# **EEVEE and GATE: Finding the right benchmarks and how to run them seamlessly**

Anonymous Author(s) Affiliation Address email

#### Abstract

Model evaluation is a cornerstone of machine learning, guiding model design and 1 2 progress measurement. Designing generalizable evaluation processes remains a challenge, however, partly due to the vast number of possible domain, task and 3 modality combinations and lack of knowledge of how informative they are. In 4 5 this paper, we propose *EEVEE* (Efficient Evaluation process Evolution Engine)<sup>1</sup>, a method that frames evaluation process design as a learning problem. By analyzing 6 a large number of evaluation metrics from diverse benchmarks and models, EEVEE 7 identifies a smaller subset of tasks with high predictive power over the full set of 8 evaluation metrics, reducing evaluation time. To find the optimal subset maximiz-9 ing signal while minimizing GPU hours, EEVEE evaluates pre-trained models of 10 various architectures, pretraining schemes, and modalities on diverse downstream 11 tasks and datasets including image classification, segmentation, relational reason-12 ing, zero-shot image-to-text tasks, medical classification and segmentation, video 13 classification, and regression. Our results identify three subsets of benchmarks, 14 15 with 8, 15 and 21 tasks, providing high quality signal for model generalization. 16 Key benchmarks selected include iWildCam, CLEVR-Math, ACDC, WinoGround, CIFAR100, Fungi, and ADE20K. We structure the subsets into three tiers for 17 12, 24, and 36 GPU-hour budgets and package them into a unified, efficient, and 18 user-friendly Python framework that we built with the researcher in mind – which 19 we refer to as the GATE engine. Our experiments reveal ConvNextV2, SigLIP 20 and CLIP as top-performing model encoders, with EfficientNetV2 and ResNext50 21 excelling in medical tasks and challenging image classification, in particular in 22 Happy Whale Individual classification, ConvNet based models seem to outperform 23 transformer models by a factor of 2.5x, which is surprising. The top performing en-24 coder being ConvNextV2 followed by CLIP seems to agree with other recent large 25 scale evaluations. We also demonstrate the framework's versatility in fine-tuning 26 models from text and audio modalities, paving the way for future cross-modal 27 evaluations. 28

# 29 1 Introduction

Increasing Complexities of Benchmarking: As we create benchmarks for expanding model capability evaluation, the growing number and complexity of these benchmarks inadvertently complicates evaluation, requiring more resources like engineering, computation, and research time. Consequently, prioritizing which benchmarks to use becomes challenging. The high costs and longer wait times of newer, complex benchmarks often deter their adoption, leading researchers to rely on older, simpler benchmarks. This risks missing valuable insights from innovative ideas that may underperform on

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

<sup>&</sup>lt;sup>1</sup>Pronounced as /'i:vi:/ EE-vee

simpler benchmarks but have broader applicability, while promoting incremental improvements that

<sup>37</sup> overfit to simpler benchmarks but underperform in comprehensive evaluations.

To illustrate the mounting increase in available benchmarks, we can look at the historical benchmarks 38 in deep learning. Few benchmarks have had as much impact as ImageNet [29], which remains a 39 rich resource for model training and evaluation, particularly in visuo-linguistic models. As key 40 capabilities for deep neural networks were discovered, more benchmarks were generated to measure 41 and stimulate progress in those areas. In natural language processing, the GLUE benchmark [65], 42 SQuAD [45], and CoNLL-2003 [48] have been instrumental. In audio processing, LibriSpeech [39], 43 TIMIT [15], and VCTK [68] are widely used. For machine translation, WMT [3], IWSLT [22], and 44 Europarl [25] have driven advancements. Relational reasoning has been advanced by benchmarks 45 such as CLEVR [23], bAbI [66], and RAVEN [71]. In segmentation, PASCAL VOC [14], Cityscapes 46 [8], and COCO [33] remain crucial. Large language models are often evaluated using benchmarks 47 like SuperGLUE [64], LAMBADA [40], and MMLU [19]. Vision-language models are typically 48 evaluated using benchmarks such as VQA [1], Visual7W [76], and Flickr30k [42]. 49

As a result, a researcher has to choose from all these options, and even more, and then find a 50 way to unify and experiment with their models across all of them. The lack of unification, and 51 the lack of guarantees for their generalization signal, quickly becomes a kind of "evaluation hell", 52 where researchers waste a lot of time just doing redudant things like fixing the same bugs to 53 download datasets, preprocess them etc, while at the same time not having any real signal as to which 54 benchmarks are more informative, other than just knowing what has been used the most - which is 55 usually a function of popularity, and not real informativeness. To elaborate, the adoption of complex 56 evaluation processes that could enhance research efficiency and impact is often hindered by the 57 engineering effort required to evaluate machine learning models. Researchers must create involved 58 pipelines across multiple datasets demanding high data engineering efforts, develop task-specific 59 adapters, and derive nuanced training recipes, which is time-consuming. As a result, researchers 60 often revert to simpler evaluation strategies instead of comprehensive assessments. 61

A good benchmark should alleviate these burdens by automating dataset handling, integrating task 62 adapters, optimizers, schedulers, and logging mechanisms seamlessly. It should provide broad and 63 meaningful signals with minimal GPU time, accommodating various computational budgets, ensuring 64 65 inclusivity. Furthermore, an increasingly important factor for a robust modern benchmark engine is its support for multi-modal learning and early fusion techniques. AI systems must seamlessly 66 integrate and reason across multiple modalities, such as text, images, audio, and more. Multi-modal 67 learning enhances self-supervised learning opportunities and provides inherent supervision through 68 natural alignments, like audio-visual synchronization in videos. Early fusion, where data from 69 different modalities is combined at the initial stages of processing, allows models to leverage shared 70 representations, improving generalization and reasoning capabilities across varied tasks and domains. 71 These key desiderata are what motivates the production of this work. 72

With the desiderata in mind, we next introduce EEVEE, a methodology developed for building
high-signal low-cost evaluation routines, and GATE, the resulting benchmark that is designed to
be extensible, readable, flexible, modular and robust, supported by a new efficient, easy to use

76 framework.

77 EEVEE, Learning Optimal Benchmarks: The ability to find which benchmarks offer the most
 rsignal with respect to a given goal, such that we can optimize our compute time, research iteration
 rspeed, and engineering time is increasingly crucial. In this work, rather than just manually designing
 a new set of benchmarks, we propose a methodology, called *EEVEE (Empirical Evaluation process Evolution Engine)* that frames evaluation design as a learning problem and then leverages machine
 learning to automate the discovery and refinement of evaluation processs.

More specifically, EEVEE operates by taking in a large set of performance metrics from diverse models applied across various benchmarks and identifies a smaller subset of benchmarks with high predictive power over the entire set. EEVEE achieves this through two main components: (a) an evolutionary algorithm to optimize the selection of benchmark combinations based on a computed score, and (b) a meta-model trained to predict a model's performance on the full set of benchmarks using performance metrics from a chosen subset. We parameterize the meta-model as as a small neural network. The meta-model receives input performance metrics from a subset of benchmarks and predicts performance on the full set of performance metrics. Through careful *k*-fold cross-validation and leveraging a diverse set of models and benchmarks, EEVEE iteratively evolves benchmark combinations that offer high information content with respect to the entire spectrum of benchmarks, ensuring robust, efficient and comprehensive evaluation that can be targeted to computational budgets ranging from more "GPU Poor" users to high-budget organizations.
Taking the desiderata explained above and the resulting understanding of what a good evaluation

engine should look like, we demonstrate the effectiveness of EEVEE by tasking it with the discovery 97 of benchmark combinations that offer good signal-to-GPU-time ratio, for the evaluation of model 98 encoders – also referred to as backbones, on their ability to adapt to new tasks, domains, and 99 modalities. For this purpose, we choose a pool of 20 models, varying in their pretraining schemes 100 (e.g CLIP, DINO, ImageNet Classification), architectures (e.g. ResNets, ViTs, ConvNext) and even 101 their source modalities (e.g. Whisper, BERT), which we adapt on 31 benchmarks ranging from image 102 classification, segmentation, relational reasoning, zero-shot image-to-text tasks, medical classification 103 and segmentation, video classification, and regression, using robust fine tuning recipes, and training 104 for 10K iterations, ensuring that the signal we get is about models that are adaptable, generalizable 105 and efficient in their adaptation. 106

By applying 20 models on 31 benchmarks and employing EEVEE on their resulting metrics, we 107 identify three subsets of benchmarks, each targeted to a specific computational budget range. Some of 108 the key benchmarks that have been selected include iWildCam, CLEVR-Math, ACDC, WinoGround, 109 mini-ImageNet, Fungi, ADE20K, and dtextures. We refer to the discovered subsets as Tiers, and 110 assign to them identifiers for their sizes, specifically, small (n=8, 12 GPU hours), base (n=15, 24 GPU 111 hours) and big (n=31, 36 GPU hours). We package these tiers into our comprehensive benchmarking 112 suite and software framework (called GATE) designed for domain, task and modality transferability 113 evaluation, which facilitates the transfer of neural network encoders to different modalities, domains, 114 and tasks. GATE's architecture caters to the research community, enabling straightforward replace-115 ment of these transferable encoders with minimal effort. With these innovations, GATE seeks to 116 evolve the landscape of model encoder evaluation, championing a deeper understanding of transfer 117 learning and model adaptability. 118

**Contributions:** 1. We introduce *EEVEE*, a machine learning approach for selecting subsets of 119 benchmarks optimized to offer maximal predictive power over a larger benchmark set. 2. We conduct 120 a comprehensive investigation of diverse benchmarks within the space of image, image+text and 121 video modalities, pinpointing those with the highest predictive value for a model's performance 122 in downstream tasks. We apply EEVEE to model encoder evaluation by training 20 models on 31 123 benchmarks, identifying subsets of 8, 15 and 21 benchmarks that offer high signal-to-GPU-hour ratios. 124 3. We pack the EEVEE-discovered subsets (of 8, 15 and 21 benchmarks out of 31 benchmarks) into 125 targeted benchmark packs, referred to as tiers, designed for specific compute budgets (of 12, 24 and 126 36 GPU hours) and project phases, and establish standard experimental settings for these tiers. We call 127 these collectively as the GATE Benchmarks. 4. We develop the GATE engine, a unified benchmark 128 suite and software framework that automates dataset downloading, preprocessing, and pipelining 129 for fine tuning and evaluation. GATE facilitates the incorporation of new model encoders, adapts 130 input modalities, fine-tunes with robust recipes, and logs critical information such as training and 131 132 evaluation metrics, power, energy, computational usage, task visualizations, and model gradients per 133 layer. 5. Through our extensive investigation, we identify foundation models demonstrating superior transferability across diverse tasks. 6. We advocate for the inclusion of modality-shifting transfer 134 experiments in the standard evaluation process for ML researchers, supported by our experimental 135 results on the performance of existing foundation models in these benchmarks. 136

#### 137 2 Related Work

**On the Diversity of Benchmarks:** There is a vast array of benchmark suites in machine learning. 138 To the best of our knowledge, the benchmark suites relating strongly to GATE are ImageNet [9], 139 VTAB [70], VLMBench [73] and WILDS [26]. ImageNet has been of tremendous importance and 140 interest to the transfer learning community. Nevertheless, there has been skepticism about overfitting 141 to such datasets resulting from implicitly qualifying models using the test set performance over 142 the years [46, 6] or the test set not being challenging enough to gauge model generalization power 143 [47]. Although ImageNet pre-training helps transfer performance to the many-shot classification 144 setting [13], it provides minimal to no gains on more challenging datasets such as fine-grained 145

Desiderata $\downarrow$ Benchmark $\rightarrow$	ImageNet	VTAB	VLMBench	WILDS	GATE (Ours)
Diversity of Tasks	<ul> <li>Image: Second sec</li></ul>	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$	~~~ · · · ·
Diversity of Domains	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$	$\checkmark\checkmark\checkmark$
Diversity of Modalities	$\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$	$\checkmark\checkmark\checkmark$
Automatic Dataset Download/Preparation	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark\checkmark\checkmark$
Code allows for easy switch of encoders	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark\checkmark\checkmark$
Optimized for fast and effective research iteration	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark\checkmark\checkmark$
Run Time	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$	$\checkmark\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$
Includes Medical Domains	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$
Includes Environmental domains	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark$	$\checkmark\checkmark\checkmark$	$\checkmark\checkmark\checkmark$
Tiered compute budgets	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark\checkmark\checkmark$
GPU poor optimized	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark\checkmark$

#### Table 1: Our Desiderata (first column) VS Benchmarks (first row)

classification [27]. Similarly, with a larger distribution shift, ImageNet pre-trained models was 146 found to offer limited benefits for medical imaging tasks due to large distribution shifts induced by 147 fundamental differences in data sizes, features, and task specifications; that is, lightweight models 148 perform comparably to standard architectures [44]. To make matters worse, ImageNet performance 149 is less correlated with and less predictive of downstream performance on diverse tasks beyond 150 classification such as object detection, few-shot classification, and segmentation [13]. On top of it all, 151 when ImageNet is extended with a perturbed temporal dimension, models performance significantly 152 worsen [52]. 153

On the Usability of Benchmarks: Beyond ImageNet, VTAB introduced a benchmark with a wider 154 diversity of tasks and domains [70]. Nevertheless, it does not offer task and domain shifts offered 155 in GATE, such as medical segmentation and video classification and regression that are known to 156 be ill-measured and gauged by ImageNet alone [44, 52]. That said, VTAB offers satellite imaging 157 and 3D tasks which GATE does not. Nevertheless, GATE as a software framework was optimized to 158 minimise usage friction, to take no more than 12 GPU hours on our smallest tier, and, to only require 159 approximately 1 hour of adding the new encoder and wrapping it into GATE wrappers for GATE to be 160 able to go away and take care of everything, including dataset downloading, task adapter integration 161 and full train/val and test cycles with logging of various key metrics. VTAB, in our experience, 162 requires a lot more manual work in getting the datasets, and integrating new models to be adapted. 163 Similarly, VLMBench [73] and WILDS [26] offer more diverse datasets beyond previous work but 164 neither offer a tiered approach that enables iterative development of models during pre-training, nor 165 produce extensible and flexible benchmarks that can be easily glued into researchers experimentation 166 code without friction. 167

On the Systematic Selection of Benchmarks: Previous work investigated the properties inherit 168 in multi-task benchmarks that trade-off diversity and sensitivity where the latter is how robust a 169 benchmark ranking is to the inclusion of irrelevant models or minute changes in the tasks themselves 170 [72]. It was found that multi-task benchmark are unstable to irrelevant changes in tasks design. 171 Nevertheless, this is related to how the benchmark ranks models; whether it compares how model often 172 ranks higher than another in cardinal benchmarks or if the performance across tasks is averaged to 173 produce a single rank in cardinal ones. Meanwhile, our benchmark produces fine-grained information 174 to model performances across diverse tasks rather than producing specific ranking which is delegated 175 to the user analysis. Another complementary thread of work investigates dynamic benchmarks where 176 model training and data collection is interleaved to continually challenge model knowledge [53]. To 177 the best of our knowledge, this is the first work that studies the selection of multi-task, multi-domain 178 benchmarks that satisfy limited compute budgets while maximizing research signal. 179

In summary, Table 1 shows the desiderata that we believe a good evaluation suite and framework
 should have such that they can both offer the community useful signal, and also balance that with
 being practical so that people can adopt it.

# **183 3 EEVEE Methodology**

EEVEE is our proposed method for automating the selection of Pareto-optimal benchmark subsets. 184 By analyzing benchmark performance metrics, EEVEE identifies a small, highly informative subset 185 that maximizes information relative to the entire benchmark pool. This ensures that, as machine 186 learning benchmark breadth and depth increases, we will always be able to identify and select few that 187 offer high information about the whole. We strike a balance between providing rich evaluation signals 188 and maintaining simplicity, reducing computational costs and human efforts required for adopting 189 new benchmarks. EEVEE enables the production of a tiered evaluation engine accommodating 190 various computational budgets, fostering an inclusive and accessible research environment, and 191 improving the quality of insights derived from machine learning research while addressing reluctance 192

towards resource-intensive evaluation processes. This balance between efficiency, simplicity, and signal richness presents EEVE's value proposition for advancing machine learning research.

Working Principle of EEVEE: EEVEE works by building a *meta-model* over the performance metrics of models sufficient both in number and diversity, on the full benchmark pool from which we want to choose our subset. With the term *benchmark* in this paper we refer to a dataset + task pairs.

Formally, given a large benchmark pool  $B = \{b_0, b_1, \ldots, b_K\}$ , where *B* is the full set of benchmarks, and  $b_i$  are individual benchmarks therein, we have a sufficiently large and diverse pool of model performance metrics  $M = \{m_0^0, m_1^0, \ldots, m_K^N\}$ . Here,  $m_i^j$  is the performance metric of model *j* on benchmark  $b_i$ . We aim to discover a subset of *B* of size *k*. This means *k* total benchmarks make up the subset. If we build a meta-model  $g(M_{selected}, \theta)$  to predict all of *M* given only the selected subset  $M_{selected}$ , it should minimize the following loss:

$$L_{EEVEE} = MSE(M, g(M_{selected}, \theta)) \tag{1}$$

In this equation, MSE is the mean squared error. M represents the full set of performance metrics of all our models on the full benchmark pool B. The term  $g(M_{selected}, \theta)$  represents the predictions of the meta-model g with parameters  $\theta$  when it is given the performance metrics of all models from the selected subset of benchmarks  $B_{selected}$ , referred to as  $M_{selected}$ .

However, our main focus lies in the selected combination of performance metrics  $M_{selected}$  that can generalize well on previously unseen models. To that end, we must split M into train, validation and test sets, each consisting of performance metrics acquired from different models (e.g. train  $\rightarrow$  ResNet50, ViT-Base, CLIP, and val  $\rightarrow$  ResNext50, DINO, DeIT), and explicitly optimize the inner loop test loss rather than the training loss, while we use the validation loss to select the best meta-model for test. Hence the loss we wish to minimize is:

$$L_{EEVEE}^{test} = MSE(M^{test}, g(M_{selected}^{test}, \theta))$$
<sup>(2)</sup>

We need a non-differentiable method for choosing the k benchmarks in  $M_{selected}$ , since brute force becomes intractable very quickly, so we employ evolutionary methods to learn the k selected benchmarks.

This results in a bi-level optimization, with an evolutionary method on the outer loop  $e(B_{selected})$ , 218 where e is the evolutionary method, and  $B_{selected}$  are the benchmarks being selected – or indeed, the 219 genes being optimized, and a small meta-model parameterized as a neural network  $g(\theta)$  that receives 220 a train/val split from B<sub>selected</sub> and trains itself to do the task described in Equation 1, after which 221 process it is scored using the val set using the loss in Equation 2. Then, once a given candidate of 222 benchmarks  $B_{selected}$  is scored, in this way, the outer loop performs a tournament selection where 223 only the top 50 candidates are preserved and mutated by removing one benchmark at random, and 224 adding another at random. Each winning candidate mutates into 10 children, and the parent is 225 also preserved in the gene pool, producing a gene pool with 550 candidates for every cycle. At 226 initialization, we sample 1000 random combinations. We have found that 1000 is a good starting 227 228 population that is both tractable to score and facilitates the necessary diversity that enables limited variation in results across several runs, showcasing convergent behaviour. diversity that our results 229 across runs have little variation from one another, pointing to a convergent behaviour. We include full 230 pseudocode showcasing all the details related to how we performed EEVEE for our experiments in 231 Algorithm 1, 2 and 3 in Figure 1 232

#### 233 Applying EEVEE on Model Encoder Generalization

Why Model Encoder Evaluation? A common practice across machine learning applications involves 234 augmenting general model encoders with task-oriented heads. The adaption of this paradigm can 235 be attributed to the computational efficiency associated with training model encoders, over more 236 expensive setups. Much of computer vision, as well as vision to text search and retrieval happen using 237 model encoders. Similarly, various applications requiring translation from one domain/modality/task 238 to another require an encoder of some sort. Even the "decoder-only" LLM models that have 239 demonstrated incredible capabilities in the last few years, internally can be seen as a series of 240 representation encoders, a series of refinement before they reach the decoding stage. 241

Multi-modal early fusion is another 242 topic closely related with model en-243 coders - as research in early fusion 244 can be done most efficiently when try-245 ing to learn data encoders rather than 246 a full encoder-decoder, or decoder-247 only models. World model research 248 in multi-modal dimensions can also 249 take place most efficiently within a 250 model-encoder context. Recent works 251 like I/VJEPA [2] for example have 252 paved the way for self-supervised 253 learning which functions using model 254 encoders, and has been demonstrated 255 to be more efficient and more gener-256 alizable than full pixel decoding vari-257 ants. 258

Algorithm 1 Scoring Require: Performance metrics M, Input metrics  $M_{\text{indext}}$ . Epochs E = 20, Hidden dimension  $d_{\text{hidden}} = 100$ , Learning rate  $\alpha = 0.01$ , Weight decay  $\lambda = 0.01$ , Opti-mizer type  $\omega = ^{\lambda} \text{dam}W$ . Ensure: Evaluation score mean(scores) 1: Convert data to tensors  $x = M_{\text{selected}}$  and y = M2: Normalize x and y = MAlgorithm 3 Evolution Require: Performance metrics  $M = \{m_1^1, m_1^2, \dots, m_N^n\}$ Benchmark set B. Combination size k. Number of wim ners W. Number of children per winner C. Number of generations G. Initial combinations size 1, Training epochs E. Hidden dimension d<sub>adden</sub> = 100. Learning rate  $\alpha = 0.01$ . Weight decay  $\lambda = 0.01$ . Optimizer type  $\sim_0 = "\lambda dnmW"$ Initialize ShuffleSplit cross-validation k f spocns rate  $\alpha = 0.01$ = "AdamW Initialize empty list scores for each train, val split in kf do re: Evolved benchmark combinations B. Divide x into  $x_{\text{train}}$  and  $x_{\text{val}}$ ; y into  $y_{\text{train}}$  and  $y_{\text{val}}$ Build meta-model  $g(\theta)$  with hidden dimension  $d_{\text{train}}$ 1: Initialize initial combinations Binitial with I random sam The finitial combinations  $B_{\text{initial}}$  with T random samples from B of size kEvaluate performance of  $B_{\text{initial}}$  using SCOR ING $(M, B_{\text{initial}}, E, d_{\text{hidden}}, \alpha, \lambda, \omega)$  and store scores in Buta intel-mode g(t) Win maden aumension dasheds
 Train g(t) on xr<sub>min</sub> and y<sub>min</sub> for E products with learn ing rate α, weight decay λ, and optimizer ω
 Predict y<sub>pend</sub> = g(x<sub>mi</sub>, θ)
 Compute mean squared error score = MSE(y<sub>pend</sub> + y<sub>min</sub>)
 Append score to scores
 end for
 return mean(scores) 2: select top W combinations from S as  $B_{\text{winners}}$ or open eration a = 1 to G do r generation g = 1 to G do Initialize a new set of combinations  $B_{new}$ for each combination  $B_{selected} \in B_{winner}$ Add  $B_{selected}$  to  $B_{new}$ new do Fraction combination  $B_{selected} \in B_{winners}$  do Add  $B_{selected}$  to  $B_{new}$ for each child c = 1 to C do Mutate  $B_{selected}$  using MUTATION( $B_{selected}$ , B) to create a new combination  $B'_{selected}$ 8: 10: Add  $B'_{selected}$  to  $B_{new}$ Algorithm 2 Mutation end for 11: 12: 13: **Require:**  $B_{\text{selected}} \subset B, B = \{b_1, b_2, \dots, b_K\}$  **Ensure:** New  $B'_{\text{celected}}$ end for Evaluate performance of  $B_{\text{new}}$  using SCOR-ING(M,  $B_{\text{new}}$ , E,  $d_{\text{hidden}}$ ,  $\alpha$ ,  $\lambda$ ,  $\omega$ ) and store scores 14 15 16 ect top W combinations from S as B<sub>winner</sub> return B<sub>w</sub> end while ted by replacing bremove with badd Create  $B'_{selec}$ return  $B'_{sele}$ 

Furthermore, model encoder evalua-259 tion has been quite diffused in the past 260 few years, with new benchmarks be-

261

277

Figure 1: (a) EEVEE Scoring algorithm, Mutation algorithm, and (b) Evolution algorithm.

ing produced in every facet of the machine learning field. Nonetheless, most of those lacked in some 262 263 key quality: they were either simply too complex to use efficiently, requiring too much compute, or, more often than the others, missing a unifying software framework that can easily, in a user-conscious 264 way, and a principled stance towards high readability, maintainability and hackability. 265

The goal of focusing on Model Encoder Eval-266 uation: By applying EEVEE to search for a 267 pareto-optimal set of benchmarks, and packag-268 ing it up in a unified framework that is built for 269 the researcher in mind from the ground up, one 270 which offers out of the box automated down-271 loading, pipeline building, task adapters, and a 272 very mature training and eval loop. Within this 273 framework, we facilitate, all relevant logging in-274 formation, including key training and eval met-275 rics, rich gradient information, power and com-276 putational information, as well as visualizations



Figure 2: GATE Framework Pipeline

278 where relevant. Finally, we support easy switching of model encoders, no matter what source modality

they come from – our framework dubbed GATE is a one stop shop for ones model representation 279 research needs, both during research, debugging, as well as at the evaluation phase. 280

GATE comes in three tiers small, base and big-GATE. Each having 8, 15 and 21 benchmarks within it, 281 and targetted towards 12/24 and 36 GPU hours on a A100 40GB. We hope that by making it very easy 282 for the end user and offering such rich signal for machine learning research, many researchers will 283 284 choose to use GATE, to enhance their research signal, whilst keeping the compute budgets relatively 285 feasible.

Preparations: Choosing Models, Benchmarks and Adaptation Processes: EEVEE will vield 286 better results if the space of models, benchmarks and adaptation processes we use is diverse, but also 287 thorough in numbers. A. Adaptation Process We wanted GATE to cover multiple domains, tasks 288 289 and modalities when shifting from the source to the target setting. For that reason we decided that if a model encoder has an input layer that does not fit the target modality, we simply remove that input 290 layer and replace it with a relevant ViT-like patchification [12] followed by a linear combination for 291 each patch. For tasks where we have text, we would tokenize the text using BPE [51], and for tasks 292 where we have video we would use the model encoder on each image, to acquire an image-level 293 vector representation, and then follow that up with a simple 4 layer transformer that receives a 294 sequence of image-vector tokens, to produce a video-level embedding, on top of which we apply the 295 task-specific head at hand. The task-adapters we used leaned on established methods, and where 296 possible we just used a transformer head, which includes segmentation, relational reasoning and 297 video classification, with everything just using a linear head, full details available at 14. After these 298

modifications, described in Figure 2, we use a fine tuning scheme – this decision was informed by preliminary experiments on both full fine tuning and linear probe with a frozen backbone, in which we found that there was a clear superiority of fine tuning over linear probing for the benchmarks we chose in our pool. Full details of these preliminary experiments can be found in Appendix 8.1. In our preliminary experiments we were able to identify three recipes, one for ConvNet-style architectures, one for ViT-style architectures and one for Hybrid architectures such as ConvNext and ResNext that worked well for all tasks, details in 8.1.

