomedical image segmentation conditioned on semantic prompts, and organ-conditioned spatial tissue segmentation under source shift. Our results also clarify when such adaptive strategy is useful. Gains arise when different corruption or allocation choices define meaningfully different prediction problems and their usefulness changes during training. The ablations show that outer and inner allocation are task dependent and can interact, with their individual contributions varying across tasks. Supplementary experiments further show that the benefit does not arise from curriculum complexity alone, but from adaptation across meaningfully heterogeneous problems.

AI USE STATEMENT

In this work, we used generative AI tools to assist in the writing of the paper, including clean and reformat floats in latex, and provide feedback to improve readability. Additionally, we used generative AI tools to provide feedback on figures and search for information. We did not use generative AI tools to design research methodology, formulate mathematical claims, propose or refine hypotheses, or interpret results. The generation of synthetic datasets and assistance with translation are not applicable to this work.

All AI-assisted outputs were reviewed by the authors before being incorporated into the work. AI-assisted suggestions were reviewed for correctness and consistency. The authors made the final decisions regarding the methodology, implementation, analysis, and presentation of the work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work is computational and does not involve new biological experiments, clinical interventions, or experiments with human participants. All biological and biomedical data used in our experiments are derived from previously published or publicly released datasets and benchmarks, and we use these data for computational model training and evaluation under their established research settings. We do not collect new biological specimens, patient data, or personally identifiable information. Refer to Appendix E for licenses of external datasets.

Our biomedical experiments, including perturbation-phenotype modeling and biomedical image and spatial segmentation tasks, are intended to evaluate multimodal machine-learning methodology rather than to provide clinical diagnoses, treatment recommendations, or experimentally validated biological conclusions. Accordingly, model-generated predictions and visualizations should be interpreted as computational research outputs and not as clinical or experimental evidence. We report dataset provenance, evaluation protocols, and relevant experimental settings to support reproducibility and appropriate interpretation of the results.

REPRODUCIBILITY STATEMENT

Reproducibility materials are shared through our anonymous code repository: `https://anonymous.4open.science/r/MaDAC-6501/`. The repository contains task configurations, ata-manifest specifications, dependency information, and data-preparation, training, and inference code used in the experiments, including utilities needed to reconstruct runs. The main paper defines the MaDAC objective and training loop in Section 3 and Algorithm 1; the assumptions, derivations, and complete proofs are provided in Appendix B.3, with controller and objective conventions in Appendices B.1 and B.2. The benchmark descriptions, task interfaces, splits, baselines, and primary metrics are specified in Sections 4.1-4.3; the appendix records task-specific training schedules, reporting and uncertainty conventions, ablations, scaling and corruption-space diagnostics, compute accounting, and additional evaluation results (Appendix C and D). Raw external datasets and pretrained checkpoints are obtained under their original terms, which are documented in Appendix E.

REFERENCES

Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630(8016):493–500, 2024.

Inclusion AI, Tiwei Bie, Haoxing Chen, Tieyuan Chen, Zhenglin Cheng, Long Cui, Kai Gan, Zhicheng Huang, Zhenzhong Lan, Haoquan Li, Jianguo Li, Tao Lin, Qi Qin, Hongjun Wang, Xiaomei Wang, Haoyuan Wu, Yi Xin, and Junbo Zhao. Llada2.0-uni: Unifying multimodal understanding and generation with diffusion large language model, 2026. URL `https://arxiv.org/abs/2604.20796`.

Sarah Alamdari, Nitya Thakkar, Rianne Van Den Berg, Neil Tenenholtz, Robert Strome, Alan M Moses, Alex X Lu, Nicolò Fusi, Ava P Amini, and Kevin K Yang. Protein generation with evolutionary diffusion: sequence is all you need. *BioRxiv*, pp. 2023–09, 2023.

Marianne Arriola, Aaron Gokaslan, Justin T. Chiu, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Subham Sekhar Sahoo, and Volodymyr Kuleshov. Block diffusion: Interpolating between autoregressive and diffusion language models, 2025. URL https://arxiv.org/abs/2503.09573.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 17981–17993. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/958c530554f78bcd8e97125b70e6973d-Paper.pdf.

Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, Zesen Cheng, Lianghao Deng, Wei Ding, Chang Gao, Chunjiang Ge, Wenbin Ge, Zhifang Guo, Qidong Huang, Jie Huang, Fei Huang, Binyuan Hui, Shutong Jiang, Zhaohai Li, Mingsheng Li, Mei Li, Kaixin Li, Zicheng Lin, Junyang Lin, Xuejing Liu, Jiawei Liu, Chenglong Liu, Yang Liu, Dayiheng Liu, Shixuan Liu, Dunjie Lu, Ruilin Luo, Chenxu Lv, Rui Men, Lingchen Meng, Xuancheng Ren, Xingzhang Ren, Sibo Song, Yuchong Sun, Jun Tang, Jianhong Tu, Jianqiang Wan, Peng Wang, Pengfei Wang, Qiuyue Wang, Yuxuan Wang, Tianbao Xie, Yiheng Xu, Haiyang Xu, Jin Xu, Zhibo Yang, Mingkun Yang, Jianxin Yang, An Yang, Bowen Yu, Fei Zhang, Hang Zhang, Xi Zhang, Bo Zheng, Humen Zhong, Jingren Zhou, Fan Zhou, Jing Zhou, Yuanzhi Zhu, and Ke Zhu. Qwen3-vl technical report, 2025. URL https://arxiv.org/abs/2511.21631.

Fan Bao, Shen Nie, Kaiwen Xue, Chongxuan Li, Shi Pu, Yaole Wang, Gang Yue, Yue Cao, Hang Su, and Jun Zhu. One transformer fits all distributions in multi-modal diffusion at scale, 2023. URL https://arxiv.org/abs/2303.06555.

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, pp. 1–8, Montreal, Quebec, Canada, 2009. ACM Press. ISBN 978-1-60558-516-1. doi: 10.1145/1553374.1553380. URL http://portal.acm.org/citation.cfm?doid=1553374.1553380.

Garyk Brixi, Matthew G. Durrant, Jerome Ku, Michael Poli, Greg Brockman, Daniel Chang, Gabriel A. Gonzalez, Samuel H. King, David B. Li, Aditi T. Merchant, Mohsen Naghipourfar, Eric Nguyen, Chiara Ricci-Tam, David W. Romero, Gwanggyu Sun, Ali Taghibakshi, Anton Vorontsov, Brandon Yang, Myra Deng, Liv Gorton, Nam Nguyen, Nicholas K. Wang, Etowah Adams, Stephen A. Baccus, Steven Dillmann, Stefano Ermon, Daniel Guo, Rajesh Ilango, Ken Janik, Amy X. Lu, Reshma Mehta, Mohammad R.K. Mofrad, Madelena Y. Ng, Jaspreet Pannu, Christopher Ré, Jonathan C. Schmok, John St. John, Jeremy Sullivan, Kevin Zhu, Greg Zynda, Daniel Balsam, Patrick Collison, Anthony B. Costa, Tina Hernandez-Boussard, Eric Ho, Ming-Yu Liu, Thomas McGrath, Kimberly Powell, Dave P. Burke, Hani Goodarzi, Patrick D. Hsu, and Brian L. Hie. Genome modeling and design across all domains of life with evo 2. *bioRxiv*, 2025. doi: 10.1101/2025.02.18.638918. URL https://www.biorxiv.org/content/early/2025/02/21/2025.02.18.638918.

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers, 2021. URL https://arxiv.org/abs/2104.14294.

Srinivas Niranj Chandrasekaran, Jeanelle Ackerman, Eric Alix, D. Michael Ando, John Arevalo, Melissa Bennion, Nicolas Boisseau, Adriana Borowa, Justin D. Boyd, Laurent Brino, Patrick J. Byrne, Hugo Ceulemans, Carolyn Ch'ng, Beth A. Cimini, Djork-Arne Clevert, Nicole Deflaux, John G Doench, Thierry Dorval, Regis Doyonnas, Vincenza Dragone, Ola Engkvist, Patrick W. Faloon, Briana Fritchman, Florian Fuchs, Sakshi Garg, Tamara J. Gilbert, David Glazer, David Gnutt, Amy Goodale, Jeremy Grignard, Judith Guenther, Yu Han, Zahra Hanifehlou, Santosh Hariharan, Desiree Hernandez, Shane R Horman, Gisela Hormel, Michael Huntley, Ilknur Icke, Makiyo Iida, Christina B. Jacob, Steffen Jaensch, Jawahar Khetan, Maria Kost-Alimova, Tomasz

Krawiec, Daniel Kuhn, Charles-Hugues Lardeau, Amanda Lembke, Francis Lin, Kevin D. Little, Kenneth R. Lofstrom, Sofia Lotfi, David J. Logan, Yi Luo, Franck Madoux, Paula A. Marin Zapata, Brittany A. Marion, Glynn Martin, Nicola Jane McCarthy, Lewis Mervin, Lisa Miller, Haseeb Mohamed, Tiziana Monteverde, Elizabeth Mouchet, Barbara Nicke, Arnaud Ogier, Anne-Laure Ong, Marc Osterland, Magdalena Otrocka, Pieter J. Peeters, James Pilling, Stefan Prechtl, Chen Qian, Krzysztof Rataj, David E Root, Sylvie K. Sakata, Simon Scrace, Hajime Shimizu, David Simon, Peter Sommer, Craig Spruiell, Iffat Sumia, Susanne E Swalley, Hiroki Terauchi, Amandine Thibaudeau, Amy Unruh, Jelle Van de Waeter, Michiel Van Dyck, Carlo van Staden, Michał Warchoł, Erin Weisbart, Amélie Weiss, Nicolas Wiest-Daessle, Guy Williams, Shan Yu, Bolek Zapiec, Marek Żyła, Shantanu Singh, and Anne E. Carpenter. Jump cell painting dataset: morphological impact of 136,000 chemical and genetic perturbations. *bioRxiv*, 2023. doi: 10.1101/2023.03.23.534023. URL https://www.biorxiv.org/content/early/2023/03/27/2023.03.23.534023.

Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation, 2017. URL https://arxiv.org/abs/1706.05587.

Wei Chen, Lin Li, Yongqi Yang, Bin Wen, Fan Yang, Tingting Gao, Yu Wu, and Long Chen. Comm: A coherent interleaved image-text dataset for multimodal understanding and generation, 2025. URL https://arxiv.org/abs/2406.10462.

Jooyoung Choi, Jungbeom Lee, Chaehun Shin, Sungwon Kim, Hyunwoo Kim, and Sungroh Yoon. Perception prioritized training of diffusion models, 2022. URL https://arxiv.org/abs/2204.00227.

Noel Codella, Veronica Rotemberg, Philipp Tschandl, M. Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, Harald Kittler, and Allan Halpern. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic), 2019. URL https://arxiv.org/abs/1902.03368.

Chaorui Deng, Deyao Zhu, Kunchang Li, Chenhui Gou, Feng Li, Zeyu Wang, Shu Zhong, Weihao Yu, Xiaonan Nie, Ziang Song, Guang Shi, and Haoqi Fan. Emerging properties in unified multimodal pretraining, 2025. URL https://arxiv.org/abs/2505.14683.

Jevgenij Gamper, Navid Koohbanani, Ksenija Benet, Ali Khuram, and Nasir Rajpoot. *PanNuke: An Open Pan-Cancer Histology Dataset for Nuclei Instance Segmentation and Classification*, pp. 11–19. 07 2019. ISBN 978-3-030-23936-7. doi: 10.1007/978-3-030-23937-4_2.

Shansan Gong, Shivam Agarwal, Yizhe Zhang, Jiacheng Ye, Lin Zheng, Mukai Li, Chenxin An, Peilin Zhao, Wei Bi, Jiawei Han, Hao Peng, and Lingpeng Kong. Scaling diffusion language models via adaptation from autoregressive models, 2025. URL https://arxiv.org/abs/2410.17891.

Alex Graves, Marc G. Bellemare, Jacob Menick, Remi Munos, and Koray Kavukcuoglu. Automated curriculum learning for neural networks, 2017. URL https://arxiv.org/abs/1704.03003.

Tiankai Hang, Shuyang Gu, Chen Li, Jianmin Bao, Dong Chen, Han Hu, Xin Geng, and Baining Guo. Efficient diffusion training via min-snr weighting strategy, 2024. URL https://arxiv.org/abs/2303.09556.

Thomas Hayes, Roshan Rao, Halil Akin, Nicholas J. Sofroniew, Deniz Oktay, Zeming Lin, Robert Verkuil, Vincent Q. Tran, Jonathan Deaton, Marius Wiggert, Rohil Badkundri, Irhum Shafkat, Jun Gong, Alexander Derry, Raul S. Molina, Neil Thomas, Yousuf A. Khan, Chetan Mishra, Carolyn Kim, Liam J. Bartie, Matthew Nemeth, Patrick D. Hsu, Tom Sercu, Salvatore Candido, and Alexander Rives. Simulating 500 million years of evolution with a language model. *Science*, 387(6736):850–858, 2025. doi: 10.1126/science.ads0018. URL https://www.science.org/doi/abs/10.1126/science.ads0018.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. doi: 10.1109/CVPR.2016.90.

Addison Howard, Cecilia Lindskog, Emma Lundberg, Katy Borner, Leah Godwin, Shriya, Sohier Dane, Trang Le, and Yashvardhan Jain. Hubmap + hpa - hacking the human body. https://www.kaggle.com/competitions/hubmap-organ-segmentation, 2022. Kaggle.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models, 2021. URL https://arxiv.org/abs/2106.09685.

John B. Ingraham, Max Baranov, Zak Costello, Karl W. Barber, Wujie Wang, Ahmed Ismail, Vincent Frappier, Dana M. Lord, Christopher Ng-Thow-Hing, Erik R. Van Vlack, Shan Tie, Vincent Xue, Sarah C. Cowles, Alan Leung, João V. Rodrigues, Claudio L. Morales-Perez, Alex M. Ayoub, Robin Green, Katherine Puentes, Frank Oplinger, Nishant V. Panwar, Fritz Obermeyer, Adam R. Root, Andrew L. Beam, Frank J. Poelwijk, and Gevorg Grigoryan. Illuminating protein space with a programmable generative model. *Nature*, 623(7989):1070–1078, 2023. doi: 10.1038/s41586-023-06728-8. URL https://doi.org/10.1038/s41586-023-06728-8.

Diederik P. Kingma, Tim Salimans, Ben Poole, and Jonathan Ho. Variational diffusion models, 2023. URL https://arxiv.org/abs/2107.00630.

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollar, and Ross Girshick. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4015–4026, October 2023.

M. Kumar, Benjamin Packer, and Daphne Koller. Self-paced learning for latent variable models. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta (eds.), *Advances in Neural Information Processing Systems*, volume 23. Curran Associates, Inc., 2010. URL https://proceedings.neurips.cc/paper_files/paper/2010/file/e57c6b956a6521b28495f2886ca0977a-Paper.pdf.

Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto. Diffusion-lm improves controllable text generation, 2022. URL https://arxiv.org/abs/2205.14217.

Fenglin Liu, Shen Ge, Yuexian Zou, and Xian Wu. Competence-based multimodal curriculum learning for medical report generation, 2023. URL https://arxiv.org/abs/2206.14579.

Shengchao Liu, Weili Nie, Chengpeng Wang, Jiarui Lu, Zhuoran Qiao, Ling Liu, Jian Tang, Chaowei Xiao, and Anima Anandkumar. Multi-modal molecule structure-text model for text-based retrieval and editing, 2024a. URL https://arxiv.org/abs/2212.10789.

Shilong Liu, Zhaoyang Zeng, Tianhe Ren, Feng Li, Hao Zhang, Jie Yang, Qing Jiang, Chunyuan Li, Jianwei Yang, Hang Su, Jun Zhu, and Lei Zhang. Grounding dino: Marrying dino with grounded pre-training for open-set object detection, 2024b. URL https://arxiv.org/abs/2303.05499.

Aaron Lou, Chenlin Meng, and Stefano Ermon. Discrete diffusion modeling by estimating the ratios of the data distribution, 2024. URL https://arxiv.org/abs/2310.16834.

Yizhen Luo, Jiahuan Zhang, Siqi Fan, Kai Yang, Yushuai Wu, Mu Qiao, and Zaiqing Nie. Biomedgpt: Open multimodal generative pre-trained transformer for biomedicine. *arXiv preprint arXiv:2308.09442*, 2023.

Timo Lüddecke and Alexander S. Ecker. Image segmentation using text and image prompts, 2022. URL https://arxiv.org/abs/2112.10003.

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024. doi: 10.1038/s41467-024-44824-z. URL https://doi.org/10.1038/s41467-024-44824-z.

Ahmed Nassar, Nikolaos Livathinos, Maksym Lysak, and Peter Staar. Tableformer: Table structure understanding with transformers, 2022. URL https://arxiv.org/abs/2203.01017.

Eric Nguyen, Michael Poli, Marjan Faizi, Armin Thomas, Callum Birch-Sykes, Michael Wornow, Aman Patel, Clayton Rabideau, Stefano Massaroli, Yoshua Bengio, Stefano Ermon, Stephen A. Baccus, and Chris Ré. Hyenadna: Long-range genomic sequence modeling at single nucleotide resolution, 2023. URL https://arxiv.org/abs/2306.15794.

Alex Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models, 2021. URL https://arxiv.org/abs/2102.09672.

Shen Nie, Fengqi Zhu, Zebin You, Xiaolu Zhang, Jingyang Ou, Jun Hu, Jun Zhou, Yankai Lin, Ji-Rong Wen, and Chongxuan Li. Large language diffusion models, 2025. URL https://arxiv.org/abs/2502.09992.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong, Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder, Mario Saltarelli, Ted Sanders, Shibani Santurkar, Girish Sastry, Heather Schmidt, David Schnurr, John Schulman, Daniel Selsam, Kyla Sheppard, Toki Sherbakov, Jessica Shieh, Sarah Shoker, Pranav Shyam, Szymon Sidor, Eric Sigler, Maddie Simens, Jordan Sitkin, Katarina Slama, Ian Sohl, Benjamin Sokolowsky, Yang Song, Natalie Staudacher, Felipe Petroski Such, Natalie Summers, Ilya Sutskever, Jie Tang, Nikolas Tezak, Madeleine B. Thompson, Phil Tillet, Amin Tootoonchian, Elizabeth Tseng, Preston Tuggle, Nick Turley, Jerry Tworek, Juan Felipe Cerón Uribe, Andrea Vallone, Arun Vijayvergiya, Chelsea Voss, Carroll Wainwright, Justin Jay Wang, Alvin Wang, Ben Wang, Jonathan Ward, Jason Wei, CJ Weinmann, Akila Welihinda, Peter Welinder, Jiayi Weng, Lilian Weng, Matt Wiethoff, Dave Willner, Clemens Winter, Samuel Wolrich, Hannah Wong, Lauren Workman, Sherwin Wu, Jeff Wu, Michael Wu, Kai Xiao, Tao Xu, Sarah Yoo, Kevin Yu, Qiming

Yuan, Wojciech Zaremba, Rowan Zellers, Chong Zhang, Marvin Zhang, Shengjia Zhao, Tianhao Zheng, Juntang Zhuang, William Zhuk, and Barret Zoph. Gpt-4 technical report, 2024. URL https://arxiv.org/abs/2303.08774.

Shubham Parashar, Shurui Gui, Xiner Li, Hongyi Ling, Sushil Vemuri, Blake Olson, Eric Li, Yu Zhang, James Caverlee, Dileep Kalathil, and Shuiwang Ji. Curriculum reinforcement learning from easy to hard tasks improves llm reasoning, 2026. URL https://arxiv.org/abs/2506.06632.

Emmanouil Antonios Platanios, Otilia Stretcu, Graham Neubig, Barnabas Poczos, and Tom M. Mitchell. Competence-based curriculum learning for neural machine translation, 2019. URL https://arxiv.org/abs/1903.09848.

Chengxuan Qian, Kai Han, Jiaxin Liu, Zhenlong Yuan, Zhengzhong Zhu, Jingchao Wang, Chongwen Lyu, Jun Chen, and Zhe Liu. Dyncim: Dynamic curriculum for imbalanced multimodal learning, 2025. URL https://arxiv.org/abs/2503.06456.

David Rogers and Mathew Hahn. Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50:742–754, 2010. doi: 10.1021/ci100050t.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. URL https://arxiv.org/abs/1505.04597.

Rohan Saha, Abrar Fahim, Alona Fyshe, and Alex Murphy. Exploring curriculum learning for vision-language tasks: A study on small-scale multimodal training, 2024. URL https://arxiv.org/abs/2410.15509.

Subham Sekhar Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin T Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models, 2024. URL https://arxiv.org/abs/2406.07524.

Gabriele Scalia, Steven T. Rutherford, Ziqing Lu, Kerry R. Buchholz, Nicholas Skelton, Kangway Chuang, Nathaniel Diamant, Jan-Christian Hütter, Jerome-Maxim Luescher, Anh Miu, Jeff Blaney, Leo Gendelev, Elizabeth Skippington, Greg Zynda, Nia Dickson, Michał Koziarski, Yoshua Bengio, Aviv Regev, Man-Wah Tan, and Tommaso Biancalani. Deep-learning-based virtual screening of antibacterial compounds. *Nature Biotechnology*, 44(7):1154–1167, 2026. doi: 10.1038/s41587-025-02814-6. URL https://doi.org/10.1038/s41587-025-02814-6.

Aditya Hemant Shahane, Anuj Kumar Sirohi, Tanmoy Chakraborty, Prathosh A P, and Sandeep Kumar. Factual and edit-sensitive graph-to-sequence generation via graph-aware adaptive noising, 2026. URL https://arxiv.org/abs/2604.24104.

Jiaxin Shi, Kehang Han, Zhe Wang, Arnaud Doucet, and Michalis K. Titsias. Simplified and generalized masked diffusion for discrete data, 2025. URL https://arxiv.org/abs/2406.04329.

Brandon Smock, Rohith Pesala, and Robin Abraham. Pubtables-1m: Towards comprehensive table extraction from unstructured documents, 2021. URL https://arxiv.org/abs/2110.00061.