**B.** Model Pool We wanted the space of models used to cover many important pretraining schemes, 306 architectures, and source modalities. The details of these choices are provided next: 1. Pretraining 307 Task and Dataset Variation: With a consistent architecture, models were subjected to various 308 pretraining tasks and datasets. Model instances representing this category include CLIPViT [43], 309 ConvNextV2 [35], Siglip, FlexViT [7], LaionViT, ImageNet1K ViT [11] with Random Aug-310 ment, SAM-ViT, DiNoViT, EfficientFormerV2 [32] and DeiT3 [59]. Further to these, we include 311 models initialized from scratch, specifically, ViT, ResNet50 [18], FlexViT, EfficientNetV2 [57], 312 and then fine-tuned on the GATE tasks. 2. Architectural Variation: We explored models having the 313 same pretraining dataset (ImageNet), but differing in their architecture. This group encompassed a 314 mix of standard CNN models such as EffNetV2, ResNet50, ResNext50 [67], ConvNextV2\_Base 315 [35] and transformer-based models like EfficientFormer [32] and FlexViT [7]. 3. Modality 316 and Dataset Variation: This axis comprised models trained on modalities other than vision such 317 as Whisper, coming from an audio to text task and Bert [10], Bart [31] and Mpnet [55] coming 318 from various text-based tasks. These models had their original input processing systems replaced by 319 320 a Vision Transformer style embedding and were subsequently fine-tuned on the GATE tasks. A more comprehensive account of these models, including their selection rationale and unique characteristics, 321 is provided in the Appendix Section 13. 322

C. Benchmark Pool The benchmark pool, detailed in the Appendix, includes Image Classification 323 (ImageNet1k [9], CIFAR100 [28], Places365 [74], Food101 [36], HappyWhale [17]), Few Shot 324 Image Classification (Aircraft [37], Fungi [50], MiniImageNet [62], CUB200 [63], Describable 325 Features [69]), Zero Shot Text-Image Classification (Flickr30K [41], New Yorker Caption Context 326 [20], Winoground [58]), Visual Relational Reasoning (CLEVR [23], CLEVRMath [34]), Image 327 Semantic Segmentation (ADE20K [75], COCO10K [33], COCO164K [33], NYU-Depth-v2 [54], 328 PascalContext [38], Cityscapes [8]), Medical Image Classification (Chexpert [21], Diabetic Retinopa-329 thy [16], HAM10000 [60]), Medical Segmentation (ACDC [5]), Video Classification (HMDB51 [30], 330 UCF-101 [56], Kinetics400 [24]) and Video Regression (iWildcam [4]). 331

Producing Diverse Model Performance Metrics: We apply our adaptation process on each and every model chosen, on every benchmark in the benchmark pool. To acquire test results we ensemble by averaging logits of the top 1, 3 and 5 validation models to produce three separate ensemble results.

**D. Experimental Approach** We wanted our research environment to reflect the end user, so we 335 can properly understand their needs, and to offer a *pragmatic* experimental setup of in-the-wild 336 researchers with little time to hyperparameter optimize, and which have to make decisions on small 337 amounts of preliminary experiments - someone choosing a model encoder off the shelf and adapting it 338 to downstream setting. For that reason, we kept any hyperparameter tuning, or human attention when 339 it came to specific models to a minimum. Instead, we relied on existing good recipes, and did some 340 preliminary experiments as explained in detail in 8.1. Briefly, we discovered specific adjustments 341 for each architecture type: for Convolutional Architectures, we used AdamW with a learning rate of 342 1e-3, and 6e-4 for segmentation tasks; for Vision Transformer Architectures, AdamW with a learning 343 rate of 1e-5; and for Convolutional + Transformer Hybrid Architectures, AdamW with a learning rate 344 of 2e-5. A plateau learning rate scheduler was configured with parameters like mode "min", factor 345 346 0.5, patience 1000, and threshold 1e-4, allowing models to effectively choose their own schedules based on their learning progress. This adaptive scheduling facilitated "good enough" learning rates 347 and enhanced performance across different architectures. 348

#### 349 4 Results

Single Benchmark Predictiveness: As demonstrated in Figure 3, using EEVEE we quantified the predictive power of each benchmark on its own, when not in a combination with others. We have found that ADE20K, Flickr30K, and the New York Caption Competition lead in their predictive power, with few-shot tasks, and relational reasoning, being very close to the best in predictive power. ImageNet1K sits squarely in the middle of the competition. Furthermore, some of the most "novel"



(a) Small-GATE (k=8, 12 GPU (b) Small-GATE (k=15, 24 GPU (c) Small-GATE (k=21, 24 GPU hour) tier hour) tier hour) tier

Figure 4: Degradation of predictive power when a given benchmark is removed and the meta-model trained from scratch, for different GATE tiers.

benchmarks like iwildcam, happy whale, ACDC, NYU and Winoground are the least predictive tasks, 355 Winoground being magnitudes less predictive. We argue that this is mainly due to the tasks being 356 "harder", and our models being less designed for those. The results in WinoGround were bearly better 357 than chance for example. However, when once we move to combinations of benchmarks, these 'less' 358 predictive benchmarks become key contributors to better predictive power, as they represent edge 359 cases, as can be seen in Figures 6g 7c, 7i, where these have the highest importance when removed 360 from a given set. 361 K>=2

Predictiveness of Discovered Combinations In 362 Figure 5, we can see how the top-50 performing 363 candidate combinations perform as we vary the 364 number of benchmarks per combination from 365 1 to 26. We can see that there is a point of di-366 minishing returns around the k = 8 point, after 367 which there appears to be some "overfitting" oc-368 369 curing. We verified that the overfitting was a 370 result of having a small sample number of 20 models, to train, val and test our meta-models 371 with, as well as the 2-layer MLP we used to 372 model Few-to-All metric predictions. We tried 373



0.04

Figure 5: Performance of Models build with Kbest datasets: We do a search over the space of all k for EEVEE and box plot the population summary statistics of the top 50 combination candidates.

our level best to find the best architecture and regularization schemes for our meta-model, and this 374 was the best we could do given available compute and (human) time. We chose 8, 15, and 21 as 375 the combination-threshold to make our packs out of as they satisfied the computational budgets 376 we set for ourselves, and they have very diverse and predictive tasks, as can be seen in Figures 6g 377 7c, 7i. For full details on all the discovered top-k combinations please look at Appendix Section 378 16.1. Best Models based on GATE: As can be seen in Table 2, or the Appendix extended Table 379 3, the best overall models are ConvNextV2, SigLIP and CLIP in that order, with SigLIP and CLIP 380 often exchanging ranks between themselves. However, it is worth noting that EfficientNetV2 381 demonstrated exceptional performance/compute across all tasks, and even outperformed all models in 382 many medical tasks. Finally, ConvNet based models, and particularly ResNext50 seem to have done 383 exceptionally well in the edge-case scenarios of ACDC, Happy Whale Individual identification, and 384

$\mathbf{Metric} \downarrow   \ \mathbf{Model} \rightarrow$	cvnxtv2	siglip	clip	flex	deit	laion	vit	dino	smvit	rnx50	effv2	r50a1	effrmr	seffv2	sflex	svit	whspr	sr50a1	bert bart	mpnet
Img Class																				
CIFAR-100 Acc@1	84.2	74.6	76.9	75.1	66.7	75.1	66.6	55.7	50.3	69.3	67.3	34.3	15.6	37.6	10.3	7.8	11.0	15.9	$14.5 \ 9.0$	1.0
Food-101 Acc@1	92.9	91.6	93.3	89.1	87.3	91.4	86.5	84.8	75.7	86.1	86.4	69.4	61.6	36.5	24.5	25.8	17.0	16.3	$18.7 \ 11.6$	8.5
HWhale Individual Acc@1	75.6	31.7	35.2	48.4	23.7	21.0	27.5	9.1	3.6	78.7	77.1	5.2	4.4	33.2	2.8	2.5	2.2	2.1	2.3 1.7	1.5
HWhale Species Acc@1	99.8	99.8	99.7	99.8	99.5	99.7	99.7	99.2	95.4	99.7	99.7	92.1	92.8	96.5	76.5	74.5	64.3	65.8	71.2 59.3	62.9
ImageNet-IK Acc@1	85.3	81.9	76.0	82.3	82.1	74.1	68.3	77.9	75.5	77.6	73.5	72.5	44.6	16.9	3.2	2.4	2.2	1.3	1.5 0.8	0.2
Diseas 265 Ass@1	54.7	90.8	93.7	90.0	94.7	59.7	89.1 47 E	93.0	90.8	51.0	91.4 E1.E	90.5	12.0	31.3	10.1	8.2	7.5	4.7	5.2 3.2	1.2
Tack Mean	84.7	75.6	75.6	77.5	49.0	72.6	41.0	41.5	50.8	70.5	78.1	40.9	45.9	40.7	10.5	18.6	16.0	15.0	17.0 12.6	11.1
Few-Shot Img Class	04.2	10.0	10.0	11.0	11.0	12.0	03.5	00.7	00.0	10.0	10.1	01.0	40.2	40.7	13.0	10.0	10.0	10.0	17.0 12.0	11.1
Aircraft Acc@1	96.7	96.6	974	95.9	95.3	96.7	96.3	94.4	92.9	91.6	90.6	86.2	78.2	59.2	54.9	50.4	55.1	58.2	61.2 60.8	57.2
CUBirds Acc@1	98.0	97.9	97.2	96.4	96.2	96.6	95.9	94.4	93.4	92.8	92.1	89.4	86.3	52.5	50.0	45.2	44.4	31.9	48.4 50.3	48.5
DTextures Acc@1	85.0	85.2	88.6	78.9	81.9	86.1	80.8	79.4	81.9	77.7	60.3	77.2	68.5	46.6	50.2	50.5	50.0	33.1	44.6 49.8	38.3
Fungi Acc@1	85.8	85.6	85.7	83.7	80.6	85.2	81.3	77.4	77.7	74.1	73.7	67.1	59.2	27.6	38.0	37.0	33.9	28.2	$32.9 \ 33.8$	7.6
Mini-Imagenet Acc@1	97.0	96.2	93.1	99.1	98.8	90.8	89.9	98.7	92.9	94.1	63.2	93.2	90.9	36.7	45.9	47.2	44.8	34.2	39.7 37.3	36.8
Omniglot Acc@1	98.6	98.9	99.0	98.9	98.7	98.9	98.8	98.6	98.6	98.5	98.7	95.5	95.8	98.2	93.4	93.6	82.9	80.5	90.2 84.1	90.7
VGG Flowers Acc@1	99.7	98.9	98.6	96.7	96.2	97.0	95.9	95.5	93.4	87.9	91.3	89.3	90.6	59.6	69.4	69.4	63.0	53.4	59.1 59.4	60.8
Task Mean	94.4	94.2	94.2	92.8	92.5	93.1	91.3	91.2	90.1	88.1	81.4	85.4	81.4	54.3	57.4	56.2	53.4	45.6	53.7 53.6	48.6
Img Seg	46.9	47.1	44.0	49.7	97.0	49.4		<u></u>	95.0	10.9	14.9	11.7	0.0	15	0.5	0.4	0.6	0.4	04.05	0.4
ADE20K mIOU	40.8	4/.1	44.0	43.1	31.8	43.4	62.0	33.3 61.4	20.9	18.2	14.2	40.9	9.8	1.5	0.0	0.4	17.1	19.6	0.4 0.5	0.4
COCO-10K mIoU	26.9	39.5	35.6	35.1	32.8	33.6	29.8	31.0	28.6	18.4	10.2	-40.2	14.0	1.1	0.9	20.0	0.4	1.6	0.1 1.3	0.1
COCO-164K mIoU	32.7	36.7	33.8	33.0	30.5	32.4	27.0	28.9	25.7	16.8	9.7	47	13.7	1.0	0.7	0.7	0.5	0.7	0.1 1.0	0.1
NYU mIoU	7.5	7.7	7.8	6.9	12.2	5.7	6.1	12.1	11.0	5.9	8.3	6.4	10.5	6.8	3.5	3.7	2.9	7.2	5.4 5.0	5.4
Pascal mIoU	32.8	34.8	35.7	30.6	31.4	28.3	27.5	29.8	24.0	16.6	11.7	6.8	14.0	1.7	1.3	1.1	1.4	2.3	1.0 1.4	0.9
Task Mean	34.8	39.3	37.4	36.2	34.8	35.2	31.3	32.8	29.1	19.5	19.7	12.6	10.8	9.8	4.9	5.0	3.8	5.1	1.6 1.9	1.6
Img Relational																				
CLEVR Acc@1	52.5	52.7	52.7	52.1	52.6	52.6	52.8	52.8	51.6	50.1	40.6	49.3	45.2	39.3	46.1	45.9	46.4	44.9	$42.6 \ 42.5$	41.2
CLEVR Colour	35.4	36.1	36.4	35.0	35.5	35.6	35.3	36.1	34.2	26.8	15.7	24.7	14.7	12.5	25.7	29.4	28.8	22.8	$13.2 \ 13.0$	13.2
CLEVR Count	45.8	45.8	45.8	45.9	45.8	45.7	45.7	45.6	45.6	45.3	39.0	45.1	44.8	37.9	45.1	44.7	44.8	44.9	44.7 44.7	43.0
CLEVR Material	60.5	60.6	60.5	60.0	60.5	60.6	61.4	61.3	60.2	58.6	52.1	57.5	53.7	49.8	53.7	51.7	54.0 96.1	53.0	49.8 50.5	49.9
CLEVR Shape	02.1 61.0	02.4 61.1	02.0	51.1 60.7	02.2	02.4 c0.9	52.9	62.2	49.9	50.2	34.3 59.5	50.2	44.8	50.6	30.8	54.9	50.1	54.0	34.0 33.7	50.1
CLEVR Size	60.7	60.5	60.8	60.6	60.5	60.7	60.4	60.4	60.9	59.8	53.3	59.9	59.6	51.4	60.1	59.2	59.5	59.8	59.5.59.3	58.6
CLEVR-Math Acc@1	79.3	65.9	68.8	59.9	73.7	62.9	60.5	59.3	58.3	55.6	44.0	56.0	56.6	30.2	46.9	46.5	46.2	45.7	44.8 42.1	36.4
Task Mean	55.9	54.4	54.9	53.1	55.2	53.9	53.9	53.6	52.6	50.8	41.6	50.1	46.9	38.1	46.2	45.9	46.4	45.0	42.9 42.5	40.7
Medical Class															-					
Chexpert APS Macro	61.6	61.0	61.2	62.6	62.3	60.9	61.2	59.9	61.5	59.8	60.2	54.1	55.2	48.0	33.9	34.1	34.3	35.7	$36.9 \ 33.7$	33.0
Chexpert AUC Macro	82.5	82.5	82.3	83.2	82.9	82.5	82.4	81.8	82.8	81.1	81.9	79.1	79.9	74.7	64.7	65.1	65.5	67.0	$67.6 \ 65.3$	64.9
Chexpert BS Macro	84.3	84.4	84.5	85.1	86.2	84.6	84.9	85.6	87.0	86.3	84.8	86.1	86.4	84.6	82.9	82.9	83.0	83.1	83.1 82.8	82.8
Diabetic APS Macro	56.9	57.2	56.4	56.3	54.2	56.4	54.4	51.9	45.2	55.6	58.7	35.5	36.6	20.6	21.6	21.5	22.5	23.3	22.4 21.2	21.3
Diabetic AUC Macro	8/.5	80.7	80.0	80.7	85.0	80.3	84.7	02.6	81.2	85.0	80.1	10.0	19.0	01.6	01.2	01.4	01.4	01.3	01 8 01 6	01.6
HAM10K APS Macro	94.5	94.0	93.9	95.9	95.8	93.0	93.7	95.0	83.4	87.9	87.1	43.7	46.9	38.8	38.0	35.0	32.2	48.5	50.6.37.6	32.6
HAM10K AUC Macro	99.1	98.6	08.7	98.5	98.6	98.6	98.7	98.5	97.8	97.9	97.5	89.3	90.1	85.6	86.1	84.6	82.8	91.0	91.1 85.9	83.3
HAM10K BS Macro	98.4	98.1	97.8	98.1	98.0	97.9	97.9	97.9	97.2	97.6	97.2	95.2	95.5	94.6	94.5	94.4	94.3	95.0	95.2 94.4	94.2
Task Mean	84.4	84.0	83.6	83.9	83.6	83.6	83.3	82.6	81.0	82.9	83.1	72.4	73.6	65.8	63.2	62.9	62.6	66.3	$66.4 \ 63.1$	62.0
Medical Seg																				
ACDC Dice Score	63.1	48.1	51.3	45.9	43.8	48.0	50.4	47.7	44.6	44.2	61.0	40.2	18.7	46.0	16.5	18.5	32.2	28.7	$23.2 \ 26.2$	25.3
Task Mean	63.1	48.1	51.3	45.9	43.8	48.0	50.4	47.7	44.6	44.2	61.0	40.2	18.7	46.0	16.5	18.5	32.2	28.7	23.2 26.2	25.3
Img to 1xt ZS	<i>c</i> 0	<b>C D</b>	7.0	5.0	F (C	0.0	r 0	r 0.	4.5	4.1	0.7	4.77	4.0	1.0	1.0		1.0		10 10	1.0
Flickr30K Img21xt	0.3 5.7	5.0	6.0	5.9	5.1	6.5	5.9 6.0	5.1	4.0	4.1	3.7	4.7	4.2	1.0	1.8	2.0	2.9	2.0	1.9 1.8	1.0
NYCC Img2Txt	6.9	6.6	6.9	5.8	6.5	6.9	6.4	6.0	4.7	4.9	4.1	4.6	4.2	1.6	2.1	1.8	1.9	2.0	20 16	1.6
NYCC Txt2Img	6.1	5.9	6.4	5.5	6.0	6.2	6.4	5.8	4.8	4.3	4.1	3.9	3.7	1.6	2.0	1.7	2.0	2.4	1.9 1.8	1.6
Winoground Img2Txt	51.0	53.4	59.5	49.7	50.0	50.3	49.5	43.5	53.8	61.9	50.0	48.9	47.3	43.9	50.0	41.3	50.0	53.2	49.6 50.1	50.4
Winoground Txt2Img	50.0	55.2	56.2	53.1	50.0	55.5	48.3	54.2	48.6	54.8	50.0	49.6	52.4	52.8	50.0	54.2	51.8	52.2	51.8 48.8	52.1
Task Mean	21.0	22.2	23.7	20.9	20.5	22.0	20.4	20.0	20.2	22.3	19.3	19.3	19.3	17.2	18.0	17.2	18.3	19.0	$18.2 \ 17.6$	18.1
Video Class																				
HMDB-51 Acc@1	52.5	40.7	40.6	32.2	39.3	24.9	27.4	32.8	33.1	5.6	11.5	1.8	2.1	3.8	8.3	7.9	6.1	5.4	6.4 7.5	4.0
Kinetics Acc@1	48.8	44.2	51.4	43.7	40.3	44.6	33.2	36.4	25.8	2.7	1.0	0.2	0.3	0.4	2.0	1.6	1.0	0.5	0.3 0.3	0.3
UCF-101 Acc@1	84.4	75.1	69.9	63.2	75.0	63.4	58.8	66.6	48.7	19.7	11.1	2.8	0.8	2.1	15.2	13.3	6.6	8.7	6.5 7.0	2.7
Task Mean	61.9	55.5	54.0	40.4	51.5	44.3	39.8	40.2	35.9	9.4	1.8	1.0	1.1	2.1	8.0	1.0	4.0	4.9	4.4 4.9	2.3
IWildCam MAE Score	55.2	53.1	56.0	54.9	54.1	46.1	52.1	49.1	45.3	34.6	35.8	37.3	13.9	29.6	41.3	30.3	36.3	40.3	27.5.38.7	20.2
Task Mean	55.2	53.1	56.0	54.9	54.1	46.1	52.1	49.1	45.3	34.6	35.8	37.3	13.9	29.6	41.3	39.3	36.3	40.3	27.5 38.7	20.2
GATE	00.2	50.1	20.0	54.0		10.1		10.1	10.0	94.9	30.0	51.5	10.0	20.0		50.0	50.0	-0.0	2.1.0 00.1	20.2
Full GATE Mean	69.0	66.8	66.8	64.6	64.3	63.4	62.1	62.2	58.5	56.3	54.4	48.4	42.8	39.6	37.5	37.2	36.2	36.9	35.0 34.9	31.8
Big GATE Mean	76.6	74.5	74.4	72.8	72.0	71.9	70.6	70.0	66.8	66.7	64.8	58.5	53.1	46.8	43.8	43.4	41.9	41.5	40.9 39.8	37.1
Base GATE Mean	68.3	65.6	65.7	62.6	63.7	60.7	60.2	60.7	58.6	55.1	53.5	48.2	42.8	38.0	36.5	36.3	35.4	36.6	$34.8 \ 34.8$	30.4
Small GATE Mean	77.7	74.9	74.6	73.3	72.4	71.2	68.9	69.1	65.3	65.7	61.7	58.5	49.3	40.5	35.7	35.4	35.9	35.3	$34.1 \ 34.4$	30.4
Full GATE Rank	1.0	3.0	2.0	4.0	5.0	6.0	8.0	7.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0	18.0	17.0	19.0 20.0	21.0
BIG GALE Kank	1.0	2.0	3.0	4.0	5.0	5.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	15.0	17.0	18.0	19.0 20.0	21.0
Small GATE Rank	1.0	2.0	3.0	4.0	5.0	6.0	8.0	7.0	10.0	9.0	11.0	12.0	13.0	14.0	16.0	17.0	15.0	18.0	20.0 19.0	21.0

Table 2: Summary of experiments: Black/Bold best model, Green second best, Blue third best, and red the worst performing model. Models prefixed with 's' refer to 'from scratch' trained models, rather than pretrained. For the full table look at Appendix Table 3

385 general medical tasks, which indicates perhaps some sort of learning efficiency advantages related to

386 their inductive biases.

**Limitations:** We empirically evaluate EEVEE on a relatively large pool of models and benchmarks,

<sup>388</sup> however, with more models, and benchmarks it could yield much more general results. Especially

with benchmarks targetting the text and audio modalities, as well as potentially offline RL.

# 390 5 Conclusion

In this paper, we propose EEVEE, an evolutionary-method-based search algorithm that can discover 391 392 out of a large collection of benchmarks, the ones that can offer the most predictive value on the original collection, for a given set of models. We apply EEVEE on the task of model-encoder 393 evaluation in the context of images, image-text, videos, and medical domains. As a result, we obtain 394 the GATE Benchmark, which consists of 3 tiers, each targeted to a particular GPU budget, from 12, 395 24 and 36 GPU hours, per model evaluation. We then introduce the GATE engine, which takes these 396 benchmarks, and offers a researcher-designed environment in which one can easily port their own 397 model encoder, and run the full GATE tiers, and automatically produce a variety of performance, 398 energy/power, hardware utilization metrics and task visualizations. We evaluated 20 representative 399 models ranging from image, image-text, text and audio pretrained models, on the GATE tiers, and we 400 discovered that ConvNextV2 and SigLIP seem to lead the pack overall, with EfficientNetV2 being an 401 exceptional, efficient alternative for the medical domain and for *unique scenario* tasks, such as Happy 402 Whale, ACDC and WinoGround. Finally, ConvNet based models, and ResNext50 in particular, seem 403 to have a lot more *learning* efficiency, as they are the best adapted models on very novel domains, 404 such as Happy Whale individual prediction challenge, ACDC and medical tasks. 405

#### 406 **References**

- 407 [1] Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra,
   408 C Lawrence Zitnick, and Devi Parikh. Vqa: Visual question answering. In *Proceedings* 409 of the IEEE International Conference on Computer Vision (ICCV), pages 2425–2433, 2015.
- [2] Adrien Bardes, Quentin Garrido, Jean Ponce, Xinlei Chen, Michael Rabbat, Yann LeCun,
   Mahmoud Assran, and Nicolas Ballas. Revisiting feature prediction for learning visual representations from video, 2024.
- [3] Loic Barrault, Ondrej Bojar, Marta R Costa-jussa, Christian Federmann, Mark Fishel, Yvette
   Graham, Barry Haddow, Matthias Huck, Philipp Koehn, Shervin Malmasi, Christof Monz, et al.
   Findings of the 2019 conference on machine translation (wmt19). In *Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1)*, pages 1–61, 2019.
- [4] Sara Beery, Grant Van Horn, and Pietro Perona. The iwildcam 2018 challenge dataset. In
   *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* Workshops, pages 54–60, 2018.
- [5] Olivier Bernard, Alain Lalande, Caio Zotti, Florence Cervenansky, Xin Yang, Pheng-Ann
   Heng, Ismail Cetin, Karim Lekadir, Oscar Camara, Miguel A Gonzalez Ballester, et al. Deep
   learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is
   the problem solved? *IEEE Transactions on Medical Imaging*, 37(11):2514–2525, 2018.
- 424 [6] Lucas Beyer, Olivier J. Hénaff, Alexander Kolesnikov, Xiaohua Zhai, and Aäron van den Oord.
   425 Are we done with imagenet?, 2020.

[7] Lucas Beyer, Pavel Izmailov, Alexander Kolesnikov, Mathilde Caron, Simon Kornblith, Xiaohua
 Zhai, Matthias Minderer, Michael Tschannen, Ibrahim M. Alabdulmohsin, and Filip Pavetic.
 Flexivit: One model for all patch sizes. 2023 IEEE/CVF Conference on Computer Vision and
 Pattern Recognition (CVPR), pages 14496–14506, 2022.

- [8] Marius Cordts, Mohamed Omran, Sebastian Ramos, Tobias Rehfeld, Markus Enzweiler, Rodrigo
   Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic
   urban scene understanding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3213–3223, 2016.
- [9] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Li Kai, and Li Fei-Fei. Imagenet: A large scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern
   recognition, pages 248–255. Ieee, 2009.
- [10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of
   deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*,
   2018.
- [11] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai,
   Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly,
   Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image
   recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [12] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai,
   Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly,
   Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image
   recognition at scale. In *International Conference on Learning Representations (ICLR)*, 2021.
- Linus Ericsson, Henry Gouk, and Timothy M. Hospedales. How well do self-supervised
   models transfer? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5414–5423, June 2021.
- [14] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman.
   The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88(2):303–338, 2010.

- Iohn S Garofolo, Lori F Lamel, William M Fisher, Jonathan G Fiscus, David S Pallett, and
   Nancy L Dahlgren. Timit acoustic-phonetic continuous speech corpus ldc93s1, 1993.
- [16] Varun Gulshan, Lily Peng, Marc Coram, Michael C Stumpe, Derek Wu, Arunachalam
  Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Travis Madams, Jorge Cuadros,
  et al. Development and validation of a deep learning algorithm for detection of diabetic
  retinopathy in retinal fundus photographs. *JAMA*, 316(22):2402–2410, 2016.
- <sup>460</sup> [17] Happywhale. Happywhale whale and dolphin identification challenge, 2022.
- [18] 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),
   pages 770–778. IEEE, 2016.
- [19] Dan Hendrycks, Collin Burns, Saurav Kadavath, Akul Arora, Steven Basart, Mantas He, Dawn
   Song, and Jacob Steinhardt. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*, 2020.
- <sup>467</sup> [20] Jack Hessel. New yorker caption contest corpus, 2023.
- [21] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute,
   Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large
   chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the* AAAI Conference on Artificial Intelligence, 33:590–597, 2019.
- [22] Niehues Jan et al. Iwslt 2017: Proceedings of the 14th international workshop on spoken
   language translation. 2017.
- Iustin Johnson, Bharath Hariharan, Laurens van der Maaten, Li Fei-Fei, C Lawrence Zitnick,
   and Ross Girshick. Clevr: A diagnostic dataset for compositional language and elementary
   visual reasoning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2901–2910, 2017.
- [24] Will Kay, Joao Carreira, Karen Simonyan, Brian Zhang, Chloe Hillier, Sudheendra Vijaya narasimhan, Fabio Viola, Tim Green, Trevor Back, Paul Natsev, et al. The kinetics human
   action video dataset. In *arXiv preprint arXiv:1705.06950*, 2017.
- [25] Philipp Koehn. Europarl: A parallel corpus for statistical machine translation. *MT summit*, 5:79–86, 2005.
- Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay
   Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Phillips, Irena Gao, et al. Wilds: A
   benchmark of in-the-wild distribution shifts. In *International Conference on Machine Learning*,
   2021.
- [27] Simon Kornblith, Jonathon Shlens, and Quoc V. Le. Do better imagenet models transfer better?
   In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), June 2019.
- [28] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images.
   Technical Report UTML TR 2009, University of Toronto, Toronto, Ontario, Canada, 2009.
- [29] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep
   convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.
- [30] Hildegard Kuehne, Hueihan Jhuang, Estibaliz Garrote, Tomaso Poggio, and Thomas Serre.
   Hmdb: A large video database for human motion recognition. In 2011 International Conference
   on Computer Vision (ICCV), pages 2556–2563. IEEE, 2011.
- [31] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed,
   Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. Bart: Denoising sequence-to-sequence
   pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*, 2020.