Xingyu Su, Xiner Li, Yuchao Lin, Ziqian Xie, Degui Zhi, and Shuiwang Ji. Language models for controllable dna sequence design, 2025. URL https://arxiv.org/abs/2507.19523.

Alexander Swerdlow, Mihir Prabhudesai, Siddharth Gandhi, Deepak Pathak, and Katerina Fragkiadaki. Unified multimodal discrete diffusion, 2025. URL https://arxiv.org/abs/2503.20853.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel

Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. Llama 2: Open foundation and fine-tuned chat models, 2023. URL https://arxiv.org/abs/2307.09288.

Xinyou Wang, Zaixiang Zheng, Fei Ye, Dongyu Xue, Shujian Huang, and Quanquan Gu. Diffusion language models are versatile protein learners, 2024a. URL https://arxiv.org/abs/2402.18567.

Xinyou Wang, Zaixiang Zheng, Fei Ye, Dongyu Xue, Shujian Huang, and Quanquan Gu. Dplm-2: A multimodal diffusion protein language model, 2024b. URL https://arxiv.org/abs/2410.13782.

Joseph L. Watson, David Juergens, Nathaniel R. Bennett, Brian L. Trippe, Jason Yim, Helen E. Eisenach, Woody Ahern, Andrew J. Borst, Robert J. Ragotte, Lukas F. Milles, Basile I. M. Wicky, Nikita Hanikel, Samuel J. Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, Valentin De Bortoli, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, Tommi S. Jaakkola, Frank DiMaio, Minkyung Baek, and David Baker. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023. doi: 10.1038/s41586-023-06415-8. URL https://doi.org/10.1038/s41586-023-06415-8.

David Weininger. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.*, 28:31–36, 1988. URL https://api.semanticscholar.org/CorpusID:5445756.

Alexander P. Wu, Heming Yao, Burkhard Hoeckendorf, Garrett Gaskins, Nont Kosaisawe, Ziqing Liu, Phil Hanslovsky, Oleg Mayba, Nicholas Skelton, Gabriele Scalia, John G. Moffat, Tommaso Biancalani, Jan-Christian Hütter, and David Richmond. Virtual phenotypic screening discovers novel scaffolds inhibiting the pi3k/mtor pathway. *bioRxiv*, 2026. doi: 10.64898/2026.06.10.731476. URL https://www.biorxiv.org/content/early/2026/06/14/2026.06.10.731476.

Yijia Xiao, Edward Sun, Yiqiao Jin, Qifan Wang, and Wei Wang. Proteingpt: Multimodal llm for protein property prediction and structure understanding, 2025. URL https://arxiv.org/abs/2408.11363.

Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers, 2021. URL https://arxiv.org/abs/2105.15203.

Jinheng Xie, Weijia Mao, Zechen Bai, David Junhao Zhang, Weihao Wang, Kevin Qinghong Lin, Yuchao Gu, Zhijie Chen, Zhenheng Yang, and Mike Zheng Shou. Show-o: One single transformer to unify multimodal understanding and generation, 2025. URL https://arxiv.org/abs/2408.12528.

Ling Yang, Ye Tian, Bowen Li, Xinchen Zhang, Ke Shen, Yunhai Tong, and Mengdi Wang. Mmada: Multimodal large diffusion language models, 2025. URL https://arxiv.org/abs/2505.15809.

Jiacheng Ye, Zhihui Xie, Lin Zheng, Jiahui Gao, Zirui Wu, Xin Jiang, Zhenguo Li, and Lingpeng Kong. Dream 7b: Diffusion large language models, 2025. URL https://arxiv.org/abs/2508.15487.

Lijun Yu, Yong Cheng, Kihyuk Sohn, José Lezama, Han Zhang, Huiwen Chang, Alexander G. Hauptmann, Ming-Hsuan Yang, Yuan Hao, Irfan Essa, and Lu Jiang. Magvit: Masked generative video transformer, 2023. URL https://arxiv.org/abs/2212.05199.

Theodore Zhao, Yu Gu, Jianwei Yang, Naoto Usuyama, Ho Hin Lee, Sid Kiblawi, Tristan Naumann, Jianfeng Gao, Angela Crabtree, Jacob Abel, Christine Moung-Wen, Brian Piening, Carlo Bifulco, Mu Wei, Hoifung Poon, and Sheng Wang. A foundation model for joint segmentation, detection and recognition of biomedical objects across nine modalities. *Nature Methods*, 22(1): 166–176, 2025. doi: 10.1038/s41592-024-02499-w. URL https://doi.org/10.1038/s41592-024-02499-w.

Xinyi Zheng, Doug Burdick, Lucian Popa, Xu Zhong, and Nancy Xin Ru Wang. Global table extractor (gte): A framework for joint table identification and cell structure recognition using visual context, 2020. URL https://arxiv.org/abs/2005.00589.

Yuwei Zhou, Xin Wang, Hong Chen, Xuguang Duan, and Wenwu Zhu. Intra- and inter-modal curriculum for multimodal learning. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, pp. 3724–3735, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701085. doi: 10.1145/3581783.3612468. URL https://doi.org/10.1145/3581783.3612468.

A Additional Related Works

Multimodal generative models for biology. Generative modeling has become an increasingly important tool for biological design and discovery. Biological foundation models are increasingly moving beyond a single representation. MoleculeSTM (Liu et al., 2024a) aligns molecular structures with natural-language descriptions, BioMedGPT (Luo et al., 2023) connects molecular, protein, and language representations, and ESM3 (Hayes et al., 2025) jointly models protein sequence, structure, and function for multimodal protein generation. In parallel, diffusion models such as RFdiffusion and Chroma (Watson et al., 2023; Ingraham et al., 2023) demonstrate that iterative denoising can generate complex biological structures while incorporating structural and functional constraints. Biological imaging provides another important multimodal setting BiomedParse (Zhao et al., 2025)jointly learns from biomedical images, textual descriptions, and segmentation structures, while the JUMP Cell Painting resource (Chandrasekaran et al., 2023) connects large numbers of chemical and genetic perturbations to high-dimensional cellular morphology. Recent systems such as PhenoCompass (Wu et al., 2026) exploit this connection by aligning compound structures and morphological phenotypes for virtual screening. These works establish the value of integrating heterogeneous biological representations. Our focus is orthogonal: rather than proposing a new modality encoder or alignment objective, we study how a unified generative model should allocate its training effort when different biological modality configurations define different prediction problems.

Diffusion language models and unified multimodal generation. Diffusion training has long optimized the distribution and weighting of noise levels. Improved DDPM (Nichol & Dhariwal, 2021) uses loss-aware timestep sampling, variational diffusion models (Kingma et al., 2023) learn a noise schedule to reduce the variance, P2 (Choi et al., 2022) emphasizes perceptually relevant noise regimes, and Min-SNR (Hang et al., 2024) balances timestep losses using signal-to-noise-based weights. Diffusion-based language modeling has developed through discrete transition models, continuous embedding diffusion, score-based formulations, and masked discrete diffusion (Austin et al., 2021; Li et al., 2022; Lou et al., 2024; Sahoo et al., 2024). Recent large diffusion language models show that masked denoising can scale to general language modeling and instruction following while avoiding a fixed left-to-right factorization (Nie et al., 2025; Gong et al., 2025). This property is especially appealing for multimodal generation. Show-o combines autoregressive and discrete-diffusion objectives within a unified Transformer (Xie et al., 2025), while UniDisc uses discrete diffusion as a joint image–text generative formulation (Swerdlow et al., 2025). MMaDA develops a modality-agnostic multimodal diffusion language model with unified post-training (Yang et al., 2025), and LLaDA2.0-Uni extends large discrete diffusion language modeling to multimodal understanding, generation, editing, and interleaved reasoning (AI et al., 2026). Biological modeling provides closely related examples of both discrete and geometric denoising. EvoDiff and DPLM (Alamdari et al., 2023; Wang et al., 2024a) apply discrete diffusion to protein sequence generation, while ESM3 extends iterative masked generative language modeling across protein sequence, structure, and function modalities (Hayes et al., 2025). DPLM-2 (Wang et al., 2024b) models sequence and discretized structure with multimodal diffusion, supporting joint and conditional generation under distinct modality noise schedules. In complementary geometric formulation, RFdiffusion and Chroma (Watson et al., 2023; Ingraham et al., 2023) generate structures through denoising, while AlphaFold 3 uses atom-coordinate diffusion to predict structures of biomolecular complexes (Abramson et al., 2024). The latter are geometric diffusion systems rather than diffusion language models, but demonstrate the relevance of denoising to heterogeneous biological inputs and outputs. Existing work primarily studies model architecture, representation, denoising objectives, decoding, or post-training. MaDAC addresses a different question: given such a multimodal diffusion model, which corruption-defined problems should it train on at each stage of learning?

Curriculum learning and adaptive multimodal training. Curriculum learning studies how the ordering or weighting of training experiences affects optimization. Classical methods range from hand-designed easy-to-hard curricula (Bengio et al., 2009) and self-paced learning (Kumar et al., 2010) to competence-based schedules and automated curricula that adapt to learning progress (Graves et al., 2017; Platanios et al., 2019). In particular, learning-progress curricula cast task selection as a nonstationary allocation problem and prioritize tasks on which the learner is currently improving (Graves et al., 2017). Curriculum ideas have also been explored in vision-language

and multimodal learning, where training is scheduled according to example difficulty, information density, or modality imbalance (Saha et al., 2024; Qian et al., 2025). MaDAC differs in what constitutes a curriculum unit. It does not primarily reorder examples from a fixed dataset; instead, the curriculum acts on the corruption process, thereby changing the observed-target configuration, modality-specific noise levels, and ultimately the denoising problem presented to the model. Its inner allocation further operates within the selected corruption state, yielding a hierarchical curriculum over both prediction problems and their targets.

B Implementation and Theoretical Analysis of MaDAC

The main text treats multimodal diffusion as an adaptive curriculum over corruption-defined training problems. This appendix makes the associated objective and policy conventions explicit, proves the main results, and analyzes when problem selection and target weighting reinforce one another. Biological instantiations appear after the analysis; the main training loop is Algorithm 1.

B.1 Notation and objective conventions

Corruption-defined problems. The number of modalities is D, the set of hidden positions is C, and the set of supervised targets is $T \subseteq C$. A route g identifies observed, predicted, and hidden-but-unsupervised streams; $a = (g, \mathbf{r}, \xi)$ identifies its corruption configuration. The problem instance $z = (x, a)$ also includes the selected training record. When an objective is normalized across examples, x denotes that minibatch and its loss-unit indices include example indices. $\mathcal{H}$ indexes normalization groups, b_h denotes their nonnegative base-loss scales, and ω_{hj} is a probability distribution over the loss units in group h.

Hidden inputs are not necessarily targets. Write $\mathcal{O}_g$, $\mathcal{P}_g$, and $\mathcal{U}_g$ for the observed, predicted, and hidden-only streams. Positions in an observed stream are retained; a predicted stream contributes targets according to its masking rule; and a hidden-only stream is unavailable to the predictor but has no loss. A route can also prescribe fully masked predicted streams. Hence rates zero and one are legitimate route specifications, even though the learned Beta factors use interior intervals. Route indicators and mask tokens are sufficient to represent this distinction; the notation does not require an additional architecture or an explicit diffusion-time input.

Loss units and normalization groups. For token prediction, a group can be the masked tokens of one modality in one record. For a continuous vector prediction, a loss unit can be the entire vector, and sample weighting is normalized across examples. An auxiliary nonseparable loss, such as an image-level Dice loss, is a whole loss unit rather than a sum of fictitious independent pixel losses. Its placement and coefficient are part of the base objective. Applying weights inside a Dice ratio is not algebraically equivalent to weighting a per-example Dice loss after it is computed.

The general loss in Eq. equation 1 separates normalized allocation from base scale. For nonempty T_m and $W_m = \sum_{j \in T_m} w_j > 0$, the rate-weighted masked-token convention is

$$\mathcal{L}_{\text{mean}} = \sum_m \frac{\lambda_m}{\max(r_m, \epsilon)} \frac{\sum_{j \in T_m} w_j \ell_j}{W_m}. \tag{13}$$

This is a well-defined supervised loss, but its likelihood interpretation does not follow solely from the factor $1/r_m$.

Relation to the masked-diffusion objective. For independent absorbing-state masking of a single stream of length L, a linear mask-probability parameterization gives the standard denoising integrand, normalized per original token,

$$\mathcal{L}_{\text{MD}}(r, C) = \frac{1}{Lr} \sum_{j \in C} \ell_j, \qquad 0 < r < 1, \tag{14}$$

inside the rate integral, with the appropriate boundary conventions (Sahoo et al., 2024; Nie et al., 2025). For $C \neq \varnothing$, writing this as a masked-token mean requires $b = |C|/(Lr)$, not $b = 1/r$. In

fact,

$$\frac{1}{r}\frac{\sum_{j\in C}\ell_j}{|C|} = \frac{L}{|C|}\mathcal{L}_{\mathrm{MD}}(r, C). \tag{15}$$

When $|C| \simeq rL$, the ratio is approximately $1/r$. The difference is rate dependent, not a harmless constant. The corresponding self-normalized weighted convention uses $b = \lambda_m|T_m|/(L_m r_m)$ and $\omega_{mj} = w_j/W_m$, whereas weighting the full-length token sum without self-normalization gives a different objective. A continuous-output regression head has no automatic masked-token likelihood interpretation either. We regard the reduction as part of the experimental contract and keep it fixed when isolating curriculum effects. Arbitrary independent multimodal rates, structured masks, and readiness weights are not claimed to preserve one diffusion ELBO.

Numerical denominator floors. An implementation may replace W_h by $\max(W_h, \epsilon_w)$. Its exact group loss is then

$$b_h \frac{W_h}{\max(W_h, \epsilon_w)} \sum_{j\in T_h} \omega_{hj}\ell_j. \tag{16}$$

Thus a denominator floor can also change the total group scale. In analyses that hold b_h fixed, we assume the floor is inactive; otherwise its detached multiplier must be included in the realized group scale. Empty groups contribute zero. Resampling a mask until it is nonempty changes K and is not an interchangeable numerical implementation.

Why the location of sample normalization matters. A continuous vector can be represented as one loss unit. Normalizing a weight inside that singleton group removes its effect: $w_j\ell_j/w_j = \ell_j$. Sample-level weighting instead spans the relevant minibatch,

$$\mathcal{L}_{\mathrm{sample}} = \frac{\sum_{n\in\mathcal{B}} w_n\ell_n}{\sum_{n\in\mathcal{B}} w_n}. \tag{17}$$

For this reduction, the selected minibatch and its joint corruption form the training unit z. The expectation is taken over the whole minibatch loss; a self-normalized ratio is not replaced by a ratio of population expectations. Token and sample weighting therefore share the same allocation principle with different normalization groups.

An exact hierarchical supervision measure. Let $B(z, C, T) = \sum_h b_h$. When $B > 0$, define $\nu_h = b_h/B$ and

$$\begin{aligned}\Pi_k(dz, dC, dT, h, j) = q_k(dz \mid s_k)K(dC, dT \mid z)\\ \times \nu_h(z, C, T)\omega_{k,hj}(z, C, T).\end{aligned} \tag{18}$$

Then the expected loss equals $\mathbb{E}_{\Pi_k}[B(z, C, T)\ell_j]$. Equivalently, $B\Pi_k$ is a nonnegative supervision measure. The factor B retains the total modality/rate scale; only the relative allocations ν and ω are normalized probabilities. Actual training sums over targets; it need not sample h, j. For zero-loss realizations one may introduce a dummy zero-loss target.

Detached optimization. At step k, the policy, cached scores, and inner weights are frozen when computing $\nabla_\theta\mathcal{L}_k$. This is a gradient of a step-specific surrogate, not the total derivative of a single objective whose sampling law and confidence weights are functions of θ. Backpropagating through $w_j(\theta)$ would add $\ell_j\nabla_\theta\omega_j$ terms and would define a different algorithm. The same distinction applies to the controller state.

Reference objective and change of measure. Let $\omega^0_{hj} = 1/|T_h|$ and define

$$J_0(\theta) = \mathbb{E}_{z\sim p_0,(C,T)\sim K}\left[\sum_h b_h \sum_j \omega^0_{hj}\ell_j(\theta)\right]. \tag{19}$$

If $p_0 \ll q_k$, $\omega^0_{hj} > 0$ implies $\omega_{k,hj} > 0$, and the kernel and base scales are unchanged, then

$$J_0(\theta) = \mathbb{E}_{q_k,K}\left[\frac{dp_0}{dq_k}(z)\sum_h b_h \sum_j \omega_{k,hj}\frac{\omega^0_{hj}}{\omega_{k,hj}}\ell_j(\theta)\right]. \tag{20}$$

The identity follows by cancellation and change of measure. In particular, correcting only the outer law does not remove inner reweighting. Including x in z also accounts for nonuniform example sampling. If the corruption kernel itself changes independently of a, its likelihood ratio is needed as well. The measure notation covers mixtures of continuous rates and discrete route atoms. MaDAC does not apply these corrections: its reallocation is intentional. The coverage mixture nevertheless gives $dp_0/dq_k \leq 1/\rho$, an upper bound on the outer importance ratio should an objective-preserving control be constructed.

B.2 Adaptive curriculum: policy and update details

B.2.1 Discrete progress and probe timing

For a discrete problem bucket b, let $V_{b,t}$ be its loss at probe event t under a fixed protocol. The feedback-driven controller uses learning progress

$$\bar{V}_{b,t} = \beta_{\text{ema}}\bar{V}_{b,t-1} + (1-\beta_{\text{ema}})V_{b,t}, \quad P_{b,t} = \bar{V}_{b,t-1} - \bar{V}_{b,t}. \tag{21}$$

A categorical factor is refreshed using

$$\pi_k(b \mid s_k) \propto \exp\left(\frac{P_{b,t} + d/(1+n_{b,k})}{\tau_o}\right), \tag{22}$$

where $n_{b,k}$ is usage, $d \geq 0$ controls exploration, and $\tau_o > 0$. Progress prioritizes currently improving problems rather than assuming that high loss alone implies usefulness. It is nevertheless a utility proxy, since progress on one problem can result from training on another.

For each probed bucket b, initialize the EMA from its first observed probe loss and initialize progress to zero. At subsequent probe events, use Eq. equation 21. A probe event need not coincide with one optimizer step. When the time since the last probe differs across buckets, the measured drops cover different intervals; either compare at common intervals or declare the elapsed-step normalization. Hold the probe data, mask-generation protocol, and loss reduction fixed. A loss change caused by changing r_m, b_h, or the readiness rule is not, by itself, learning progress on the same problem.

Usage bonuses in Eq. equation 22 provide an exploration heuristic. Stale progress values can additionally be decayed between probes, with the decay convention fixed for each controller. Neither a usage bonus nor a clipped logit establishes an action-probability floor. In particular, clamping probabilities below by $p_{\min}$ and then renormalizing does not generally retain $p_{\min}$ as the final lower bound. Use an explicit mixture or a properly constrained probability map when a floor is needed. Learning-progress task allocation has prior precedent (Graves et al., 2017); the multimodal curriculum here concerns the problem definition and the coupled levels of allocation.

B.2.2 Bounded Beta factors

For fixed bounds $l_m < u_m$ and $v_m = (r_m - l_m)/(u_m - l_m)$, the transformed density is

$$\pi_m(r_m) = \frac{v_m^{\alpha_m - 1}(1-v_m)^{\beta_m - 1}}{(u_m - l_m)\mathrm{B}(\alpha_m, \beta_m)}. \tag{23}$$

Write $\phi_m = (\log \alpha_m, \log \beta_m)$. With ψ the digamma function,

$$\begin{aligned} \partial_{\log \alpha_m} \log \pi_m &= \alpha_m[\log v_m - \psi(\alpha_m) \\ &\qquad + \psi(\alpha_m + \beta_m)], \end{aligned} \tag{24}$$

$$\begin{aligned} \partial_{\log \beta_m} \log \pi_m &= \beta_m[\log(1-v_m) - \psi(\beta_m) \\ &\qquad + \psi(\alpha_m + \beta_m)]. \end{aligned} \tag{25}$$

Conditional products of these factors permit different modality rates but do not represent arbitrary within-route rate correlations. The route mixture can represent additional regimes; it should not be described as an unrestricted joint density. Affinely transforming a Beta random variable is also different from clipping or truncating it. Clipping creates boundary atoms; truncation changes the normalizing constant. A score-function estimator must match the actual sampling construction. Bounding log-parameters is a stabilization choice, not a proof that the learned policy finds the optimal rate density.

B.2.3 COVERAGE MIXTURES AND CONTROLLER GRADIENTS

Fix the pre-action state and suppress its notation. Let $q_\phi = (1-\rho)\pi_\phi + \rho p_0$, where p_0 and ρ do not depend on ϕ. Let R have finite conditional mean $\overline{R}(z)$ and depend on ϕ only through the sampled z, for this local controller objective. Under the usual differentiation-under-the-integral conditions,

$$\nabla_\phi \mathbb{E}_{q_\phi} R = (1-\rho)\mathbb{E}_{\pi_\phi}[\overline{R}(z)\nabla_\phi \log \pi_\phi(z)]. \tag{26}$$

There are two equivalent unbiased estimators. If the mixture component is logged, sample $A \sim \mathrm{Bernoulli}(1-\rho)$ and use

$$\widehat{g}_\phi = A(R-b)\nabla_\phi \log \pi_\phi(z). \tag{27}$$

Alternatively, for a sample from the marginal mixture use

$$\widehat{g}_\phi = (R-b)\frac{(1-\rho)\pi_\phi(z)}{q_\phi(z)}\nabla_\phi \log \pi_\phi(z). \tag{28}$$

The multiplier is the posterior probability of an adaptive draw. A baseline b depending only on pre-action history has zero contribution because $\mathbb{E}_{\pi_\phi}\nabla \log \pi_\phi = 0$. Using the pre-update baseline is therefore the clean convention. In contrast, applying $(R-b)\nabla \log \pi_\phi(z)$ to every draw from q_ϕ is generally biased. Equations equation 27–equation 28 are local bandit identities; they omit derivatives through future model states and do not provide a policy-gradient guarantee for the complete training run.

Derivation of the score estimators. Differentiating the mixture while holding the pre-action state fixed gives

$$\begin{aligned} &\nabla_\phi \int \overline{R}(z) q_\phi(z)\, dz \\ &\quad = (1-\rho)\int \overline{R}(z)\nabla_\phi \pi_\phi(z)\, dz \\ &\quad = (1-\rho)\mathbb{E}_{\pi_\phi}[\overline{R}(z)\nabla_\phi \log \pi_\phi(z)]. \end{aligned} \tag{29}$$

The expectation of Eq. equation 27 is this integral because $\Pr(A=1) = 1-\rho$ and $z \mid A=1 \sim \pi_\phi$. For Eq. equation 28, multiplication by its sampling density q_ϕ cancels the denominator. In either case the baseline term integrates to $(1-\rho)b\nabla_\phi \int \pi_\phi(z)\, dz = 0$. For mixed discrete/continuous laws the same derivation uses a common dominating measure. Scores are evaluated at the policy that generated the action; a stale policy requires its corresponding off-policy correction.

Factorwise mixing is a different joint policy. Suppose a sequential factorization is deployed as

$$q_j(z_j \mid z_{<j}, s) = (1-\rho_j)\pi_j(z_j \mid z_{<j}, s) + \rho_j p_{0,j}(z_j \mid z_{<j}). \tag{30}$$

Then the product law satisfies

$$q(z \mid s) = \prod_j q_j(z_j \mid z_{<j}, s) \geq \left(\prod_j \rho_j\right) p_0(z). \tag{31}$$

Fixed reference factors contribute a factor of one. The joint floor is not the per-modality mixing mass. For a parameter block exclusive to one conditional factor, its gradient uses that factor's logged component indicator or marginal responsibility and its conditional score. Shared parameters require summing these contributions over all factors that depend on them. Floors and score corrections must match this factorization. If a purportedly protected route or endpoint has zero reference mass, coverage gives no protection for it.

B.2.4 WHAT THE IMPROVEMENT REWARD MEASURES

Fix a common state distribution μ, a validation objective J, and the inner rule used to evaluate each candidate problem b. Define its one-step intervention utility

$$Q(s,b) = J(\theta_s) - \mathbb{E}[J(\theta_s - \eta \widehat{d}_b) \mid s, b], \tag{32}$$

where $\widehat{d}_b$ includes the candidate's fixed sampling protocol. This compares future model performance, not the current difficulty of the corrupted input. On this candidate set, p_0 denotes the induced reference law. Write $V_\mu(q) = \mathbb{E}_{S\sim\mu, b\sim q(\cdot|S)} Q(S,b)$ and

$$\Gamma_\mu = \mathbb{E}_\mu \max_b Q(S,b) - \max_b \mathbb{E}_\mu Q(S,b). \tag{33}$$

Three quantities must be distinguished. A matched bucket probe measures how that bucket's loss changes under the training actually performed. A shared-checkpoint intervention measures before/after loss on the same validation population after a declared action or action sequence. An adjacent-batch training-loss difference compares losses that can differ because of data, masks, reduction coefficients, or target weights. Only the second is a direct experimental estimate of the intervention utility in Eq. equation 32; the others may be useful proxies but are not interchangeable unbiased estimates. For multi-step utility, specify which action is used initially and which continuation policy is held fixed. Controller selection uses training or calibration data rather than final test labels. Reward scaling is part of the protocol when metrics or loss units differ across problems.

B.2.5 Inner weighting and continuous targets

The positive target priority

$$w_{k,hj} = (1 + \gamma u_j + \kappa e_j)\, \sigma\left(\frac{u_j - b_0}{\tau_w}\right), \tag{34}$$

where $\gamma, \kappa \geq 0$, $\tau_w > 0$, and σ is the logistic function. The normalized weights

$$\omega_{k,hj} = \frac{w_{k,hj}}{\sum_{v\in T_h} w_{k,hv}} \tag{35}$$

then redistribute supervision within group h without changing its base scale.

The probe sees the same corrupted inputs as the supervised forward pass; clean labels are used only to evaluate target probabilities or agreement. Scores are computed with stopped gradients. A second forward pass is unnecessary when scores can be detached from the same logits without changing the mask, although a separate probe is permitted and its compute must be counted. The candidate hidden set C and supervised set T are unchanged by inner weighting.

For continuous predictions, a possible bounded agreement convention is $c = (1 + \cos(\widehat{y}, y))/2$. Mapping, clipping, treatment of zero norms, and reduction across vector components define different scores and are specified by the task configuration. A simpler sample-level variant uses $w = 1 + c$. Neither this rule nor Eq. equation 34 equates confidence with validation utility. Unit weights are the explicit fallback. Setting only $\gamma = \kappa = 0$ in the sigmoid rule does not disable weighting, because its gate remains.

B.3 Proofs and quantitative allocation analysis

B.3.1 Proof of Proposition 3.2

Let $Z_m = |C_m|$ and let the conditional token masks be independent given the selected rates. The route and stream lengths are fixed throughout this comparison.

Proof. Condition on a shared rate r. The random difference $Y = Z_1/n_1 - Z_2/n_2$ has conditional mean zero. It is a sum of independent bounded variables: n_1 terms with range length $1/n_1$ and n_2 terms with range length $1/n_2$. Thus the sum of squared range lengths is $n_1^{-1} + n_2^{-1} = n_{\text{eff}}^{-1}$. For a centered random variable with range length c, the bounded-variable exponential-moment inequality gives $\mathbb{E}e^{tY_i} \leq e^{t^2c^2/8}$. Multiplying these bounds and applying Markov's inequality yields, for $t > 0$,

$$\Pr(Y \geq u - v \mid r) \leq \exp\left[-t(u-v) + \frac{t^2}{8n_{\text{eff}}}\right]. \tag{36}$$

Minimizing at $t = 4n_{\text{eff}}(u - v)$ proves

$$\Pr(E_{u,v} \mid r) \leq \Pr(Y \geq u - v \mid r) \leq e^{-2n_{\text{eff}}(u-v)^2}. \tag{37}$$

The bound is uniform in r, so integrating against any shared-rate law preserves it. In particular, that law may depend on the training history, provided the fresh masks are conditionally independent Bernoulli draws.

For independent uniform rates, integration of the binomial law gives

$$\begin{aligned}\Pr(Z_m = j) &= \binom{n_m}{j} \int_0^1 r^j (1-r)^{n_m - j}\, dr \\ &= \binom{n_m}{j} \mathrm{B}(j+1, n_m - j + 1) = \frac{1}{n_m + 1}.\end{aligned} \tag{38}$$

The two counts are independent. Counting admissible integers therefore gives

$$\begin{aligned}\Pr_{\text{independent}}(E_{u,v}) &= \frac{n_1 - \lceil u n_1 \rceil + 1}{n_1 + 1} \frac{\lfloor v n_2 \rfloor + 1}{n_2 + 1} \\ &\longrightarrow (1-u)v.\end{aligned} \tag{39}$$

This proves both claims. □

Bounded-variable exponential moment. For completeness, let a centered variable Y lie in $[a, b]$, with $c = b - a$. Convexity of the exponential bounds its moment-generating function by that of a variable on the two endpoints with the same mean. Writing $p = -a/c$ and $s = tc$, the logarithm of this upper bound is $f(s) = \log(1 - p + pe^s) - ps$. It satisfies $f(0) = f'(0) = 0$ and $f''(s) = p_s(1 - p_s) \le 1/4$, where $p_s = pe^s/(1 - p + pe^s)$. Hence $f(s) \le s^2/8$ by integration, proving $\mathbb{E}e^{tY} \le e^{t^2c^2/8}$. Degenerate endpoints follow by continuity.

Consequence for finite training exposure. If each of K steps uses the same joint route and lengths, but can choose its shared rate using the entire preceding history, let $E_{u,v}^{(k)}$ denote the directional event at step k. A union bound gives

$$\Pr\left[\bigcup_{k=1}^{K} E_{u,v}^{(k)}\right] \le \min\{1, Ke^{-2n_{\text{eff}}(u-v)^2}\}. \tag{40}$$

This is a finite-budget limitation of the shared-rate parameterization, not merely of uniform scalar scheduling. No independence across steps is needed. Independent-rate exposure is nonvanishing for fixed fractional thresholds; a protected directional route can allocate probability even more directly. These conclusions motivate richer corruption choices without asserting that any particular directional problem improves downstream evaluation.

Exact all-or-nothing corner. Let $E = \{Z_1 = n_1, Z_2 = 0\}$. For a shared uniform rate,

$$\begin{aligned}\Pr_{\text{shared}}(E) &= \int_0^1 t^{n_1}(1-t)^{n_2}\, dt \\ &= \frac{n_1! n_2!}{(n_1 + n_2 + 1)!},\end{aligned} \tag{41}$$

whereas independent uniform rates give

$$\Pr_{\text{independent}}(E) = \frac{1}{(n_1 + 1)(n_2 + 1)}. \tag{42}$$

For $n_1 = n_2 = n$, the shared expression is $1/[(2n+1)\binom{2n}{n}]$. Stirling's formula gives $\Theta(4^{-n}/\sqrt{n})$ versus $\Theta(n^{-2})$. The corner probability vanishes even for independent rates; the partial-region result explains why useful *families* of directional problems can retain nonvanishing exposure.

Rate bounds and scope. The shared-rate bound holds for every shared distribution, including one supported on an interior interval. For independently sampled rates with continuous distribution functions F_1, F_2, the partial-region probability converges to $(1 - F_1(u))F_2(v)$ when neither distribution has an atom at the thresholds. This follows from the conditional law of large numbers and dominated convergence. For interior uniform laws, exposure is therefore positive only when their supports include the relevant thresholds. Exact corner expressions become the corresponding truncated integrals. These results do not apply unchanged to fixed-cardinality or structured masks, nor do they diagnose a baseline that already samples separate directional routes.

B.3.2 THE EXACT FIXED-CHECKPOINT GRADIENT RESTRICTION

Fix the data law, mask kernel, base-loss convention, inner allocation, and parameters θ. Let $G_\theta(a)$ be the expected supervised gradient under configuration a, and let $\mathcal{A}_{\mathrm{diag}}$ contain only shared-rate configurations on the active streams, with other route choices fixed or explicitly included in the comparison.

Proposition B.1 (Restricted rate mixtures)**.** *If G_θ is integrable under a probability law q supported on $\mathcal{A}_{\mathrm{diag}}$, then*

$$\mathbb{E}_q G_\theta(a) \in \overline{\mathrm{conv}}\{G_\theta(a) : a \in \mathcal{A}_{\mathrm{diag}}\}. \tag{43}$$

If a configuration gradient lies outside this set, no such probability mixture reproduces it at that checkpoint.

Proof. Let the displayed closed convex hull be H. If the integrable expectation y were outside H, finite-dimensional strict separation would give a vector v and scalar c with $v^\top y > c$ and $v^\top h \le c$ for every $h \in H$. But $v^\top y = \mathbb{E}_q[v^\top G_\theta(a)] \le c$, a contradiction. □

The claim is local in parameter space. It does not exclude the same endpoint being reached by another multi-step path, and it does not cover arbitrary unbounded loss reweighting, which can change the feasible cone rather than only its convex hull. Existence of an out-of-hull gradient also does not imply that it aligns with the chosen validation objective. Those conditions must be examined separately.

B.3.3 A SMOOTH LOCAL UTILITY BOUND

Proposition B.2 (Alignment and second moments)**.** *Let J have L-Lipschitz gradient on the segments from θ to all sampled $\theta - \eta\widehat{g}$, and let $\mathbb{E}\|\widehat{g}\|^2 < \infty$. Then*

$$|J(\theta) - \mathbb{E}J(\theta - \eta\widehat{g}) - \eta\langle \nabla J(\theta), \mathbb{E}\widehat{g}\rangle| \le \frac{L\eta^2}{2}\mathbb{E}\|\widehat{g}\|^2. \tag{44}$$

Proof. For deterministic d, the fundamental theorem of calculus gives

$$\begin{aligned} &J(\theta - \eta d) - J(\theta) + \eta \nabla J(\theta)^\top d \\ &\quad = -\eta \int_0^1 [\nabla J(\theta - t\eta d) - \nabla J(\theta)]^\top d \, dt. \end{aligned} \tag{45}$$

Its absolute value is at most $\eta \int_0^1 Lt\eta\|d\|^2 dt = L\eta^2\|d\|^2/2$. Substitute $d = \widehat{g}$ and take expectations. □

This strengthens a formal first-order expansion by making the remainder assumption explicit. Differentiability alone gives an $o(\eta)$ remainder, not an $O(\eta^2)$ bound. Moreover, $\mathbb{E}\|\widehat{g}\|^2 = \|\mathbb{E}\widehat{g}\|^2 + \operatorname{tr}\operatorname{Cov}(\widehat{g})$, so gradient variance affects the guarantee. The result is an SGD-style local analysis. For a practical optimizer it applies to the actual scaled update direction when its moment and state dependence are included, not automatically to the raw gradient. Thresholded Dice, retrieval ranks, and AUROC need not satisfy the smoothness assumption; J is a declared smooth validation surrogate unless otherwise justified.

B.3.4 PROOF OF THEOREM 3.3

Proof. For a fixed state s, the uniform error assumption implies

$$\max_b Q(s,b) \le \max_b \widehat{Q}(s,b) + \varepsilon(s), \tag{46}$$

$$\mathbb{E}_\pi Q(s,b) \ge \mathbb{E}_\pi \widehat{Q}(s,b) - \varepsilon(s). \tag{47}$$

Subtracting gives the pointwise bound

$$\max_b Q(s,b) - \mathbb{E}_\pi Q(s,b) \le 2\varepsilon(s) + \zeta(s). \tag{48}$$

Let $V_0 = \mathbb{E}_{S,b\sim p_0} Q(S,b)$. Linearity yields

$$\begin{aligned} V_\mu(q) &\geq (1-\rho)\big[\mathbb{E}_\mu \max_b Q(S,b) - 2\mathbb{E}_\mu \varepsilon - \mathbb{E}_\mu \zeta\big] + \rho V_0, \\ V^\rho_{\mathrm{stat}} &= (1-\rho)\max_b \mathbb{E}_\mu Q(S,b) + \rho V_0. \end{aligned} \tag{49}$$

The second equality holds because maximizing a linear functional over a finite probability simplex attains its maximum at a point mass. Subtraction proves Eq. equation 9. Nonnegativity of Γ_μ follows from $Q(s,b) \leq \max_v Q(s,v)$ for each fixed b and taking expectations.

For the softmax claim, let $v_b = \widehat{Q}(s,b)$, $Z = \sum_b \exp(v_b/\tau_o)$, and $H(\pi) = -\sum_b \pi_b \log \pi_b$. Substitution gives

$$\tau_o \log Z = \sum_b \pi_b v_b + \tau_o H(\pi). \tag{50}$$

Since $\max_b v_b \leq \tau_o \log Z$ and $H(\pi) \leq \log|\mathcal{B}|$, $\max_b v_b - \mathbb{E}_\pi v_b \leq \tau_o \log|\mathcal{B}|$. □

When the oracle gap is strict. For finite $\mathcal{B}$, $\Gamma_\mu = 0$ exactly when at least one fixed action is optimal μ-almost surely. To see the nontrivial direction, let b^* maximize $\mathbb{E}_\mu Q(S,b)$. Zero gap means that the integrable nonnegative random variable $\max_b Q(S,b) - Q(S,b^*)$ has expectation zero and hence is zero almost surely. With a common coverage mass, the exact oracle adaptive-versus- stationary gap is $(1-\rho)\Gamma_\mu$. Coverage reduces allocation freedom linearly; it does not eliminate the need for accurate utility estimates.

What the theorem does not assume. The distribution μ is fixed before comparing policies, for example by selecting shared checkpoints. It is not silently identified with the different state-occupancy distributions generated by two complete adaptive training runs. The error ε includes bias from progress proxies, changing loss scales, partial state observation, and any exploration bonus not measuring utility. The selection error ζ also includes restrictions of the policy family or incomplete controller optimization. No bound on these quantities is asserted from endpoint improvements alone.

Fixed-horizon utility. The selection theorem is not restricted to one-step utility. Fix a positive integer H and a continuation policy ν. Starting from the same state s, let $\theta^{b,\nu}_{s,H}$ result from one update using candidate b, followed by $H-1$ updates using ν. Define

$$Q_H(s,b) = J(\theta_s) - \mathbb{E}[J(\theta^{b,\nu}_{s,H})]. \tag{51}$$

The first candidate is imposed rather than inferred from an observational association. The continuation rule, optimizer, and data protocol are shared across compared candidates; realized future states can differ. Substituting this integrable utility into the selection proof gives the identical bound with Q_H, its estimation error, and its oracle gap. This does not extend the one-step Taylor bound to an entire trajectory without additional assumptions.

B.3.5 Continuous action spaces

The same argument applies on a compact metric action space and a standard Borel state space when $Q(s,a)$ and $\widehat{Q}(s,a)$ are measurable in s, continuous in a, and dominated by an integrable function of s. Measurable maximizing selections then exist, and expectation defines a continuous utility for stationary actions. Define $\varepsilon(s) = \sup_a |\widehat{Q}(s,a) - Q(s,a)|$ and $\zeta(s) = \sup_a \widehat{Q}(s,a) - \int \widehat{Q}(s,a)\pi(da \mid s)$. The identical pointwise argument proves the bound. The finite-cardinality entropy bound must not be applied to a Beta density. For that family, ζ remains its approximation and optimization error.

For a finite h-net of a compact action space and an $L_Q(s)$-Lipschitz utility, the grid maximum is within $L_Q(s)h$ of the continuous maximum. Accordingly, comparison of a grid-based adaptive policy with the best continuous stationary policy incurs at most an additional $(1-\rho)h\mathbb{E}_\mu L_Q(S)$ on the lower bound. This expresses a genuine resolution–estimation tradeoff: a finer action space can reduce approximation error while making progress estimates noisier or more expensive.

B.3.6 WHY OBSERVED PROGRESS IS NOT ACTION UTILITY

Let J_u be smooth probe losses and $d_v = \mathbb{E}\widehat{g}_v$ the expected training direction for action v at a fixed state. Define the local transfer matrix

$$A_{uv} = \langle \nabla J_u(\theta), d_v \rangle. \tag{52}$$

If training draws actions from q, Proposition B.2 shows that the expected decrease of probe u is

$$\Delta_u = \eta \sum_v q_v A_{uv} + e_u, \qquad |e_u| \le \frac{L_u \eta^2}{2} \sum_v q_v \mathbb{E}\|\widehat{g}_v\|^2. \tag{53}$$

By contrast, for $J = \sum_u \nu_u J_u$ with fixed nonnegative weights ν_u, the first-order utility of action v is

$$Q(v) = \eta \sum_u \nu_u A_{uv} + e'_v, \tag{54}$$

with the analogous smoothness remainder. Probe progress averages a *row* using the actual training mixture; intervention utility averages a *column* using evaluation weights. An EMA cannot by itself remove this credit-assignment difference.

For example, with $A = \mathrm{diag}(2, 1)$, $q = (0.01, 0.99)$, and $\nu = (0.9, 0.1)$, the first-order observed progress is proportional to $(0.02, 0.99)$, whereas the action utilities are proportional to $(1.8, 0.1)$. Prioritizing the largest observed decrease selects the less useful action for the stated evaluation objective. This is a mathematical example, not a reported result. It motivates shared-state interventions and route-wise validation of the progress signal; it does not preclude progress being an effective low-cost heuristic in the actual tasks.

B.3.7 PROOF OF PROPOSITION 3.4

Fix a checkpoint, a realized problem, and one nonempty group. Write $n_h = |T_h|$, $g_{hj} = \nabla \ell_j(\theta)$, $\chi_{hj} = \langle \nabla J(\theta), g_{hj} \rangle$, and $\bar{w}_h = n_h^{-1} \sum_j w_{hj}$. Weights are positive and detached.

Proof. With expectations taken uniformly over this group,

$$\begin{aligned} &\sum_j \frac{w_{hj}}{\sum_v w_{hv}} \chi_{hj} - \frac{1}{n_h} \sum_j \chi_{hj} \\ &= \frac{\mathbb{E}_h[w\chi]}{\mathbb{E}_h w} - \mathbb{E}_h \chi = \frac{\mathrm{Cov}_h(w, \chi)}{\bar{w}_h}. \end{aligned} \tag{55}$$

Multiply by the unchanged group scales b_h, sum over groups, and take the conditional expectation over K to obtain Eq. equation 10. The expected full alignment is $\mathcal{A}_\omega = \mathcal{A}_0 + \Delta\mathcal{A}$. Adding and subtracting $\mathbb{E}_q \mathcal{A}_0$ and $\mathbb{E}_{p_0} \Delta\mathcal{A}$ gives

$$\begin{aligned} &\mathbb{E}_q \mathcal{A}_\omega - \mathbb{E}_{p_0} \mathcal{A}_0 \\ &= (\mathbb{E}_q - \mathbb{E}_{p_0}) \mathcal{A}_0 + \mathbb{E}_{p_0} \Delta\mathcal{A} + (\mathbb{E}_q - \mathbb{E}_{p_0}) \Delta\mathcal{A}, \end{aligned} \tag{56}$$

which proves the interaction decomposition. □

The interaction has an interpretable sign. The final term is positive when outer allocation increases exposure to problems on which inner weighting improves alignment, and negative when it emphasizes problems on which weighting harms alignment. If $q \ll p_0$ with density $r(z) = dq/dp_0$, it can also be written

$$(\mathbb{E}_q - \mathbb{E}_{p_0}) \Delta\mathcal{A} = \mathrm{Cov}_{p_0}(r, \Delta\mathcal{A}), \tag{57}$$

since $\mathbb{E}_{p_0} r = 1$. This alternative requires $q \ll p_0$, which is stronger than the reverse domination guaranteed by coverage. The expectation form in the main proposition does not require it.

From alignment to finite-step improvement. For one group with unit base scale, define the realized directions $d_w = \sum_j w_j g_j / \sum_j w_j$ and $d_0 = n^{-1} \sum_j g_j$.