- [32] Xiuyu Li, Yutong Yuan, Shu Chen, Martin Danelljan, Radu Timofte, and Luc Van Gool.
   Efficientformer: Vision transformers at mobilenet speed. *arXiv preprint arXiv:2206.01191*, 2022.
- [33] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr
   Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European Conference on Computer Vision*, pages 740–755, 2014.
- [34] Adam Dahlgren Lindström and Savitha Sam Abraham. Clevr-math: A dataset for compositional language, visual and mathematical reasoning. *ArXiv*, abs/2208.05358, 2022.
- [35] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining
   Xie. A convnet for the 2020s. *arXiv preprint arXiv:2201.03545*, 2022.
- [36] Matthieu Guillaumin Lukas Bossard and Luc Van Gool. Food-101 mining discriminative
   components with random forests. In *European Conference on Computer Vision*, 2014.
- [37] Subhransu Maji, Esa Rahtu, Juho Kannala, Matthew Blaschko, and Andrea Vedaldi. Fine grained visual classification of aircraft. In 2013 IEEE Conference on Computer Vision and
   Pattern Recognition (CVPR), pages 554–561. IEEE, 2013.
- [38] Roozbeh Mottaghi, Xiaobai Chen, Xiaofeng Liu, Nam-Gyu Cho, Seong-Whan Lee, Sanja
   Fidler, Raquel Urtasun, and Alan Yuille. The role of context for object detection and semantic
   segmentation in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 891–898, 2014.
- [39] Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. Librispeech: An
   asr corpus based on public domain audio books. In 2015 IEEE International Conference on
   Acoustics, Speech and Signal Processing (ICASSP), pages 5206–5210, 2015.
- [40] Denis Paperno, Germán Kruszewski, Angeliki Lazaridou, Ngoc Quan Pham, Raffaella Bernardi,
   Sandro Pezzelle, Marco Baroni, Gemma Boleda, and Raquel Fernández. The lambada dataset:
   Word prediction requiring a broad discourse context. In *Proceedings of the 54th Annual Meeting* of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1525–1534,
   2016.
- [41] Bryan Plummer, Liwei Wang, Chris Cervantes, Juan Caicedo, Julia Hockenmaier, and Svetlana
   Lazebnik. Flickr30k entities: Collecting region-to-phrase correspondences for grounded image
   descriptions. *arXiv preprint arXiv:1505.04870*, 2015.
- [42] Bryan A Plummer, Liwei Wang, Christopher M Cervantes, Juan C Caicedo, Julia Hockenmaier,
   and Svetlana Lazebnik. Flickr30k entities: Collecting region-to-phrase correspondences for
   richer image-to-sentence models. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2641–2649, 2015.
- [43] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal,
   Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual
   models from natural language supervision. In *International conference on machine learning*,
   pages 8748–8763. PMLR, 2021.
- [44] Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, and Samy Bengio. Transfusion: Understanding
   transfer learning for medical imaging. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*,
   volume 32. Curran Associates, Inc., 2019.
- [45] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. Squad: 100,000+ questions
   for machine comprehension of text. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392, 2016.
- 547 [46] Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, and Vaishaal Shankar. Do cifar-10
   548 classifiers generalize to cifar-10?, 2018.

- [47] Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, and Vaishaal Shankar. Do ImageNet
   classifiers generalize to ImageNet? In *Proceedings of the 36th International Conference on Machine Learning*, Proceedings of Machine Learning Research. PMLR, 2019.
- [48] Erik F Sang and Fien De Meulder. Introduction to the conll-2003 shared task: Language independent named entity recognition. In *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003*, pages 142–147, 2003.
- [49] Adam Santoro, David Raposo, David G.T. Barrett, Mateusz Malinowski, Razvan Pascanu, Peter
   Battaglia, and Timothy Lillicrap. A simple neural network module for relational reasoning. In
   *Advances in Neural Information Processing Systems (NeurIPS)*, pages 4967–4976, 2017.
- [50] Dirk Schroeder, Yin Cui, Yang Chai, Daniel Kristensen, Evangelos Kalogerakis, and Serge
   Belongie. The fgvcx fungi classification challenge. In *CVPR Workshops*, 2018.
- [51] Rico Sennrich, Barry Haddow, and Alexandra Birch. Neural machine translation of rare
   words with subword units. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1715–1725, 2016.
- [52] Vaishaal Shankar, Achal Dave, Rebecca Roelofs, Deva Ramanan, Benjamin Recht, and Ludwig
   Schmidt. Do image classifiers generalize across time? In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9661–9669, October 2021.
- <sup>566</sup> [53] Ali Shirali, Rediet Abebe, and Moritz Hardt. A theory of dynamic benchmarks, 2023.
- [54] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and
   support inference from rgbd images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 746–760, 2012.
- [55] Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. Mpnet: Masked and permuted
   pre-training for language understanding. *arXiv preprint arXiv:2004.09297*, 2020.
- [56] Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah. Ucf101: A dataset of 101 human
   actions classes from videos in the wild. In *arXiv preprint arXiv:1212.0402*, 2012.
- [57] Mingxing Tan and Quoc V Le. Efficientnetv2: Smaller models and faster training. arXiv
   preprint arXiv:2104.00298, 2021.
- [58] Tristan Thrush, Hongyu Jiang, Goutham Prasad, and Jacob Andreas. Winoground: Prob ing vision and language models for visio-linguistic compositionality. *arXiv preprint arXiv:2204.03162*, 2022.
- [59] Hugo Touvron, Piotr Bojanowski, Mathilde Caron, Matthieu Cord, Alaaeldin El-Nouby,
   Edouard Grave, Armand Joulin, Gabriel Synnaeve, and Jakob Verbeek. Deit iii: Revenge
   of the vit. *arXiv preprint arXiv:2204.07118*, 2022.
- [60] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection
   of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*,
   5:180161, 2018.
- [61] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez,
   Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Informa- tion Processing Systems (NeurIPS)*, pages 5998–6008, 2017.
- [62] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Koray Kavukcuoglu, and Daan Wierstra.
   Matching networks for one shot learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3630–3638, 2016.
- [63] Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The caltech ucsd birds-200-2011 dataset. Technical Report CNS-TR-2011-001, California Institute of
   Technology, 2011.

- [64] Alex Wang, Yada Pruksachatkun, Nikita Nangia, Amanpreet Singh, Julian Michael, Felix Hill,
   Omer Levy, and Samuel R Bowman. Superglue: A stickier benchmark for general-purpose
   language understanding systems. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 3266–3280, 2019.
- [65] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel R Bowman.
   Glue: A multi-task benchmark and analysis platform for natural language understanding. In
   *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, pages 353–355, 2018.
- [66] Jason Weston, Antoine Bordes, Sumit Chopra, Alexander M Rush, Bart van Merrienboer,
   Armand Joulin, and Tomas Mikolov. Towards ai-complete question answering: A set of
   prerequisite toy tasks. *arXiv preprint arXiv:1502.05698*, 2015.
- [67] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual
   transformations for deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1492–1500. IEEE, 2017.
- [68] Junichi Yamagishi, Christophe Veaux, and Kirsten MacDonald. Cstr vctk corpus: English
   multi-speaker corpus for cstr voice cloning toolkit (version 0.92). University of Edinburgh. The
   *Centre for Speech Technology Research (CSTR)*, 2019.
- [69] Fisher Yu, Yinda Zhang, Shuran Song, Ari Seff, and Jianxiong Xiao. Fine-grained visual
   comparisons with local learning. In 2014 IEEE Conference on Computer Vision and Pattern
   *Recognition (CVPR)*, pages 192–199. IEEE, 2014.
- [70] Xiaohua Zhai, Joan Puigcerver, Alexander Kolesnikov, Pierre Ruyssen, Carlos Riquelme, Mario
   Lucic, Josip Djolonga, Andre Susano Pinto, Maxim Neumann, Alexey Dosovitskiy, et al. A
   large-scale study of representation learning with the visual task adaptation benchmark. In
   *International Conference on Learning Representations*, 2020.
- [71] Chi Zhang, Feng Gao, Baoxiong Jia, Song-Chun Zhu, and Yixin Zhu. Raven: A dataset for
   relational and analogical visual reasoning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5317–5327, 2019.
- [72] Guanhua Zhang and Moritz Hardt. Inherent trade-offs between diversity and stability in multi-task benchmarks, 2024.
- [73] Kaizhi Zheng, Xiaotong Chen, Odest Chadwicke Jenkins, and Xin Eric Wang. VLMbench: A
   compositional benchmark for vision-and-language manipulation. In *Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks*, 2022.
- [74] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. Places: A 10
   million image database for scene recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 40, pages 1452–1464. IEEE, 2017.
- [75] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba.
   Scene parsing through ade20k dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5122–5130, 2017.
- [76] Yuke Zhu, Olaf Groth, Michael S Bernstein, and Li Fei-Fei. Visual7w: Grounded question
   answering in images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4995–5004, 2016.

# 635 6 End-user Guidelines

636 For an end-user to use GATE, they need to
---

- 1. Install the GATE framework python package, as described in the Github repo's readme page.
- Choose a path for implementing the new foundation model encoder they wish to evaluate.
  This is either cloning the full GATE repo and modifying existing components directly,
  or, importing the GATEncoder and GATEModel classes from GATE, and wrapping up
  their model within it. Doing so requires the researcher to implement a relevant forward
  function that can take in the modalities their model needs to process, as well as defining a
  configuration that tells GATE what modalities a model can receive and output features on,
  as well as any transforms needed for a batch to be ready for their model.
- 3. The user chooses a GATE tier to use (from smallGATE, baseGATE and bigGATE). Based
   on the configuration defined by the user in step 2.
- GATE generates a list of commands, each representing an experiment that needs to be run, and can then run these commands on your local GPU box, parallelizing the tasks, one on each available GPU, or, can provide a list of commands or json file that one can use to run these commands on a GPU cluster, or other hardware.
- GATE emits a wandb project, with metrics, visualizations and other measures, allowing easy
   tracking of experiments, and sharing thereof, as well as huggingface model weights for each
   model being trained which is also used to achieve a *stateless* execution.
- 654 6. Once the experiments are completed, one can invoke the produce-analysis.py file within 655 GATE to get tables and figures that analyse the data, similar to what appears in this paper. 656 Those results can then be used to report results in a paper, or, be used to make decisions for 657 production models.

This process ensures the GATE framework is aware of what a model's supported modalities are, as 658 well as how to produce modality-specific features, given the model. Once this is completed, the user, 659 with a single line of code, can select a GATE tier, and launch all jobs needed to produce results for that 660 tier. Importantly, GATE is made to facilitate and encourage foundation models that are diverse in their 661 capabilities, and allow the researchers to focus on what matters - that is, designing and training their 662 foundation model - rather than spending the majority of their time building and optimizing evaluation 663 boilerplate. Furthermore, the diversity of signal that GATE provides allows better understanding of a 664 given model's strengths and weaknesses, which as a result makes the research, review and iteration 665 process of the field as a whole more efficient. This is because there is a consistent boilerplate that 666 runs all models, with broad signal that reduces probability of making erroneous conclusions - both in 667 the overly optimistic, or overly pessimistic side of things. 668

# 669 6.1 Principal Use Cases

- 1. Model Development and Iteration: GATE serves as a valuable tool during the model 670 research and development phase. By integrating the model into GATE and running either 671 the smallGATE or baseGATE tiers, developers can obtain a comprehensive and robust 672 performance evaluation of their model across diverse domains, tasks, and modalities. Worth 673 noting that GATE allows easy inclusion of foundation models pretrained on images, video, 674 audio, text, etc, to be fine-tuned on pixel-based tasks. It achieves this by replacing a 675 model's root layer / embedding layer, with one appropriate for a given task's modality, and 676 adding on top a relevant task adapter head. 677
- 2. Model Evaluation for Machine Learning Research: GATE enhances the communication 678 of research findings and their potential applications, a vital aspect of scientific collabo-679 ration. By using GATE as a benchmark, even at the most cost-efficient GPU hour level 680 of smallGATE, the clarity and depth of future ML papers can be significantly improved. 681 GATE's explicit evaluation of modality, domain, and task shifts in a given foundation model 682 provides a nuanced and informative perspective on a model's true capabilities, offering a 683 more detailed understanding of a model's strengths and weaknesses than optimizing a single 684 metric, such as ImageNet validation error. 685

# 686 7 Result Extras

The results were logged in WandB, and then further processed after all experiments were completed to generate the tables and figures in this paper. Much of the logged information outside of testing metrics were not used for any of the figures and tables in this paper. The full set of experiments and all the logged results can be found at our wandb gate project repo<sup>2</sup>.

# 691 7.1 Result Processing

Once all experiments were completed, we queried our wandb project repository and returned test results from all our experiments, if an experiment name was duplicated, we used the latest entries, and, for each experiment type there existed three independent runs. We averaged the results of any metrics across such independent runs to acquire a better approximation to the true performance of those models.

# **697** 8 Preliminary Experiments Details

# 698 8.1 Preliminary Experiments

First, we trained models on ImageNet1k, CIFAR100, CLEVR, ADE20K, CityScapes, and, ACDC
 for 5K iterations, using cosine annealing learning schedule or plateau annealing, with AdamW,
 weight decays varying from 0.1 - 0.0001, and applied models from each major architecture category –
 specifically, the CLIPViT, ImageNet pretrained ViT, ResNext, ResNet and ConvNextV2. The results
 from these experiments pointed to the fact that there exists one general and good recipe for each
 architecture style. The recipes that we discovered were as follows:

#### 705 8.1.1 Across Architecture Settings

- <sup>706</sup> Unless otherwise stated, the settings here are applied universally in all experiments.
- **Optimizer:** AdamW, weight decay 0.01, plateau annealing with patience 1000, relative scaling and scale factor 0.5, and, threshold 0.0001.
- **Training Details**: Training iterations: 10K, validate every 500 iterations.
- **Test Details**: Top-3 validation models (across all validated checkpoints) are ensembled by prediction
   averaging.

#### 712 8.1.2 Architecture Specific Settings

- Convolutional Architectures: Optimizer: AdamW, learning rate 1e-3, and for segmentation tasks
   only, we used learning rate 6e-4
- 715 Vision Transformer Architectures: Optimizer: AdamW, learning rate 1e-5
- 716 Convolutional + Transformer Hybrid Architectures Optimizer: AdamW, learning rate 2e-5
- The above recipes were what we used throughout all our experiments unless otherwise stated.

# 718 **9** GATE Guiding Principles

- The fundamental values driving the design decisions behind GATE are the following:
- Maximizing Generalization Signal: GATE is designed to provide a high signal-to-noise ratio concerning a model's ability to generalize in diverse downstream contexts, that vary in domain, task and modality. This allows for a more robust assessment of a model's capacity for adaptation and versatility. By noise here we refer to how clear a given signal response is. For example, an image classification test accuracy signal on ImageNet, would provide clear

<sup>&</sup>lt;sup>2</sup>omitted until double blind is over

- r25 signal with respect to the natural domain and the classification task, but would be blurry for
   r26 more compositional, object disentanglement and relational tasks, such as segmentation, or,
   r27 visual question answering.
- 728
   2. Time Efficiency: Acknowledging the importance of computational resources and time, GATE operates within set benchmarks of 12, 24, and 36 GPU hours (established on A100 @ 40GB). These set timeframes ensure GATE's assessments are both thorough and expedient.
- Minimizing Usage Friction: The framework supporting GATE is designed to be user-friendly,
   enabling easy integration of new backbones and facilitating smooth experimentation. This
   low-friction approach ensures a streamlined experience when using GATE, making the
   process of evaluation more efficient.

We argue that a good balance of the above can generate a pragmatic, yet thorough foundation model evaluation suite, that will, importantly, be of real use to most researchers in the field.

# **10 Defining the GATE Benchmark**

GATE is a comprehensive evaluation engine designed to advance the development of more general
 machine learning models. It improves on existing benchmarks by enabling the evaluation of models
 across diverse modalities, domains, and tasks.

GATE is composed of three key components. The first is a benchmark *pool*, a broad collection of 741 datasets, tasks, and processes that measure a model's performance across various domains, tasks, 742 and modalities. The second component is a set of benchmark *tiers*, which are meticulously curated 743 subsets from the GATE benchmark pool, tailored to specific compute budgets and project phases. 744 The final, and is a software framework, designed to seamlessly integrate new foundation models and 745 execute the GATE tiers, thereby enabling efficient performance evaluation across a diverse range of 746 downstream modalities, domains, and tasks. Practically, GATE is directed towards machine learning 747 researchers and developers as a means to efficiently, and with little friction, get broad signal about 748 how their model performs after transfer in diverse contexts, specifically selected for their empirically 749 evaluated high signal-to-noise ratio with respect to predictive power in how a model performs in 750 previously unseen contexts. 751

<sup>752</sup> Building GATE was a careful balancing act. We needed to respect specific time budgets while also <sup>753</sup> aiming for a wide variety of evaluation scenarios. Our approach was as follows:

- Select a diverse set of learning contexts, spanning multiple domains, tasks and modalities.
   We refer this as the *Benchmark Pool*.
- 2. Select a broad set of key foundation models, varying in their architecture, pretraining scheme
   and source modality. We refer to this as the *Model Pool*.
- Fine tune each of the models in the model pool, on each of the contexts in the benchmark
   pool. Evaluate trained models on each context's test sets.
- 4. Use the test set results acquired to quantify the predictive power each benchmark holds with
   respect to previously unseen benchmarks, both at the individual level and the collection
   level. We call this measure, the *downstream generalization predictability measure* (DGPM).
- 5. Use the DGPM values of the various combinations of benchmarks to build the three GATE tiers, selecting combinations of benchmarks that can provide the most information within a target time budget.
- <sup>766</sup> We elaborate on each of the above steps in the following subsections.

# 767 **11 Benchmark Pool Selection Details**

Medical Image Classification: Medical data are known to present a substantial shift in both domain and even modality depending on their format. We have selected datasets that not only pose significant challenges for foundation models but also align with the broader imperative to deliver real-world benefits downstream. *Chexpert*: A dataset comprising a challenging array of chest x-rays annotated with findings critical to
 diagnosing thoracic diseases. It tests models on their ability to navigate complex, multi-label medical
 data, encapsulating the kind of nuanced decision-making that AI must augment in clinical settings.

Diabetic Retinopathy Classification: Early detection of diabetic retinopathy from retinal images
 is a public health priority; models fine-tuned on this dataset can have immediate implications for
 preventing vision loss on a global scale. This dataset requires models to decipher fine-grained,
 progressive changes indicative of the disease, reflecting the precision necessary for medical AI
 applications.

HAM10000 (Human Against Machine with 10000 dermatoscopic images): The dataset provides
 a diverse spectrum of skin lesion images vital for differentiating between benign and malignant
 conditions. Incorporating this dataset not only challenges the pattern recognition prowess of AI but
 also contributes to the advancement of dermatology through machine learning technologies.

Metrics: We collect Average Precision Score (APS), Area Under the Receiver Operating Char acteristics Curve (AUC), and Brier Score (BS) both overall (i.e. macro) as well as for individual
 pathologies/classes.

Medical Segmentation: This category evaluates foundational models' ability to generalize from
 natural to medical image modalities and to perform domain-specific tasks that require precision and
 complex spatial understanding:

ACDC (Automated Cardiac Diagnosis Challenge): This dataset is aimed at assessing models' generalization to the medical domain, particularly the transferability of representations for segmenting anatomical structures in cardiac MRI images. By focusing on the heart's intricate anatomy, ACDC tests the models' ability to adapt to clinically relevant shapes and patterns—a shift from common visual recognition tasks to precise medical delineation. Metrics: We collect dice loss, mIoU, mean accuracy and overall accuracy.

# 796 **12 Benchmark Pool Details**

Having a set of diverse benchmarks ranging in challenge factor, as well as modality, task and domain
 shift was key. We explain in more detail why why consider these factors important in Appendix in
 more detail. We refer to this as our *benchmark pool*, and it consists of the following:

**Image Classification:** We employ **ImageNet1k** [9], **CIFAR100** [28], **Places365** [74], and **Food101** [36] to cover diverse natural image domains. Additionally, we include **HappyWhale** [17] for a more challenging domain shift, aiding in wildlife research and providing an interesting test case for model evaluation.

Few Shot Image Classification: We use the MetaDataset task recipe on the Aircraft [37], Fungi
 [50], MiniImageNet [62], CUB200 [63], and Describable Features [69] datasets to evaluate task
 and domain shift robustness for an evaluation model.

Zero Shot Text-Image Classification: Another key setting is that of zero-shot text-image classification, on which many current key models were trained and evaluated [43]. We utilize Flickr30K, New
 Yorker Caption Context (a challenging humor task), and Winoground–a task requiring the model to match two texts with their corresponding images, focusing on compositional differences.

**Visual Relational Reasoning**: A context where earlier models, such as ResNet50 [18] had low performance without layers with associative inductive biases (e.g., relational neural networks or transformers [49, 61]). This ensures we are aware of any trade-offs in relational compositional abilities in our models. We use **CLEVR** [23] and **CLEVRMath** [34].

Image Semantic Segmentation: Essential for various real-world applications, serving as an indicator
of a model's ability to retain spatial information and identify objects at a per-pixel level. ADE20K
[75], COCO10K [33], COCO164K [33], NYU-Depth-v2 [54], PascalContext [38], and Cityscapes
[8].

Medical Image Classification: Medical data exhibit substantial domain and modality shifts, posing
 significant challenges for machine learning models while aligning with the imperative to deliver
 real-world benefits.*Chexpert* [21] (chest X-rays annotated for thoracic disease diagnosis), *Dia*-

*betic Retinopathy Classification* [16] (retinal images for early detection of diabetic retinopathy),
 *HAM10000* [60] (dermatoscopic images for differentiating skin lesions).

Medical Segmentation  $\rightarrow$  *ACDC* (*Automated Cardiac Diagnosis Challenge*) [5]: This dataset assesses models' generalization to the medical domain, particularly the transferability of representations for segmenting anatomical structures in cardiac MRI images. By focusing on the heart's intricate anatomy, ACDC tests the models' ability to adapt to clinically relevant shapes and patterns.

Video Classification: Video classification tasks test models on their temporal generalization abilities
 and require an understanding of not only individual frame content but also the transition and context
 between frames. *HMDB51 (Human Motion Database)* [30], *UCF-101 (University of Central Florida - 101 action categories)* [56], *Kinetics400* [24].

Video Regression: Where classification tasks gauge categorical distinctions, video regression tasks
 assess models' ability to make continuous numerical predictions from temporal data, serving as an
 indicator of a model's capability to process and quantify dynamic content. *iWildcam (International Wildlife Camera Trap Challenge)* [4]: This dataset targets estimating animal species abundance from
 videos and is a direct test of modality and task shift, and showcases a models' potential impact on
 ecological monitoring and species conservation efforts.

- Modality shifting contexts: Contexts where the foundation model is asked to learn to do
   well at a task that requires understanding of a previously unseen modality. More specifically,
   assuming a foundation model has been trained on natural images, this would be transferring
   to medical imaging, video, audio and test contexts. This would shed light on the performance
   of a model's middle layers.
- 2. Task shifting contexts: Contexts where a model is tasked with performing a previously
   unseen task, for example, transferring from classification to segmentation or relational
   reasoning.
- 3. Domain shifting contexts: Contexts where a model is required to perform a task on a domain that is different from the one it was trained on. For example moving from natural images on ImageNet at 224x224 resolution to black and white Omniglot characters at 28x28 resolution, or, moving from ImageNet to images of fungi. More extreme domain shifts would be going from natural images to medical images for example.

# **851 13 Model Pool Details**

# **14 Task Adapter Details**

# **853 15** Experimental Details

Experimental Environment Details: GPUs: 4 x A6000 Ada @ 48GB, CPUs: 128 Core AMD
EPYC 7713 64-Core Processor, RAM: 1 TB, HD: 15TB NVME. All experiments were done with
BF16 precision.

# **857 16 Additional Results**

# **16.1** Full details on discovered combinations





(a) Best k=2 discovered combina-(b) Best k=3 discovered combina- (c) Best k=4 discovered combination tion





tion



(d) Best k=5 discovered combina- (e) Best k=6 discovered combination tion





(f) Best k=7 discovered combination



(g) Best k=8 discovered combina- (h) Best k=9 discovered combina- (i) Best k=10 discovered combination tion tion



(j) Best k=11 discovered combina- (k) Best k=12 discovered combina- (l) Best k=13 discovered combination tion tion

Figure 6: Degradation of predictive power when a given benchmark is removed and the meta-model trained from scratch, for different best combinations in varying k.