Proposition B.3 (A sufficient condition for useful target weighting)**.** *If J has L-Lipschitz gradient along both update segments, then*

$$\begin{aligned} &J(\theta - \eta d_0) - J(\theta - \eta d_w) \\ &\quad \geq \eta \frac{\mathrm{Cov}(w, \chi)}{\bar{w}} - \frac{L\eta^2}{2}(\|d_w\|^2 + \|d_0\|^2). \end{aligned} \tag{58}$$

For several groups, replace the covariance ratio by $\sum_h b_h \mathrm{Cov}_h(w, \chi)/\bar{w}_h$ and use the full update directions in the remainder.

Proof. Proposition B.2 bounds each deterministic update's remainder in absolute value by $L\eta^2\|d\|^2/2$. Subtract the two expansions and use Eq. equation 55. The worst-case difference of the remainders is bounded by their sum. □

Positive covariance therefore guarantees improvement for sufficiently small positive η, provided the directions are finite. Good probability calibration alone concerns prediction correctness, not the sign of this covariance. Readiness and uncertainty are practical choices to be validated, not universally ordered by the analysis. For a nonlinear internally weighted auxiliary loss, the smoothness bound can use the actual resulting direction, but the linear covariance identity does not automatically apply.

A finite-step bound for the combined curriculum. Let $\widehat{d}_{q,\omega}$ be the sampled full direction under the composed curriculum and $\widehat{d}_{0,0}$ the direction under p_0 and uniform targets. Suppose both have finite second moments and J is L-smooth along their update segments. Denote the three alignment terms in Eq. equation 11 by $G_{\mathrm{out}}, G_{\mathrm{in}}, G_{\mathrm{int}}$. Then

$$\begin{aligned} &\mathbb{E}J(\theta - \eta\widehat{d}_{0,0}) - \mathbb{E}J(\theta - \eta\widehat{d}_{q,\omega}) \\ &\quad \geq \eta(G_{\mathrm{out}} + G_{\mathrm{in}} + G_{\mathrm{int}}) \\ &\qquad - \frac{L\eta^2}{2}\big(\mathbb{E}\|\widehat{d}_{0,0}\|^2 + \mathbb{E}\|\widehat{d}_{q,\omega}\|^2\big). \end{aligned} \tag{59}$$

This follows by applying Proposition B.2 to each policy and substituting the exact alignment decomposition. It connects the hierarchy to a local improvement condition while retaining the effect of stochastic update magnitudes. Outer-only and inner-only gains need not imply a positive composed gain, because their interaction and second moments can change. The practical controller is consequently evaluated with its executed inner rule rather than borrowing a utility ranking from another weighting rule.

B.3.8 Proof of Proposition 3.5

Proof. Condition on the full history $\mathcal{F}_{k-1}$ before the next draw. The mixture gives $q_k(A \mid \mathcal{F}_{k-1}) \geq p$, where $p = \rho p_0(A)$. Let $E_k = \{N_k(A) = 0\}$. Then

$$\begin{aligned} \Pr(E_k) &= \mathbb{E}[\mathbf{1}_{E_{k-1}} \Pr(z_k \notin A \mid \mathcal{F}_{k-1})] \\ &\leq (1-p)\Pr(E_{k-1}). \end{aligned} \tag{60}$$

Induction gives $(1-p)^K$, and $1 - p \leq e^{-p}$ gives the exponential bound. For target allocation, $\sum_{v \in T_h} w_v \leq n_h w_+$ and $w_j \geq w_-$, so $\omega_{hj} \geq w_-/(n_h w_+)$. □

For Eq. equation 34, $u, e \in [0, 1]$ imply the valid bounds

$$w_- = \sigma(-b_0/\tau_w) > 0, \qquad w_+ = 1 + \gamma + \kappa. \tag{61}$$

They can be numerically weak when the gate is very sharp. Numerical underflow or hard thresholding invalidates a claimed positive implemented floor unless an explicit lower weight bound is enforced. No new hard floor is assumed to have been used in the experiments.

More than one visit. The same conditional lower bound gives $\mathbb{E}N_K(A) \geq Kp$ and, for $0 < \delta < 1$,

$$\Pr[N_K(A) \leq (1-\delta)Kp] \leq \exp(-\delta^2 Kp/2). \tag{62}$$

Indeed, for $t > 0$, $\mathbb{E}[e^{-t\mathbf{1}\{z_k \in A\}} \mid \mathcal{F}_{k-1}] \leq 1 - p + pe^{-t}$. Iteration bounds the moment-generating function of $-N_K$ by $(1 - p + pe^{-t})^K$. Markov's inequality and $1 + u \leq e^u$ give

$$\Pr[N_K \leq (1-\delta)Kp] \leq \exp\{Kp[t(1-\delta) + e^{-t} - 1]\}. \tag{63}$$

Choosing $t = -\log(1-\delta)$ and using $\delta + (1-\delta)\log(1-\delta) \geq \delta^2/2$ proves the claim. A union bound applies to finitely many protected regions. Continuous singleton actions usually have zero mass; coverage protects regions or explicit atoms, not every exact floating-point rate. Target-weight positivity preserves a relative contribution only when its group scale is positive; it does not prevent a zero target gradient, gradient cancellation, or downstream forgetting.

B.4 Cross-modal sensitivity and denoising information

Cross-modal sensitivity. For two active streams m, n, let $G_m(\theta, \mathbf{r})$ and $G_n(\theta, \mathbf{r})$ be matched probe losses at a fixed checkpoint. We define

$$\Omega^{m,n}_{\text{cross}} = \frac{\partial G_m}{\partial r_n} - \frac{\partial G_n}{\partial r_m}. \tag{64}$$

Matched examples, target definitions, and coupled mask randomness isolate the cross-effects of removing context. This is a diagnostic of the current model's loss sensitivity under the stated normalization, not an oracle for future learning progress.

Definition and estimation. For two streams m, n, let $G_m(\theta, \mathbf{r})$ and $G_n(\theta, \mathbf{r})$ be losses measured under a fixed probe protocol. For brevity, write $\Omega_{mn} = \Omega^{m,n}_{\text{cross}}$ from Eq. equation 64:

$$\Omega_{mn}(\theta, \mathbf{r}) = \partial_{r_n} G_m(\theta, \mathbf{r}) - \partial_{r_m} G_n(\theta, \mathbf{r}). \tag{65}$$

Estimate each derivative by matched centered differences. Use common examples, model state, target-loss definition, and coupled mask randomness. For independent erasures, common uniforms U_j and masks $\mathbf{1}\{U_j < r_{m(j)}\}$ provide nested contexts. Changing a context rate need not change the evaluated target units: a target-forced probe can keep them masked. Multiplying one modality's loss by a constant changes Ω; normalization and stream lengths are therefore part of the diagnostic. No sign of Ω is assumed to identify a beneficial training intervention.

A precise matched-probe result. To separate information effects from model and normalization effects, we analyze an idealized matched probe. Conditional mutual information yields an exact reciprocity relation for Bayes prediction. Let all variables be discrete with finite entropies, and let masking be independent of the data and independent across positions. For a target $i \in \mathcal{I}_m$, force i to be hidden and let V_i be the visible subset of all other positions. Define the Bayes probe

$$G^*_m(\mathbf{r}) = \frac{1}{n_m} \sum_{i \in \mathcal{I}_m} \mathbb{E}_{V_i} H(X_i \mid X_{V_i}), \quad n_m = |\mathcal{I}_m|. \tag{66}$$

This is a full-stream average of target-forced token losses, not a mean normalized by a random realized number of masked targets.

Proposition B.4 (Bayes cross-modal reciprocity for independent erasures)**.** *For distinct modalities m, n at interior rates,*

$$\begin{aligned} n_m \partial_{r_n} G^*_m &= \sum_{i \in \mathcal{I}_m} \sum_{j \in \mathcal{I}_n} \mathbb{E}_{V_{ij}} I(X_i; X_j \mid X_{V_{ij}}) \\ &= n_n \partial_{r_m} G^*_n \geq 0, \end{aligned} \tag{67}$$

where V_{ij} contains visible positions other than i, j.

Proof. Give each context position its own erasure probability. Conditional on the visible positions V_{ij} other than i, j, the contribution for target i as a function of the erasure probability r_j is

$$r_j H(X_i \mid X_{V_{ij}}) + (1 - r_j) H(X_i \mid X_{V_{ij}}, X_j). \tag{68}$$

Its derivative is $I(X_i; X_j \mid X_{V_{ij}})$. The expectation over the remaining masks is a finite multilinear polynomial, so derivatives and expectations can be interchanged. Setting all probabilities in modality n equal to r_n and applying the chain rule sums this derivative over $j \in \mathcal{I}_n$. Summing over $i \in \mathcal{I}_m$ establishes the first equality. Exchanging i and j preserves the distribution of V_{ij}, and conditional mutual information is symmetric and nonnegative, proving the rest. □

For the corresponding model probe,

$$G_m(\theta, \mathbf{r}) = G_m^*(\mathbf{r}) + \mathcal{E}_m(\theta, \mathbf{r}), \tag{69}$$

where $\mathcal{E}_m \geq 0$ is the average conditional KL divergence between the true token conditional and the predictor. Whenever derivatives exist, the size-adjusted antisymmetry is

$$\begin{aligned} \widetilde{\Omega}_{mn} &= n_m \partial_{r_n} G_m - n_n \partial_{r_m} G_n \\ &= n_m \partial_{r_n} \mathcal{E}_m - n_n \partial_{r_m} \mathcal{E}_n. \end{aligned} \tag{70}$$

Thus, under this probe, size-adjusted asymmetry reflects excess-risk sensitivities, not asymmetric mutual information. Unadjusted Ω can be nonzero even at Bayes optimality if $n_m \neq n_n$. Arbitrary modality scales, random masked-count denominators, structured or data-dependent masks, continuous losses, and auxiliary losses need not obey the proposition. The interpretation of Ω is therefore tied to its probe definition; an unnormalized scalar difference does not identify an intrinsic direction of biological information flow. The result also does not negate the utility of asymmetric routes: reciprocal information effects do not imply equal training gradients, optimization speeds, or evaluation weights.

A local off-diagonal criterion is a different statement. Let $V_\theta(\mathbf{r})$ be a differentiable scalar probe risk, and let $\mathbf{r}_0 = t\mathbf{1}$ lie in the interior of a fixed route's rate domain. For a feasible transverse direction d with $d^\top \mathbf{1} = 0$, if $\nabla_{\mathbf{r}} V_\theta(\mathbf{r}_0)^\top d < 0$, then

$$V_\theta(\mathbf{r}_0 + hd) < V_\theta(\mathbf{r}_0) \quad \text{for all sufficiently small } h > 0. \tag{71}$$

To prove this, expand as $V_\theta(\mathbf{r}_0) + h\nabla V_\theta(\mathbf{r}_0)^\top d + o(h)$; the negative linear term dominates the remainder. This is a statement about changing the *current probe problem*, not the future model produced by training on that problem. Lower current loss can merely indicate more visible information. Equation equation 32, rather than this derivative or Ω, defines training utility.

B.5 Instantiations and training procedure

Exposure under bounded example priorities. For N training examples, normalized priorities h_n with $0 < h_- \leq h_n \leq h_+$ satisfy

$$q(n) = \frac{h_n}{\sum_v h_v} \geq \frac{h_-}{h_+} \frac{1}{N}. \tag{72}$$

Capping priorities only from above does not yield this positive lower bound. An explicit reference mixture provides one when the priorities can vanish. The exact priority mapping and cap are implementation parameters, distinct from this lower-bound condition.

Algorithm 1 uses a global mixture for clarity. Categorical factors are refreshed from their probe statistics, whereas score-trained factors use the estimator matched to their sampling law. The two updates need not both act on the same factor. At delayed controller updates, stored actions retain their generating policy and pre-action baseline; changing the policy before evaluating a score requires the corresponding correction.

Optimization level. The denoiser receives gradients of supervised target losses. Policy-score updates apply only to the curriculum. Readiness weighting, continuous-vector heads, and auxiliary mask losses do not imply a likelihood bound for all outputs, and the local theory does not establish global convergence or performance preservation. Its purpose is to identify what each part of the allocation policy controls and the conditions under which that control can improve training.

C Experimental Details

C.1 Dataset Details

C.1.1 JUMP

We use the compound subset of JUMP Cell Painting distributed with PhenoCompass (Chandrasekaran et al., 2023; Wu et al., 2026). A packed record contains a SMILES string, a morphology vector, a molecular-structure vector, and a binary BH-corrected bioactivity label. RAW experiments use 1,536-dimensional DINO morphology (Caron et al., 2021) and 3,000-dimensional GNEprop structure (Scalia et al., 2026); co-embedding experiments use the fixed 256-dimensional PhenoCompass `combined_fp` views. After removing the DMSO control, the released scaffold-cluster split contains 96,393 training, 5,497 validation, and 6,945 test compounds.

The evaluation contains SMILES→morphology, structure→morphology, and morphology→compound retrieval, each scored by Hits@10 against 100 candidates; structure-based morphological bioactivity prediction, scored by standardized partial AUROC at FPR 0.1; and morphology-guided MoA virtual screening, scored by the equally weighted partial AUROC across six mechanism families. The first four endpoints use the 6,945 test compounds. The morphology-guided MoA screen ranks held-out actives against 95,573 background compounds using measured morphology anchors and SMILES-conditioned predictions.

C.1.2 BioMed

BioMed uses 14 sources from BiomedParseData (Zhao et al., 2025), covering pathology, chest radiography, dermoscopy, endoscopy, fundus imaging, mammography, and ultrasound. Each record contains an image, a harmonized text prompt, and a binary segmentation mask. Native masks are area-reduced to a 16×16 occupancy grid; a cell is foreground when foreground covers at least 25% of its native pixels.

BioMed-Large contains 100,268 records and uses two alternative, task-specific partitions of that same population. For IID training and evaluation, 80,213 records fit the IID adapter, a disjoint 3,799-record calibration set selects checkpoints, thresholds, and decode steps, and 16,256 records form the IID test. For source-shift evaluation, the population is repartitioned by source: the source-shift adapter trains on 94,998 records, while all 5,270 examples from PolypGen, GlaS, BreastUS, and REFUGE form the held-out test. These two sets of counts are not five mutually disjoint subsets to be added together; non-held-out records assigned to IID calibration or testing may belong to the source-shift training partition. No record from the four held-out sources enters source-shift fitting or IID calibration, and the IID and source-shift results use separately trained adapters. The modality analysis maps the sources to dermoscopy, endoscopy, fundus, mammography, pathology, ultrasound, and X-ray groups and applies one calibration-selected threshold to every subgroup.

C.1.3 HuBMAP

We use all 880 labeled images in the post-competition 2022 HuBMAP+HPA Hacking the Human Body release. HPA provides 306 training images from 31 donors, 61 calibration images from six donors, and 65 IID-test images from nine donors; source-image and donor assignments are disjoint across these partitions. All 448 HuBMAP images form the source-held-out test. The five prompts identify the functional tissue unit for kidney, large intestine, lung, prostate, or spleen.

Images are divided into non-overlapping 512×512 tiles and stitched to native resolution for evaluation. The resulting train/calibration/IID-test/ source-held-out manifests contain 7,911/2,185/2,340/7,688 tiles. Training keeps all tissue-qualified tiles, including 3,727 negative tiles; calibration and test retain complete spatial coverage. Each model predicts a 16×16 occupancy grid, which is expanded, cropped at padded image boundaries, and stitched before computing source-image metrics.

C.2 Setup Details

JUMP models and baselines. LLaDA2.0-Uni uses its native tokenizer, a 96-token SMILES budget, rank-32 LoRA (Hu et al., 2021), and a continuous input/output module that maps each vector

to four transformer slots. Its uniform and MaDAC arms train for 27k updates. UniDisc uses a 54-symbol SMILES vocabulary and the same four-slot continuous interface; the entire scratch model trains for 36k updates. The main comparison uses the PhenoCompass co-embedding so PhenoCompass and our model share the target representation. GNEprop/Morgan kNN and linear predictors and a matched ridge regressor provide classical references. Frozen LLaDA2, uniform corruption, and easy-to-hard corruption provide internal controls. The checkpoint selection uses the validation mean of the three retrieval endpoints and structure-based morphological bioactivity prediction.

BioMed models and baselines. LLaDA2.0-Uni receives a spatially aligned image-token sequence and predicts dedicated foreground/background mask tokens. We adapt it with rank-32 LoRA and a trainable mask-token head. UniDisc places 256 image tokens and the 256-cell binary answer in a 512-token block; its 147M transformer receives full parameter updates from random initialization. Frozen, fixed-uniform, and easy-to-hard LLaDA2 arms share the task interface, while the UniDisc comparison uses fixed-uniform and adaptive arms.

BiomedParse, SAM ViT-B, and MedSAM ViT-B are trained on the corresponding BioMed split. At evaluation, Grounding DINO maps each image and text prompt to the box consumed by SAM or MedSAM. The released BiomedParse checkpoint and zero-shot CLIPSeg (Lüddecke & Ecker, 2022) appears in Table 7 as a pretrained reference. Checkpoints and one foreground threshold are selected on the IID calibration set and then fixed for both test sets.

BiomedParse optimization. The matched-data specialist uses the official `microsoft/BiomedParse` code (commit `750eb40`) through its own `entry.py train` entry point and `biomed_seg_lang` recipe. We sweep learning rates 10^{-4}, 2×10^{-5}, and 10^{-5} with gradient-clip values 5.0 and 1.0. Rare batches trigger abrupt `loss_mask_bce` and gradient-norm growth across these settings. A guard that skips updates whose pre-clip norm exceeds $12\times$ the running median yields the selected checkpoint. Table 2 reports that validation-selected checkpoint.

HuBMAP models and baselines. LLaDA2.0-Uni processes all 1,024 visual tokens from its 32×32 grid. Hidden states from layers 4, 12, and 20 feed a residual spatial decoder that emits 16×16 logits. The decoder trains at 3×10^{-4}, while rank-8 LoRA in transformer layers 18–19 trains at 2×10^{-5}. The UniDisc model contains 146,833,409 transformer parameters in 12 width-768, 12-head blocks. Its pretrained MagViT tokenizer (Yu et al., 2023) remains fixed; hidden states from blocks 4, 8, and 12 feed a four-block spatial decoder. The scratch backbone and decoder train at 10^{-4} and 3×10^{-4}, respectively.

All backbone arms use 12k-update cosine schedules, an effective batch size of eight, identical data and spatial losses, and checkpoints every 1,000 updates. The uniform and MaDAC arms update the same parameters. SegFormer MiT-B4, DeepLabV3/ResNet-50, and U-Net/ResNet-34 are trained as dense segmentation references. Checkpoints and the global foreground threshold are selected by IID-calibration mean native-image Dice and then fixed for both tests.

C.3 Policy and reporting conventions

All task-level comparisons use the splits, checkpoint-selection rules, and metrics declared for that benchmark. Internal frozen/uniform/MaDAC rows within each backbone family use the same data and task interface. External systems provide endpoint references under the protocols stated above.

JUMP. Routes determine which streams are visible, hidden, and supervised. Continuous-output readiness uses a detached cosine-agreement score with sample weighting normalized across the relevant minibatch. Retrieval, morphological bioactivity prediction, and morphology-guided MoA virtual screening are scored independently from the training regression loss.

BioMed. Image, text prompt, and occupancy-mask streams define the multimodal record. The IID calibration split determines the checkpoint and foreground threshold.

HuBMAP. The initial decoder-only phase stabilizes the spatial head before the adaptive sampler and readiness weights activate. The IID calibration split determines the checkpoint and foreground threshold.

Table 5: JUMP matched-policy controls across RAW and PhenoCompass co-embedding representations. Retrieval columns are 100-way Hits@10; structure-based morphological bioactivity prediction and morphology-guided MoA virtual screening are standardized pAUROC@0.1. RAW and co-embedding are distinct target representations and should be interpreted separately. Values are mean ± SD; higher is better.

Backbone / representation	Policy	SMILES→ Morphology	Structure→ Morphology	Morphology→ Compound	Morphological bioactivity	Morphology-guided MoA
LLaDA2 / RAW	frozen	.152 ± .004	.165 ± .005	.192 ± .005	.540 ± .004	.505 ± .010
LLaDA2 / RAW	uniform	.321 ± .006	.246 ± .005	.289 ± .005	.621 ± .006	.523 ± .013
LLaDA2 / RAW	easy-to-hard	.315 ± .006	.224 ± .005	.272 ± .005	.618 ± .006	.524 ± .012
LLaDA2 / RAW	MaDAC	.322 ± .006	**.271 ± .005**	**.298 ± .006**	**.651 ± .007**	**.527 ± .014**
UniDisc / RAW	uniform	.354 ± .002	.345 ± .003	.392 ± .005	.628 ± .001	.544 ± .010
UniDisc / RAW	easy-to-hard	.316 ± .006	.330 ± .006	.372 ± .006	.619 ± .006	.510 ± .013
UniDisc / RAW	MaDAC	**.372 ± .006**	**.376 ± .006**	**.401 ± .006**	**.633 ± .006**	**.563 ± .017**
LLaDA2 / co-emb.	uniform	**.535 ± .005**	.596 ± .005	**.637 ± .006**	.676 ± .007	.746 ± .024
LLaDA2 / co-emb.	MaDAC	.529 ± .006	**.615 ± .006**	.636 ± .006	.676 ± .007	**.747 ± .024**
UniDisc / co-emb.	uniform	.486 ± .006	.606 ± .006	.632 ± .006	.683 ± .007	.755 ± .023
UniDisc / co-emb.	MaDAC	**.564 ± .006**	**.623 ± .006**	**.642 ± .006**	**.684 ± .007**	**.757 ± .024**

C.4 Compute resources

All experiments were conducted on shared high-performance computing systems equipped with NVIDIA B200 and H200 GPUs. Within each controlled comparison, the baseline and proposed method used the same accelerator type, training budget, and resource configuration. The principal LLaDA2 and UniDisc environment used Python 3.12.3, PyTorch 2.8.0 with CUDA 12.8 (`torch==2.8.0+cu128`), torchvision 0.23.0, Transformers 5.9.0, PEFT 0.19.1, Accelerate 1.13.0, NumPy 1.26.4, and WebDataset 1.0.2; external baselines with incompatible upstream requirements used their corresponding task-specific environments.