$\mathbf{Metric} \downarrow    \mathbf{Model} \rightarrow$	cvnxtv2	siglip	clip	flex	deit	laion	vit	dino	smvit	rnx50	effv2	r50a1	effrmr	seffv2	sflex	svit	whspr	sr50a1	bert	bart	mpnet
Img Class																					
CIFAR-100 Acc@1	84.2	74.6	76.9	75.1	66.7	75.1	66.6	55.7	50.3	69.3	67.3	34.3	15.6	37.6	10.3	7.8	11.0	15.9	14.5	9.0	1.0
CIFAR-100 Acc@5	97.4	93.8	95.1	94.4	90.9	93.9	89.7	83.6	80.1	91.9	90.7	65.9	42.3	67.6	30.6	25.5	31.6	40.2	38.1	29.2	5.0
Food-101 Acc@1	0.0	0.9 91.6	0.8	0.9 89 1	1.2	0.9 91.4	1.2	1.0 84.8	75.7	1.2 86 1	1.5 86.4	2.5	5.5 61.6	2.4	3.9 24 5	25.8	3.9	3.0	3.7 18.7	4.0	4.0
Food-101 Acc@5	99.0	98.7	99.1	98.1	97.8	91.4 98.7	97.4	97.0	93.5	97.2	97.1	91.0	86.6	66.1	51.0	52.8	41.1	38.9	43.0	32.2	26.1
Food-101 Loss	0.3	0.3	0.2	0.4	0.4	0.3	0.5	0.5	1.0	0.6	0.6	1.1	1.5	2.6	3.2	3.1	3.6	3.6	3.5	3.9	4.1
HWhale Individual Acc@1	75.6	31.7	35.2	48.4	23.7	21.0	27.5	9.1	3.6	78.7	77.1	5.2	4.4	33.2	2.8	2.5	2.2	2.1	2.3	1.7	1.5
HWhale Individual Acc@5	84.6	49.5	53.9	64.5	40.9	37.9	46.0	22.0	11.0	86.7	83.6	14.8	11.9	52.5	9.2	8.1	6.9	6.8	7.6	5.7	5.4
HWhale Individual Loss	1.6	4.6	4.3	3.6	4.9	5.1	4.7	5.9	6.7	1.3	1.5	6.4	6.6	3.9	7.0	7.1	7.3	7.3	7.2	7.5	7.4
Hwhale Species Acc@1	99.8	<b>99.8</b>	99.7	99.8	99.5	99.7	99.7	99.2	95.4	100.0	99.7 100.0	92.1	92.8	96.5	70.0 06.1	05.8	04.3	02.8	04.2	59.3 80.8	02.9
HWhale Species Loss	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.3	0.2	0.1	0.8	0.8	1.2	1.1	0.9	1.4	1.2
ImageNet-1K Acc@1	85.3	81.9	76.0	82.3	82.1	74.1	68.3	77.9	75.5	77.6	73.5	72.5	44.6	16.9	3.2	2.4	2.2	1.3	1.5	0.8	0.2
ImageNet-1K Acc@5	96.8	95.8	93.7	95.5	94.7	93.1	89.1	93.0	90.8	93.3	91.4	90.5	72.5	37.3	10.1	8.2	7.7	4.7	5.2	3.2	1.2
ImageNet-1K Loss	0.6	0.8	1.0	0.8	0.8	1.1	1.3	1.0	2.3	1.0	1.2	1.1	2.8	4.3	6.0	6.1	6.1	6.5	6.4	6.6	6.8
Places365 Acc@1	54.7	53.5	54.1	52.1	49.0	53.7	47.5	47.3	27.1	51.8	51.5	40.9	25.2	26.6	9.0	8.6	7.5	5.0	5.2	3.0	2.2
Places365 Acc@5	85.3	84.1	84.7	83.3	80.8	84.3	79.9	79.5	59.9 2 1	82.9	82.6	73.5	55.2	55.5	26.3	25.0	22.4	16.4	16.4	5.2	9.0
Task Mean	88.0	79.6	80.1	81.9	76.1	76.9	2.0 74.8	70.8	63.5	84.6	83.4	62.4	51.0	52.2	4.5 29.1	28 1	25.5	25.5	26.5	21.4	17.8
Few-Shot Img Class	0010	1010	00.1	0110	1011	1010	1 110	10.0	00.0	01.0	00.1	02.1	0110	02.2	2011	2011	20.0	2010	20.0	2111	1110
Aircraft Acc@1	96.7	96.6	97.4	95.9	95.3	96.7	96.3	94.4	92.9	91.6	90.6	86.2	78.2	59.2	54.9	50.4	55.1	58.2	61.2	60.8	57.2
Aircraft Loss	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.3	0.3	0.4	1.2	0.4	311.5	44.1	2.1	2.1	1.6	2.3	2.5	1.2	1.6
CUBirds Acc@1	98.0	97.9	97.2	96.4	96.2	96.6	95.9	94.4	93.4	92.8	92.1	89.4	86.3	52.5	50.0	45.2	44.4	31.9	48.4	50.3	48.5
CUBirds Loss	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.3	0.3	0.4	0.5	0.4	33.7	2.5	3.6	3.5	2.3	8.8	3.2	2.0	1.6
D lextures Acc@l	85.0	85.2	88.6 0 5	78.9	81.9	86.1	80.8	79.4	81.9	77.7	60.3	77.2	68.5 3.6	46.6	50.2 2 E	50.5	50.0 24	$\frac{33.1}{5.0}$	44.6	49.8	38.3
Fungi Acc@1	85.8	0.7 85.6	0.5 85.7	1.1 83.7	0.9 80 6	0.7 85.2	1.1 81 3	1.2 77 4	0.9 77 7	74 1	14.3 73.7	0.0 67 1	59.2 59.2	$^{1.8}_{27.6}$	⊿.ə 38.0	2.1 37.0	⊿.4 33.0	$\frac{3.0}{28.2}$	⊿.0 32.9	1.9 33 8	1.4 7.6
Fungi Loss	0.6	0.6	0.6	0.7	0.8	0.6	0.8	0.9	0.8	1.1	5.8	1.1	1031.2	2.6	2.2	2.2	2.2	2.4	2.4	2.3	2.9
Mini-Imagenet Acc@1	97.0	96.2	93.1	99.1	98.8	90.8	89.9	98.7	92.9	94.1	63.2	93.2	90.9	36.7	45.9	47.2	44.8	34.2	39.7	37.3	36.8
Mini-Imagenet Loss	0.1	0.1	0.3	0.0	0.0	0.3	0.4	0.1	0.2	0.3	23.7	0.3	0.6	2.4	1.6	1.6	1.6	2.1	1.8	1.9	1.9
Omniglot Acc@1	98.6	98.9	99.0	98.9	98.7	98.9	98.8	98.6	98.6	98.5	98.7	95.5	95.8	98.2	93.4	93.6	82.9	80.5	90.2	84.1	90.7
Omniglot Loss	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.1	0.3	0.2	0.6	0.7	0.4	0.6	0.3
VGG Flowers Acc@1	99.7	98.9	98.6	96.7	96.2	97.0	95.9	95.5	93.4	87.9	91.3	89.3	90.6	59.6	69.4	69.4	63.0	53.4	59.1	59.4	60.8
VGG Flowers Loss	0.1	0.1	0.1	0.2	0.2	0.1	0.2	0.2	0.2	0.4	0.5	0.4	0.3	1.6	1.8	1.0	1.4 52.4	4.2	2.5	1.0	1.5
Img Seg	74.4	94.4	94.2	92.0	92.0	93.1	91.5	91.2	90.1	86.1	01.4	00.4	01.4	54.5	57.4	30.2	55.4	45.0	55.7	55.0	40.0
ADE20K CE Loss	1.1	1.0	1.1	1.1	1.3	1.0	1.3	1.4	1.7	2.0	2.2	2.8	2.8	3.3	3.8	3.8	3.7	3.7	3.7	3.7	3.8
ADE20K Focal Loss	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.5	0.6	0.6	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9
ADE20K Mean Acc@	59.8	60.8	57.5	56.0	49.1	57.3	44.2	45.1	36.3	26.8	20.4	17.9	15.2	3.6	1.6	1.6	1.8	1.8	1.8	1.8	1.8
ADE20K Overall Acc@	71.8	74.4	72.6	71.4	66.9	72.4	64.2	63.5	57.5	49.6	43.9	34.6	39.7	21.3	11.7	11.9	13.1	14.1	14.0	14.4	14.2
ADE20K mIoU	46.8	47.1	44.0	43.7	37.8	43.4	33.2	33.3	25.9	18.2	14.2	11.7	9.8	1.5	0.5	0.4	0.6	0.4	0.4	0.5	0.4
Cityscapes Eocal Loss	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.4	0.2	0.4	4.1	0.3	0.7	0.7	0.9	0.9	3.9	4.0	3.8
Cityscapes Overall Acc@	92.5	94.2	93.9	93.6	93.1	93.7	93.4	93.1	92.8	88.5	93.2	87.4	41.5	90.4	78.1	78.6	72.2	75.4	47 4	37.7	47.3
Cityscapes mIoU	62.3	69.8	67.6	67.5	63.9	67.7	63.9	61.4	59.5	40.8	64.2	40.2	2.5	46.7	22.8	23.5	17.1	18.6	2.7	2.0	2.7
COCO-10K CE Loss	3.0	1.3	1.5	1.4	1.5	1.4	1.5	1.6	1.6	2.1	2.6	3.3	3.5	3.6	4.5	3.8	4.0	3.6	4.1	3.7	4.1
COCO-10K Focal Loss	0.7	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.6	0.8	0.8	0.8	1.1	0.9	0.9	0.8	1.0	0.9	1.0
COCO-10K Mean Acc@	38.8	50.6	47.2	46.0	43.4	44.9	41.2	43.5	40.7	27.0	15.8	8.2	20.9	2.2	1.7	1.9	1.3	2.9	0.6	2.5	0.6
COCO-10K Overall Acc@	57.9	69.8	66.4	66.0	64.4	65.9	62.8	63.1	61.2	51.3	40.1	23.3	45.2	20.9	15.2	20.5	14.7	24.6	9.4	22.5	9.3
COCO-10K mIoU	26.9	39.5	35.6	35.1	32.8	33.6	29.8	31.0	28.6	18.4	10.2	5.7	14.0	1.1	0.9	0.8	0.4	1.6	0.1	1.3	0.1
COCO-164K Eccal Loss	1.9	1.4	1.0	1.0	1.0	1.5	1.0	1.7	1.8	2.2	2.7	0.8	1.0	0.0	4.5	0.0	4.0	4.0	4.2	0.0	4.2
COCO-164K Mean Acc@	45.9	50.1	46.9	45.3	42.6	44.5	38.6	43.0	38.7	25.4	14.7	7.0	21.3	2.0	1.5	1.9	1.5	1.8	0.6	2.5	0.7
COCO-164K Overall Acc@	60.9	65.8	63.5	63.0	60.3	63.2	59.5	59.1	55.6	47.9	39.3	20.3	39.3	19.2	13.6	19.4	15.6	18.3	9.5	21.7	9.6
COCO-164K mIoU	32.7	36.7	33.8	33.0	30.5	32.4	27.0	28.9	25.7	16.8	9.7	4.7	13.7	1.0	0.7	0.7	0.5	0.7	0.1	1.1	0.1
NYU CE Loss	2.5	1.5	2.0	2.3	1.5	2.5	2.3	1.5	1.6	1.6	1.8	1.6	1.4	1.6	1.6	1.6	1.7	1.5	1.5	1.5	1.5
NYU Dice Score	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.7	0.8	0.8	0.8	0.8
NYU Focal Loss	0.5	0.2	0.4	0.5	0.3	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.2	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2
NYU Mean Acc@	19.7	21.5	20.8	19.6	128	19.4	19.7	23.0	22.9	18.5	18.3	12.7	18.9	21.0	24.6	24.6	24.2	13.0	27.2	11.7	27.4
NYU mIoU	7.5	7.7	7.8	6.9	12.2	5.7	6.1	12.1	11.0	5.9	8.3	6.4	10.5	6.8	3.5	3.7	2.9	7.2	5.4	5.0	5.4
Pascal CE Loss	1.0	0.5	0.5	0.6	0.9	0.5	0.8	0.8	0.9	1.4	1.5	2.2	3.1	2.3	2.4	2.4	2.4	2.4	2.6	2.5	2.6
Pascal Dice Loss	0.8	0.6	0.4	0.5	0.5	0.4	0.5	0.5	0.4	0.5	0.4	0.5	0.5	0.2	0.4	0.4	0.4	0.4	0.5	0.5	0.4
Pascal Focal Loss	0.2	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.2	0.2	0.3	0.4	0.7	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Pascal Loss	1.4	0.5	0.1	0.3	0.3	0.4	0.3	0.6	0.6	0.5	0.4	1.4	4.2	1.6	1.6	1.6	1.6	1.6	3.4	1.7	3.5
Pascal Mean Acc@	42.2	43.5	44.2	39.6	38.8	37.4	34.7	40.3	29.1	20.7	16.2	10.6	18.0	3.5	3.1	2.8	3.3	4.5	2.6	3.3	2.5
Pascal mIoU	(0.1	8/.0 3/ 8	357	80.0 30.6	11.0 31.4	80.0	78.9 27.5	79.5	24.0	16.6	11 7	49.7	14.0	37.3	34.2	35.4	37.4	39.0	34.4	30.3	32.3
Task Mean	44.1	49.6	47.1	46.4	45.1	45.7	41.8	43.6	39.9	31.9	28.5	21.2	24.0	17.0	13.1	13.9	12.7	14.7	10.0	11.2	9.9
Img Relational			-	-	-		-						-		-						
CLEVR Acc@1	52.5	52.7	52.7	52.1	52.6	52.6	52.8	52.8	51.6	50.1	40.6	49.3	45.2	39.3	46.1	45.9	46.4	44.9	42.6	42.5	41.2
CLEVR Colour Acc@1	35.4	36.1	36.4	35.0	35.5	35.6	35.3	36.1	34.2	26.8	15.7	24.7	14.7	12.5	25.7	29.4	28.8	22.8	13.2	13.0	13.2
CLEVR Colour Loss	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.6	1.9	2.1	2.0	2.1	2.1	2.0	1.9	1.9	2.0	2.1	2.1	2.1
CLEVR Count Acc@1	45.8	45.8	45.8	45.9	45.8	45.7	45.7	45.6	45.6	45.3	39.0	45.1	44.8	37.9	45.1	44.7	44.8	44.9	44.7	44.7	43.0
CLEVE Count Loss	1.1 60 5	1.2	1.1 60 5	1.2	1.2 60 5	1.2	1.2	1.2	1.2	1.2 58.6	1.3 59 1	1.2 57 5	1.2 52.7	1.4 40 °	1.2	1.2	1.2 54 0	1.2	1.2 40.9	1.2 50 5	1.2 40.0
CLEVR Material Loss	0.7	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
CLEVR Shape Acc@1	52.1	52.4	52.5	51.1	52.2	52.4	52.9	51.2	49.9	50.2	34.3	50.2	44.8	33.3	35.8	34.9	36.1	34.6	34.6	33.7	33.4
CLEVR Shape Loss	0.9	0.9	0.9	1.0	0.9	0.9	0.9	1.0	1.0	1.0	1.1	1.0	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
CLEVR Size Acc@1	61.0	61.1	61.3	60.7	61.1	60.8	62.0	62.3	60.9	59.6	53.5	58.3	55.7	50.6	56.2	55.2	55.2	54.6	54.2	54.1	50.1
CLEVR Size Loss	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
CLEVR Yes/No Acc@1	60.7	60.5	60.8	60.6	60.5	60.7	60.4	60.4	60.2	59.8	53.3	59.9	59.6	51.4	60.1	59.2	59.5	59.8	59.5	59.3	58.6
CLEVR Yes/No Loss	U.6 70.2	0.6	0.6	U.6	U.6	0.6	0.6	U.6	0.6	U.6	0.7	0.6	U.6 56.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6 2€ 4
CLEVR-Math Acc@5	19.5	00.9 00 5	8.6U 9.0U	09.9 08 0	13.1	02.9 00 3	00.0 99.2	09.3 08.0	08.0 08.0	0.00	44.U 97.7	0.00	0.00 98 8	30.2 86 1	40.9	40.0	40.2	45.7	44.8 97 5	42.1 96 0	30.4 92.8
CLEVR-Math Loss	0.5	0.8	0.7	0.9	0.6	0.8	0.9	0.9	1.0	1.0	1.3	1.0	1.0	1.7	1.2	1.2	1.2	1.2	1.3	1.3	1.5
Task Mean	60.8	59.4	59.8	58.2	60.2	59.0	58.9	58.7	57.8	56.1	47.8	55.5	52.7	43.5	52.0	51.7	52.1	50.9	49.0	48.5	46.5
Medical Class																					
Chexpert 0 APS	75.7	76.5	76.6	76.8	76.8	74.7	76.0	75.8	76.3	75.1	75.2	69.1	70.3	65.3	20.6	22.3	21.9	29.4	31.6	25.2	23.2