D Additional Experimental Results

This appendix supplements the main results with aggregate visualizations, expanded metric and baseline comparisons, curriculum-space and component analyses, scaling and capability studies, general-domain evaluations, computational-efficiency results, qualitative galleries, and reproducibility details. Each topic is introduced in its corresponding subsection below.

D.1 JUMP

Figure 5 provides a task-by-task view of the aggregate results in Table 1. Its first three panels show 100-way cross-modal retrieval, while the last two show morphological bioactivity prediction and morphology-guided MoA screening. The RAW and co-embedding groups are separated because they predict different target representations. Within each group, the paired bars isolate the curriculum effect under a fixed backbone and representation. The largest visual gains occur for UniDisc-based MaDAC retrieval; the comparison between LLaDA2-based MaDAC and uniform training is more endpoint dependent.

Matched internal controls and RAW features. Table 5 extends the main co-embedding comparison with the representation-independent RAW setting and the complete matched policy controls available in the supplied JUMP runs. The co-embedding rows use the fixed PhenoCompass representation; RAW rows use independently derived DINO/GNEprop features.

The RAW comparison shows that curriculum gains are not contingent on the PhenoCompass alignment: UniDisc-based MaDAC training improves every reported endpoint over its uniform counterpart, while LLaDA2-based MaDAC training improves the cross-modal retrieval, morphological bioactivity, and morphology-guided MoA endpoints over LLaDA2-based uniform training.

Table 6 complements the aggregate MoA endpoint with a mechanism-level view. The leading method varies across mechanism families, while the balanced column summarizes performance without allowing a single family to determine the comparison.

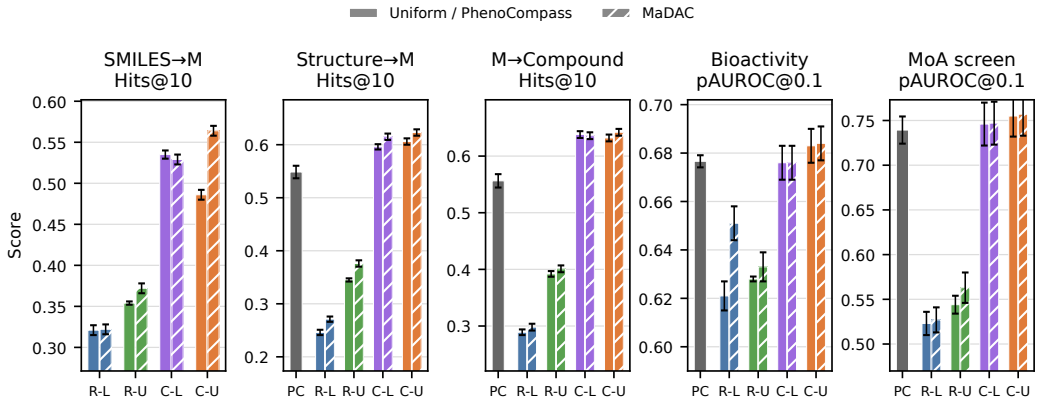

Figure 5: JUMP aggregate result visualization corresponding to Table 1. The panels compare PhenoCompass with uniform and MaDAC training built on LLaDA2 and UniDisc in the RAW-feature and PhenoCompass co-embedding settings. Matched training variants are paired within each backbone–representation group. Axes are task-specific and do not start at zero; error bars show ± SD.

Table 6: Morphology-guided MoA virtual screening by mechanism. Entries are pAUROC@0.1; Balanced is the equal-weight mean across the six MoA groups shown in Figure 2. Higher is better.

Method / representation	PI3K/mTOR	HSP90	JAK/ROCK	HDAC	MEK/MAPK	CDK	Balanced
PhenoCompass	**.941 ± .006**	.672 ± .055	.755 ± .038	.603 ± .042	**.568 ± .029**	**.898 ± .009**	.739 ± .015
GNEprop + kNN	.474	.532	.565	.506	.507	.538	.520 ± .010
Morgan + kNN	.503	.505	.502	.515	.476	.542	.507 ± .008
GNEprop + logistic reg.	.679	.507	.488	.483	.494	.525	.529 ± .003
Morgan + logistic reg.	.611	.565	.474	.509	.567	.474	.533 ± .001
GNEprop + linear reg.	.626	.474	.543	.617	.535	.607	.567 ± .019
Morgan + linear reg.	.642	.474	.608	.506	.561	.564	.559 ± .009
MaDAC (LLaDA2, co-emb.)	.896 ± .051	.744 ± .082	.773 ± .075	.648 ± .048	.554 ± .052	.869 ± .046	.747 ± .024
MaDAC (UniDisc, co-emb.)	.892 ± .050	**.766 ± .088**	**.790 ± .071**	**.680 ± .047**	.555 ± .052	.862 ± .045	**.757 ± .024**

D.2 BioMed

Figure 6 shows both overlap quality (Dice) and threshold-free discrimination (AUROC) under IID and source shift. The pattern is consistent across all four panels: LLaDA2-based MaDAC training improves over LLaDA2-based uniform training, with the separation largest for source-held-out Dice.

Extended baseline context. The main table uses same-data specialists to keep the comparison interpretable. Table 7 restores released checkpoints and the zero-shot CLIPSeg reference. The released BiomedParse checkpoint is trained on substantially broader medical data and is therefore external context, not a matched-data comparator.

Table 7 shows the distinction between matched-data comparisons and external pretrained references. Among the matched LLaDA2-based training variants, MaDAC leads the aggregate IID and source-held-out columns; the released BiomedParse model provides broader-pretraining context.

Table 8 shows that the modality-level Dice pattern is more heterogeneous than the aggregate results. LLaDA2-based MaDAC training consistently improves over its matched uniform counterpart.

D.3 HuBMAP

Complete aggregate results. Table 9 reports the complete aggregate comparison on both evaluation splits. The source-held-out columns reproduce the primary comparison, while the IID columns show the complementary in-distribution behavior: SegFormer leads the IID Dice measures, whereas MaDAC is the strongest policy within each diffusion-model backbone family.

Table 9 also reports the empty-image rate, which distinguishes low overlap caused by inaccurate masks from complete prediction failure. On the source-held-out split, LLaDA2-based MaDAC

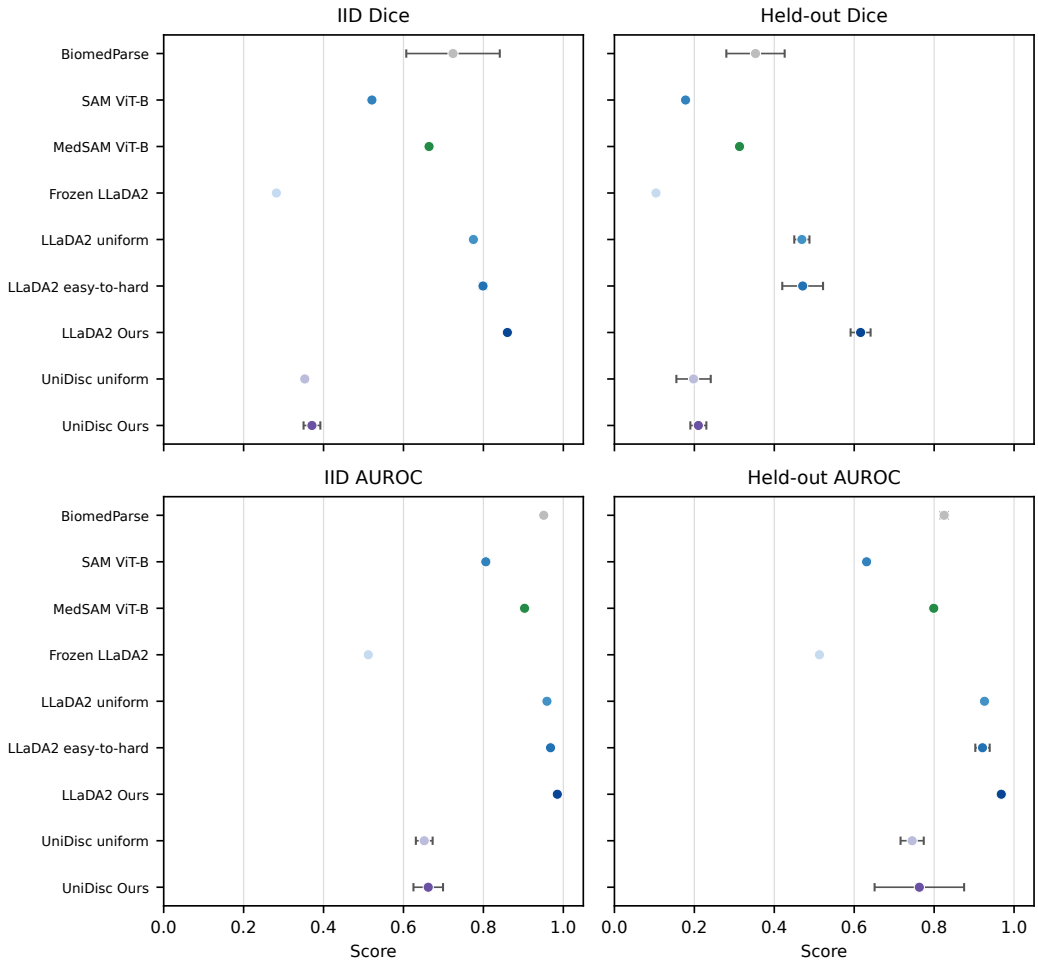

Figure 6: BioMed-Large aggregate result visualization corresponding to Table 2. The panels compare Dice and AUROC on the IID and source-held-out tests across the principal curriculum controls and segmentation baselines.

Table 7: Extended BioMed-Large baseline comparison on IID and source-held-out tests. Dice and AUROC are higher-is-better. Released checkpoints use different pretraining data and are included as external context rather than matched-data baselines.

	IID		Source-held-out	
Method	Dice	AUROC	Dice	AUROC
BiomedParse released	.900 ± .062	.995 ± .000	.891 ± .025	.987 ± .004
BiomedParse same data	.724 ± .117	.951 ± .001	.353 ± .073	.825 ± .009
CLIPSeg zero-shot	.269 ± .001	.525 ± .002	.145 ± .002	.600 ± .004
SAM released	.289 ± .001	.514 ± .002	.206 ± .003	.717 ± .003
MedSAM released	.345 ± .002	.626 ± .002	.155 ± .002	.538 ± .003
SAM same data	.521 ± .002	.806 ± .002	.178 ± .002	.631 ± .004
MedSAM same data	.664 ± .002	.903 ± .001	.313 ± .004	.799 ± .004
Frozen LLaDA2	.282 ± .001	.512 ± .000	.104 ± .002	.513 ± .002
Uniform (LLaDA2)	.775 ± .006	.959 ± .002	.469 ± .019	.926 ± .003
Easy-to-Hard (LLaDA2)	.799 ± .008	.968 ± .003	.471 ± .051	.921 ± .018
MaDAC (LLaDA2)	.860 ± .002	.985 ± .001	.616 ± .025	.968 ± .003
Uniform (UniDisc)	.353 ± .004	.652 ± .021	.198 ± .043	.745 ± .029
MaDAC (UniDisc)	.371 ± .021	.662 ± .037	.210 ± .020	.763 ± .112

Table 8: BioMed Dice by imaging modality. Values are mean ± SD. The LLaDA2-based Frozen, Uniform, Easy-to-Hard, and MaDAC variants are matched curriculum controls.

			IID test				
Method	Derm.	Endo.	Fundus	Mammo.	Path.	US	X-Ray
BiomedParse (same data)	.649 ± .008	.162 ± .008	.336 ± .008	.241 ± .018	.432 ± .005	.229 ± .013	.857 ± .002
SAM (same data)	.678 ± .014	.216 ± .009	.041 ± .001	.010 ± .003	.505 ± .006	.219 ± .016	.572 ± .002
MedSAM (same data)	.668 ± .014	.236 ± .010	.072 ± .002	.393 ± .021	.550 ± .007	.282 ± .016	.749 ± .002
Frozen LLaDA2	.067 ± .014	.087 ± .016	.035 ± .025	.000 ± .000	.038 ± .005	.009 ± .008	.189 ± .005
Uniform (LLaDA2)	.693 ± .030	.462 ± .058	.568 ± .028	.198 ± .067	.247 ± .022	.398 ± .102	.803 ± .008
Easy-to-Hard (LLaDA2)	.676 ± .039	.380 ± .068	.634 ± .031	.118 ± .040	.211 ± .020	.457 ± .085	.806 ± .008
MaDAC (LLaDA2)	.814 ± .025	.512 ± .065	.767 ± .033	.359 ± .071	.256 ± .018	.566 ± .104	.885 ± .006
			Source-held-out test				
Method	Derm.	Endo.	Fundus	Mammo.	Path.	US	X-Ray
BiomedParse (same data)	–	.179 ± .013	.433 ± .010	–	.072 ± .014	.351 ± .015	–
SAM (same data)	–	.304 ± .008	.052 ± .001	–	.722 ± .011	.198 ± .005	–
MedSAM (same data)	–	.317 ± .008	.230 ± .003	–	.696 ± .013	.258 ± .007	–
Frozen LLaDA2	–	.082 ± .006	.023 ± .006	–	.038 ± .008	.003 ± .001	–
Uniform (LLaDA2)	–	.166 ± .017	.563 ± .010	–	.013 ± .010	.197 ± .021	–
Easy-to-Hard (LLaDA2)	–	.181 ± .017	.567 ± .011	–	.008 ± .005	.114 ± .015	–
MaDAC (LLaDA2)	–	.508 ± .022	.701 ± .009	–	.722 ± .030	.498 ± .022	–

Table 9: HuBMAP organ-conditioned segmentation results. We report mean stitched-image Dice, macro-organ Dice, and empty-image rate on IID and source-held-out tests. Dice is higher-is-better and empty-image rate is lower-is-better. Values are mean ± SD; bold marks the best-performing policy within each backbone family.

	IID			Source-held-out		
Method	Image	Macro	Empty (%; ↓)	Image	Macro	Empty (%; ↓)
SegFormer MiT-B4	.755 ± .005	.702 ± .006	.000 ± .000	.415 ± .050	.418 ± .037	21.540 ± 1.420
DeepLabV3 ResNet-50	.611 ± .018	.548 ± .019	.000 ± .000	.151 ± .007	.152 ± .008	24.550 ± 1.310
U-Net ResNet-34	.596 ± .024	.553 ± .021	.000 ± .000	.235 ± .007	.251 ± .008	15.400 ± 1.070
Frozen LLaDA2	.557 ± .012	.511 ± .005	.770 ± 1.090	.396 ± .003	.390 ± .000	20.310 ± 6.940
Uniform (LLaDA2)	.575 ± .015	.528 ± .016	**.000 ± .000**	.387 ± .004	.382 ± .012	8.370 ± 3.000
MaDAC (LLaDA2)	**.578 ± .003**	**.531 ± .002**	**.000 ± .000**	**.431 ± .039**	**.418 ± .042**	**.450 ± .630**
Frozen UniDisc	.187 ± .005	.183 ± .008	41.540 ± .000	.173 ± .083	.213 ± .077	54.350 ± 15.310
Uniform (UniDisc)	.278 ± .013	.274 ± .008	6.150 ± 4.350	.337 ± .005	.368 ± .006	14.060 ± 6.000
MaDAC (UniDisc)	**.289 ± .018**	**.284 ± .020**	**1.540 ± 2.180**	**.345 ± .002**	**.375 ± .003**	**11.500 ± 1.100**

training raises both image-averaged and organ-balanced Dice over LLaDA2-based uniform training while sharply reducing empty predictions. UniDisc-based MaDAC training gives smaller but consistent aggregate gains over its uniform counterpart. The dense specialists remain strong IID references, but their source-held-out results show that high in-distribution Dice does not imply robust transfer.

Organ-level behavior. Table 10 expands the main image- and organ-balanced metrics. It shows that the aggregate gain is not uniform across organs and that the source shift changes which organs dominate the error profile. Each cell is computed separately within one evaluation split and one organ. Because the organs contribute different numbers of images, their simple arithmetic mean does not recover the image-averaged Dice in Table 9; that aggregate weights images according to their empirical frequencies. The unweighted mean of the five organ means is instead the macro-organ Dice reported there.

Table 10 localizes the aggregate behavior: LLaDA2-based MaDAC training improves over LLaDA2-based uniform training most clearly for selected organs, while other organs change little or decline. UniDisc-based MaDAC training likewise has a mixed organ-level pattern relative to its uniform counterpart, clarifying why the aggregate and organ-balanced summaries should be read together.

Table 10: HuBMAP stitched-image Dice computed separately for each organ within the IID and source-held-out splits. The table expands the aggregate main result for the principal specialist and backbone-specific uniform/MaDAC training comparisons. Organ counts differ, so an unweighted mean of these cells gives macro-organ Dice, not the image-averaged Dice in Table 9.

Method	Kidney	Intestine	Lung	Prostate	Spleen
IID test					
SegFormer	$.916 \pm .003$	$.848 \pm .006$	$.181 \pm .006$	$.850 \pm .004$	$.715 \pm .012$
Uniform (LLaDA2)	$.622 \pm .040$	$.794 \pm .013$	$.058 \pm .022$	$.788 \pm .029$	$.451 \pm .058$
MaDAC (LLaDA2)	$.630 \pm .023$	$.781 \pm .012$	$.164 \pm .029$	$.780 \pm .030$	$.428 \pm .047$
Uniform (UniDisc)	$.068 \pm .018$	$.499 \pm .041$	$.054 \pm .014$	$.456 \pm .036$	$.259 \pm .036$
MaDAC (UniDisc)	$.113 \pm .025$	$.541 \pm .037$	$.046 \pm .019$	$.463 \pm .035$	$.299 \pm .038$
Source-held-out test					
SegFormer	$.722 \pm .070$	$.372 \pm .057$	$.361 \pm .084$	$.021 \pm .009$	$.613 \pm .078$
Uniform (LLaDA2)	$.492 \pm .024$	$.159 \pm .028$	$.090 \pm .011$	$.628 \pm .017$	$.492 \pm .015$
MaDAC (LLaDA2)	$.616 \pm .016$	$.163 \pm .028$	$.338 \pm .011$	$.358 \pm .027$	$.480 \pm .014$
Uniform (UniDisc)	$.209 \pm .005$	$.590 \pm .016$	$.109 \pm .013$	$.641 \pm .016$	$.365 \pm .014$
MaDAC (UniDisc)	$.227 \pm .006$	$.567 \pm .015$	$.095 \pm .011$	$.642 \pm .016$	$.374 \pm .014$

D.4 Two-dimensional curriculum-space analysis

Common geometry. Figure 4 contrasts a fixed point, coupled random and linear diagonals, and adaptive two-dimensional supports. Only the adaptive policies move off the diagonal and redirect their temporal centers. BioMed coordinates control image-mask and prompt corruption; JUMP coordinates allocate effort to morphology and structure. We therefore compare geometry and direction, not coordinate values. JUMP actions are requested actions retained from the training logs. The historical BioMed runs did not retain per-step requested or executed actions, so we replay the declared 6,000-step rate laws and use fixed-seed representative draws from the specified distributions. We call these reconstructions *schedule-exact* because they preserve the schedule's phases, rate laws, and two-dimensional support; they do not recover the original random-number states, minibatch sequence, or an executed feedback trace. Accordingly, the BioMed panels support claims about the policy's available geometry and intended temporal organization, not the exact empirical frequency or sequence of actions in the historical run. Their linked markers average 200 representative draws per window, whereas the JUMP markers average 200 logged steps. Thus, "realized training actions" in Figure 4 denotes actions instantiated from the declared BioMed schedule, not actions recovered from a retained BioMed execution log; the JUMP points, by contrast, come from retained requested-action logs. Geometry shows use of independent controls; endpoint results establish their benefit.

Full support and temporal movement. Figure 7 separates the complete action support from the narrower path of temporal-window centers. BioMed combines broad support with a narrower, direction-changing center path; JUMP combines compact support with a more directed allocation shift. Thus, coverage and temporal redirection describe distinct curriculum functions.

BioMed Joint-only effective rates. BioMed first selects base image-mask and prompt corruption rates and then chooses the Seg, Rec, or Joint prediction pattern. Seg activates only the image-mask rate, Rec activates only the prompt rate, and Joint activates both. Consequently, a zero on an inactive axis in a mixed-pattern view does not necessarily represent a rate chosen by the policy. Restricting the analysis to Joint examples removes this ambiguity: both streams are active, so the two plotted coordinates are the effective rates actually applied to the example.

Figure 8 shows that coupled-random training remains one-dimensional even after conditioning on Joint examples: one sampled rate determines both active streams. In contrast, the LLaDA2-based MaDAC curriculum occupies a two-dimensional region and its windowed center changes direction over training. Thus, the off-diagonal structure in the main BioMed visualization is not an artifact of mixing mask-only Seg examples with prompt-only Rec examples; it persists when mask and prompt prediction are simultaneously active. Because the historical BioMed actions were not retained as an execution log, the plotted actions are schedule-exact representative reconstructions rather than recovered logged samples. As elsewhere, this geometry demonstrates independent control and temporal redirection, while the matched endpoint results establish whether that mechanism improves performance.