| Chexpert 0 AUC<br>Chexpert 0 BS   | 91.3  | 92.1<br>7.4   | $\frac{92.5}{7.3}$   
   | $92.3 \\ 7.4$  | 92.6   | 91.4   | 92.2   
   | 92.3<br>73   | 92.6<br>6 9   | 91.0   | $91.6 \\ 7.3$   | $\frac{89.9}{7.9}$                
   | 90.5   | $\frac{88.5}{8.4}$  | $61.5 \\ 12.6$   | 64.0<br>12.5  
   | 65.2<br>12.5   | 71.3  | 72.3  | 66.4  | 65.9<br>12.4   
   |
|---|---|---
--	--	--	--
---	---	--	---
--	---	--	---
---	--		
Chexpert 1 APS	55.3	55.2	55.5
   | 55.8   | 54.2   | 54.4   | 54.2   
   | 52.1   | 55.9  | 53.1   | 53.3  | 44.2                              
   | 43.0   | 33.5  | 28.9   | 30.1  
   | 31.0   | 28.9  | 30.1  | 29.9  | 28.5   
   |
| Chexpert 1 AUC<br>Chexpert 1 BS   | 75.7<br>18.8  | $\frac{76.0}{18.5}$   | 75.3<br>19.4   
   | 77.0<br>20.2   | 75.4<br>18.7   | $75.3 \\ 20.6$   | $\frac{75.3}{20.6}$  
   | $73.8 \\ 18.6$   | 76.1<br>16.3  | $74.9 \\ 17.6$   | $75.2 \\ 20.3$  | $69.4 \\ 17.2$                    
   | $69.8 \\ 17.2$   | $64.1 \\ 18.3$  | $\frac{56.3}{18.7}$  | $56.9 \\ 18.6$  
   | 57.7<br>18.6   | $57.1 \\ 18.6$  | $57.6 \\ 18.5$  | 57.5<br>18.6  | 57.0<br>18.6   
   |
| Chexpert 2 APS  | 43.8  | 43.8  | 43.5   
   | 45.1   | 45.5   | 44.8   | 43.9   
   | 42.3   | 43.6  | 43.5   | 44.4  | 41.8                              
   | 42.8   | 35.3  | 30.1   | 31.1  
   | 30.4   | 32.3  | 32.6  | 31.2  | 30.8   
   |
| Chexpert 2 AUC<br>Chexpert 2 BS   | 71.8<br>18.5  | $71.2 \\ 17.8$  | 71.8<br>21.0   
   | 72.4<br>21.1   | 72.1<br>18.6   | $\frac{72.0}{21.2}$  | 71.7<br>20.5   
   | 71.3<br>191  | 71.7<br>17.0  | 70.5<br>16.0   | 71.1<br>20.4  | 69.9<br>16 2                      
   | 70.9<br>16.2   | 63.1<br>174   | 58.6<br>18.4   | 59.0<br>18.2  
   | 58.7<br>18 1   | 60.7<br>179   | 60.5<br>179   | 60.1<br>17.8  | 58.9<br>17.9   
   |
| Chexpert 3 APS  | 80.7  | 80.9  | 80.8   
   | 82.1   | 81.7   | 80.5   | 79.7   
   | 78.6   | 79.1  | 79.2   | 80.6  | 73.5                              
   | 75.3   | 58.6  | 51.7   | 50.3  
   | 52.4   | 53.2  | 54.0  | 48.8  | 49.4   
   |
| Chexpert 3 AUC<br>Chexpert 3 BS   | $86.8 \\ 17.4$  | $86.8 \\ 16.4$  | $86.5 \\ 16.4$   
   | 87.9<br>15.6   | 87.2<br>15.3   | $86.6 \\ 16.2$   | $85.9 \\ 16.9$   
   | $84.6 \\ 17.2$   | $85.8 \\ 15.9$  | $84.9 \\ 17.6$   | $\frac{87.0}{16.3}$   | 82.3<br>18.1                      
   | $83.5 \\ 17.1$   | $73.0 \\ 23.5$  | 65.6<br>26.1   | 65.5<br>26.0  
   | $65.2 \\ 26.0$   | $65.6 \\ 25.2$  | $67.2 \\ 24.8$  | 64.2<br>26.1  | 64.2<br>26.1   
   |
| Chexpert 4 APS  | 53.4  | 49.5  | 50.1   
   | 53.4   | 54.5   | 50.9   | 52.6   
   | 50.8   | 52.3  | 49.9   | 50.7  | 41.7                              
   | 44.9   | 47.3  | 38.4   | 36.7  
   | 36.0   | 39.2  | 37.9  | 35.9  | 33.2   
   |
| Chexpert 4 AUC<br>Chexpert 4 BS   | $87.5 \\ 10.4$  | $86.7 \\ 10.0$  | $87.0 \\ 10.9$   
   | <b>88.1</b><br>10.2  | $\frac{88.0}{9.1}$   | $87.0 \\ 10.9$   | $87.3 \\ 10.2$   
   | $\frac{86.8}{9.9}$   | 87.7<br>8.8   | $\frac{86.0}{9.4}$   | $86.4 \\ 11.6$  | $84.1 \\ 10.1$                    
   | 85.1<br>9.6  | 84.8<br>9.4   | 81.7<br>9.7  | $80.3 \\ 10.0$  
   | $80.8 \\ 9.9$  | $81.5 \\ 9.9$   | $81.3 \\ 10.7$  | $79.4 \\ 10.1$  | $79.3 \\ 10.4$   
   |
| Chexpert APS Macro  | 61.6  | 61.0  | 61.2   
   | 62.6   | 62.3   | 60.9   | 61.2   
   | 59.9   | 61.5  | 59.8   | 60.2  | 54.1                              
   | 55.2   | 48.0  | 33.9   | 34.1  
   | 34.3   | 35.7  | 36.9  | 33.7  | 33.0   
   |
| Chexpert AUC Macro<br>Chexpert BS Macro   | $82.5 \\ 15.7$  | $82.5 \\ 15.6$  | $\frac{82.3}{15.5}$  
   | <b>83.2</b><br>14.9  | $\frac{82.9}{13.8}$  | $82.5 \\ 15.4$   | $82.4 \\ 15.1$   
   | $81.8 \\ 14.4$   | 82.8<br>13.0  | $81.1 \\ 13.7$   | $81.9 \\ 15.2$  | $79.1 \\ 13.9$                    
   | $79.9 \\ 13.6$   | $74.7 \\ 15.4$  | $\frac{64.7}{17.1}$  | $65.1 \\ 17.1$  
   | $65.5 \\ 17.0$   | $67.0 \\ 16.9$  | $67.6 \\ 16.9$  | 65.3<br>17.2  | $64.9 \\ 17.2$   
   |
| Chexpert Loss   | 0.3   | 0.4   | 0.5  
   | 0.3  | 0.3  | 0.3  | 0.4  
   | 0.4  | 0.3   | 0.3  | 0.4   | 0.3                               
   | 0.4  | 0.4   | 0.5  | 0.5   
   | 0.5  | 0.5   | 0.4   | 0.5   | 0.5  
   |
| Diabetic 0 APS<br>Diabetic 0 AUC  | 93.0<br>86.3  | $91.8 \\ 84.6$  | $\frac{91.5}{84.0}$  
   | $91.3 \\ 83.9$   | 90.9<br>83.0   | $91.3 \\ 83.6$   | $90.6 \\ 81.7$   
   | $90.4 \\ 80.9$   | $\frac{88.3}{77.2}$   | $90.8 \\ 83.9$   | $91.5 \\ 84.3$  | $85.4 \\ 72.2$                    
   | $87.2 \\ 75.1$   | $\frac{75.5}{52.4}$   | $\frac{76.3}{54.3}$  | $75.6 \\ 53.6$  
   | $77.4 \\ 56.5$   | $79.8 \\ 60.4$  | $79.4 \\ 58.6$  | $76.4 \\ 54.7$  | $77.2 \\ 55.3$   
   |
| Diabetic 0 BS   | 10.7  | 11.9  | 12.3   
   | 12.4   | 12.6   | 12.6   | 13.0   
   | 13.0   | 14.6  | 12.1   | 11.7  | 16.5                              
   | 15.7   | 19.0  | 19.5   | 19.4  
   | 19.3   | 19.1  | 18.6  | 19.0  | 19.0   
   |
| Diabetic 1 APS<br>Diabetic 1 AUC  | 14.0<br>69.6  | $13.6 \\ 67.2$  | $14.0 \\ 67.4$   
   | $13.0 \\ 66.0$   | $13.0 \\ 65.3$   | $12.9 \\ 66.1$   | 14.5<br>66.5   
   | $10.8 \\ 65.3$   | $9.0 \\ 59.7$   | $12.6 \\ 66.5$   | $13.5 \\ 66.4$  | $8.4 \\ 54.4$                     
   | $9.0 \\ 59.5$  | $\frac{7.2}{51.4}$  | $\frac{8.4}{54.9}$   | $\frac{8.8}{56.9}$  
   | $\frac{8.9}{54.5}$   | $\frac{8.4}{53.9}$  | $7.7 \\ 54.9$   | $7.4 \\ 52.1$   | $7.3 \\ 53.3$  
   |
| Diabetic 1 BS   | 6.1   | 6.4   | 6.5  
   | 6.1  | 6.0  | 6.8  | 6.4  
   | 5.8  | 5.8   | 6.0  | 6.4   | 6.9                               
   | 5.3  | 6.5   | 6.7  | 6.5   
   | 6.9  | 6.4   | 6.3   | 6.4   | 6.3  
   |
| Diabetic 2 APS<br>Diabetic 2 AUC  | 65.5<br>88.5  | $\frac{61.6}{86.9}$   | $\frac{60.7}{86.3}$  
   | $61.4 \\ 86.0$   | $\frac{58.4}{84.7}$  | $57.1 \\ 85.3$   | $54.2 \\ 84.3$   
   | $\frac{51.1}{82.5}$  | $44.3 \\ 79.6$  | $59.7 \\ 85.5$   | $63.1 \\ 87.4$  | $28.9 \\ 71.6$                    
   | $32.2 \\ 73.8$   | $14.6 \\ 50.9$  | $17.0 \\ 53.4$   | $16.7 \\ 52.2$  
   | $17.9 \\ 55.8$   | $20.2 \\ 61.2$  | $17.8 \\ 57.7$  | $17.0 \\ 54.1$  | $17.3 \\ 55.5$   
   |
| Diabetic 2 BS   | 8.0   | 8.5   | 9.0  
   | 9.3  | 9.7  | 9.0  | 9.5  
   | 9.9  | 10.7  | 9.2  | 8.3   | 11.7                              
   | 11.7   | 12.1  | 12.7   | 12.8  
   | 12.6   | 12.7  | 11.9  | 12.4  | 12.5   
   |
| Diabetic 3 APS<br>Diabetic 3 AUC  | $\frac{41.6}{94.8}$   | 49.7<br>96.5  | $47.6 \\ 95.7$   
   | $48.4 \\ 95.6$   | $45.3 \\ 93.9$   | 53.1<br>95.1   | $46.5 \\ 95.0$   
   | $38.8 \\ 94.1$   | $37.1 \\ 93.5$  | $47.2 \\ 95.1$   | $50.7 \\ 96.2$  | $22.4 \\ 87.2$                    
   | 32.0<br>92.3   | $\frac{2.8}{56.0}$  | $\frac{3.1}{56.1}$   | $\frac{3.1}{57.2}$  
   | $\frac{4.1}{59.1}$   | $4.6 \\ 64.2$   | $4.0 \\ 64.9$   | $3.4 \\58.4$  | $\frac{2.6}{52.3}$   
   |
| Diabetic 3 BS   | 1.9   | 1.6   | 1.6  
   | 1.6  | 1.7  | 1.9  | 1.7  
   | 1.8  | 1.8   | 1.7  | 1.5   | 2.0                               
   | 2.1  | 2.4   | 2.4  | 2.5   
   | 2.3  | 2.2   | 2.3   | 2.1   | 2.1  
   |
| Diabetic 4 APS<br>Diabetic 4 AUC  | 73.9<br>98.7  | $\frac{74.3}{98.2}$   | 73.0<br>97.7   
   | 75.3<br>98.7   | $67.5 \\ 98.0$   | 68.7<br>97.4   | 70.2<br>98.4   
   | 72.3<br>98.3   | $47.5 \\ 96.9$  | $67.5 \\ 97.2$   | $74.6 \\ 97.9$  | 32.4<br>94.7                      
   | 23.7<br>94.3   | $\frac{2.9}{56.4}$  | $3.1 \\ 60.1$  | $3.0 \\ 58.6$   
   | $4.4 \\ 63.0$  | $3.9 \\ 68.1$   | $3.7 \\ 64.3$   | $\frac{2.5}{56.9}$  | $2.6 \\ 57.8$  
   |
| Diabetic 4 BS   | 1.0   | 1.1   | 1.0  
   | 0.9  | 1.1  | 1.1  | 0.9  
   | 0.9  | 1.3   | 1.1  | 0.8   | 1.4                               
   | 1.8  | 1.9   | 1.9  | 1.8   
   | 1.7  | 1.7   | 1.8   | 1.9   | 1.8  
   |
| Diabetic APS Macro<br>Diabetic AUC Macro  | 56.9<br>87.5  | $57.2 \\ 86.7$  | $56.4 \\ 86.0$   
   | $\frac{56.3}{85.7}$  | $54.2 \\ 85.0$   | $\frac{56.4}{85.3}$  | $54.4 \\ 84.7$   
   | $\frac{51.9}{83.8}$  | $\frac{45.2}{81.2}$   | $55.6 \\ 85.6$   | 58.7<br>86.1  | $35.5 \\ 76.0$                    
   | $36.6 \\ 79.0$   | $20.6 \\ 53.4$  | 21.6<br>55.7   | 21.5<br>55.7  
   | 22.5<br>57.8   | $23.3 \\ 61.3$  | $22.4 \\ 59.4$  | 21.2<br>55.1  | $21.3 \\ 54.0$   
   |
| Diabetic BS Macro   | 5.5   | 6.0   | 6.1  
   | 6.1  | 6.2  | 6.4  | 6.3  
   | 6.4  | 7.0   | 6.1  | 5.8   | 7.7                               
   | 7.4  | 8.4   | 8.7  | 8.6   
   | 8.6  | 8.5   | 8.2   | 8.4   | 8.4  
   |
| Diabetic Loss<br>HAM10K 0 APS   | 0.2<br>94.3   | $\frac{0.1}{90.0}$  | $0.2 \\ 90.3$  
   | $0.1 \\ 88.7$  | 0.1<br>89.2  | $0.1 \\ 90.9$  | $0.2 \\ 89.8$  
   | $0.2 \\ 89.0$  | 0.2 83.3  | $0.2 \\ 88.2$  | $0.2 \\ 84.1$   | $0.2 \\ 47.4$                     
   | $\frac{0.3}{58.0}$   | $0.2 \\ 30.4$   | $0.3 \\ 32.8$  | $0.3 \\ 25.8$   
   | $0.3 \\ 25.0$  | $0.3 \\ 41.2$   | $0.3 \\ 46.2$   | 0.2<br>34.4   | 0.3<br>33.8  
   |
| HAM10K 0 AUC  | 99.1  | 98.2  | 98.3   
   | 97.6   | 97.8   | 98.2   | 97.7   
   | 98.0   | 96.7  | 97.6   | 97.0  | 89.0                              
   | 91.7   | 80.6  | 81.2   | 78.5  
   | 79.4   | 85.2  | 86.9  | 82.1  | 79.7   
   |
| HAM10K 0 BS<br>HAM10K 1 APS   | 2.1<br>99.2   | $\frac{2.9}{99.2}$  | $3.5 \\ 99.1$  
   | $2.8 \\ 99.2$  | $3.1 \\ 99.2$  | $3.1 \\ 99.1$  | $3.1 \\ 99.1$  
   | $3.4 \\ 99.2$  | $3.8 \\98.7$  | $3.4 \\98.9$   | $4.0 \\ 98.1$   | $7.0 \\ 96.2$                     
   | $6.3 \\ 96.5$  | $\frac{8.2}{94.2}$  | 8.1<br>93.9  | $8.5 \\ 93.7$   
   | $8.6 \\ 93.1$  | $7.6 \\ 95.5$   | $7.3 \\ 96.0$   | 8.1<br>94.0   | 8.3<br>93.7  
   |
| HAM10K 1 AUC  | 98.9  | 98.7  | 98.4   
   | 98.5   | 98.4   | 98.4   | 98.4   
   | 98.4   | 97.3  | 98.1   | 97.1  | 92.7                              
   | 93.5   | 89.7  | 88.7   | 88.1  
   | 87.8   | 91.0  | 91.9  | 88.3  | 87.3   
   |
| HAM10K 1 BS<br>HAM10K 2 APS   | 3.1<br>95.5   | 3.7<br>98.6   | $4.5 \\ 89.0$  
   | $\frac{4.2}{94.4}$   | $\frac{4.5}{88.7}$   | $4.4 \\ 92.1$  | $4.6 \\ 92.4$  
   | $\frac{4.4}{95.3}$   | $6.2 \\ 69.7$   | $5.0 \\ 81.6$  | $6.3 \\ 89.0$   | $10.0 \\ 11.3$                    
   | $9.4 \\ 5.0$   | $\frac{11.7}{5.7}$  | $\frac{12.5}{5.2}$   | 12.8<br>8.1   
   | 12.9<br>2.2  | $11.3 \\ 19.5$  | $10.7 \\ 12.2$  | $13.0 \\ 7.4$   | 13.9<br>3.6  
   |
| HAM10K 2 AUC  | 99.9  | 100.0   | 99.7   
   | 99.9   | 99.7   | 99.8   | 99.8   
   | 99.9   | 99.3  | 98.4   | 99.8  | 81.1                              
   | 75.6   | 79.4  | 79.6   | 73.0  
   | 68.2   | 90.8  | 87.2  | 81.2  | 78.3   
   |
| HAMIOK 2 BS   | 0.3   | 0.3   | 0.4  
   | 0.3  | 0.5  | 0.3  | 0.3  
   | 0.3  | 0.8   | 0.5  | 0.4   | 1.3                               
   | 1.3  | 1.3   | 1.3  | 1.3   
   | 1.3  | 1.2   | 1.3   | 1.3   | 1.3  
   |
| HAMIUK 3 APS  | 88.0  | 85.5  | 83.9   
   | 85.2   | 86.2   | 83.0   | 84.0   
   | 82.5   | 74.2  | 80.8   | 74.3  | 41.9                              
   | 46.7   | 34.7  | 35.1   | 33.2  
   | 31.5   | 42.5  | 48.4  | 42.4  | 35.2   
   |
| HAM10K 3 APS<br>HAM10K 3 AUC  | 88.0<br>96.7  | $\frac{85.5}{95.5}$   | 83.9<br>95.6   
   | 85.2<br>95.9   | 86.2<br>96.1   | 83.0<br>95.3   | 84.0<br>95.9   
   | 82.5<br>96.1   | 74.2<br>94.4  | 80.8<br>95.4   | 74.3<br>92.5  | 41.9<br>83.8                      
   | $46.7 \\ 84.9 \\ 7.6$  | 34.7<br>81.7  | 35.1<br>80.0   | 33.2<br>80.3  
   | 31.5<br>80.7   | $42.5 \\ 85.9 \\ 7.7$   | 48.4<br>88.1  | 42.4<br>84.2  | 35.2<br>82.6   
   |
| HAM10K 3 APS<br>HAM10K 3 AUC<br>HAM10K 3 BS<br>HAM10K 4 APS   | 88.0<br>96.7<br>3.5<br>99.5   | 85.5<br>95.5<br>3.7<br><b>100.0</b>   | $83.9 \\ 95.6 \\ 4.2 \\ 99.7$  
   | $85.2 \\ 95.9 \\ 3.9 \\ 98.2$  | $86.2 \\ 96.1 \\ 3.5 \\ 100.0$   | $83.0 \\ 95.3 \\ 4.1 \\ 98.5$  | 84.0<br>95.9<br>4.2<br>100.0   
   | 82.5<br>96.1<br>4.4<br>98.5  | 74.2<br>94.4<br>5.0<br>98.7   | $80.8 \\ 95.4 \\ 4.7 \\ 96.4$  | $74.3 \\ 92.5 \\ 5.1 \\ 96.9$   | 41.9<br>83.8<br>7.9<br>26.8       
   | $46.7 \\ 84.9 \\ 7.6 \\ 21.9$  | $34.7 \\ 81.7 \\ 8.4 \\ 33.6$   | 35.1<br>80.0<br>8.4<br>32.3  | 33.2<br>80.3<br>8.4<br>24.6   
   | 31.5<br>80.7<br>8.5<br>26.4  | 42.5<br>85.9<br>7.7<br>52.8   | 48.4<br>88.1<br>7.2<br>73.8   | 42.4<br>84.2<br>8.0<br>34.8   | 35.2<br>82.6<br>8.2<br>11.5  
   |
| HAM10K 3 APS<br>HAM10K 3 AUC<br>HAM10K 3 BS<br>HAM10K 4 APS<br>HAM10K 4 AUC<br>HAM10K 4 PS  | 88.0<br>96.7<br>3.5<br>99.5<br>100.0  | 85.5<br>95.5<br>3.7<br><b>100.0</b><br>100.0  | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b>  
   | 85.2<br>95.9<br>3.9<br>98.2<br>100.0   | 86.2<br>96.1<br>3.5<br>100.0<br>100.0  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0   | 84.0<br>95.9<br>4.2<br>100.0<br>100.0  
   | 82.5<br>96.1<br>4.4<br>98.5<br>100.0   | $74.2 \\94.4 \\5.0 \\98.7 \\100.0 \\0.2$  | 80.8<br>95.4<br>4.7<br>96.4<br>99.9  | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2   |
41.9<br>83.8<br>7.9<br>26.8<br>92.3   | 46.7<br>84.9<br>7.6<br>21.9<br>94.5<br>1.2   | 34.7<br>81.7<br>8.4<br>33.6<br>84.2<br>1.2  | 35.1<br>80.0<br>8.4<br>32.3<br>89.0  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5  
  | 31.5<br>80.7<br>8.5<br>26.4<br>87.1  | 42.5<br>85.9<br>7.7<br>52.8<br>97.7   | 48.4<br>88.1<br>7.2<br>73.8<br>97.6<br>0.5  | 42.4<br>84.2<br>8.0<br>34.8<br>92.0   | 35.2<br>82.6<br>8.2<br>11.5<br>78.7   
  |
| HAM10K 3 APS<br>HAM10K 3 AUC<br>HAM10K 3 BS<br>HAM10K 4 AS<br>HAM10K 4 AUC<br>HAM10K 4 BS<br>HAM10K 5 APS   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6   | 85.5<br>95.5<br>3.7<br><b>100.0</b><br>100.0<br>0.0<br>94.8   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5   
   | $\begin{array}{c} 85.2 \\ 95.9 \\ 3.9 \\ 98.2 \\ 100.0 \\ 0.1 \\ 91.5 \end{array}$   | $\begin{array}{c} 86.2 \\ 96.1 \\ 3.5 \\ 100.0 \\ 100.0 \\ 0.0 \\ 90.3 \end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br><b>0.0</b><br>91.6  
   | $\begin{array}{c} 82.5 \\ 96.1 \\ 4.4 \\ 98.5 \\ 100.0 \\ 0.0 \\ 90.8 \end{array}$   | $74.2 \\94.4 \\5.0 \\98.7 \\100.0 \\0.2 \\83.2$   | $80.8 \\ 95.4 \\ 4.7 \\ 96.4 \\ 99.9 \\ 0.1 \\ 88.0$   | $74.3 \\92.5 \\5.1 \\96.9 \\100.0 \\0.2 \\91.4$   | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1 \end{array}$   | $\begin{array}{r} 46.7 \\ 84.9 \\ 7.6 \\ 21.9 \\ 94.5 \\ 1.2 \\ 67.0 \end{array}$  | 34.7<br>81.7<br>8.4<br>33.6<br>84.2<br>1.2<br>41.8  | 35.1<br>80.0<br>8.4<br>32.3<br>89.0<br>1.0<br>36.4   | $33.2 \\ 80.3 \\ 8.4 \\ 24.6 \\ 89.5 \\ 1.1 \\ 36.8$   
  | 31.5<br>80.7<br>8.5<br>26.4<br>87.1<br>1.2<br>22.3   | $\begin{array}{c} 42.5 \\ 85.9 \\ 7.7 \\ 52.8 \\ 97.7 \\ 0.9 \\ 48.1 \end{array}$   | $\begin{array}{c} 48.4 \\ 88.1 \\ 7.2 \\ 73.8 \\ 97.6 \\ 0.5 \\ 41.3 \end{array}$   | 42.4<br>84.2<br>8.0<br>34.8<br>92.0<br>1.1<br>36.6  | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8  
  |
| HAMIOK 3 APS<br>HAMIOK 3 AUC<br>HAMIOK 3 BS<br>HAMIOK 4 APS<br>HAMIOK 4 AUC<br>HAMIOK 4 BS<br>HAMIOK 5 APS<br>HAMIOK 5 AUC<br>HAMIOK 5 SPS  | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7   | 85.5<br>95.5<br>3.7<br><b>100.0</b><br>100.0<br>0.0<br>94.8<br>99.7   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6   
   | 85.2<br>95.9<br>3.9<br>98.2<br>100.0<br>0.1<br>91.5<br>99.5<br>1.1   | $\begin{array}{c} 86.2 \\ 96.1 \\ 3.5 \\ 100.0 \\ 100.0 \\ 0.0 \\ 90.3 \\ 99.2 \\ 1.3 \end{array}$   | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4  
   | 82.5 96.1 4.4 98.5 100.0 $0.090.899.01 3$  | $74.2 \\94.4 \\5.0 \\98.7 \\100.0 \\0.2 \\83.2 \\98.8 \\1.7 $   | 80.8<br>95.4<br>4.7<br>96.4<br>99.9<br>0.1<br>88.0<br>98.0<br>1.5  | $\begin{array}{c} 74.3 \\ 92.5 \\ 5.1 \\ 96.9 \\ 100.0 \\ 0.2 \\ 91.4 \\ 99.4 \\ 1.4 \end{array}$   | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4 \end{array}$  | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\end{array}$  | 34.7<br>81.7<br>8.4<br>33.6<br>84.2<br>1.2<br>41.8<br>92.9<br>3.8   | 35.1<br>80.0<br>8.4<br>32.3<br>89.0<br>1.0<br>36.4<br>92.0<br>3.9  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>3.9  
  | 31.580.78.526.487.11.222.387.54 4  | $\begin{array}{c} 42.5\\85.9\\7.7\\52.8\\97.7\\0.9\\48.1\\94.0\\3.6\end{array}$   | $\begin{array}{r} 48.4 \\ 88.1 \\ 7.2 \\ 73.8 \\ 97.6 \\ 0.5 \\ 41.3 \\ 92.5 \\ 3.9 \end{array}$  | 42.4<br>84.2<br>8.0<br>34.8<br>92.0<br>1.1<br>36.6<br>90.3<br>4.2   | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8<br>88.2<br>4.3   
  |
| HAMIOK 3 APS<br>HAMIOK 3 AUC<br>HAMIOK 3 BS<br>HAMIOK 4 APS<br>HAMIOK 4 AUC<br>HAMIOK 4 BS<br>HAMIOK 5 APS<br>HAMIOK 5 SB<br>HAMIOK 6 APS   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2  | 85.5<br>95.5<br>3.7<br><b>100.0</b><br>100.0<br>94.8<br>99.7<br>1.1<br>85.2   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6<br>1.1<br>83.9  
   | $\begin{array}{c} 85.2 \\ 95.9 \\ 3.9 \\ 98.2 \\ 100.0 \\ 0.1 \\ 91.5 \\ 99.5 \\ 1.1 \\ 86.3 \end{array}$  | $\begin{array}{c} 86.2 \\ 96.1 \\ 3.5 \\ 100.0 \\ 100.0 \\ 90.3 \\ 99.2 \\ 1.3 \\ 88.0 \end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7   
   | $\begin{array}{c} 82.5 \\ 96.1 \\ 4.4 \\ 98.5 \\ 100.0 \\ 0.0 \\ 90.8 \\ 99.0 \\ 1.3 \\ 83.3 \end{array}$  | $74.2 \\94.4 \\5.0 \\98.7 \\100.0 \\0.2 \\83.2 \\98.8 \\1.7 \\75.8$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3 \end{array}$  | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2 \end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4 \end{array}$   | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4 \end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4 \end{array}$   | $\begin{array}{r} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5 \end{array}$  | $\begin{array}{c} 33.2 \\ 80.3 \\ 8.4 \\ 24.6 \\ 89.5 \\ 1.1 \\ 36.8 \\ 91.5 \\ 3.9 \\ 29.5 \end{array}$   
  | $\begin{array}{r} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5 \end{array}$  | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6 \end{array}$   | $\begin{array}{r} 48.4 \\ 88.1 \\ 7.2 \\ 73.8 \\ 97.6 \\ 0.5 \\ 41.3 \\ 92.5 \\ 3.9 \\ 36.6 \end{array}$  | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\end{array}$  | $\begin{array}{c} 35.2 \\ 82.6 \\ 8.2 \\ 11.5 \\ 78.7 \\ 1.2 \\ 26.8 \\ 88.2 \\ 4.3 \\ 23.5 \end{array}$  
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 AUC<br>HAMIUK 6 BS  | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>10  | 85.5<br>95.5<br>3.7<br>100.0<br>100.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1 1 \end{array}$  | $\begin{array}{c} 86.2 \\ 96.1 \\ 3.5 \\ 100.0 \\ 100.0 \\ 90.3 \\ 99.2 \\ 1.3 \\ 88.0 \\ 98.9 \\ 1.1 \end{array}$   | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3   | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br><b>99.3</b><br>1 4   
   | $\begin{array}{c} 82.5 \\ 96.1 \\ 4.4 \\ 98.5 \\ 100.0 \\ 0.0 \\ 90.8 \\ 99.0 \\ 1.3 \\ 83.3 \\ 98.6 \\ 1.3 \end{array}$   | $74.2 \\94.4 \\5.0 \\98.7 \\100.0 \\0.2 \\83.2 \\98.8 \\1.7 \\75.8 \\98.0 \\1.7 \\$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\end{array}$  | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ \end{array}$  | $\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ \end{array}$  | $\begin{array}{c} 35.1 \\ 80.0 \\ 8.4 \\ 32.3 \\ 89.0 \\ 1.0 \\ 36.4 \\ 92.0 \\ 3.9 \\ 30.5 \\ 91.9 \\ 3.0 \end{array}$  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>3.9<br>29.5<br>91.4<br>3.0   
  | $\begin{array}{c} 31.5 \\ 80.7 \\ 8.5 \\ 26.4 \\ 87.1 \\ 1.2 \\ 22.3 \\ 87.5 \\ 4.4 \\ 24.5 \\ 89.1 \\ 3.2 \end{array}$  | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\end{array}$   | $\begin{array}{c} 48.4 \\ 88.1 \\ 7.2 \\ 73.8 \\ 97.6 \\ 0.5 \\ 41.3 \\ 92.5 \\ 3.9 \\ 36.6 \\ 93.4 \\ 2.9 \end{array}$   | $\begin{array}{c} 42.4 \\ 84.2 \\ 8.0 \\ 34.8 \\ 92.0 \\ 1.1 \\ 36.6 \\ 90.3 \\ 4.2 \\ 25.8 \\ 90.5 \\ 3.2 \end{array}$   | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8<br>88.2<br>4.3<br>23.5<br>88.5<br>3.2  
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 AUC<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK APS Macro   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5  | $\begin{array}{c} 85.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 0.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ \end{array}$   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>91.4   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2 \end{array}$   | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ \end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1   | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br><b>99.3</b><br>1.4<br>91.6   
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ \end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4 \end{array}$  | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\end{array}$   | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1 \end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\end{array}$  | $\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\end{array}$   | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8 \end{array}$   | $\begin{array}{c} 35.1 \\ 80.0 \\ 8.4 \\ 32.3 \\ 89.0 \\ 1.0 \\ 36.4 \\ 92.0 \\ 3.9 \\ 30.5 \\ 91.9 \\ 3.0 \\ 38.0 \end{array}$  | $\begin{array}{c} 33.2 \\ 80.3 \\ 8.4 \\ 24.6 \\ 89.5 \\ 1.1 \\ 36.8 \\ 91.5 \\ 3.9 \\ 29.5 \\ 91.4 \\ 3.0 \\ 35.9 \end{array}$  
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2 \end{array}$  | $\begin{array}{c} 42.5\\85.9\\7.7\\52.8\\97.7\\0.9\\48.1\\94.0\\3.6\\39.6\\92.3\\2.8\\48.5\end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\end{array}$  | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\end{array}$  | $\begin{array}{c} 35.2 \\ 82.6 \\ 8.2 \\ 11.5 \\ 78.7 \\ 1.2 \\ 26.8 \\ 88.2 \\ 4.3 \\ 23.5 \\ 88.5 \\ 3.2 \\ 32.6 \end{array}$   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK APS Macro<br>HAMIUK S Macro<br>HAMIUK S Macro   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6  | 85.5<br>95.5<br>3.7<br>100.0<br>100.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9  | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>91.4<br>98.7<br>2.2  
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\end{array}$   | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ \end{array}$   | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br><b>99.3</b><br>1.4<br>91.6<br>98.7<br>2.1  
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ \end{array}$  | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8 \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4 \end{array}$   | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8 \end{array}$  | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4 \end{array}$  | $\begin{array}{r} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 86.1\\ 5.5\end{array}$  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>3.9<br>29.5<br>91.4<br>3.0<br>35.9<br>84.6<br>5.6  
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7 \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ \end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8 \end{array}$  | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\end{array}$   | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8<br>88.2<br>4.3<br>23.5<br>88.5<br>3.2<br>32.6<br>83.3<br>5.8   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 AVC<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK AUC<br>HAMIUK AUS<br>HAMIUK AUS<br>HAMIUK AUS<br>HAMIUK AUS<br>HAMIUK AUS<br>HAMIUK AUS<br>HAMIUK AUS   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3   | 85.5<br>95.5<br>3.7<br>100.0<br>0.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2   | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.1\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 3.9\end{array}$   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ \end{array}$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ \end{array}$   | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br>1.0<br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1<br>0.2  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br>99.3<br>1.4<br>91.6<br>98.7<br>2.1<br>0.2  
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ \end{array}$  | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ \end{array}$   | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ \end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$  | $\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ \end{array}$   | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ \end{array}$  | $\begin{array}{r} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 86.1\\ 5.5\\ 0.2\\ \end{array}$   | $\begin{array}{c} 33.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.0\\ 35.9\\ 84.6\\ 5.6\\ 0.2\\ \end{array}$   
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ 0.2\\ \end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$  | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ \end{array}$  | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8<br>88.2<br>4.3<br>23.5<br>88.5<br>3.2<br>32.6<br>83.3<br>5.8<br>0.2  