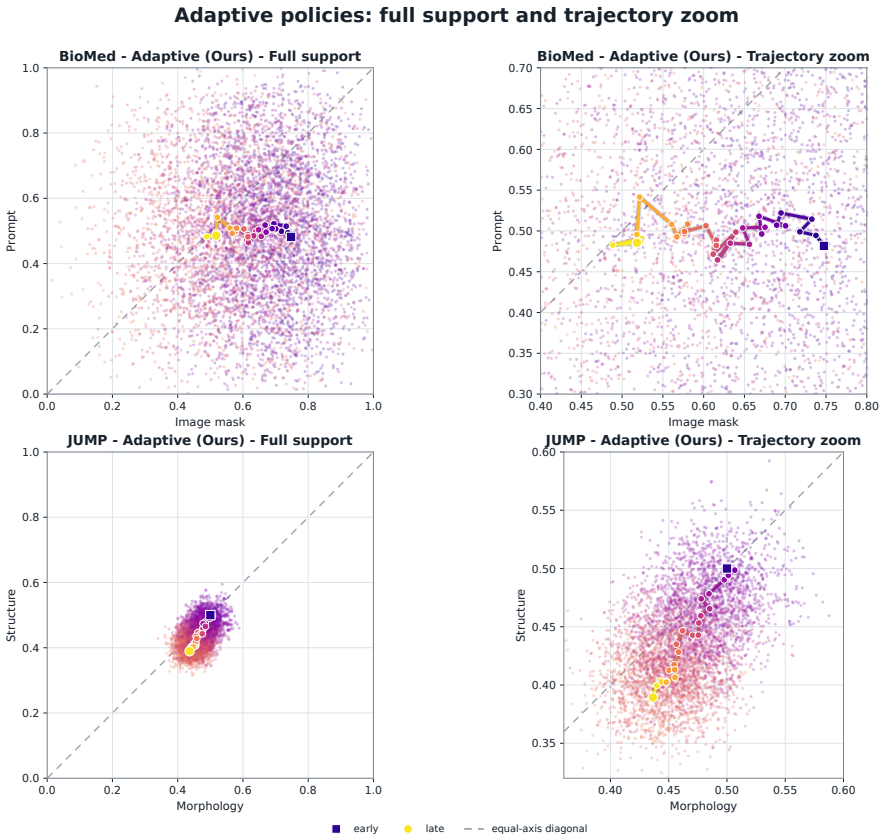

Figure 7: Adaptive-policy support and trajectory zoom. Full-domain panels (left) show requested-action support; zooms (right) show temporal-window centers. BioMed covers broad mask–prompt combinations, whereas JUMP occupies a compact region while shifting morphology–structure allocation. The dashed line marks equal coordinates; color progresses from early (purple) to late (yellow).

Table 11: Task challenges, measurements, and curriculum diagnostics.

Challenge	BioMed measure	JUMP measure	Figure diagnostic	Implication
Target heterogeneity	Mask size, topology, boundaries, prompt diversity	Morphology/structure diversity by retrieval stratum	Full support shows action range	Heterogeneity motivates coverage, not a broad path
Cross-modal dependence	Prompt drop/shuffle; image-only and mask-prior controls	Route ablations; cross-modal retrieval errors	Joint-only rates isolate simultaneous prediction	Separate controls address unequal information sources
Learning asymmetry	Stream loss and segmentation under controlled rates	Target loss, readiness, morphology–structure imbalance	Zooms show off-diagonal redirection	Changing imbalance favors independent control
Temporal response	Phase support, direction changes, checkpoint errors	Response lag, allocation movement, route errors	Linked centers separate motion from support	Useful motion tracks a changing bottleneck
Generalization	Held-out results by mask/prompt properties	Held-out retrieval by compound, scaffold, route	Geometry shows mechanism, not outcome	Held-out endpoints test benefit

Connecting task challenge to curriculum. Table 11 aligns each challenge with a task-specific measure, figure diagnostic, and curriculum implication.

Matched longitudinal cases connect policy movement to task bottlenecks: BioMed holds image and mask fixed while varying prompt information, and JUMP follows the same compounds across direct and cross-modal retrieval. Held-out endpoints then determine whether the observed redirection is useful.

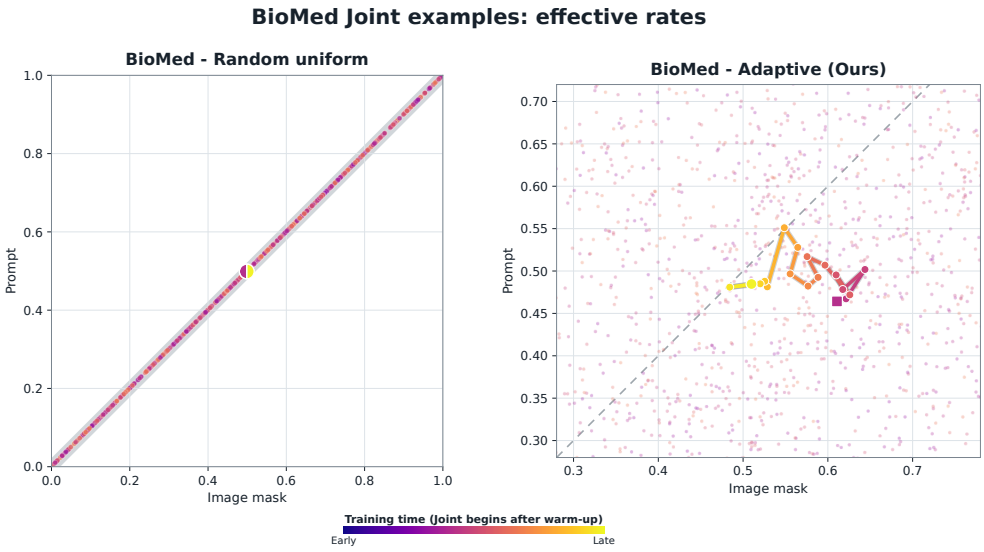

Figure 8: Effective image-mask and prompt rates for BioMed Joint examples. Faint points are optimizer-step means for steps containing Joint examples, and linked markers summarize successive training windows from early (purple) to late (yellow). Coupled random applies the same sampled rate to both streams and therefore remains on the equal-rate diagonal with a stationary temporal center. MaDAC varies the two effective rates independently and redirects its temporal center through off-diagonal regions. Joint is inactive during the shared Seg-only warm-up, so the displayed actions begin after that phase.

Table 12: JUMP policy-factor ablation for UniDisc-based co-embedding training. Columns report 100-way Hits@10 for SMILES→morphology, structure→morphology, and morphology→compound retrieval; higher is better.

Policy intervention	SMILES→M	Str.→M	M→Comp.
Fixed allocation	.486 ± .006	.606 ± .006	.632 ± .006
Outer allocation only	.490 ± .006	.572 ± .006	.621 ± .006
Route-decoupled hierarchy (base)	.527 ± .006	.618 ± .006	.634 ± .006
MaDAC without focal objective	.543 ± .006	.619 ± .006	.637 ± .006
Undifferentiated hierarchy	.555 ± .006	.619 ± .006	.634 ± .006
Route-decoupled, no inner readiness	.532 ± .006	.620 ± .006	.638 ± .006
MaDAC	**.564 ± .006**	**.623 ± .006**	**.642 ± .006**

Table 13: JUMP policy-factor ablation for LLaDA2-based RAW-feature training. Retrieval columns report 100-way Hits@10 for SMILES→morphology, structure→morphology, and morphology→compound retrieval. The last two columns report structure-based morphological bioactivity prediction and morphology-guided MoA virtual screening as standardized pAUROC@0.1. Higher is better.

Policy intervention	SMILES→M	Str.→M	M→Comp.	Morphological bioactivity	Morphology-guided MoA
Frozen representation control	.152 ± .004	.165 ± .005	.192 ± .005	.540 ± .004	.505 ± .010
Fixed allocation	.321 ± .006	.246 ± .005	.289 ± .005	.621 ± .006	.523 ± .013
Easy-to-hard	.315 ± .006	.224 ± .005	.272 ± .005	.618 ± .006	.524 ± .012
Direct-route inner readiness only	**.349 ± .006**	.099 ± .004	.140 ± .004	.633 ± .007	**.529 ± .013**
Outer allocation only	.297 ± .006	.236 ± .005	.289 ± .005	.606 ± .006	.509 ± .011
Undifferentiated hierarchy	.334 ± .006	.189 ± .005	.238 ± .005	.636 ± .007	.515 ± .012
Route-decoupled, no inner readiness	.318 ± .006	.237 ± .005	.279 ± .005	.629 ± .006	.525 ± .014
MaDAC	.322 ± .006	**.271 ± .005**	**.298 ± .006**	**.651 ± .007**	.527 ± .014

Table 14: BioMed scaling for LLaDA2-based training across dataset size and source diversity. Small and Medium hold out PolypGen; Large holds out four sources.

		IID Dice			Held-out Dice		
Scale	Sources/train	Frozen	Uniform	MaDAC	Frozen	Uniform	MaDAC
Small	4 / 2,477	.280	$.342 \pm .005$	$\mathbf{.401 \pm .023}$	.131	$.150 \pm .029$	$\mathbf{.214 \pm .035}$
Medium	14 / 18,566	.264	$.414 \pm .000$	$\mathbf{.559 \pm .079}$	.167	$.126 \pm .168$	$\mathbf{.423 \pm .032}$
Large	14 / 80,213	$.282 \pm .001$	$.775 \pm .006$	$\mathbf{.860 \pm .002}$	$.104 \pm .002$	$.469 \pm .019$	$\mathbf{.616 \pm .025}$

D.5 Ablation studies

We ablate the hierarchy in Eq. equation 2 at the levels defined in Sections 3.2.1 and 3.2.2. *Fixed allocation* replaces the adaptive outer policy with its base law; *outer allocation only* adapts the sampled problem but removes inner readiness weights; and *inner readiness only* retains readiness weighting without adaptive outer allocation. For multi-route JUMP, *route-decoupled, no inner readiness* preserves route-specific outer allocation and coverage but sets the inner weights to one, while the *undifferentiated hierarchy* applies one inner rule across direct and cross-modal routes. These interventions test the roles of off-diagonal support, state-dependent allocation, and protected coverage formalized in Proposition B.1, Theorem 3.3, and Proposition 3.5.

We examine this shared decomposition on JUMP in Tables 12 and 13. JUMP exposes multiple prediction routes and continuous targets, so its ablations include route decoupling, outer rate allocation, and direct-route readiness.

For Jump, the UniDisc-based co-embedding ablation favors MaDAC on all three retrieval directions. The LLaDA2-based RAW-feature study is slightly worse: MaDAC is strongest on the two cross-modal retrieval directions and morphological bioactivity, while direct-route readiness only is slightly stronger on SMILES→M and the morphology-guided MoA endpoint. The result establishes backbone- and route-specific effects for the curriculum factors.

D.6 Scaling and corruption-space analyses

BioMed scale ladder. Table 14 compares the progressive BioMed builds. They change both sample count and source composition, so they form a capability ladder rather than a controlled data-only scaling law. Nevertheless, the selected MaDAC policy improves over uniform at each scale and on both IID and held-out evaluation.

Because the source mix, held-out definition, and budget change across the three builds, we do not interpret the margins as a monotonic scaling law. The supported conclusion is that the curriculum advantage survives substantial changes in dataset size and composition.

HuBMAP training trajectories. Figure 9 follows the matched LLaDA2-based frozen, uniform, and MaDAC variants at 1k-step intervals through the shared 12k budget. Here, *frozen* refers to the pretrained LLaDA2 backbone, not to the entire segmentation system: the backbone and its LoRA parameters remain fixed, but the randomly initialized spatial decoder is trained throughout. The frozen trajectory therefore rises as this decoder learns to map fixed pretrained features to HuBMAP masks. Uniform training updates the same decoder plus rank-8 LoRA in the final two transformer layers using fixed balanced sampling and ordinary loss averaging. MaDAC updates exactly the same decoder and LoRA parameters, but begins with decoder-only stabilization and then activates adaptive example sampling and readiness weighting.

These trajectories expose the rate at which useful segmentation performance is acquired, rather than only the final selected values. All three variants improve as their spatial decoder trains, but updating the backbone through LoRA lets uniform and MaDAC move beyond the decoder-only control. The curves are similar early in training; MaDAC separates more clearly around the middle of the budget, especially on the source-held-out test. Uniform training later narrows the IID gap, whereas the held-out advantage of MaDAC remains visible near the end of training. Thus, the comparison distinguishes decoder learning, backbone adaptation, and curriculum-dependent improvement speed under the same data, training budget, checkpoint cadence, and inference pipeline; uniform

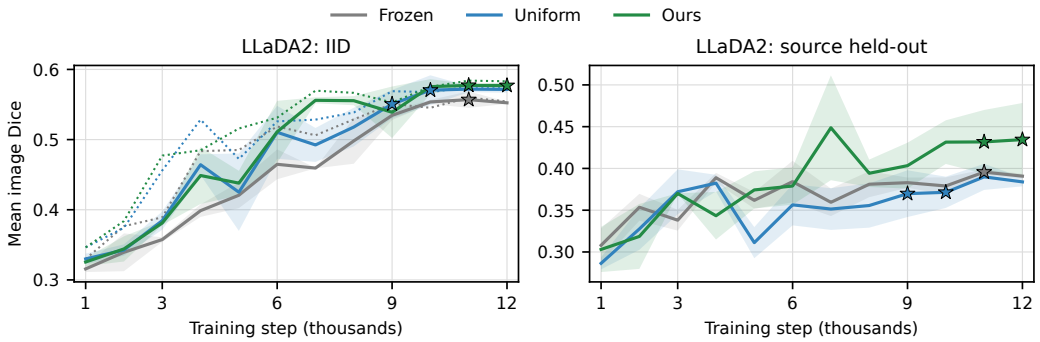

Figure 9: HuBMAP mean-image Dice for matched LLaDA2-based frozen, uniform, and MaDAC training over 12k updates. "Frozen" fixes the pretrained backbone while continuing to train the task-specific spatial decoder. Dotted curves in the IID panel are calibration scores used for checkpoint selection; solid curves are locked IID and source-held-out test evaluations. Stars mark steps selected from calibration for at least one seed. Test labels do not participate in selection.

and MaDAC additionally share the same trainable parameter set. It also confirms that the main HuBMAP result uses calibration-only checkpoint selection rather than test-set selection.

D.7 General-domain multimodal tasks

The biological benchmarks are the focus of the paper. We additionally test whether the same curriculum principles transfer to non-biological multimodal settings through table reconstruction and interleaved-document generation.

D.7.1 Table structure reconstruction

Dataset and task. We construct a table-recognition benchmark from PubTables-1M and FinTabNet (Smock et al., 2021; Zheng et al., 2020). The model observes a clean table image and predicts normalized HTML structure. PubTables provides the IID evaluation and FinTabNet measures transfer to a distinct financial-table domain.

Setting and baselines. We evaluate LLaDA2-based and UniDisc-based MaDAC training with matched fixed-allocation controls. The LLaDA2-based comparison also includes a frozen checkpoint; Qwen3-VL-8B (Bai et al., 2025), TableFormer (Nassar et al., 2022), and BAGEL (Deng et al., 2025) provide external references. The two backbone tables use different evaluation populations. LLaDA2 and its external references are evaluated on all 1,500 PubTables-1M test records, whereas UniDisc's 512-token interface retains 1,227 PubTables-1M records (and 1,371 FinTabNet records) after excluding structures longer than 1,100 characters. External baselines such as BAGEL are reevaluated on the corresponding population, so their absolute scores can differ between the two tables and must not be compared across tables as though they came from identical examples.

Metrics. All six metrics lie in $[0, 1]$ and are higher-is-better, but they test different aspects of a reconstructed table. *Struct* is structure-TEDS: normalized tree-edit similarity between the predicted and reference table structures after cell text is removed. *Full* is the corresponding full-TEDS score with cell content retained, and therefore measures structure and text recovery jointly; we report it only for the content-bearing LLaDA2 evaluation. *GriTS* denotes GriTS-Top, which compares the tables as cell grids and scores topology without relying on one particular HTML-tree serialization. *Adj.* is adjacency-F1 over the horizontal and vertical neighbor relations between cells. *Exact* is the fraction of examples whose normalized predicted structure exactly matches the reference structure. Finally, *Grid* is grid consistency: the fraction of predictions that expand, including row and column spans, into a valid rectangular grid. Unlike the other metrics, Grid is a reference-free structural-validity check; a consistent grid need not be the correct grid.

Results. Tables 15 and 16 report the primary backbone comparisons.

Table 15: Table-structure reconstruction with LLaDA2-based training on all 1,500 PubTables-1M test records. All metrics are higher-is-better; bold marks the best internal LLaDA2-based training result. External baselines are scored on this same 1,500-record population.

Method	Struct	Full	GriTS	Adj.	Exact	Grid
Frozen LLaDA2	0.729	0.573	0.742	0.782	0.045	0.201
Joint uniform	0.760	0.591	0.781	0.804	0.130	0.342
MaDAC (LLaDA2)	**0.800**	**0.625**	**0.818**	**0.819**	**0.141**	**0.431**
Qwen3-VL-8B	0.790	0.729	–	–	–	–
TableFormer	0.847	0.799	–	–	–	–
BAGEL zero-shot	0.806	0.759	–	–	–	–
BAGEL SFT	0.943	0.933	–	–	–	–

Table 15 shows that LLaDA2-based MaDAC training improves every reported structural metric over its joint-uniform counterpart on PubTables IID. The external systems provide architecture- and training-data context on the shared output metrics.

Table 16 shows that UniDisc-based MaDAC training improves Structure-TEDS and GriTS over its joint-uniform counterpart on both domains, while exact-match, adjacency, and grid metrics remain endpoint dependent.

Table 16: Table-structure reconstruction with UniDisc-based training on its length-filtered populations: 1,227 PubTables-1M IID and 1,371 FinTabNet OOD records. All metrics are higher-is-better; bold marks the best internal UniDisc-based training result. BAGEL is reevaluated on these retained records, not the 1,500-record population in Table 15.

Method	Struct	GriTS	Adj.	Exact	Grid
PubTables IID					
BAGEL SFT	0.964	–	–	–	–
Joint uniform	0.629	0.648	0.660	**0.065**	**0.297**
MaDAC (UniDisc)	**0.707**	**0.727**	**0.728**	0.047	0.210
FinTabNet OOD					
BAGEL SFT	0.925	–	–	–	–
Joint uniform	0.403	0.410	**0.420**	0.006	**0.287**
MaDAC (UniDisc)	**0.459**	**0.454**	0.412	**0.007**	0.217

Table 17: Effect of joint multimodal corruption on LLaDA2-based table-reconstruction training. Each cell reports IID/OOD Structure-TEDS; Δ is MaDAC minus uniform.

Regime	Frozen	Uniform	MaDAC	Δ
Single-clean i2t	.664/.682	.739/.684	.738/.649	−.001
Joint image+text	–	.760/.680	**.800**/.676	+.040

Corruption-regime analysis. Table 17 isolates the theoretical significance of joint multimodal corruption by comparing it with a single-clean recognition regime under the same dataset and evaluation environment.

Theory-level interpretation. Table 17 separates the two regimes predicted by the corruption-space analysis in Section 3.3. Because the IID comparison holds the data and evaluation environment fixed, it isolates the effect of corruption-task allocation. In the single-clean image-to-text task, uniform rate sampling already realizes the standard Monte Carlo masked-diffusion ELBO across its one-dimensional corruption continuum; reallocating points along that same continuum does not create a new conditional prediction problem. Accordingly, MaDAC saturates at the uniform level on IID. In the joint image–text task, the action space instead contains distinct modality-specific rate pairs. MaDAC can allocate training to asymmetric "observe one modality, recover the other" conditionals that a shared-rate uniform policy underexposes, as formalized by Proposition 3.2. Its IID gain over uniform therefore links the empirical advantage to genuinely multimodal problem allocation, rather than to a generally stronger scalar masking schedule.

D.7.2 Interleaved document generation

Dataset and tasks. CoMM (Chen et al., 2025) evaluates four conditionals over interleaved image-text documents: image-to-text generation (T1), text-to-image generation (T2), document continuation (T3), and title-conditioned document generation (T4).

Setting and baselines. We instantiate LLaDA2-based MaDAC training. The Frozen and Uniform variants are matched LLaDA2-based controls, while SEED-8B/14B provide external endpoint references on their available evaluation populations. No UniDisc variant is included in the reported CoMM experiments.

Results. Tables 18–21 report the four CoMM conditionals. Relative to matched LLaDA2-based uniform training, LLaDA2-based MaDAC training improves every reported T1–T4 metric. The large gap from the frozen control to both trained variants shows the importance of task adaptation, whereas the consistent MaDAC–uniform ordering isolates the additional curriculum effect. For the text-producing T1, T3, and T4 conditionals, gains appear at both the aligned-step and whole-document levels. They are most pronounced in CIDEr—including more than a twofold increase over uniform

Table 18: CoMM T1 image-to-text generation. Step-level and whole-document METEOR, ROUGE-L, and CIDEr are computed after aligning generated and reference segments; higher is better. LLaDA2-based Frozen and Uniform training are matched controls, and LLaDA2-based MaDAC training uses the same backbone.

Method	Step-level			Whole-document		
	METEOR	ROUGE-L	CIDEr	METEOR	ROUGE-L	CIDEr
SEED-8B	2.688	6.998	0.648	3.745	5.197	0.056
SEED-14B	3.459	8.858	1.133	4.738	6.536	0.236
Frozen (LLaDA2)	2.964	7.796	0.438	3.703	7.619	0.045
Uniform (LLaDA2)	3.493	8.855	0.509	4.413	8.489	0.082
MaDAC (LLaDA2)	**3.628**	**8.913**	**0.533**	**4.583**	**8.508**	**0.173**

Table 19: CoMM T2 text-to-image generation. FID and near-blank fraction are lower-is-better; Inception Score (IS) and generated-panel coverage are higher-is-better. The LLaDA2-based training variants share the same 1,052-panel evaluation set; SEED uses different generated/reference panel counts and is shown only as external context.

Method	FID↓	IS↑	Coverage↑	Blank↓
SEED-8B	81.82	9.161	–	–
SEED-14B	76.28	9.462	–	–
Frozen (LLaDA2)	104.41	8.195	1.00	0.008
Uniform (LLaDA2)	88.63	8.565	**1.00**	0.002
MaDAC (LLaDA2)	**87.52**	**8.917**	**1.00**	**0.001**

for T1 whole-document CIDEr—while the smaller METEOR and ROUGE-L gains show that the ordering is not confined to one overlap metric. On T2, MaDAC gives a modest FID reduction, a larger Inception Score increase, and a lower near-blank rate while retaining complete panel coverage, so the image gain is not obtained by dropping requested panels. These results support transfer of the curriculum benefit across different multimodal conditionals, but not a blanket external-SOTA claim: SEED remains stronger on several endpoints and is included as external context because its model scale and available evaluation populations differ from the matched internal comparison.