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK AUC<br>HAMIUK AUC Macro<br>HAMIUK BS Macro<br>HAMIUK Loss<br><b>Task Mean</b><br>Medical Seg   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0   | 85.5<br>95.5<br>3.7<br><b>100.0</b><br>0.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br><b>0.2</b><br>56.7   | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.1\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 56.5\\ \end{array}$   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7 \end{array}$   | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ \end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1<br>0.2<br>56.4   | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br><b>99.3</b><br>1.4<br>91.6<br>98.7<br>2.1<br>0.2<br>56.2   
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\end{array}$   | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ \end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ \end{array}$   | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4 \end{array}$   | $\begin{array}{r} 35.1\\ \textbf{80.0}\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 86.1\\ 5.5\\ 0.2\\ 37.4 \end{array}$   | $\begin{array}{c} 33.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.0\\ 35.9\\ 84.6\\ 5.6\\ 0.2\\ 37.0\\ \end{array}$  
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ \end{array}$  | $\begin{array}{r} 42.5\\85.9\\7.7\\52.8\\97.7\\0.9\\48.1\\94.0\\3.6\\39.6\\92.3\\2.8\\48.5\\91.0\\5.0\\0.2\\40.6\end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ \end{array}$   | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\end{array}$  | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ 36.2\\ \end{array}$  
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 SB<br>HAMIUK 6 APS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK APS Macro<br>HAMIUK BS Macro<br>HAMIUK BS Macro<br>HAMIUK BS Macro<br>HAMIUK Loss<br><b>Task Mean</b><br>Medical Seg<br>ACDC Dice Score   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>0.6  | 85.5<br>95.5<br>3.7<br>100.0<br>100.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2<br>56.7<br>0.5  | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.1\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 56.5\\ \end{array}$   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ 0.5\\ 78.5\\ \end{array}$  | 86.2<br>96.1<br>3.5<br>100.0<br>0.0<br>90.3<br>99.2<br>1.3<br>88.0<br>98.9<br>1.1<br>91.3<br>98.6<br>2.0<br>0.3<br>56.2  | $\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ \textbf{1.0}\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 98.6\\ 2.1\\ 0.2\\ 56.4\\ \end{array}$  | 84.0<br>95.9<br>4.2<br>100.0<br>100.0<br>91.6<br>99.4<br>1.2<br>84.7<br><b>99.3</b><br>1.4<br>91.6<br>98.7<br>2.1<br>0.2<br>56.2   
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ \hline 0.4\\ 0.4\\ \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ \end{array}$  | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 99.4\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ \end{array}$  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ \end{array}$   | $\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ \end{array}$   | 35.1<br>80.0<br>8.4<br>32.3<br>89.0<br>1.0<br>36.4<br>92.0<br>3.9<br>30.5<br>91.9<br>3.0<br>36.1<br>5.5<br>0.2<br>37.4<br>0.2  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>3.9<br>29.5<br>91.4<br>3.0<br>35.9<br>84.6<br>5.6<br>0.2<br>37.0<br>0.2  
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 39.6\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ \end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$   | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ \hline 0.3\\ 7.7\\ \hline 0.3\\ 7.7\\ \hline \end{array}$  | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \end{array}$   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK APS Macro<br>HAMIUK AS Macro<br>HAMIUK AS Macro<br>HAMIUK AS Macro<br>HAMIUK AS Macro<br>HAMIUK AS Macro<br>HAMIUK Coss<br>Task Mean<br>Medical Seg<br>ACDC Dice Score<br>ACDC Mean Acc@<br>ACDC Overall Acc@  | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>0.6<br>86.3<br>86.5  | 85.5<br>95.5<br>3.7<br>100.0<br>100.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2<br>56.7<br>0.5<br>85.8<br>86.2  | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.1\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 56.5\\ \end{array}$   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ 0.5\\ 78.5\\ 78.7\end{array}$   | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ \end{array}$  | $\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ \textbf{1.0}\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 98.6\\ 2.1\\ 0.2\\ 56.4\\ \end{array}$  | $\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 100.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ \end{array}$  
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ \hline 0.4\\ 74.0\\ 73.5\\ \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ \hline \\ 0.4\\ 93.4\\ 93.5\\ \end{array}$   | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2<br>91.4<br>83.2<br>98.6<br>87.1<br>97.5<br>2.8<br>0.6<br>56.0<br>0.6<br>94.1<br>94.2  | $\begin{array}{c} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ \end{array}$   | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \end{array}$  | $\begin{array}{r} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ \hline \end{array}$  | $\begin{array}{r} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 86.1\\ 5.5\\ 0.2\\ 37.4\\ \hline 0.2\\ 46.7\\ 47.2\\ \end{array}$  | $\begin{array}{r} 33.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.0\\ 35.9\\ 84.6\\ 5.6\\ 0.2\\ 37.0\\ \hline \\ 0.2\\ 53.7\\ 53.4\\ \end{array}$  
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ \end{array}$   | $\begin{array}{r} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ \end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ \hline 0.2\\ 56.1\\ 55.5\\ \end{array}$  | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ \hline 0.3\\ 50.8\\ 51.4 \end{array}$  | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \end{array}$   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 BS<br>HAMIUK 6 BS<br>HAMIUK AUC<br>Macro<br>HAMIUK AUC Macro<br>HAMIUK AUS<br>Macro<br>HAMIUK BS Macro<br>HAMIUK BS Macro<br>HAMIUK Loss<br><b>Task Mean</b><br>Medical Seg<br>ACDC Dice Score<br>ACDC Dice Score<br>ACDC Mean Acc@<br>ACDC Oureal Acc@<br>ACDC Oureal  | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>0.6<br>86.3<br>86.5<br>57.9   | 85.5<br>95.5<br>3.7<br>100.0<br>0.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2<br>56.7<br>0.5<br>85.8<br>86.2<br>57.0  | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.6\\ 1.1\\ 83.9\\ 99.6\\ 1.1\\ 83.9\\ 99.6\\ 1.1\\ 83.9\\ 99.6\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 56.5\\ \hline 0.5\\ 83.4\\ 83.2\\ 57.4\\ 83.2\\ 57.4\\ 83.2\\ 57.4\\ 83.4\\ 83.2\\ 57.4\\ 83.4\\ 83.2\\ 57.4\\ 83.4\\ 83.2\\ 57.4\\ 83.4\\ 83.2\\ 57.4\\ 83.4\\ 83.2\\ 57.4\\ 83.4\\
83.4\\ $ | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ \hline 0.5\\ 78.5\\ 78.7\\ 53.1\\ 52.7\\ \hline \end{array}$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ \hline 0.4\\ 75.5\\ 75.1\\ 50.2\\ 50.2\\ \end{array}$   | $\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ 1.0\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 98.6\\ 2.1\\ 0.2\\ 56.4\\ \hline \end{array}$  | $\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 100.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.4\\ 1.2\\ 84.7\\ 2.1\\ 0.2\\ 56.2\\ \hline 0.5\\ 76.9\\ 77.0\\ 47.7\\ 5.6 \\ \epsilon \end{array}$  
  | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \hline 0.5\\ 79.4\\ 79.0\\ 54.3\\ 2.2\\ 79.4\\ 79.0\\ 54.3\\ 2.2\\ 79.4\\ 79.0\\ 54.3\\ 2.2\\ 79.4\\ 79.0\\ 54.3\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.4\\ 79.0\\ 79.0\\ 79.4\\ 79.0\\ 79.0\\ 79.4\\ 79.0\\$ | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 83.4\\ 97.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 0.2\\ 53.6\\ 0.2\\ 53.6\\ 0.4\\ 74.0\\ 73.5\\ 50.1\\ 40.5\\ \end{array}$  | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 87.9\\ 97.9\\ 2.4\\ 93.5\\ 6.9\\ 93.4\\ 93.5\\ 66.9\\ c2\\ c2\\ c2\\ c2\\ c2\\ c2\\ c2\\ c2\\ c2\\ c2$  | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2<br>91.4<br>99.4<br>1.4<br>83.2<br>98.6<br>87.1<br>97.5<br>2.8<br>0.6<br>56.0<br>0.6<br>94.1<br>94.2<br>67.2<br>67.2   | $\begin{array}{r} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ \hline 0.4\\ 71.7\\ 71.5\\ 47.5\\ 47.5\\ \end{array}$   | $\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ 0.5\\ 76.0\\ 76.0\\ 76.0\\ 50.8\\ $             | $\begin{array}{r} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 86.1\\ 5.5\\ 0.2\\ 37.4\\ \hline 0.2\\ 46.7\\ 47.2\\ 27.6\\ 27.6\\ 27.6\\ \end{array}$  | 33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>3.9<br>29.5<br>91.4<br>3.0<br>35.9<br>84.6<br>0.2<br>37.0<br>0.2<br>53.7<br>53.4<br>30.4<br>30.4   
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ 35.6\\ 26.2\\ \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ \hline 0.3\\ 60.3\\ 60.3\\ 35.1\\ 20.0\\ \end{array}$  | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ \hline 0.2\\ 56.1\\ 55.5\\ 32.1\\ 26.0\\ \end{array}$  | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 0.2\\ 37.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 21.7\\ \end{array}$   | 35.2<br>82.6<br>8.2<br>11.5<br>78.7<br>1.2<br>26.8<br>88.2<br>4.3<br>23.5<br>88.5<br>3.2<br>32.6<br>83.3<br>5.8<br>0.2<br>36.2<br>0.3<br>51.4<br>26.9<br>51.4<br>22.5   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AVC<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 AVC<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 AVC<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIOK AUC MACRO<br>HAMI   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>0.6<br>86.3<br>86.5<br>57.9<br>57.8   | 85.5<br>95.5<br>3.7<br>100.0<br>0.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2<br>56.7<br>0.5<br>85.8<br>86.2<br>57.0<br>57.3  | $\begin{array}{c} 83.9\\ 95.6\\ 4.2\\ 99.7\\ 100.0\\ 0.1\\ 94.5\\ 99.6\\ 1.1\\ 83.9\\ 99.1\\ 1.5\\ 91.4\\ 98.7\\ 2.2\\ 0.3\\ 56.5\\ \hline 0.5\\ 83.4\\ 83.2\\ 57.4\\ 83.2\\ 57.4\\ 56.1\\ \hline \end{array}$   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ 0.5\\ 78.5\\ 78.7\\ 53.1\\ 52.7\\ \end{array}$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 100.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ \hline \end{array}$   | $\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ \textbf{1.0}\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 98.6\\ 2.1\\ 0.2\\ 56.4\\ \hline \end{array}$   | $\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ \hline \end{array}$   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \hline \\ 0.5\\ 79.4\\ 79.0\\ 54.3\\ 53.3\\ \hline \end{array}$   | $\begin{array}{r} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ \hline 0.4\\ 74.0\\ 73.5\\ 50.1\\ 49.5\\ \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ 0.4\\ 93.4\\ 93.5\\ 66.9\\ 63.6\\ \end{array}$   | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2<br>91.4<br>99.4<br>1.4<br>83.2<br>98.6<br>87.1<br>97.5<br>2.8<br>0.6<br>56.0<br>0.6<br>94.1<br>94.2<br>67.2<br>64.0   | $\begin{array}{r} 41.9\\ 83.8\\ 7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 0.4\\ 71.7\\ 71.5\\ 47.5\\ 47.8\\ 47.8\\ \end{array}$   | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.0\\ 38.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ 0.2\\ 39.4\\ 0.5\\ 76.0\\ 76.0\\ 50.8\\ 50.8\\ 50.8\\ \end{array}$   | $\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 38.0\\ 38.0\\ 38.0\\ 27.4\\ 47.2\\ 27.6\\ 30.4\\ \end{array}$   | $\begin{array}{c} 33.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.0\\ 35.9\\ 91.4\\ 3.0\\ 35.9\\ 35.9\\ 35.9\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 34.4\\ \end{array}$   | $\begin{array}{r} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ 35.6\\ 36.2\\ \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 39.6\\ 2.8\\ 48.5\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 39.0\\ \end{array}$   | $\begin{array}{r} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 50.6\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ 0.2\\ 56.1\\ 55.5\\ 32.1\\ 36.0\\ \end{array}$   | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ \end{array}$   | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \end{array}$  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 AVC<br>HAMIUK 6 APS<br>HAMIUK 6 APS<br>HAMIUK 6 AUC<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK LOSS<br><b>Task Mean</b><br><b>Medical Seg</b><br>ACDC Dice Score<br>ACDC Dice Score<br>ACDC Dice Score<br>ACDC Dice Score<br>ACDC Coreall Acc@<br>ACDC Oreal Acc@<br>ACDC Oreal Task Mean<br><b>Ing to Txt ZS</b><br>Flickr30K Img2Txt Acc@1   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>0.6<br>86.3<br>86.5<br>57.9<br>57.8   | 85.5<br>95.5<br>3.7<br>100.0<br>0.0<br>94.8<br>99.7<br>1.1<br>85.2<br>98.3<br>1.5<br>98.3<br>1.5<br>93.3<br>98.6<br>1.9<br>0.2<br>56.7<br>0.5<br>85.8<br>86.2<br>57.3<br>57.3<br>57.3   | 83.9<br>95.6<br>4.2<br>99.7<br>100.0<br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>91.4<br>98.7<br>2.2<br>0.3<br>56.5<br>0.5<br>83.4<br>83.2<br>57.4<br>83.2<br>57.4<br>257.4<br>100<br>21.0   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ 0.5\\ 78.5\\ 78.7\\ 53.1\\ 52.7\\ 53.1\\ 52.7\\ 5.9\\ 20.0\\ \end{array}$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ \hline 0.4\\ 75.5\\ 75.1\\ 50.2\\ 50.3\\ 50.3\\ \hline 5.6\\ 19.3\\ \end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>93.7<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1<br>98.6<br>2.1<br>98.6<br>2.1<br>0.2<br>56.4<br>0.5<br>78.0<br>78.3<br>53.0<br>4.2<br>56.4<br>0.5<br>78.0<br>78.3<br>53.0<br>6.8<br>221  | $\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ \hline 0.5\\ 76.9\\ 77.0\\ 47.7\\ 50.5\\ 5.9\\ 20.4 \end{array}$  
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ \hline \\ 79.4\\ 79.0\\ 54.3\\ 53.3\\ \hline \\ 5.2\\ 18.8 \end{array}$  | $\begin{array}{r} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 0.4\\ 74.0\\ 73.5\\ 50.1\\ 49.5\\ 18.0\\ \end{array}$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.4\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.4\\ 93.4\\ 93.5\\ 66.9\\ 63.6\\ 4.1\\ 16.0\\ \end{array}$  | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2<br>91.4<br>99.4<br>1.4<br>83.2<br>98.6<br>1.6<br>87.1<br>97.5<br>2.8<br>0.6<br>94.1<br>94.2<br>67.2<br>64.0<br>3.7<br>16  | $\begin{array}{r} 41.9\\ 83.8\\
7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ \hline 0.4\\ 71.7\\ 71.5\\ 47.5\\ 47.5\\ 47.8\\ 4.7\\ 16.9\\ \end{array}$  | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ \hline \end{array}$   | 34.7<br>8.4<br>33.6<br>8.4<br>33.6<br>84.2<br>1.2<br>92.9<br>3.8<br>31.4<br>41.8<br>92.9<br>3.8<br>31.4<br>3.0<br>38.8<br>85.6<br>0.2<br>39.4<br>0.5<br>76.0<br>76.0<br>76.0<br>76.0<br>50.8<br>50.8<br>1.6   | $\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 91.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 86.1\\ 5.5\\ 0.2\\ 37.4\\ 47.2\\ 27.6\\ 46.7\\ 47.2\\ 27.6\\ 30.4\\ 48.8\\ 8.6\\ \end{array}$  | $\begin{array}{c} 333.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 91.5\\ 3.9\\ 91.4\\ 3.0\\ 35.9\\ 91.4\\ 3.0\\ 35.9\\ 91.4\\ 3.0\\ 35.9\\ 0.2\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 30.4\\ 30.4\\ 30.4\\ 2.0\\ 8.4 \end{array}$   
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ 35.6\\ 235.6\\ 36.2\\ 1.9\\ 8.9\end{array}$  | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 35.1\\ 35.1\\ 35.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 0.2\\ 0.3\\ 9.1\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3$   | 48.4<br>488.1<br>7.2<br>73.8<br>97.6<br>0.5<br>3.9<br>36.6<br>91.1<br>4.8<br>0.2<br>40.8<br>0.2<br>40.8<br>0.2<br>55.5<br>32.1<br>32.6<br>0.2<br>54.1<br>55.5<br>32.1<br>1.9<br>9.1   | 42.4<br>84.2<br>8.0<br>34.8<br>92.0<br>1.1<br>1.1<br>36.6<br>90.3<br>4.2<br>25.8<br>90.5<br>3.2<br>37.6<br>0.2<br>37.7<br>0.3<br>50.8<br>51.4<br>24.3<br>31.7<br>1.8<br>8 5.9   | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline \end{array}$  
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 BS<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 AVC<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 AVC<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIOK AUC Macro<br>HAMIOK DES<br>Macro<br>HAMIOK AUC Macro<br>HAMIOK AUC Macro<br>HICK AUC MINING<br>HICK AUC MINI   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.0<br>86.3<br>86.5<br>57.9<br>57.8<br>6.3<br>20.9<br>3.8   | 85.5         95.5           95.5         3.7           100.0         0.0           94.8         99.7           91.1         85.2           98.3         1.5           93.3         98.6           1.5         93.3           98.6         1.9           0.2         56.7           0.5.8         886.2           57.0         57.3           6.3         21.3           3.8         3.8   | 83.9<br>95.6<br>4.2<br>99.7<br><b>100.0</b><br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>83.9<br>91.4<br>98.7<br>2.2<br>0.3<br>56.5<br>83.4<br>83.2<br>57.4<br>56.7<br><b>2</b><br>7.0<br>21.0<br>3.8   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 0.3\\ 56.7\\ 0.3\\ 56.7\\ 578.5\\ 78.7\\ 53.1\\ 52.7\\ 5.9\\ 20.0\\ 3.8 \end{array}$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 99.2\\ 1.3\\ 88.0\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.1\\ 91.3\\ 56.2\\ 2.0\\ 0.3\\ 56.2\\ 50.3\\ \hline \\ 50.2\\ 50.3\\ \hline \\ 5.6\\ 19.3\\ 3.9 \end{array}$  | $\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 99.5\\ 1.0\\ 87.6\\ 99.5\\ 1.0\\ 87.6\\ 99.5\\ 1.3\\ 92.1\\ 98.6\\ 2.1\\ 0.2\\ 56.4\\ 0.5\\ 78.0\\ 52.4\\ 6.8\\ 22.1\\ 0.5\\ 78.0\\ 52.4\\ 3.7\\ \end{array}$   | $\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 100.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ 0.5\\ 76.9\\ 77.0\\ 47.7\\ 50.5\\ 5.9\\ 20.4\\ 3.8\\ \end{array}$   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 0.5\\ 79.4\\ 79.0\\ 54.3\\ 53.3\\ 5.2\\ 18.8\\ 3.9 \end{array}$   | $\begin{array}{r} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 0.2\\ 53.6\\ 0.4\\ 74.0\\ 73.5\\ 50.1\\ 49.5\\ 18.0\\ 3.9\\ \end{array}$  | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 85.3\\ 98.0\\ 0.2\\ 55.3\\ 0.4\\ 93.4\\ 93.4\\ 93.5\\ 66.9\\ 63.6\\ 66.9\\ 63.6\\ 116.0\\ 3.9\\ \end{array}$   | 74.3<br>92.5<br>5.1<br>96.9<br>100.0<br>0.2<br>91.4<br>99.4<br>1.4<br>83.2<br>98.6<br>1.6<br>56.0<br>0.6<br>94.1<br>94.2<br>67.2<br>64.0<br>64.0<br>3.7<br>16.1<br>3.9  | $\begin{array}{r} 41.9\\ 81.8\\ 7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 4.8\\ 0.2\\ 4.5\\ 4.7\\ 71.5\\ 47.5\\ 47.5\\ 47.5\\ 4.7\\ 16.9\\ 4.0\\ \end{array}$  | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 92.9\\ 3.8\\ 31.4\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 0.2\\ 39.4\\ 0.2\\ 39.4\\ 0.5\\ 76.0\\ 76.0\\ 0.5\\ 50.8\\ 85.6\\ 1.6\\ 8.1\\ 4.2\\ \end{array}$  | $\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 3.9\\ 30.5\\ 3.0\\ 38.0\\ 86.1\\ 5.5\\ 0.2\\ 37.4\\ 47.2\\ 27.6\\ 47.2\\ 27.6\\ 30.4\\ 1.8\\ 8.6\\ 4.1 \end{array}$   | $\begin{array}{c} 333.2\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.0\\ 35.9\\ 91.4\\ 3.0\\ 35.9\\ 91.4\\ 3.0\\ 35.9\\ 0.2\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 30.4\\ 30.4\\ 4.1\\ \end{array}$   | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ 35.6\\ 2\\ 35.6\\ 2\\ 1.9\\ 8.9\\ 4.1 \end{array}$   | $\begin{array}{c} 42.5\\ 85.9\\ 7.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 2.0\\ 9.1\\ 4.1\\ \end{array}$  | $\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 1.1\\ 55.5\\ 32.1\\ 36.0\\ 91.1\\ 4.8\\ 0.2\\ 40.8\\ 0.2\\ 1.9\\ 9.1\\ 4.1\\ 1.9\\ 9.1\\ 4.1\\ \end{array}$   | $\begin{array}{c} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2 \end{array}$  | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 4.3\\ 23.5\\ 88.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 88.3\\ 3.2\\ 36.2\\ 36.2\\ 36.2\\ 1.6\\ 8.4\\ 4.1\\ \end{array}$  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 AVC<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 AVC<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIOK AUC Macro<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK LOSS<br>Task Mean<br>Medical Seg<br>ACDC Dice Score<br>ACDC Mean Acc@<br>ACDC Overall Acc@<br>ACDC Overall Acc@<br>Flickr30K Img2Txt Acc@1<br>Flickr30K Img2Txt Acc@1   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.9<br>57.8<br>6.3<br>86.5<br>57.9<br>57.8<br>6.3<br>80.5<br>57.9<br>57.8   | $\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ \end{array}$  | 83.9<br>95.6<br>4.2<br>99.7<br>100.0<br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>56.3<br>56.5<br>0.5<br>83.4<br>83.2<br>56.4<br>0.5<br>83.4<br>83.2<br>56.1<br>1<br>7.0<br>2.2<br>0.3<br>56.5   
  | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 20.3\\ 3.8\\ 5.3\\ 20.8\\ $   | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 0.3\\ 56.2\\ 0.4\\ 75.5\\ 75.1\\ 50.2\\ 50.3\\ 55.3\\ 3.9\\ 5.1\\ 9.3\\ 3.9\\ 5.1\\ 20.0\\ \end{array}$   | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1<br>0.2<br>56.4<br>0.2<br>56.4<br>0.5<br>78.0<br>78.3<br>53.0<br>52.4<br>6.5<br>23.0  | 84.0<br>95.9<br>4.2<br>100.0<br>0.0<br>91.6<br>99.4<br>1.2<br>84.7<br>99.3<br>1.4<br>91.6<br>98.7<br>2.1<br>0.5<br>76.9<br>77.0<br>47.7<br>50.5<br>5.9<br>20.4<br>3.8<br>6.0  
  | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 82.1\\ 0.2\\ 55.5\\ 52.1\\ 0.5\\ 79.4\\ 79.0\\ 54.3\\ 53.3\\ 53.3\\ 53.3\\ 53.3\\ 19.8\\ \end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7$   | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 87.9\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ 0.2\\ 66.9\\ 63.6\\ 66.9\\ 63.6\\ 63.6\\ 3.9\\ 3.8\\ 3.9\\ 3.8\\ 16.5\\ 5.5\\ 3.9\\ 3.8\\ 16.5\\ 5.5\\ 3.9\\ 3.8\\ 16.5\\ 5.5\\ 16.5$ | 74.3<br>92.5<br>5.1<br>96.9<br>9100.0<br>0.2<br>91.4<br>83.2<br>98.6<br>87.1<br>97.5<br>2.8<br>0.6<br>87.1<br>97.5<br>2.8<br>0.6<br>94.1<br>94.2<br>64.0<br>94.2<br>64.0<br>3.7<br>16.1<br>3.9<br>4.0<br>0<br>7<br>3.9<br>4.0<br>0<br>7<br>3.9<br>17<br>3.9   | $\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 73.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 45.0\\ 45.0\\ 45.0\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\
47.5\\ 47.$    | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 1.5\\ 67.5\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ \end{array}$   | $\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ 0.5\\ 76.0\\ 76.0\\ 76.0\\ 50.8\\ 81.1\\ 4.2\\ 1.7\\ 7\\ 8\end{array}$   | $\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 38.0\\ 38.0\\ 38.0\\ 38.0\\ 237.4\\ 47.2\\ 27.6\\ 30.4\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ \end{array}$   |
33.2<br>80.3<br>8.4<br>24.6<br>89.5<br>1.1<br>36.8<br>91.5<br>29.5<br>91.4<br>3.9<br>29.5<br>91.4<br>3.0<br>29.5<br>91.4<br>3.0<br>29.5<br>37.0<br>0.2<br>53.7<br>53.4<br>30.4<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>2.0<br>8.4<br>4.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1<br>3.1   | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 4.4\\ 24.5\\ 89.1\\ 3.2\\ 82.2\\ 82.8\\ 8.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.3\\ 54.5\\ 54.2\\ 35.6\\ 2\\ 35.6\\ 2\\ 1.9\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ \end{array}$   | $\begin{array}{c} 42.5 \\ 85.9 \\ 7.7 \\ 52.8 \\ 97.7 \\ 0.9 \\ 48.1 \\ 94.0 \\ 3.6 \\ 39.6 \\ 92.3 \\ 2.8 \\ 91.0 \\ 5.0 \\ 0.2 \\ 40.6 \\ \hline 0.3 \\ 60.3 \\ 35.1 \\ 60.3 \\ 35.1 \\ 91.0 \\ 4.1 \\ 2.3 \\ 9.4 \\ 9.4 \\ \end{array}$  | $\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9$  | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 1.7\\ 8.8\\ \end{array}$   | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 84.2\\ 4.3\\ 23.5\\ 88.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ 36.2\\ 1.6\\ 8.3\\ 5.8\\ 0.2\\ 36.2\\ 1.6\\ 8.3\\ 4.1\\ 1.6\\ 8.3\\ 8.3\\ 3.2\\ 3.2\\ 3.2\\ 3.2\\ 3.2\\ 3.2\\ 3.2\\ 3$   
  |
| HAMIUK 3 APS<br>HAMIUK 3 AUC<br>HAMIUK 3 BS<br>HAMIUK 4 APS<br>HAMIUK 4 APS<br>HAMIUK 4 AUC<br>HAMIUK 4 AUC<br>HAMIUK 5 APS<br>HAMIUK 5 APS<br>HAMIUK 5 AVC<br>HAMIUK 5 BS<br>HAMIUK 6 APS<br>HAMIUK 6 AVC<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIUK AUC Macro<br>HAMIOK AUC Macro<br>HAMIOK AUC Macro<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK DE SMACRO<br>HAMIOK LOSS<br>Task Mean<br>Ing to Txt ZS<br>Flickr30K Img2Txt Acc@1<br>Flickr30K Img2Txt Acc@1<br>Flickr30K Txt2Img Acc@1<br>Flickr30K Txt2Img Acc@1<br>Flickr30K Txt2Img Acc@1<br>Flickr30K Txt2Img Acc@1   | 88.0<br>96.7<br>3.5<br>99.5<br>100.0<br>0.0<br>95.6<br>99.7<br>1.1<br>89.2<br>99.3<br>1.0<br>94.5<br>99.1<br>1.6<br>0.3<br>57.9<br>57.8<br>6.3<br>86.5<br>57.9<br>57.8<br>6.3<br>80.5<br>57.9<br>57.8   | $\begin{array}{c} 85.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 0.5\\ 85.8\\ 86.2\\ 57.3\\ 0.5\\ 7.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ \end{array}$   | 83.9<br>95.6<br>4.2<br>99.7<br>100.0<br>0.1<br>94.5<br>99.6<br>1.1<br>83.9<br>99.1<br>1.5<br>56.5<br>0.5<br>83.4<br>83.2<br>0.3<br>56.5<br>0.5<br>83.4<br>83.2<br>2.2<br>0.3<br>56.5<br>0.5<br>83.4<br>83.2<br>21.0<br>3.8<br>6.0<br>21.6<br>3.8   