D.8 Computational Efficiency

The adaptive controller updates sampling probabilities and curriculum statistics without adding another backbone forward or backward pass in the matched backbone implementations. Its cost is therefore small relative to model training: the matched LLaDA2-based and JUMP UniDisc-based comparisons range from 2.1% to 3.4%, although the HuBMAP UniDisc-based implementation reaches 10.5%. The overhead is thus implementation-dependent rather than a fixed property of the method. These measurements come from matched short-budget timing runs extracted from the same implementations and configurations as the full experiments. Their 30k- or 6k-unit budgets differ from the principal experiments, so they estimate runtime overhead only and are not full-experiment training times or endpoint results.

Equal-budget wall time. Table 22 divides each matched short timing job's mean wall time by its training budget. It reports the practical cost of 1,000 common training units for uniform sampling and MaDAC.

The BioMed counter advances once per microbatch and uses gradient accumulation of eight. Its first row is therefore per 1,000 logged microbatches, not per 1,000 optimizer updates. Expressed per 1,000 optimizer updates, the estimated times are 40:07 for uniform and 41:29 for MaDAC, a difference of about 1:22; the percentage overhead is unchanged.

Table 20: CoMM T3 interleaved document continuation. Step-level and whole-document METEOR, ROUGE-L, and CIDEr are reported after segment alignment; higher is better. The LLaDA2-based Frozen, Uniform, and MaDAC training variants use the same task contract.

Method	Step-level			Whole-document		
	METEOR	ROUGE-L	CIDEr	METEOR	ROUGE-L	CIDEr
SEED-8B	3.987	9.593	1.339	4.856	9.241	0.275
SEED-14B	4.291	9.950	1.865	5.360	9.821	0.447
Frozen (LLaDA2)	2.769	6.337	0.851	3.028	6.461	0.190
Uniform (LLaDA2)	3.613	8.759	0.982	3.592	7.392	0.285
MaDAC (LLaDA2)	**3.727**	**8.929**	**1.210**	**3.699**	**7.489**	**0.314**

Table 21: CoMM T4 title-conditioned document generation. Step-level and whole-document METEOR, ROUGE-L, and CIDEr are computed after segment alignment; higher is better. The LLaDA2-based Frozen, Uniform, and MaDAC training variants use the same task contract.

Method	Step-level			Whole-document		
	METEOR	ROUGE-L	CIDEr	METEOR	ROUGE-L	CIDEr
SEED-8B	2.527	6.149	0.457	3.477	5.577	0.168
SEED-14B	3.561	8.881	1.039	4.758	6.843	0.089
Frozen (LLaDA2)	1.886	4.978	0.272	2.411	5.020	0.049
Uniform (LLaDA2)	3.301	7.979	0.517	3.518	6.202	0.117
MaDAC (LLaDA2)	**3.404**	**8.477**	**0.631**	**3.572**	**6.411**	**0.127**

These values are normalized from complete short matched timing jobs rather than separate 1,000-step launches. They include startup, data loading, checkpointing, and shutdown, and should be read as end-to-end averages for the short timing protocol rather than a kernel-only controller benchmark or the duration of a principal experiment. Table 23 reports the corresponding complete short-run measurements.

Overall, MaDAC requires fewer training samples on BioMed and HuBMAP. On JUMP, it improves average learning-curve area and reaches the T2 and T7 targets earlier, but does not improve aggregate checkpoint-level time to target. The small matched LLaDA2-based wall-clock overhead leaves the BioMed and HuBMAP sample savings net-positive, while avoiding a blanket efficiency claim across implementations.

D.9 Complete qualitative galleries

The main paper shows four representative cases for space. We retain the remaining original BioMed and HuBMAP galleries here because qualitative counterexamples, source-shift behavior, and morphology of segmentation errors are complementary to aggregate metrics. These are selected cases and should not be interpreted as population-level win frequencies or as a single best-of-gallery model-selection rule.

BioMed gallery. Additional 34 cases are consolidated into six dense full-page sheets, with five or six original gallery rows per sheet. The title above each row preserves the original split, source, and sample identifier; rows are ordered from top to bottom by the original gallery figure number. Figures 10–15 show these cases in a consistent image, target, and prediction layout, making recurring boundary and source-shift errors directly comparable across samples.

HuBMAP gallery. Additional 16 cases are consolidated into three dense full-page sheets, with five or six original gallery rows per sheet. As above, each row retains the original sample title and the rows follow the original gallery figure order from top to bottom. Figures 16–18 provide the

Table 22: Normalized wall-clock cost from matched short-budget timing runs, not the full principal experiments. The runs use the corresponding full experiment's implementation and configuration; the training unit follows the counter used by each implementation.

Matched implementation	Normalized counter	Uniform per 1,000	MaDAC per 1,000	Additional time	Overhead
BioMed LLaDA2-based	logged microbatches	5:00.9	5:11.2	+10.3 s	+3.4%
HuBMAP LLaDA2-based	optimizer steps	50:49.2	51:54.1	+1:04.9	+2.1%
JUMP UniDisc-based co-embedding	optimizer updates	1:45.5	1:48.3	+2.8 s	+2.7%
HuBMAP UniDisc-based	optimizer steps	55:30.0	1:01:20.0	+5:50.0	+10.5%

Table 23: End-to-end wall time for the complete matched short-budget timing runs. These are overhead measurements extracted from the full experimental implementations, not full-budget principal experiments.

Short-run timing comparison	Uniform	MaDAC	Overhead
BioMed LLaDA2-based, 30k microbatches	2:30:27	2:35:35	+3.4%
HuBMAP LLaDA2-based, 6k steps	5:04:55	5:11:25	+2.1%
JUMP UniDisc-based co-embedding, 6k updates	10:33	10:50	+2.7%
HuBMAP UniDisc-based, 6k steps	5:33:00	6:08:00	+10.5%

corresponding organ-segmentation comparisons. The sheets expose variation in mask extent and morphology without assigning population-level frequency to the selected examples.

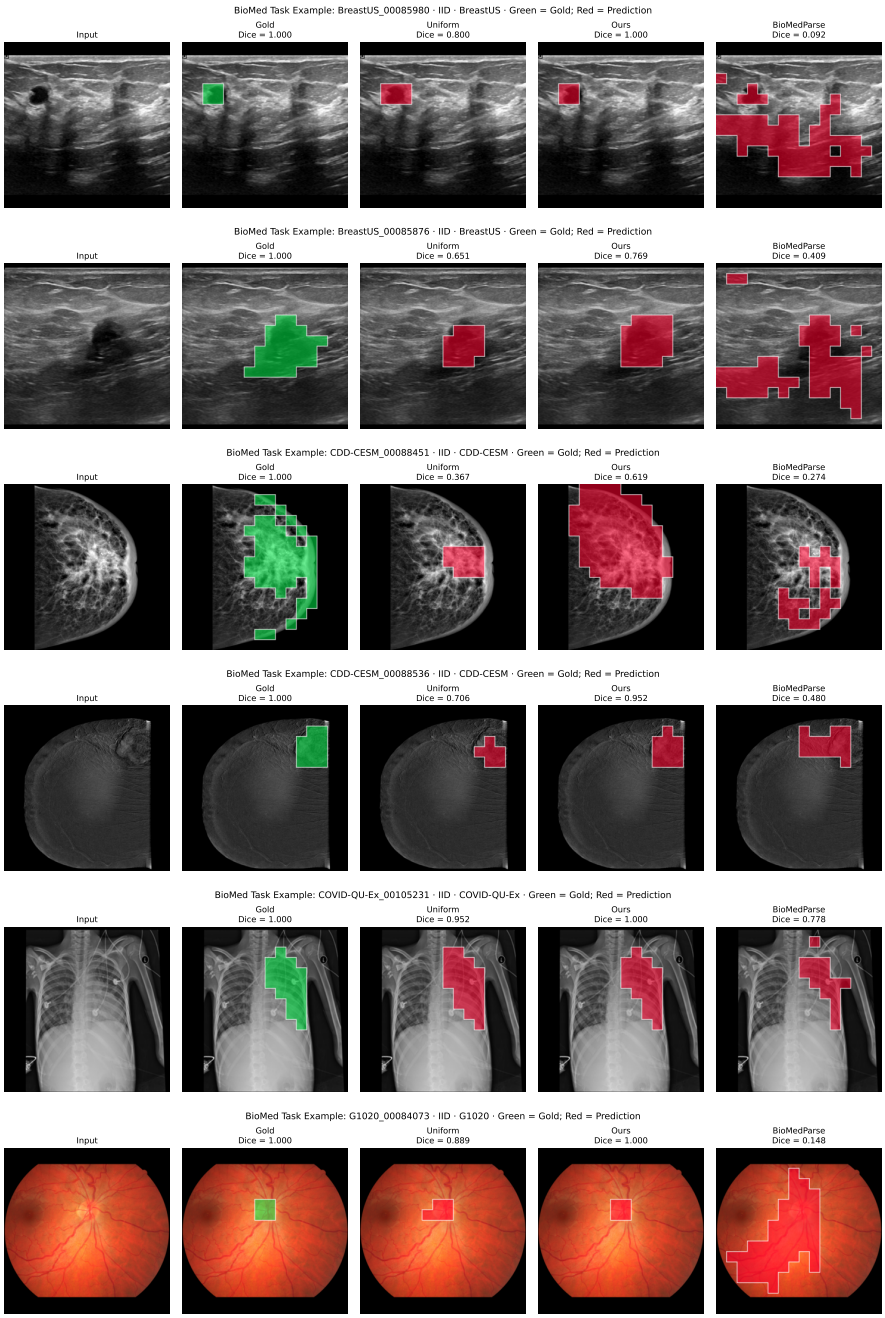

Figure 10: BioMed sample-pack-v2 qualitative gallery sheet 1 of 6. Rows, top to bottom, are original gallery Figures 1–6.

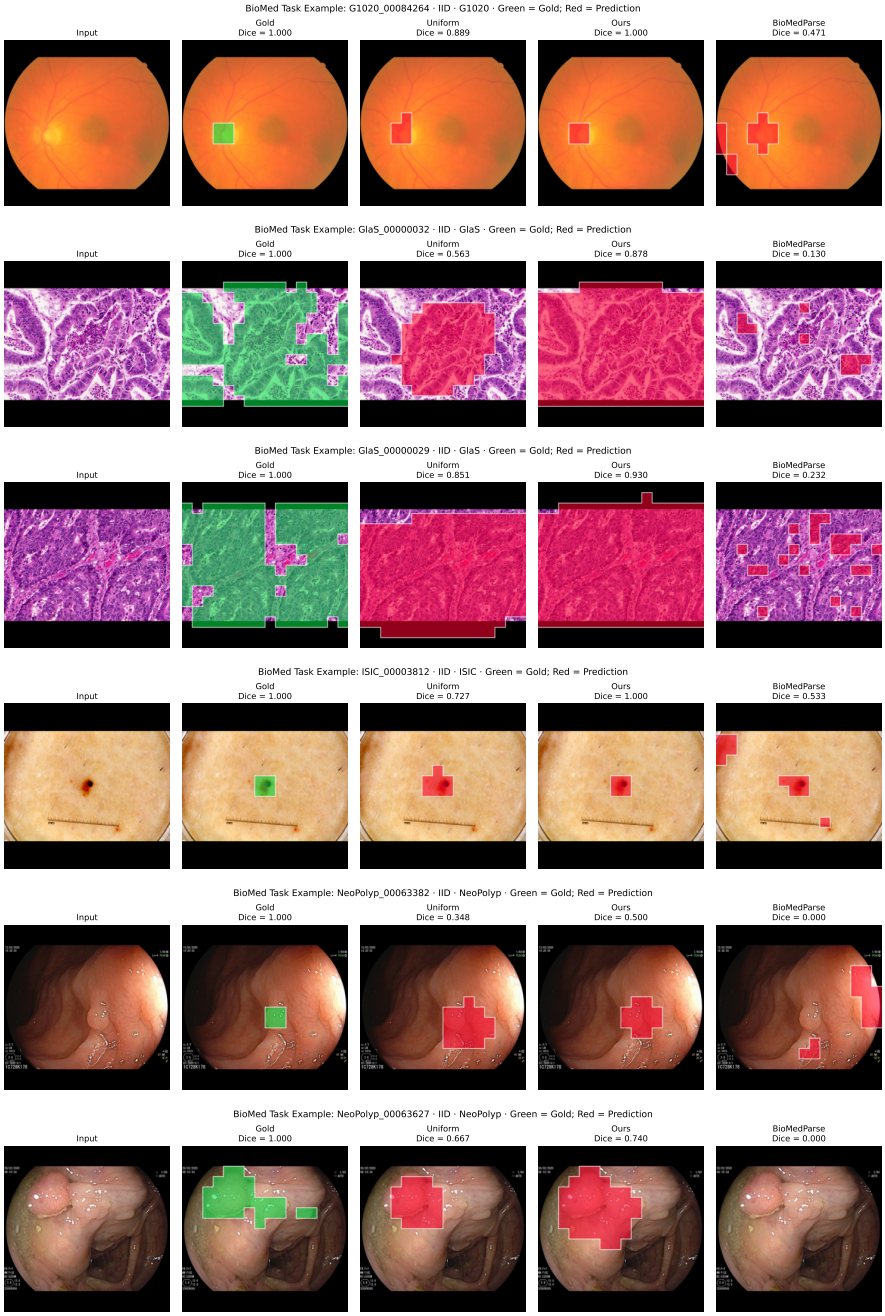

Figure 11: BioMed sample-pack-v2 qualitative gallery sheet 2 of 6. Rows, top to bottom, are original gallery Figures 7–9 and 11–13.

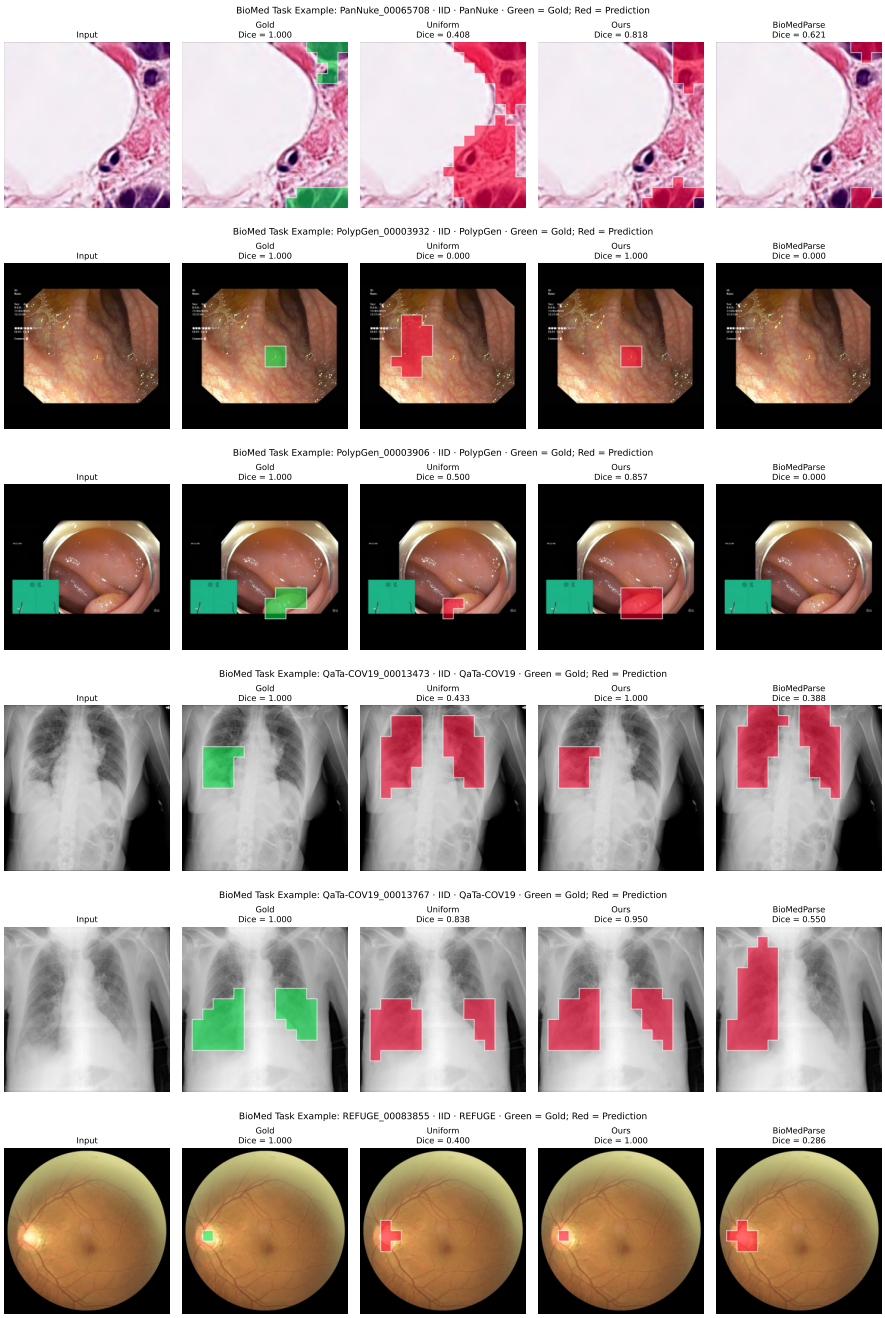

Figure 12: BioMed sample-pack-v2 qualitative gallery sheet 3 of 6. Rows, top to bottom, are original gallery Figures 14, 16–20.

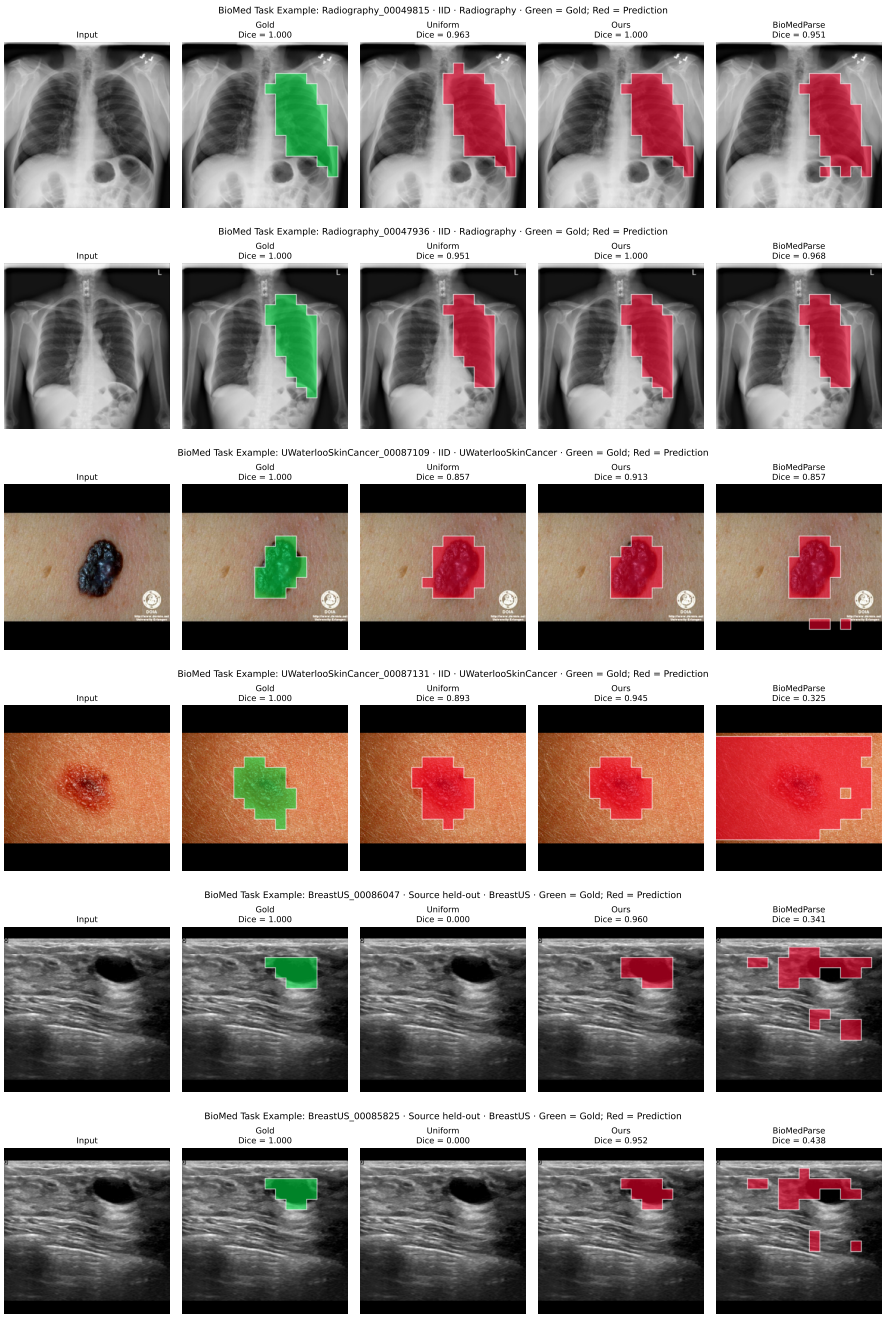

Figure 13: BioMed sample-pack-v2 qualitative gallery sheet 4 of 6. Rows, top to bottom, are original gallery Figures 21–26.