   | $\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 20.3\\ 3.9\\ 1.1\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0$  | $\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 0.3\\ 56.2\\ 0.4\\ 75.5\\ 75.1\\ 50.2\\ 50.3\\ 55.3\\ 19.3\\ 3.9\\ 5.1\\ 20.0\\ 3.9\\ 5.1\\ 20.0\\ 3.9\end{array}$  | 83.0<br>95.3<br>4.1<br>98.5<br>100.0<br>0.1<br>99.5<br><b>1.0</b><br>87.6<br>99.1<br>1.3<br>92.1<br>98.6<br>2.1<br>0.2<br>56.4<br>0.2<br>56.4<br>0.5<br>78.0<br>78.3<br>53.0<br>52.4<br>6.5<br><b>23.0</b><br>3.8  | $\begin{array}{c} 84.0 \\ 95.9 \\ 4.2 \\ 100.0 \\ 0.0 \\ 91.6 \\ 99.4 \\ 1.2 \\ 84.7 \\ 99.3 \\ 1.4 \\ 91.6 \\ 98.7 \\ 2.1 \\ 0.2 \\ 56.2 \\ 2.1 \\ 0.5 \\ 77.0 \\ 47.7 \\ 50.5 \\ 5.9 \\ 20.4 \\ 3.8 \\ 6.0 \\ 21.0 \\ 3.8 \\ \end{array}$  
   | $\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.6\\ 2.1\\ 0.2\\ 55.5\\ 5.2\\ 1.3\\ 90.8\\ 52.1\\ 0.5\\ 79.4\\ 79.0\\ 54.3\\ 53.3\\ 53.3\\ 19.8\\ 3.9\\ 5.1\\ 19.8\\ 3.9\end{array}$   | $\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7$  | $\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 88.13\\ 98.0\\ 1.5\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 97.9\\ 2.4\\ 0.2\\ 55.3\\ 0.4\\ 93.4\\ 93.5\\ 66.9\\ 63.6\\ 63.6\\ 66.9\\ 63.6\\ 16.5\\ 3.9\\ 3.8\\ 16.5\\ 3.9\\ \end{array}$  | $\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 99.4\\ 99.4\\ 83.2\\ 98.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 94.1\\ 97.5\\ 2.8\\ 0.6\\ 94.1\\ 97.5\\ 2.8\\ 0.6\\ 94.1\\ 97.5\\ 2.8\\ 0.6\\ 0.6\\ 94.1\\ 3.7\\ 4.0\\ 17.3\\ 4.0\\ \end{array}$   | $\begin{array}{c} 41.9\\ 87.9\\
26.8\\ 92.3\\ 1.1\\ 54.1\\ 73.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 45.0\\ 45.0\\ 45.0\\ 47.5\\ 47.$    | $\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 1.5\\ 0.2\\ 67.6\\ 67.5\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ \end{array}$  | $\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 85.6\\ 5.4\\ 8.5\\ 65.6\\ 8.5\\ 1.6\\ 8.1\\ 4.2\\ 1.7\\ 7.8\\ 4.2 \end{array}$   | $\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 3.9\\ 30.5\\ 91.9\\ 3.0\\ 3.9\\ 3.0\\ 86.1\\ 5.5\\ 3.7.4\\ 47.2\\ 27.6\\ 47.2\\ 27.6\\ 47.2\\ 27.6\\ 41.1\\ 8.8\\ 8.4\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ \end{array}$   | $\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 0.2\\ 53.7\\ 0.2\\ 53.7\\ 0.2\\ 53.7\\ 0.2\\ 53.7\\ 0.2\\ 53.7\\ 0.2\\ 8.4\\ 4.1\\ 2.0\\ 8.4\\ 4.2\end{array}$   
  | $\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 87.5\\ 4.4\\ 24.5\\ 24.5\\ 37.2\\ 22.3\\ 37.5\\ 37.5\\ 37.5\\ 37.5\\ 37.5\\ 54.2\\ 35.6\\ 23.5\\ 35.6.2\\ 1.9\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ 4.1 \end{array}$   | $\begin{array}{c} 42.5 \\ 85.9 \\ 7.7 \\ 52.8 \\ 97.7 \\ 0.9 \\ 48.1 \\ 94.0 \\ 3.6 \\ 39.6 \\ 92.3 \\ 2.8 \\ 91.0 \\ 5.0 \\ 0.2 \\ 40.6 \\ 0.3 \\ 60.3 \\ 35.1 \\ 9.4 \\ 4.1 \\ 2.3 \\ 9.4 \\ 4.1 \end{array}$   | $\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.6\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 56.1\\ 55.5\\ 32.1\\ 36.0\\ 0.2\\ 56.1\\ 1.9\\ 9.1\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ \end{array}$   | $\begin{array}{r} 42.4\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 1.7\\ 8.8\\ 8.4\\ 4.2 \end{array}$   | $\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 84.2\\ 4.3\\ 23.5\\ 88.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 88.5\\ 3.2\\ 32.6\\ 88.3\\ 3.2\\ 36.2\\ 36.2\\ 36.2\\ 1.6\\ 8.3\\ 4.1\\ 1.6\\ 8.3\\ 4.2\\ \end{array}$   
  |
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AUC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AVC HAMIUK 5 BS HAMIUK 6 APS HAMIUK 6 AVC HAMIUK AUC Macro HAMIUK AUC Macro HAMIOK DES Macro HAMIOK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK	88.0 96.7 3.5 99.5 100.0 0.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.1 1.6 0.3 86.5 57.9 57.8 6.3 86.5 57.9 57.8 6.3 80.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 5.8 5.7 20.9 5.8 5.7 20.9 5.7 5.8 5.7 20.9 5.7 5.7 5.8 5.7 20.9 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	$\begin{array}{c} 85.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 0.5\\ 85.8\\ 86.2\\ 57.3\\ 0.5\\ 85.8\\ 86.2\\ 57.3\\ 0.2\\ 57.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 21.4\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.5 56.5 0.5 83.4 83.2 91.4 98.7 2.2 0.3 56.5 57.4 56.1 7.0 2.10 0.3 8 8.4 83.2 91.4 91.4 91.4 91.4 91.4 91.4 91.4 91.4	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 20.3\\ 56.7\\ 53.7\\ 53.7\\ 53.7\\ 53.7\\ 53.7\\ 53.7\\ 53.7\\ 53.9\\ 20.0\\ 3.8\\ 5.3\\ 3.9\\ 20.8\\ 3.9\\ 20.0\\ 3.8\end{array}$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 0.3\\ 56.2\\ 0.3\\ 55.0\\ 250.3\\ 3.9\\ 5.1\\ 20.0\\ 3.9\\ 21.2\\ 20.0\\ 3.8\\ \end{array}$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 98.6 2.1 0.2 56.4 0.5 78.0 78.0 78.0 78.3 53.0 52.4 6.8 22.1 3.7 6.5 23.0 3.8 22.1 3.8	$\begin{array}{c} 84.0 \\ 95.9 \\ 4.2 \\ 100.0 \\ 0.0 \\ 91.6 \\ 99.4 \\ 1.2 \\ 84.7 \\ 99.3 \\ 1.4 \\ 91.6 \\ 98.7 \\ 2.1 \\ 0.2 \\ 56.2 \\ 2.1 \\ 0.5 \\ 77.0 \\ 47.7 \\ 50.5 \\ 5.9 \\ 20.4 \\ 3.8 \\ 6.0 \\ 21.0 \\ 3.8 \\ 21.4 \\ 3.8 \\ 21.4 \\ 3.8 \end{array}$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 98.6\\ 2.1\\ 0.2\\ 55.5\\ 5.3\\ 53.3\\ 53.3\\ 53.3\\ 53.3\\ 53.3\\ 90.8\\ 5.1\\ 19.8\\ 3.9\\ 20.0\\ 3.8\\ \end{array}$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 63.6\\ 63.6\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 66.9\\ 63.6\\ 16.5\\ 3.9\\ 17.1\\ 3.9\\ \end{array}$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ \hline 0.6\\ 94.1\\ 94.2\\ 67$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 73.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 45.0\\ 45.0\\ 45.0\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.5\\ 47.4\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 14.0\\ 15.9\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 15$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ $	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.8\\ 8.0\\ 4.2\\ 27.6\\ 30.4\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.4\\ 4.2\\ 8.7\\ 8.7\\ 4.2\\ 8.7\\ 8.7\\ 8.7\\ 8.7\\ 8.7\\ 8.7\\ 8.7\\ 8.7$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 33.4\\ 4\\ 2.0\\ 8.4\\ 4.1\\ 2.0\\ 8.4\\ 4.2\\ 8.4\\ 8.4\\ 4.2\\ 8.4\\ 8.4\\ 8.4\\ 8.4\\ 8.4\\ 8.4\\ 8.4\\ 8.4$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 87.5\\ 87.1\\ 24.5\\ 89.1\\ 32.2\\ 22.2\\ 36.7\\ 35.6$	$\begin{array}{c} 42.5 \\ 85.9 \\ 7.7 \\ 52.8 \\ 97.7 \\ 0.9 \\ 48.1 \\ 94.0 \\ 39.6 \\ 92.3 \\ 2.8 \\ 91.0 \\ 5.0 \\ 0.2 \\ 40.6 \\ 0.3 \\ 60.3 \\ 60.3 \\ 35.1 \\ 35.1 \\ 39.0 \\ \end{array}$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 3.9\\ 3.9\\ 3.6\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 56.1\\ 55.5\\ 32.1\\ 336.0\\ 0.2\\ 56.1\\ 52.1\\ 336.0\\ 0.2\\ 56.1\\ 54.1\\ 4.1\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 8.9\\ 4.1\\ 8.9\\ 1.5\\ 4.1\\ 8.9\\ 1.5\\ 4.1\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1$	$\begin{array}{c} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 25.8\\ 90.5\\ 3.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 55.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 1.7\\ 8.8\\ 8.4\\ 4.2\\ 8.5\\ 4.2\\ 1.7\\ 1.8\\ 8.8\\ 4.2\\ 1.7\\ 1.8\\ 8.8\\ 4.2\\ 1.7\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \hline 1.6\\ 8.4\\ 4.1\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ \hline \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AVC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AS HAMIUK 6 APS HAMIUK 6 APS HAMIUK 6 AVC HAMIUK AUC Macro HAMIUK AUC Macro HAMIOK DES Macro HAMIOK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK AUC HAMIOK HAMI	<b>88.0</b> <b>96.7</b> <b>3.5</b> <b>99.5</b> <b>100.0</b> <b>0.0</b> <b>95.6</b> <b>99.7</b> <b>1.1</b> <b>89.2</b> <b>99.3</b> <b>1.0</b> <b>94.5</b> <b>99.1</b> <b>1.6</b> <b>0.3</b> <b>57.0</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>8.6</b> ,5 <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>8.6</b> ,5 <b>57.9</b> <b>5.7</b> <b>8.5</b> <b>5.7</b> <b>2.0</b> ,9 <b>3.8</b> <b>5.7</b> <b>2.0</b> ,9 <b>3.8</b> <b>5.7</b> <b>2.0</b> ,9 <b>3.8</b> <b>5.7</b> <b>2.0</b> ,9 <b>3.8</b> <b>5.7</b> <b>2.0</b> ,9 <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b>	$\begin{array}{c} 85.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.9\\ 3.8\\ 5.9\\ 3.3\\ 86.2\\ 55.8\\ 86.2\\ 57.0\\ 57.3\\ \hline \begin{array}{c} 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ \hline \begin{array}{c} 3.8\\ 5.9\\ 22.1\\ 3.8\\ 51.4\\ 3.8\\ 5.6\\ \end{array}$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.5 56.5 0.5 83.4 83.2 91.4 98.7 2.2 0.3 56.5 57.4 56.1 7.0 21.0 3.8 6.0 3.8 6.9	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 78.7\\ 56.7\\ 8.5\\ 78.7\\ 53.1\\ 52.7\\ 5.9\\ 20.0\\ 3.8\\ 5.3\\ 3.9\\ 20.8\\ 3.9\\ 20.8\\ 3.9\\ 20.8\\ 3.8\\ 5.8\\ \end{array}$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 0.3\\ 56.2\\ 0.3\\ 55.2\\ 0.3\\ 3.9\\ 5.1\\ 50.2\\ 20.0\\ 3.9\\ 21.2\\ 20.0\\ 3.9\\ 21.2\\ 20.0\\ 3.8\\ 6.5\\ \end{array}$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 98.6 2.1 0.2 56.4 0.5 78.0 78.0 78.0 78.0 78.3 53.0 52.4 6.8 22.1 3.7 6.5 23.0 3.8 6.9	$\begin{array}{c} 84.0 \\ 95.9 \\ 4.2 \\ 100.0 \\ 0.0 \\ 91.6 \\ 99.4 \\ 1.2 \\ 84.7 \\ 99.3 \\ 1.4 \\ 91.6 \\ 98.7 \\ 2.1 \\ 0.2 \\ 56.2 \\ 2.1 \\ 0.5 \\ 77.0 \\ 47.7 \\ 50.5 \\ 5.9 \\ 20.4 \\ 3.8 \\ 6.0 \\ 21.0 \\ 3.8 \\ 21.4 \\ 3.8 \\ 6.4 \\ \end{array}$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 98.6\\ 1.3\\ 90.8\\ 92.1\\ 0.2\\ 55.5\\ 52.5\\ 52.5\\ 52.3\\ 53.3\\ 53.3\\ 53.3\\ 53.3\\ 90.8\\ 5.1\\ 19.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ \end{array}$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 97.8\\ 2.8\\ 0.2\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 83.4\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 63.6\\ 63.6\\ 63.6\\ 63.6\\ 16.5\\ 3.9\\ 17.1\\ 3.9\\ 4.9\\ \end{array}$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 100.0\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ \hline 0.6\\ 94.1\\ 94.2\\ 67.2\\ 67.2\\ 67.2\\ 67.2\\ 64.0\\ \hline 3.7\\ 16.1\\ 3.9\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 73.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 43.7\\ 71.5\\ 47.5\\ 47.5\\ 47.5\\ 47.8\\ 4.7\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 14.0\\ 15.9\\ 4.0\\ 4.2\\ 4.6\\ \end{array}$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 14.2\\ \end{array}$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 1.2\\ 1.2\\ 1.2\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 3.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ 8.7\\ 4.1\\ 2.1\\ \end{array}$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 30.4\\ 30.4\\ 4.1\\ 2.0\\ 8.4\\ 4.1\\ 2.0\\ 8.4\\ 4.1\\ 1.8\\ \end{array}$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 87.5\\ 87.1\\ 24.5\\ 89.1\\ 32.2\\ 22.2\\ 36.7\\ 35.6\\ 35.6\\ 35.6\\ 35.6\\ 35.6\\ 35.6\\ 35.6\\ 1.9\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ 4.1\\ 8.7\\ 4.1\\ 1.9\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 3.6\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 33.9.0\\ \hline \\ 2.0\\ 9.1\\ 4.1\\ 2.3\\ 9.4\\ 4.1\\ 9.5\\ 4.1\\ 2.1\\ \end{array}$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 56.1\\ 55.5\\ 32.1\\ 32.1\\ 36.0\\ 1.9\\ 9.1\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 8.9\\ 4.1\\ 2.0\\ \end{array}$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 25.8\\ 90.5\\ 3.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ \hline 0.3\\ 50.8\\ 51.4\\ 8.5\\ 4.2\\ 1.7\\ 1.8\\ 8.8\\ 4.2\\ 8.5\\ 4.2\\ 1.7\\ 1.6\\ \end{array}$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \hline 1.6\\ 8.4\\ 4.1\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AUC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AVC HAMIUK 5 BS HAMIUK 6 APS HAMIUK 6 AVC HAMIUK AUC Macro HAMIUK AU	88.0 96.7 3.5 99.5 100.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.1 1.6 0.3 57.0 0.6 86.3 86.5 57.9 57.8 6.3 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 21.9	$\begin{array}{c} 85.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.9\\ 38.8\\ 86.2\\ 57.0\\ 57.3\\ \hline \begin{array}{c} \\ 6.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 21.4\\ 3.8\\ 6.6\\ 3.8\\ 21.6\\ \end{array}$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.5 56.5 0.3 83.4 83.2 0.3 56.5 0.5 83.4 83.2 0.3 57.4 56.1 7.0 21.0 3.8 6.0 21.6 3.8 6.0 21.6 3.8 6.9 3.8 6.9 3.8 22.5	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.2\\ \end{array}$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 1.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 1.3\\ 56.2\\ 50.3\\ 5.1\\ 2.0\\ 0.3\\ 5.1\\ 2.0\\ 0.3\\ 3.9\\ 5.1\\ 2.0.0\\ 3.9\\ 5.1\\ 2.0.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.5\\ 3.8\\ 21.9\\ 21.9\\ 21.9\\ 21.9\\ 3.8\\ 3.8\\ 21.9\\ 21.9\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 92.1 92.1 92.1 92.1 92.1 92.1 0.2 56.4 0.2 56.4 0.2 56.4 0.2 56.4 0.2 56.4 0.2 56.4 0.5 78.0 78.0 78.0 78.0 78.0 78.0 78.0 78.0	$\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 89.4\\ 1.2\\ 84.7\\ 99.3\\ 1.2\\ 84.7\\ 99.4\\ 91.6\\ 99.4\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 79.4\\ 79.0\\ 0.2\\ 55.3\\ 3.3\\ 53.3\\ \hline \\ 5.2\\ 18.8\\ 3.9\\ 5.1\\ 19.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 3.8\\ 6.0\\ 3.8\\ 20.7\\ \end{array}$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 9.8\\ 0.2\\ 53.6\\ 1.7\\ 83.4\\ 9.5\\ 0.1\\ 4.5\\ 18.0\\ 3.9\\ 5.0\\ 18.9\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 18.4\\ \end{array}$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 97.9$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 1.7.4$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 11.1\\ 54.1\\ 73.4\\ 28.4\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 43.7\\ 71.5\\ 47.8\\ 4.8\\ 0.2\\ 45.0\\ 4.2\\ 4.7\\ 11.5\\ 47.8\\ 4.7\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 14.0\\ 15.9\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 15.9\\ 16.0\\ 16.0\\ 16.0\\ 16.0\\ 16.0\\ 16.0\\ 16.0\\ 16.0\\ 10$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 46.9\\ 90.1\\ 4.5\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.3\\ \end{array}$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 1.2\\ 1.2\\ 1.6\\ 50.8\\ 50.8\\ \hline \end{array}$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ 8.7\\ 4.1\\ 2.1\\ 4.1\\ 9.4\\ \end{array}$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 51.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 4.1\\ 2.0\\ 8.4\\ 4.1\\ 2.0\\ 8.4\\ 4.1\\ 1.8\\ 8.4\\ 4.1\\ 1.8\\ 8.4\end{array}$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 37.2\\ 32.2\\ 38.5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 1.2\\ 9.2\\ 4.1\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ 4.1\\ 8.7\\ 4.1\\ 1.9\\ 4.1\\ 9.4\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 33.9.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 3.2\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ \hline 0.3\\ 50.8\\ 51.4\\ 1.6\\ 85.9\\ 4.2\\ 1.7\\ 1.8\\ 8.8\\ 4.2\\ 1.7\\ 8.8\\ 4.2\\ 8.5\\ 4.1\\ 1.6\\ 8.8\\ 4.2\\ 8.5\\ 8.5\\ 4.2\\ 8.5\\ 8.5\\ 4.2\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 32.6\\ 83.3\\ 5.8\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 8.3\\ 4.2\\ 35.9\\ 32.4\\ \hline 1.6\\ 8.3\\ 4.2\\ 7.9\\ 1.6\\ 4.2\\ 1.6\\ 4.2\\ 7.9\\ \hline \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AVC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AS HAMIUK 5 AS HAMIUK 6 APS HAMIUK 6 APS HAMIUK 6 AC HAMIUK AUC Macro HAMIUK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK AUC HAMIOK	88.0 96.7 3.5 99.5 100.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.1 1.6 0.3 57.0 0.6 86.3 86.5 57.9 57.8 6.3 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 21.9 3.8 6.9 3.8	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.9\\ 53.3\\ 98.6\\ 1.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ \hline \begin{array}{c} 6.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 5.8\\ 21.4\\ 3.8\\ 6.6\\ 3.8\\ 21.6\\ 3.8\\ 21.6\\ 3.8\\ \end{array}$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.5 56.5 2.2 0.3 56.5 83.4 83.2 0.5 83.4 83.2 0.5 83.4 56.5 7.4 56.1 7.0 21.0 3.8 6.0 21.6 6.0 21.6 3.8 8.22.5 3.8	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 20.3\\ 56.7\\ 5.7\\ 8.5\\ 78.$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 91.3\\ 88.0\\ 98.9\\ 91.3\\ 88.6\\ 2.0\\ 0.3\\ 56.2\\ 191.3\\ 56.2\\ 50.3\\ 55.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.5\\ 3.8\\ 21.9\\ 3.8\\ \end{array}$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 92.1 92.1 92.1 0.2 56.4 0.2 56.4 0.5 78.0 78.0 78.0 78.0 78.0 52.4 0.5 78.0 52.4 6.8 22.1 3.7 6.5 23.0 3.8 6.9 3.8 6.9 3.8 6.9 3.8	$\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 89.4\\ 1.2\\ 84.7\\ 99.4\\ 1.2\\ 84.7\\ 99.4\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.4\\ 3.8\\ 21.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 8.21.4\\ 3.8\\ 1.4\\ 3$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 83.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 79.4\\ 79.0\\ 0.2\\ 55.3\\ 3.3\\ 53.3\\ 5.2\\ 18.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 0.8\\ 8.20.7\\ 3.9\\ \end{array}$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 9.8\\ 0.2\\ 53.6\\ 0.4\\ 73.5\\ 50.1\\ 49.5\\ 18.0\\ 3.9\\ 5.0\\ 18.9\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 18.4\\ 3.9\end{array}$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 93.5\\ 1.5\\ 3.9\\ 1.1\\ 16.5\\ 3.9\\ 17.1\\ 3.9\\ 17.3\\ 3.9\\ \end{array}$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 3.9\\ 17.4\\ 3.9\end{array}$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 11.1\\ 54.1\\ 73.4\\ 28.4\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 43.7\\ 71.5\\ 47.8\\ 4.8\\ 0.2\\ 45.0\\ 4.2\\ 4.7\\ 11.5\\ 47.8\\ 4.7\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 14.0\\ 15.9\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0\\ 4.0$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ \end{array}$	$\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 1.2\\ 1.2\\ 1.2\\ 1.0\\ 55.8\\ \end{array}$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 84.0\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ 8.7\\ 4.1\\ 2.1\\ 4.1\\ 9.4\\ 4.1\end{array}$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 37.0\\ 0.2\\ 53.7\\ 30.4\\ 34.4\\ 2.0\\ 8.4\\ 4.1\\ 2.0\\ 8.4\\ 4.1\\ 1.8\\ 8.4\\ 4.1\\ 1.8\\ 4.1\\ 1.8\\ 4.1\\ \end{array}$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 22.3\\ 87.5\\ 89.1\\ 232.2\\ 36.7\\ 36.7\\ 36.7\\ 36.7\\ 0.2\\ 36.7\\ 36.7\\ 1.2\\ 2.2\\ 9.2\\ 4.1\\ 8.7\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 39.0\\ 0.2\\ 40.6\\ 1.3\\ 39.0\\ 0.2\\ 4.1\\ 2.1\\ 4.1\\ 9.9\\ 4.1\\ \end{array}$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 40.8\\ 40.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 4.1\\ 0.2\\ 4.1\\ 0.5\\ 4.1\\ 2.0\\ 0.4\\ 1.8,9\\ 4.1\\ 0.5\\ 4.1\\ 0.5\\ 1.8,9\\ 4.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.8,9\\ 1.1\\ 0.5\\ 1.8,9\\ 1.8$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 8.5\\ 4.2\\ 1.7\\ 8.8\\ 4.2\\ 8.5\\ 4.2\\ 8.5\\ 4.2\\ 1.6\\ 8.9\\ 4.2\\ \end{array}$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 32.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \hline 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 4.2\\ 7.9\\ 4.2\\ \hline \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AUC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AS HAMIUK 6 APS HAMIUK 6 APS HAMIUK 6 AC HAMIUK AUC Macro HAMIUK AUC Macro HAMIOK AUC Macro HAMIOK DES Macro HAMIOK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIO	<b>88.0</b> <b>96.7</b> <b>3.5</b> <b>99.5</b> <b>100.0</b> <b>95.6</b> <b>99.7</b> <b>1.1</b> <b>89.2</b> <b>99.3</b> <b>99.3</b> <b>1.0</b> <b>94.5</b> <b>99.1</b> <b>1.6</b> <b>0.3</b> <b>57.0</b> <b>0.6</b> <b>86.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>21.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.193.8</b> <b>6.193.8</b> <b>6.1193.75.75.75.75.75.75.75.75</b>	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.3\\ 38.6\\ 25.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 21.4\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 0.7\\ \end{array}$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.1 83.9 99.1 1.5 2.2 2.3 56.5 56.5 0.3 83.4 83.2 0.3 56.5 7.4 56.1 7.0 21.0 3.8 6.0 21.6 3.8 6.2 2.2 5.8 8 3.8 6.9 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 3.8 6.4 2.5 5.5 5.5 5.5 5.5 5.5 5.5 5.5 5.5 5.5	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 298.5\\ 1.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.2\\ 3.8\\ 5.8\\ 3.8\\ 5.5\\ 0.7\\ \end{array}$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 91.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 2.0\\ 0.3\\ 56.2\\ 50.3\\ 55.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ \end{array}$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 92.1 92.1 92.1 92.1 0.2 56.4 0.2 56.4 0.5 78.0 78.0 78.0 78.0 52.4 0.5 52.4 6.8 22.1 3.7 6.5 23.0 3.8 6.9 3.8 7 6.9 7 7 7 7 7 8 7 7 7 7 7 7 7 7 7 7 7 7 7	$\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 84.7\\ 99.4\\ 1.2\\ 84.7\\ 99.4\\ 1.2\\ 84.7\\ 99.4\\ 1.2\\ 56.2\\ 56.2\\ 1.2\\ 56.2\\ 56.2\\ 1.2\\ 56.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 83.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 52.1\\ 0.2\\ 55.5\\ 53.3\\ 53.3\\ 53.3\\ 5.2\\ 18.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 0.3.8\\ 6.0\\ 0.7\\ 5.8\\ 0.7\\ 0.7\\ 0.7\\ 0.8\\ 0.8\\ 0.7\\ 0.8\\ 0.8\\ 0.8\\ 0.8\\ 0.8\\ 0.8\\ 0.8\\ 0.8$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 9.8\\ 0.2\\ 53.6\\ 0.4\\ 73.5\\ 50.1\\ 49.5\\ 18.0\\ 3.9\\ 5.0\\ 18.9\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 18.4\\ 3.9\\ 4.8\\ 3.9\\ 4.8\\ 3.9\\ 4.8\\ 0.7\\ \end{array}$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 63.6\\ 0.2\\ 93.5\\ 63.6\\ 0.4\\ 93.4\\ 93.5\\ 66.9\\ 63.6\\ 0.4\\ 93.4\\ 93.5\\ 16.5\\ 3.9\\ 17.1\\ 3.9\\ 4.9\\ 3.9\\ 4.3\\ 3.9\\ 4.3\\ 0.7\\ \end{array}$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 3.9\\ 4.1\\ 3.9\\ 4.1\\ 3.9\\ 4.1\\ 0.7\\ \end{array}$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 11.1\\ 54.1\\ 73.4\\ 28.4\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 43.7\\ 71.5\\ 47.8\\ 4.8\\ 0.2\\ 45.0\\ 4.2\\ 4.7\\ 11.5\\ 47.8\\ 4.7\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 16.9\\ 4.0\\ 4.0\\ 3.9\\ 0.7\\ 16.0\\ 0.7\\ 10.0\\ 10.$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 46.9\\ 90.1\\ 4.5\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 3.7\\ 0.7\\ \end{array}$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 8.4\\ 2.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 3$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ 8.7\\ 4.1\\ 2.1\\ 4.1\\ 2.0\\ 0.7\\ \end{array}$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 29.5\\ 3.9\\ 29.5\\$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 22.3\\ 87.5\\ 22.3\\ 87.5\\ 87.1\\ 22.3\\ 87.5\\ 87.1\\ 22.3\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 33.5$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 0.3\\ 6\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 35.1\\ 39.0\\ 0.2\\ 40.6\\ 4.1\\ 2.1\\ 4.1\\ 2.1\\ 4.1\\ 2.1\\ 4.1\\ 2.1\\ 4.1\\ 2.0\\ 9.9\\ 4.1\\ 2.0\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 37.6\\ 85.9\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 1.8\\ 8.5\\ 4.2\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 0,7\\ \end{array}$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 5.8\\ 0.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \hline 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 0.7\\ \hline \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 BS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AVC HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AS HAMIUK 5 AS HAMIUK 6 APS HAMIUK 6 APS HAMIUK 6 AC HAMIUK AUC Macro HAMIUK AUC MACR	<b>88.0</b> <b>96.7</b> <b>3.5</b> <b>99.5</b> <b>100.0</b> <b>95.6</b> <b>99.7</b> <b>1.1</b> <b>89.2</b> <b>99.3</b> <b>1.0</b> <b>94.5</b> <b>99.1</b> <b>1.6</b> <b>0.3</b> <b>57.0</b> <b>0.6</b> <b>86.3</b> <b>86.5</b> <b>57.9</b> <b>57.