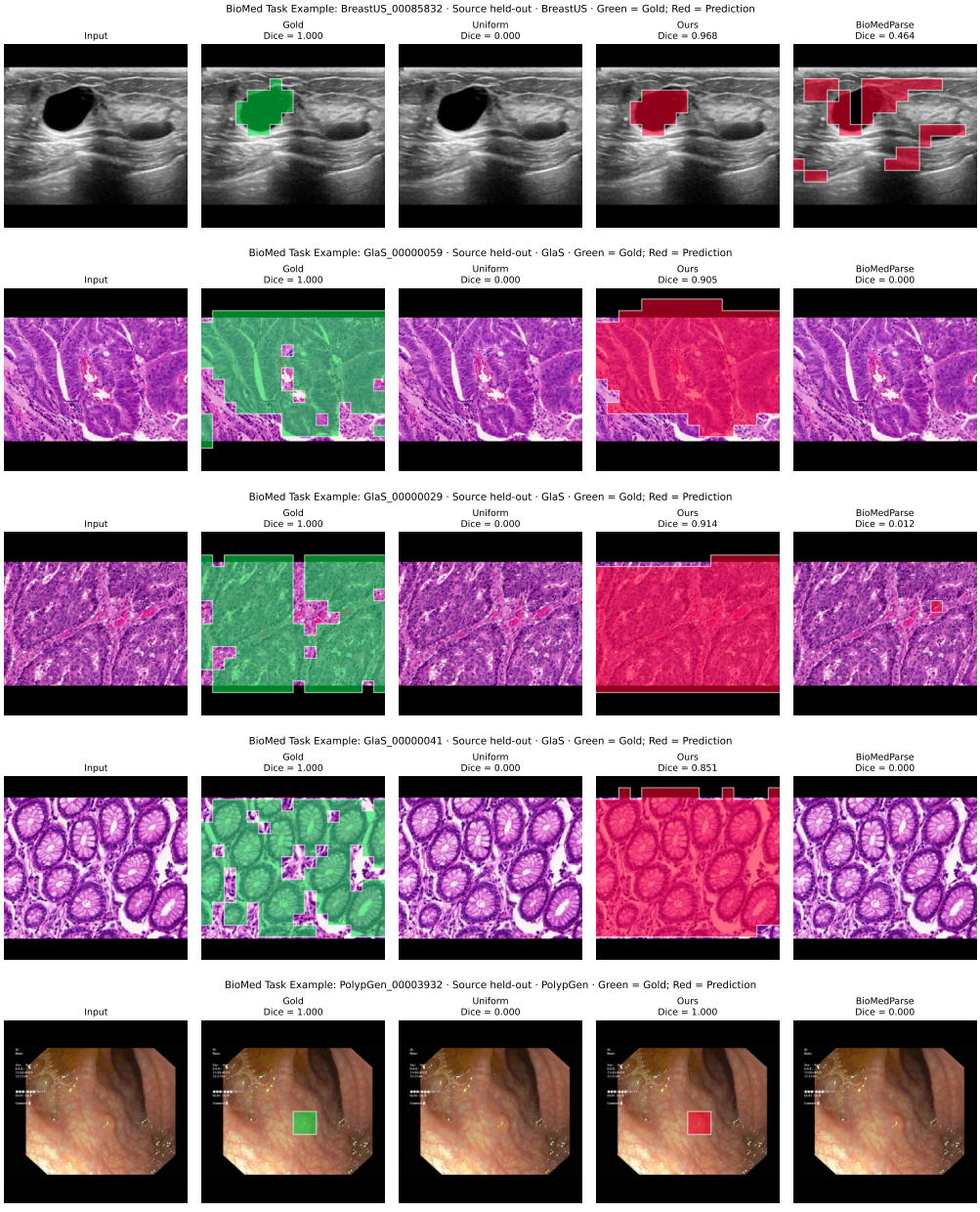

Figure 14: BioMed sample-pack-v2 qualitative gallery sheet 5 of 6. Rows, top to bottom, are original gallery Figures 27–31.

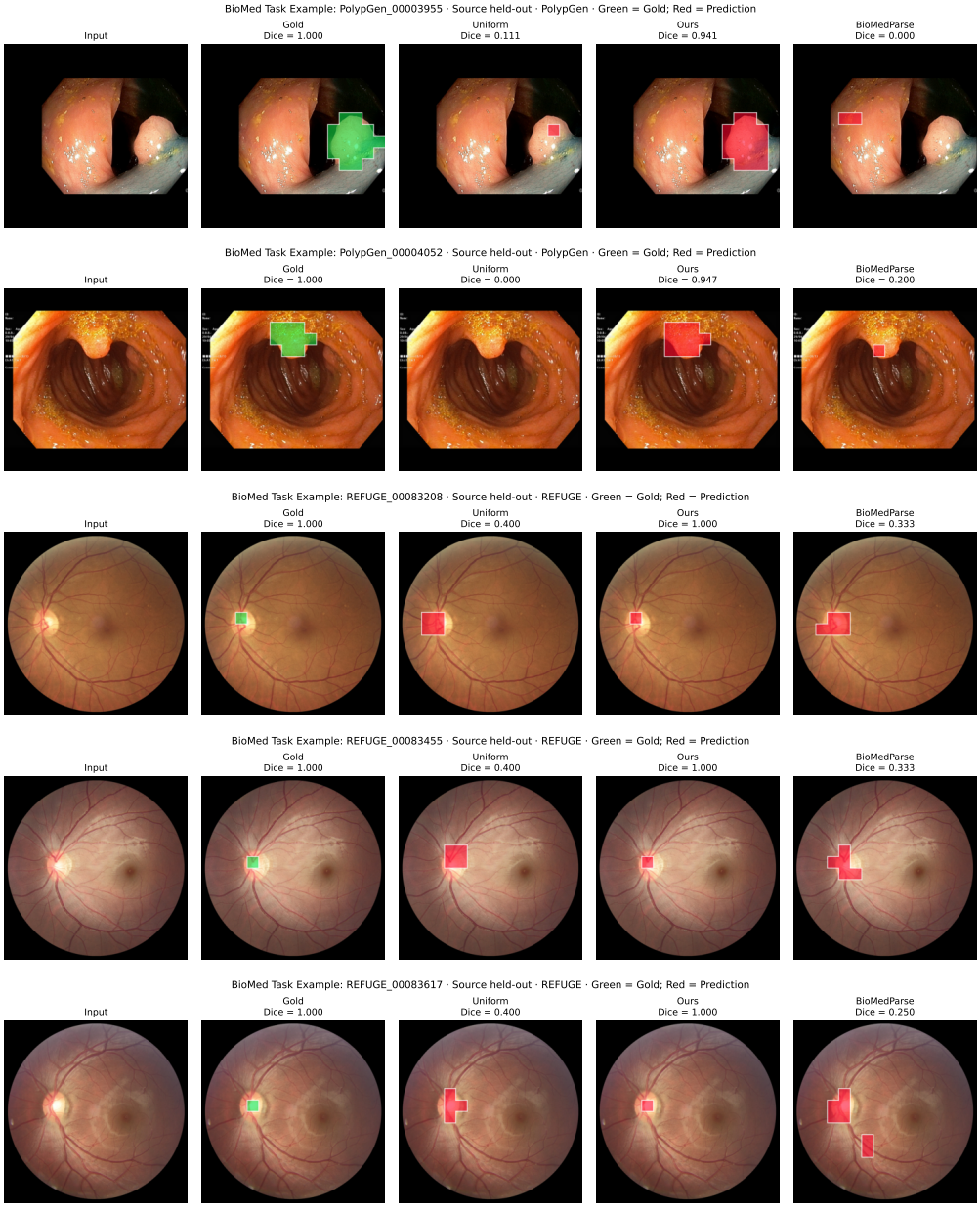

Figure 15: BioMed sample-pack-v2 qualitative gallery sheet 6 of 6. Rows, top to bottom, are original gallery Figures 32–36.

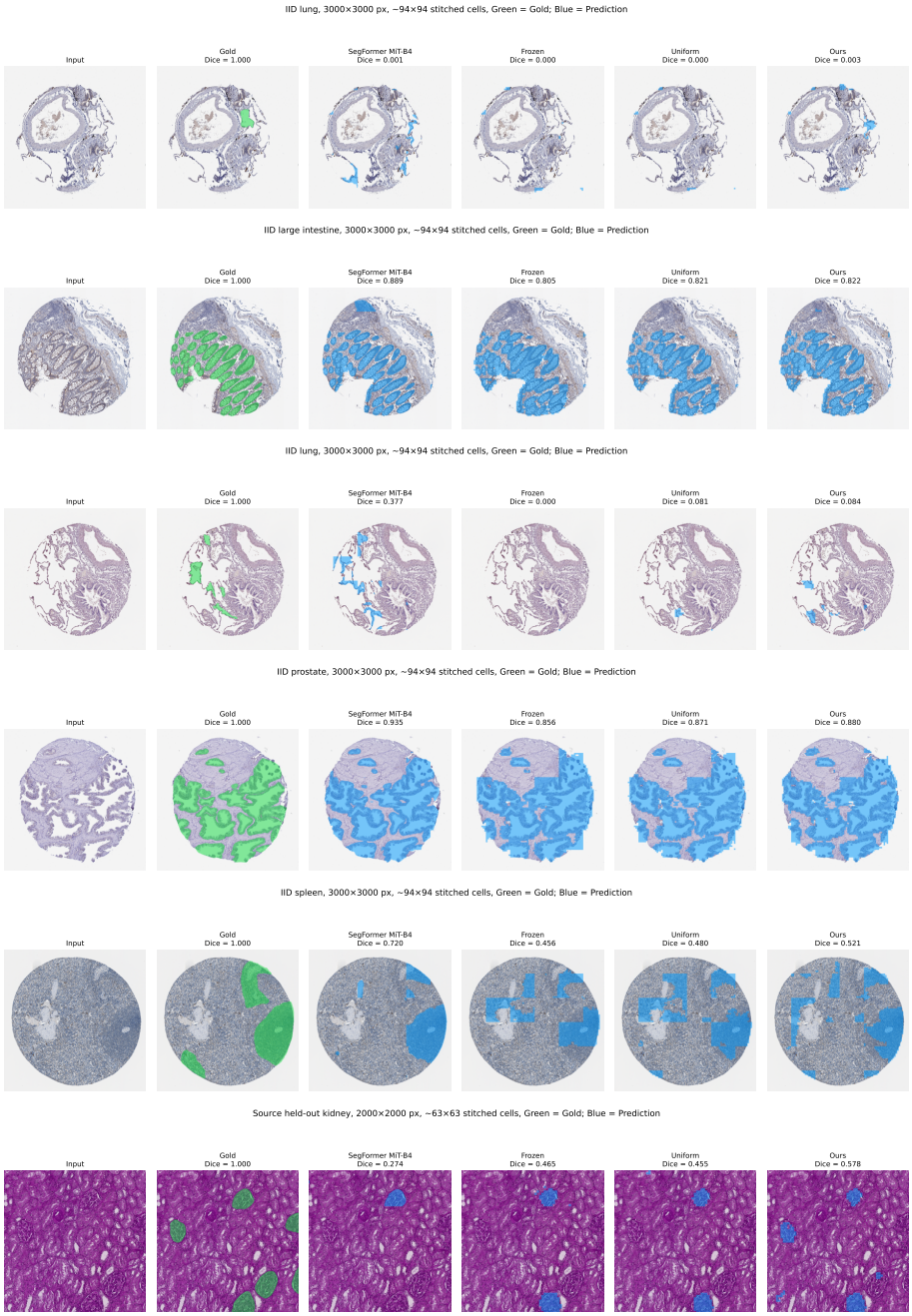

Figure 16: HuBMAP qualitative gallery sheet 1 of 3. Rows, top to bottom, are original gallery Figures 1, 3–7.

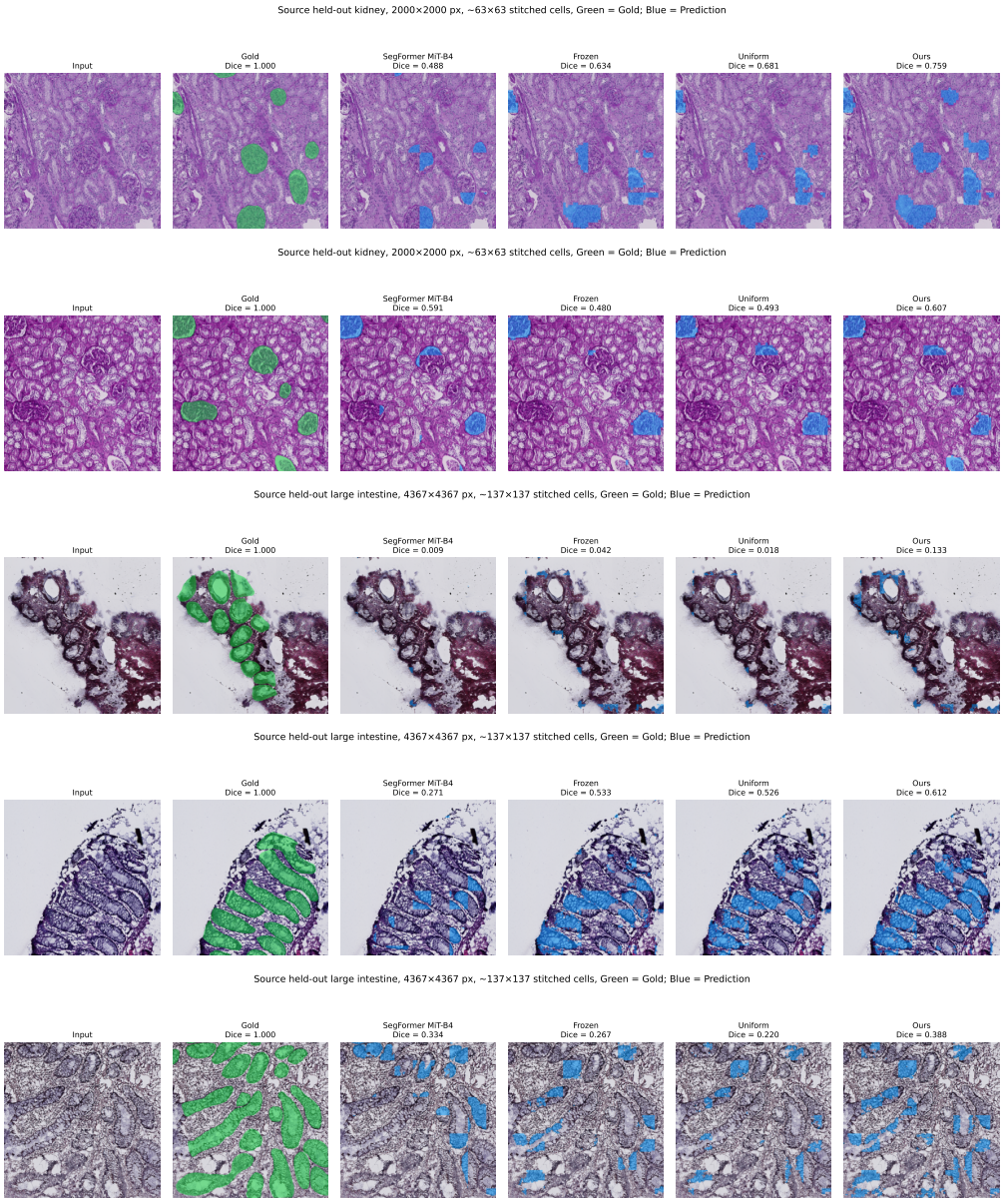

Figure 17: HuBMAP qualitative gallery sheet 2 of 3. Rows, top to bottom, are original gallery Figures 8–12.

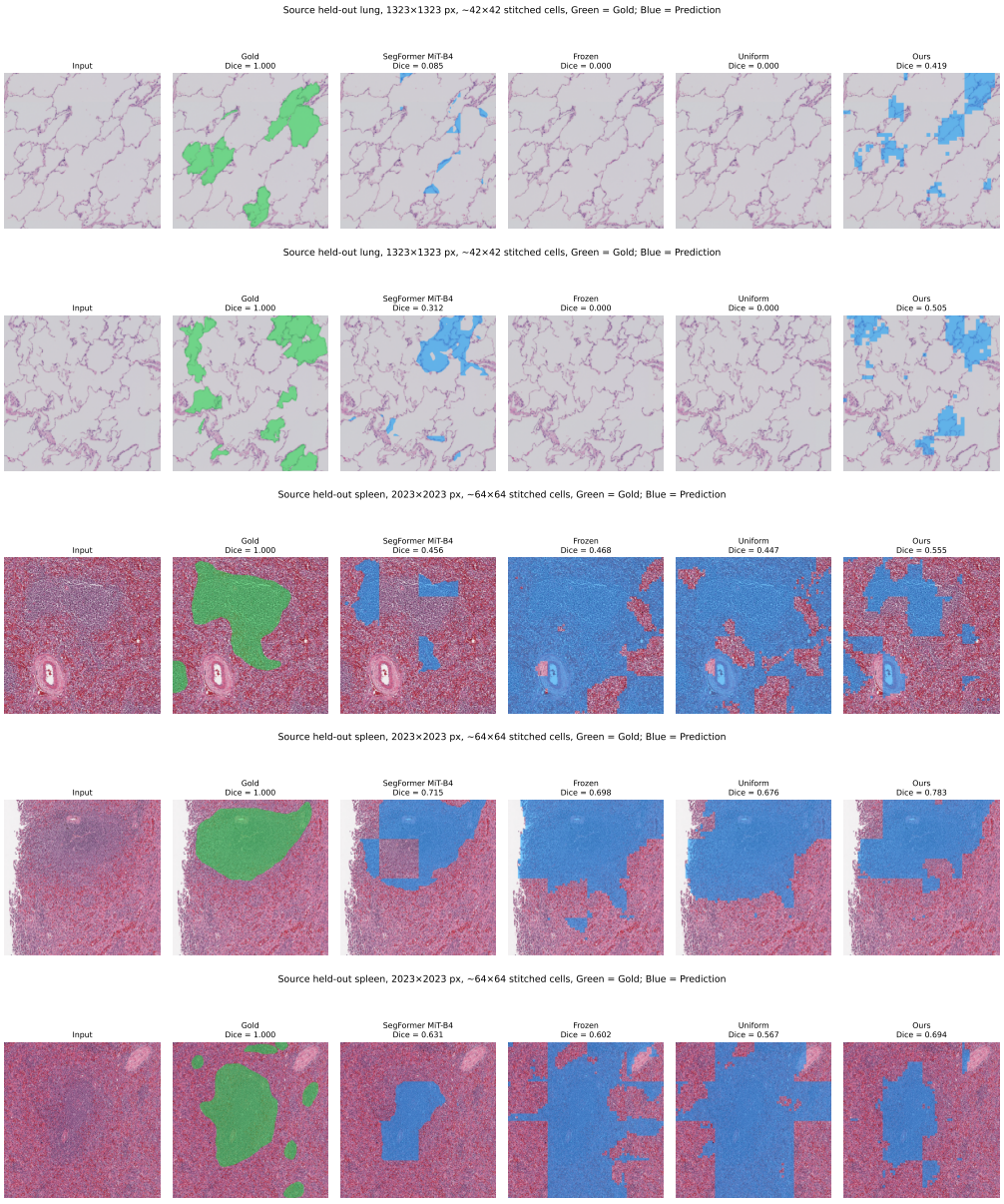

Figure 18: HuBMAP qualitative gallery sheet 3 of 3. Rows, top to bottom, are original gallery Figures 13–14 and 16–18.

E LICENSES

This section records the licenses or use terms for external datasets, checkpoints, and software/model baselines retained in the paper. A repository license applies only to the identified artifact and does not relicense upstream images, papers, or checkpoints.

E.1 DATASETS AND DERIVED FEATURES

BioMed. BioMed is our processed subset of `microsoft/BiomedParseData`, whose dataset card designates the collection CC BY-NC-SA 4.0 (`https://huggingface.co/datasets/microsoft/BiomedParseData`). Component datasets retain their original terms and restrictions. The released BiomedParse checkpoint is likewise designated CC BY-NC-SA 4.0 and is intended for research/development rather than clinical decision making (`https://huggingface.co/microsoft/BiomedParse`).

JUMP. The JUMP Cell Painting images and profiles are released under CC0 1.0 (`https://github.com/jump-cellpainting/datasets/blob/main/LICENCE.md`). We use the compound-level package released with PhenoCompass, including its DINO morphology, GNEprop structure, Morgan-fingerprint, activity-label, and scaffold-split artifacts. The audited Zenodo record states Genentech copyright but does not name a standard license (`https://doi.org/10.5281/zenodo.20367744`); upstream GNEprop data are CC BY-NC 4.0 and the DINO implementation/weights are Apache 2.0.

HuBMAP/HPA. The 2022 HuBMAP+HPA competition release is governed by the competition rules accepted at download time (`https://www.kaggle.com/competitions/hubmap-organ-segmentation/rules`). HuBMAP additionally applies its data-use terms where relevant, while Human Protein Atlas content is CC BY-SA 3.0 unless stated otherwise.

Tables and CoMM. PubTables-1M is distributed under CDLA-Permissive-2.0 and FinTabNet under CDLA-Permissive-1.0; Docling/TableFormer code is MIT licensed. CoMM combines material from several source websites and does not provide one license that supersedes the original sources; we use it under the official research-release terms and retain source attribution.

E.2 PRETRAINED CHECKPOINTS AND EXTERNAL BASELINES

LLaDA2.0-Uni is Apache 2.0. The released BiomedParse checkpoint is covered by the terms above; CLIPSeg is MIT licensed. SegFormer is distributed under the NVIDIA Source Code License; torchvision is BSD-3-Clause, while ImageNet images retain their non-commercial research/educational access terms. BAGEL and Qwen3-VL are Apache 2.0; Docling/TableFormer is MIT. SEED code is Apache 2.0 and its 8B/14B weights use the Llama 2 Community License. In our UniDisc variants, the transformer backbone is trained from scratch on task data; this statement does not apply to fixed input tokenizers. In particular, the HuBMAP UniDisc pipeline uses the fixed pretrained MagViT visual tokenizer described in Appendix C. The classical ridge/kNN/linear models introduce no external pretrained weights beyond their software dependencies.

F LIMITATIONS AND FUTURE WORK

Several limitations remain. Current implementations are practical proxies for training utility rather than optimal intervention estimates, and some supplementary studies use historical checkpoints, unequal budgets, or task-specific output contracts. Adaptation can also trade in-domain gains for robustness, as seen in the LLaDA2 Tables results, while the CoMM experiments show that improvements on multimodal text tasks do not automatically transfer to image-generation quality. Future work should develop more reliable utility estimators, explicitly preserve pretrained capabilities and robustness, and scale adaptive allocation to richer modality spaces and longer training horizons.