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>21.4</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.9</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.8</b> <b>6.19</b> <b>3.5</b> <b>5.10</b> <b>5.10</b> <b>5.10</b> <b>5.10</b> <b>5.10</b>	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 21.3\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 21.4\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 3.8\\ 5.9\\ 0.7\\ 53.4\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6\\ 1.6$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.1 83.9 99.1 1.5 5.5 5.5 83.4 83.2 0.3 56.5 57.4 56.1 7.0 21.0 3.8 6.0 21.6 3.8 6.2 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.4 22.5 3.8 6.5 7 59.5 5 5.5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.5\\ 78.5\\ $	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 91.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 2.0\\ 0.3\\ 56.2\\ 50.3\\ 55.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 3.9\\ 3.8\\ 5.0\\ 5.0\\ 3.8\\ 5.0\\ 5.0\\ 5.0\\ 5.0\\ 5.0\\ 5.0\\ 5.0\\ 5.0$	83.0 95.3 4.1 98.5 100.0 0.1 93.7 99.5 1.0 87.6 99.1 1.3 92.1 92.1 92.1 92.1 92.1 92.1 0.2 56.4 0.2 56.4 0.5 78.0 78.0 78.0 78.0 52.4 0.5 78.0 52.4 6.8 22.1 3.7 52.4 6.8 22.1 3.8 6.9 3.8 7 5.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 5.0 7 8.0 7 8.0 7 8.0 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 7 8.0 7 8.0 7 8.0 7 8.0 7 8.0 7 8.0 7 8.0 7 8.0 7 8.0 8 8 8 8.0 7 7 8.0 8 8 8 8 8.0 9 8.5 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	$\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.2\\ 56.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.2\\ 56.2\\ 1.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 1.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 1.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 1.4\\ 1.4\\ 1.4\\ 1.4\\ 1.4\\ 1.4\\ 1.4\\ 1.4$	$\begin{array}{c} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 99.0\\ 1.3\\ 90.8\\ 99.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 5.2\\ 18.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 0.3.8\\ 6.0\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 43.5\\ 1.3\\ 3.9\\ 5.8\\ 0.7\\ 43.5\\ 1.3\\ 3.9\\ 5.8\\ 0.7\\ 43.5\\ 1.3\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 5.8\\ 0.7\\ 3.9\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.9\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 98.0\\ 1.7\\ 83.4\\ 98.0\\ 1.7\\ 83.4\\ 2.8\\ 0.2\\ 53.6\\ 1.7\\ 83.9\\ 1.7\\ 83.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 8.9\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7\\ 1.7$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 55.3\\ 0.2\\ 493.5\\ 0.2\\ 55.3\\ 0.2\\ 65.3\\ 0.2\\ 65.3\\ 0.2\\ 65.3\\ 0.2\\ 65.3\\ 0.2\\ 10.2$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 3.7\\ 16.1\\ 3.9\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 3.9\\ 4.1\\ 17.4\\ 3.9\\ 4.1\\ 0.7\\ 50.0\\ 7\\ 50.0\\ 0.6\\ 17.4\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5$	$\begin{array}{c} 41.9\\ 87.9\\ 26.8\\ 92.3\\ 11.1\\ 54.1\\ 73.4\\ 28.4\\ 89.3\\ 4.8\\ 0.2\\ 45.0\\ 43.7\\ 71.5\\ 47.8\\ 4.8\\ 0.2\\ 45.0\\ 4.2\\ 4.7\\ 11.5\\ 47.8\\ 4.7\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 16.9\\ 4.0\\ 4.2\\ 17.1\\ 16.9\\ 4.0\\ 4.0\\ 4.0\\ 3.9\\ 4.6\\ 4.0\\ 0.7\\ 48.9\\ 10.7\\ 10.7\\ 10.5\\$	$\begin{array}{c} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 2.9\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ 15.3\\ 4.0\\ 3.7\\ 15.3\\ $	$\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ $	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 8.4\\ 32.3\\ 89.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 92.0\\ 36.4\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.6\\ 4.1\\ 1.8\\ 8.9\\ 4.2\\ 8.7\\ 4.1\\ 2.0\\ 4.1\\ 2.1\\ 4.1\\ 2.0\\ 0.7\\ 50.0\\ 7\\ 50.0\\ \end{array}$	33.2 80.3 80.3 8.4 24.6 89.5 1.1 36.8 91.5 29.5 91.4 3.9 29.5 91.4 3.9 29.5 37.0 0.2 53.7 53.4 2.0 8.4 4.1 2.0 8.4 4.1 2.0 8.4 4.1 1.8 3.9 4.1 1.2 2.0 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.3.9 2.2 1.4 3.5.9 1.4 3.5.9 1.4 3.2 1.4 3.2 1.4 3.2 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.2 1.4 3.4 4.1 1.8 3.4 4.1 1.8 3.4 4.1 1.8 3.3 4.1 1.7 1.4 3.2 1.4 3.4 4.1 1.6 1.7 1.5 3.4 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 4.1 1.8 3.5 3	$\begin{array}{c} 31.5\\ 80.7\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 32.2\\ 87.5\\ 89.1\\ 32.2\\ 82.8\\ 5.7\\ 0.2\\ 36.7\\ 0.2\\ 36.7\\ 36.2\\ 1.9\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ 4.1\\ 8.7\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 2.0\\ 0.7\\ 50.0\\ 7\\ 50.0\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 0.3\\ 60.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 2.9\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$	$\begin{array}{c} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 37.7\\ 37.6\\ 85.9\\ 5.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 1.8\\ 8.5\\ 4.2\\ 8.5\\ 4.2\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 9.9\\ 4.2\\ 1.8\\ 9.9\\ 4.2\\ 1.8\\ 8.9\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 3.2\\ 32.6\\ 83.3\\ 3.2\\ 36.2\\ \hline 0.3\\ 50.9\\ 51.4\\ 26.9\\ 32.4\\ \hline 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 0.7\\ 50.4\\ \hline \end{array}$
HAMIUK 3 APS HAMIUK 3 AUC HAMIUK 3 AUC HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 APS HAMIUK 4 AUC HAMIUK 5 APS HAMIUK 5 APS HAMIUK 5 AS HAMIUK 5 AS HAMIUK 6 APS HAMIUK 6 APS HAMIUK 6 AC HAMIUK AUC Macro HAMIUK AUC Macro HAMIOK AUC HAMIOK AUC MACRO HAMIOK AUC HAMIOK AUC MACRO HAMIOK AUC HAMIOK AUC MACRO HAMIOK AUC HAMIOK AUC MACRO HAMIOK AUC MACRO HAMI	88.0 96.7 3.5 99.5 100.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.1 1.6 0.3 57.0 0.6 86.3 86.5 57.9 57.8 6.3 20.9 3.8 5.7 20.9 3.8 21.4 3.8 6.9 3.8 21.9 3.8 21.9 3.8 6.9 3.8 21.9 3.8 6.9 3.8 21.9 3.8 6.9 5.7 20.9 3.8 21.9 3.8 5.7 20.9 3.8 21.9 3.8 6.9 5.7 20.9 3.8 21.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 3.8 5.7 20.9 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.3\\ 1.5\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 21.3\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 0.7\\ 55.2\\ 100\\ 100\\ 100\\ 100\\ 100\\ 100\\ 100\\ 10$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.5 5.5 2.2 2.3 83.4 83.2 91.4 98.7 2.2 2.5 83.4 83.2 0.5 83.4 83.2 2.6 57.4 56.1 21.0 3.8 6.0 21.6 3.8 6.2 22.5 3.8 6.4 0.7 59.5 0.7 59.5 0.7 59.5 0.7 59.5 56.2	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.1\\ 52.7\\ 0.3\\ 85.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.2\\ 3.8\\ 5.8\\ 3.8\\ 5.5\\ 0.7\\ 49.7\\ 0.7\\ 93.1\\ 100 \\ 100$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 91.3\\ 88.0\\ 98.9\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 2.0\\ 0.3\\ 56.2\\ 2.0\\ 0.3\\ 56.2\\ 50.3\\ 55.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ $	$\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ 1.0\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 92$	$\begin{array}{r} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 84.7\\ 99.3\\ 84.7\\ 99.4\\ 1.2\\ 84.7\\ 91.6\\ 99.4\\ 1.2\\ 56.2\\ 21.0\\ 0.2\\ 56.2\\ 256.2\\ 0.2\\ 56.2\\ 21.0\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 48.3\\ 0.7\\ 0.7\\ 10\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.$	$\begin{array}{r} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 99.0\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 0.2\\ 55.5\\ 2.1\\ 19.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 3.8\\ 6.0\\ 7.4\\ 3.9\\ 5.8\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 1.3\\ 0.3\\ 0.7\\ 1.3\\ 0.7\\ 1.3\\ 0.7\\ 1.3\\ 0.3\\ 0.7\\ 1.3\\ 0.3\\ 0.7\\ 1.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0.3\\ 0$	$\begin{array}{r} 74.2\\ 94.4\\ 5.0\\ 98.7\\ 100.0\\ 0.2\\ 83.2\\ 98.8\\ 1.7\\ 75.8\\ 98.0\\ 1.7\\ 83.4\\ 98.0\\ 1.7\\ 83.4\\ 98.0\\ 1.7\\ 83.4\\ 9.5\\ 0.2\\ 3.6\\ 1.7\\ 83.4\\ 9.5\\ 0.1\\ 49.5\\ 18.9\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 17.8\\ 3.9\\ 18.4\\ 3.9\\ 18.4\\ 3.9\\ 4.8\\ 0.7\\ 53.8\\ 0.7\\ 53.8\\ 0.7\\ 48.6\\ \end{array}$	$\begin{array}{c} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.15\\ 81.3\\ 98.0\\ 0.15\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 4.3\\ 93.5\\ 65.3\\ 93.5\\ 66.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 63.6\\ 16.5\\ 3.9\\ 17.1\\ 3.8\\ 16.5\\ 3.9\\ 17.1\\ 3.9\\ 4.3\\ 4.3\\ 4.3\\ 4.3\\ 4.3\\ 4.3\\ 4.3\\ 4.3$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 99.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 94.1\\ 94.2\\ 64.0\\ 3.7\\ 16.1\\ 3.9\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.4\\ 3.9\\ 4.1\\ 17.4\\ 3.9\\ 4.1\\ 0.7\\ 50.0\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ $	$\begin{array}{c} 41.9\\ 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 73.4\\ 28.4\\ 8.9\\ 3.4\\ 28.4\\ 8.9\\ 3.4\\ 28.4\\ 8.9\\ 3.4\\ 4.8\\ 0.2\\ 4.5\\ 4.5\\ 4.5\\ 4.5\\ 4.5\\ 4.5\\ 4.5\\ 4.5$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 4.2\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 3.9\\ 15.5\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 3.7\\ 47.3\\ 0.7\\ 47.3\\ 0.7\\ 52.4\\ \end{array}$	$\begin{array}{c} 34.7\\ 81.7\\ 8.4\\ 33.6\\ 84.2\\ 1.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ $	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 30.4\\ 1.0\\ 30.4\\ 1.0\\ 30.4\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9\\ 3.9$	$\begin{array}{c} 31.5\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 22.3\\ 87.5\\ 87.5\\ 87.1\\ 22.3\\ 87.5\\ 87.5\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 33.5\\ 54.2\\ 33.5\\ 54.2\\ 33.5\\ 54.2\\ 33.5\\ 54.2\\ 33.5\\ 54.2\\ 33.5$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 0.3\\ 6\\ 39.6\\ 92.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2\\ 1.2$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 3.9\\ 36.6\\ 93.4\\ 92.5\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 5.8\\ 90.5\\ 5.6\\ 0.2\\ 37.7\\ \hline 0.3\\ 50.8\\ 51.4\\ 224.3\\ 31.7\\ \hline 1.8\\ 8.5\\ 4.2\\ 8.5\\ 8.5\\ 4.2\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5\\ 8.5$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 23.5\\ 88.5\\ 3.2\\ 32.6\\ 83.3\\ 3.2\\ 32.6\\ 83.3\\ 3.2\\ 36.2\\ 36.2\\ 36.2\\ 36.2\\ 36.2\\ 36.2\\ 1.6\\ 8.4\\ 4.1\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 0.7\\ 50.4\\ 0.7\\ 50.4\\ 0.7\\ 52.1\\ \end{array}$
HAMIOK 3 APS HAMIOK 3 AUC HAMIOK 3 AUC HAMIOK 4 APS HAMIOK 4 APS HAMIOK 4 APS HAMIOK 4 AUC HAMIOK 5 APS HAMIOK 5 APS HAMIOK 5 AVC HAMIOK 6 APS HAMIOK 6 APS HAMIOK 6 AVC HAMIOK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HA	<b>88.0</b> <b>96.7</b> <b>3.5</b> <b>99.5</b> <b>100.0</b> <b>95.6</b> <b>99.7</b> <b>1.1</b> <b>89.2</b> <b>99.3</b> <b>1.0</b> <b>94.5</b> <b>99.1</b> <b>1.0</b> <b>94.5</b> <b>99.7</b> <b>1.1</b> <b>89.2</b> <b>99.3</b> <b>8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>5.7</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.5</b> <b>7.7</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.3</b> <b>20.9</b> <b>3.8</b> <b>6.5</b> <b>7.7</b> <b>20.9</b> <b>3.8</b> <b>6.1</b> <b>0.7</b> <b>5.7</b> <b>5.7</b> <b>2.0</b> <b>3.8</b> <b>6.1</b> <b>3.6</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>3.9</b> <b>3.8</b> <b>6.1</b> <b>0.7</b> <b>5.1</b> .0 <b>0.7</b> <b>5.1</b> .0 <b>0.7</b> <b>5.1</b> .0 <b>0.7</b> <b>5.1</b> .0 <b>2.1</b> .1	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 0.5\\ 85.8\\ 86.2\\ 0.5\\ 57.3\\ 21.3\\ 3.8\\ 86.6\\ 3.8\\ 21.6\\ 3.8\\ 21.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 22.0\\ 5.2\\ 5.2\\ 22.0\\ 5.2\\ 5.2\\ 5.2\\ 5.2\\ 5.2\\ 5.2\\ 5.2\\ 5.2$	83.9 95.6 4.2 99.7 100.0 0.1 94.5 99.6 1.1 83.9 99.1 1.1 83.9 99.1 50.5 83.4 83.2 0.3 56.5 57.4 56.1 7.0 21.0 3.8 6.0 21.6 3.8 6.0 21.6 3.8 6.2 2.5 5.5 5 5.5 5 7.5 5.5 5 0.7 5 9.5 5 6.2 22.5 3.8 8.4 20.7 7 5 9.5 5 5 6.2 22.5 3.8 8 6.4 22.5 5 5 6.5 22.5 3.8 8 6.4 22.5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 8.3\\ 56.7\\ 0.3\\ 8.3\\ 20.0\\ 3.8\\ 5.8\\ 3.9\\ 20.0\\ 3.8\\ 5.8\\ 3.8\\ 20.2\\ 3.8\\ 5.8\\ 3.8\\ 5.5\\ 0.7\\ 49.7\\ 0.7\\ 49.7\\ 0.7\\ 53.1\\ 20.6\\ \end{array}$	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 99.2\\ 88.0\\ 98.9\\ 98.9\\ 98.9\\ 88.0\\ 98.9\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 19.3\\ 3.9\\ 56.2\\ 50.2\\ 50.2\\ 50.2\\ 50.3\\ 3.9\\ 92.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 21.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 20.6\\ 19.3\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 20.6\\ 19.3\\ 3.8\\ 10.2\\ 10$	$\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ 1.0\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 92.1\\ 92.1\\ 92.1\\ 0.2\\ 56.4\\ 2.1\\ 0.2\\ 56.4\\ 0.2\\ 56.4\\ 2.1\\ 3.2\\ 3.0\\ 3.8\\ 22.1\\ 3.8\\ 6.9\\ 3.8\\ 6.2\\ 0.7\\ 50.3\\ 8.8\\ 21.9\\ 3.8\\ 6.2\\ 0.7\\ 55.5\\ 22.1\\ \end{array}$	$\begin{array}{r} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 2.1\\ 0.2\\ 56.2\\ 0.5\\ 76.9\\ 20.4\\ 3.8\\ 6.4\\ 21.0\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 0.7\\ 48.3\\ 20.8\\$	$\begin{array}{r} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 99.0\\ 1.3\\ 83.3\\ 99.6\\ 1.3\\ 90.8\\ 99.0\\ 1.3\\ 90.8\\ 99.0\\ 1.3\\ 90.8\\ 99.0\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 0.2\\ 1.3\\ 53.3\\ 3.9\\ 5.8\\ 0.0\\ 5.8\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 43.5\\ 0.7\\ 19.9\\ 10.7\\ $	$74.2 \\ 94.4 \\ 5.0 \\ 98.7 \\ 100.0 \\ 0.2 \\ 98.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 97.8 \\ 3.4 \\ 97.8 \\ 0.2 \\ 53.6 \\ 0.4 \\ 74.0 \\ 73.5 \\ 1.7 \\ 83.4 \\ 97.8 \\ 0.2 \\ 53.6 \\ 1.7 \\ 1.$	$\begin{array}{l} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.1\\ 5.5\\ 39.0\\ 0.2\\ 55.3\\ 0.2\\ 4.3\\ 93.5\\ 66.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 63.6\\ 16.5\\ 3.9\\ 17.1\\ 3.8\\ 16.5\\ 3.9\\ 4.3\\ 3.9\\ 4.3\\ 3.9\\ 17.3\\ 3.9\\ 4.3\\ 0.7\\ 61.9\\ 0.7\\ 0.7\\ 61.9\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 99.4\\ 83.2\\ 98.6\\ 1.4\\ 83.2\\ 98.6\\ 56.0\\ 0.6\\ 56.0\\ 0.6\\ 94.1\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.1\\ 0.7\\ 50.0\\ 0.7\\ 50.0\\ 0.8\\ 4.1\\ 1.4\\ 3.9\\ 17.4\\ 3.9\\ 4.1\\ 1.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 17.4\\ 3.9\\ 10.7\\ 10.7\\ 10.7\\ 10.8\\ 1$	$\begin{array}{c} 41.9\\ 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 43.7\\ 89.3\\ 0.2\\ 43.7\\ 89.3\\ 0.2\\ 43.7\\ 10.2\\ 45.0\\ 0.2\\ 45.0\\ 0.2\\ 45.0\\ 10.2\\ 17.1\\ 16.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 10.7\\ 48.9\\ 0.7\\ 48.9\\ 0.7\\ 48.9\\ 0.7\\ 48.9\\ 0.7\\ 18.2\\ 10$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 47.9\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.5\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 4.0\\ 15.3\\ 0.7\\ 47.3\\ 0.7\\ 0.7\\ 47.3\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 8.4\\ 2.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 0.3\\ 85.6\\ 5.4\\ 0.2\\ 39.4\\ 1.2\\ 76.0\\ 76.0\\ 55.8\\ 1.6\\ 8.1\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.5\\ 1.6\\ 1.7\\ 7.8\\ 1.3\\ 1.5\\ 1.6\\ 1.3\\ 1.3\\ 1.5\\ 1.5\\ 1.6\\ 1.3\\ 1.3\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 0.2\\ 37.4\\ 47.2\\ 27.6\\ 30.4\\ 47.2\\ 27.6\\ 30.4\\ 47.2\\ 27.6\\ 30.4\\ 4.1\\ 2.1\\ 4.1\\ 2.0\\ 0.7\\ 50.0\\ 0.7\\ 50.0\\ 0.7\\ 50.0\\ 0.14.3\\ \end{array}$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 37.0\\ 0.2\\ 53.7\\ 37.0\\ 0.2\\ 53.7\\ 33.4\\ 4.1\\ 1.2\\ 0.7\\ 41.3\\ 4.1\\ 1.7\\ 0.7\\ 41.3\\ 0.7\\ 13.7\\ 13.7\\ \end{array}$	$\begin{array}{c} 31.5\\ 80.7\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 22.3\\ 87.5\\ 87.5\\ 37.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 32.2\\ 33.5$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 0.3\\ 6\\ 39.6\\ 92.3\\ 2.8\\ 48.5\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 1.2\\ 39.0\\ 9.1\\ 4.1\\ 9.5\\ 4.1\\ 2.3\\ 9.9\\ 4.1\\ 2.3\\ 9.5\\ 4.1\\ 2.3\\ 0.7\\ 55.2\\ 0.7\\ 55.2\\ 15.2\\ 15.2\\ \end{array}$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 99.1\\ 1.9\\ 9.1\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 1.9\\ 1.9\\ 1.9\\ 1.9\\ 1.9\\ 1.9\\ 1$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 25.8\\ 90.5\\ 25.8\\ 90.5\\ 25.8\\ 90.3\\ 31.7\\ 0.2\\ 37.6\\ 90.3\\ 37.6\\ 0.2\\ 37.7\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 8.5\\ 4.2\\ 1.8\\ 0.7\\ 50.1\\ 0.7\\ 48.8\\ 14.1\\ 0.7\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 0.7\\ 50.1\\ 0.7\\ 1.8\\ 14.1\\ 0.7\\ 1.8\\ 14.1\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1.8\\ 1$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 32.5\\ 88.5\\ 32.5\\ 32.5\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 32.4\\ 32.6\\ 9\\ 32.4\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 0.7\\ 50.4\\ 0.7\\ 52.1\\ 14.1\\ \end{array}$
HAMIOK 3 APS HAMIOK 3 AUC HAMIOK 3 BS HAMIOK 4 APS HAMIOK 4 APS HAMIOK 4 AVC HAMIOK 4 AUC HAMIOK 5 APS HAMIOK 5 APS HAMIOK 5 AVC HAMIOK 5 AVC HAMIOK 6 APS HAMIOK 6 APS HAMIOK 6 AVC HAMIOK AUC Macro HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK AUC HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK AUC HAMIOK HAMIOK AUC HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIOK HAMIO	88.0 96.7 3.5 99.5 100.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.5 1.0 94.5 99.7 1.1 89.2 91.1 0.3 57.0 0.6 86.3 57.9 57.8 6.3 20.9 3.8 6.3 20.9 3.8 5.7 20.9 5.7 20.9 3.8 5.7 20.9 3.8 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 0.5\\ 85.8\\ 86.2\\ 0.5\\ 57.3\\ 21.3\\ 3.8\\ 86.6\\ 3.8\\ 21.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 0.7\\ 55.2\\ 22.0\\ 40.7\\ 55.2\\ 22.0\\ 40.7\\ 55.2\\ 22.0\\ 40.7\\ 55.2\\ 3.8\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.2\\ 2.20\\ 5.8\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.8\\ 5.9\\ 5.9\\ 5.2\\ 2.0\\ 5.8\\ 5.9\\ 5.2\\ 5.2\\ 5.9\\ 5.8\\ 5.9\\ 5.2\\ 5.2\\ 5.2\\ 5.2\\ 5.8\\ 5.9\\ 5.8\\ 5.8\\ 5.9\\ 5.8\\ 5.8\\ 5.8\\ 5.9\\ 5.8\\ 5.8\\ 5.8\\ 5.9\\ 5.8\\ 5.8\\ 5.8\\ 5.8\\ 5.8\\ 5.8\\ 5.8\\ 5.8$	83.9 95.6 4.2 99.7 100.0 1 94.5 99.6 1.1 83.9 99.1 1.1 83.9 99.1 1.1 83.9 99.1 56.5 51.4 98.7 2.2 0.3 56.5 51.4 98.7 2.2 0.3 56.5 57.4 57.4 57.4 57.4 57.4 57.4 57.4 57	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ $	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 99.2\\ 88.0\\ 98.9\\ 98.9\\ 98.9\\ 88.0\\ 98.9\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 19.3\\ 3.9\\ 56.2\\ 50.2\\ 50.2\\ 50.2\\ 50.3\\ 5.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 21.9\\ 3.8\\ 6.0\\ 0.7\\ 50.0\\ 20.6\\ 39.3\\ 39.3\\ \end{array}$	$\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ 1.0\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 92.1\\ 92.1\\ 92.1\\ 0.2\\ 56.4\\ 2.1\\ 0.2\\ 56.4\\ 2.1\\ 0.2\\ 56.4\\ 2.1\\ 3.8\\ 22.1\\ 3.8\\ 6.9\\ 22.1\\ 3.8\\ 6.9\\ 3.8\\ 22.1\\ 3.8\\ 6.9\\ 3.8\\ 21.9\\ 3.8\\ 6.2\\ 0.7\\ 55.5\\ 22.1\\ 24.9\\ \end{array}$	$\begin{array}{r} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ 0.5\\ 76.9\\ 20.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 48.3\\ 20.8\\ 22.7\\ 48.3\\ 20.8\\ 27.4\\ \end{array}$	$\begin{array}{r} 82.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 99.0\\ 1.3\\ 83.3\\ 99.6\\ 1.3\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 99.5\\ 1.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 79.4\\ 79.0\\ 3.3\\ 5.2\\ 18.8\\ 3.9\\ 0.2\\ 5.5\\ 5.1\\ 19.8\\ 3.9\\ 20.0\\ 3.8\\ 6.0\\ 0.7\\ 43.5\\ 5.8\\ 0.7\\ 43.5\\ 5.8\\ 0.7\\ 43.5\\ 20.7\\ 3.9\\ 5.8\\ 0.7\\ 43.5\\ 20.7\\ 3.8\\ 6.0\\ 0.7\\ 43.5\\ 5.8\\ 0.7\\ 43.5\\ 20.7\\ 3.8\\ 6.0\\ 0.7\\ 43.5\\ 20.7\\ 3.8\\ 6.0\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 20.7\\ 3.8\\ 6.0\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 20.7\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 43.5\\ 2.8\\ 3.8\\ 0.7\\ 3.8\\ 0.7\\ 3.8\\ 0.7\\ 3.8\\ 0.7\\ 3.8\\ 0.7\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 0.7\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8$	$74.2 \\ 94.4 \\ 5.0 \\ 98.7 \\ 100.0 \\ 0.2 \\ 98.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.8 \\ 97.8 \\ 97.8 \\ 98.8 \\ 0.2 \\ 53.6 \\ 0.4 \\ 74.0 \\ 75.5 \\ 1.7 \\ 83.4 \\ 97.8 \\ 0.2 \\ 53.6 \\ 1.7 \\ 1.7 \\ 83.4 \\ 1.7 \\ 1.7 \\ 53.8 \\ 0.7 \\ 0.7$	$\begin{array}{l} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.15\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 4.3\\ 93.5\\ 66.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 4.3\\ 3.8\\ 16.5\\ 3.9\\ 4.3\\ 3.9\\ 4.3\\ 3.9\\ 4.3\\ 3.9\\ 4.3\\ 0.7\\ 61.9\\ 0.7\\ 61.9\\ 0.7\\ 5.6\\ \end{array}$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 99.4\\ 83.2\\ 98.6\\ 1.6\\ 87.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ 94.1\\ 94.2\\ 64.0\\ 3.7\\ 16.1\\ 3.9\\ 94.2\\ 64.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.0\\ 17.3\\ 4.1\\ 0.7\\ 50.0\\ 0.7\\ 55.0\\ 0.18.4\\ 11.5\\ \end{array}$	$\begin{array}{c} 41.9\\ 41.9\\ 83.8\\ 7.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 43.7\\ 89.3\\ 0.2\\ 43.7\\ 89.3\\ 0.2\\ 45.0\\ 0.4\\ 45.0\\ 0.2\\ 45.0\\ 0.2\\ 45.0\\ 16.9\\ 4.0\\ 16.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 15.9\\ 4.0\\ 16.0\\ 16.0\\ 18.2\\ 1.8\\ 1.8\\ \end{array}$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 4.2\\ 15.5\\ 4.0\\ 15.3\\ 15.5\\ 15.3\\ 15.5\\ 15.3\\ 15.5\\ 15.3\\ 15.5\\ 1$	$\begin{array}{c} 34.7\\ 8.4\\ 33.6\\ 8.4\\ 2.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 41.8\\ 92.9\\ 3.8\\ 31.4\\ 91.2\\ 3.8\\ 31.4\\ 91.2\\ 0.2\\ 39.4\\ 1.0\\ 76.0\\ 50.8\\ 50.8\\ 1.6\\ 8.1\\ 4.2\\ 7.9\\ 5.5\\ 1.6\\ 4.9\\ 7.9\\ 5.8\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 1.6\\ 1.7\\ 7.8\\ 4.2\\ 7.9\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8\\ 3.8$	$\begin{array}{c} 35.1\\ 80.0\\ 8.4\\ 32.3\\ 89.0\\ 1.0\\ 36.4\\ 92.0\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 91.9\\ 30.5\\ 0.2\\ 37.4\\ 47.2\\ 27.6\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 30.4\\ 47.2\\ 47.2\\ 30.4\\ 47.2\\ 47.2\\ 30.4\\ 47.2\\ 47$	$\begin{array}{c} 33.2\\ 80.3\\ 80.3\\ 8.4\\ 24.6\\ 89.5\\ 1.1\\ 36.8\\ 91.5\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 91.4\\ 3.9\\ 29.5\\ 37.0\\ 0.2\\ 53.7\\ 37.0\\ 0.2\\ 53.7\\ 37.0\\ 0.2\\ 37.0\\ 0.2\\ 33.4\\ 4.1\\ 1.8\\ 8.4\\ 4.1\\ 1.7\\ 0.7\\ 41.3\\ 0.7\\ 13.7\\ 7.9\\ \end{array}$	$\begin{array}{c} 31.5\\ 80.7\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 3.2\\ 32.2\\ 36.7\\ 36.2\\ 36.2\\ 36.6\\ 36.2\\ 1.9\\ 8.9\\ 4.1\\ 1.9\\ 9.2\\ 4.1\\ 1.9\\ 9.2\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 6.1\\ 1.6\\ 6.1\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 0.9\\ 48.1\\ 94.0\\ 0.9\\ 48.1\\ 92.3\\ 2.8\\ 91.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 10.2\\ 0.2\\ 0.2\\ 10.2\\ 0.2\\ 0.2\\ 10.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ 0.2\\ $	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 93.4\\ 93.4\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 1.9\\ 9.1\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 1.9\\ 5.5\\ 5.5\\ 5.5\\ 5.5\\ 5.5\\ 5.5\\ 5.5\\ 5$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 25.8\\ 90.5\\ 25.8\\ 90.5\\ 25.8\\ 90.3\\ 37.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 8.5\\ 4.2\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 0.7\\ 50.1\\ 0.7\\ 48.8\\ 14.1\\ 0.7\\ 7.5\\ \end{array}$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 32.5\\ 88.5\\ 32.5\\ 32.5\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 35.8\\ 0.2\\ 36$
HAMIOK 3 APS HAMIOK 3 AUC HAMIOK 3 AUC HAMIOK 4 APS HAMIOK 4 APS HAMIOK 4 APS HAMIOK 4 AUC HAMIOK 5 APS HAMIOK 5 APS HAMIOK 5 AVC HAMIOK 5 AS HAMIOK 6 APS HAMIOK 6 APS HAMIOK 6 AVC HAMIOK AUC Macro HAMIOK AUC M	88.0 96.7 3.5 99.5 100.0 95.6 99.7 1.1 89.2 99.3 1.0 94.5 99.7 1.1 89.2 99.3 1.0 94.5 99.7 1.1 89.2 91.6 0.3 57.0 0.6 86.3 20.9 3.8 5.7 20.9 5.7 5.7 20.9 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7 5.7	$\begin{array}{c} 85.5\\ 95.5\\ 95.5\\ 3.7\\ 100.0\\ 100.0\\ 94.8\\ 99.7\\ 1.1\\ 85.2\\ 98.6\\ 1.9\\ 93.3\\ 98.6\\ 1.9\\ 0.2\\ 56.7\\ 0.5\\ 85.8\\ 86.2\\ 57.0\\ 57.3\\ 6.3\\ 21.3\\ 3.8\\ 86.6\\ 3.8\\ 86.2\\ 1.6\\ 3.8\\ 21.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 22.1\\ 3.8\\ 6.6\\ 3.8\\ 5.9\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\ 22.0\\ 0.7\\ 55.2\\$	83.9 95.6 4.2 99.7 100.0 1 94.5 99.6 1.1 83.9 99.1 1.1 83.9 99.1 1.1 83.9 99.1 56.5 51.4 98.7 2.2 0.3 56.5 51.4 98.7 2.2 0.3 56.5 57.4 57.4 57.4 21.0 3.8 6.0 21.6 3.8 6.9 22.0 3.8 6.4 0.7 59.5 3.8 6.4 0.7 59.5 50.7 59.5 50.7 50.7 50.7 50.7 50.7 50.7 50.7 50	$\begin{array}{c} 85.2\\ 95.9\\ 3.9\\ 98.2\\ 100.0\\ 0.1\\ 91.5\\ 99.5\\ 1.1\\ 86.3\\ 98.6\\ 1.1\\ 92.2\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.9\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 98.5\\ 1.1\\ 92.2\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ 1.1\\ $	$\begin{array}{c} 86.2\\ 96.1\\ 3.5\\ 100.0\\ 0.0\\ 90.3\\ 99.2\\ 99.2\\ 88.0\\ 98.9\\ 98.9\\ 98.9\\ 88.0\\ 98.9\\ 98.9\\ 1.1\\ 91.3\\ 98.6\\ 2.0\\ 0.3\\ 56.2\\ 19.3\\ 3.9\\ 5.1\\ 20.0\\ 3.9\\ 5.1\\ 20.0\\ 3.9\\ 5.1\\ 20.0\\ 3.9\\ 21.2\\ 3.8\\ 6.5\\ 5.1\\ 20.0\\ 3.9\\ 3.8\\ 6.5\\ 5.1\\ 20.0\\ 3.9\\ 3.8\\ 6.5\\ 5.1\\ 20.0\\ 3.9\\ 3.8\\ 6.5\\ 5.1\\ 20.0\\ 3.9\\ 3.8\\ 6.5\\ 5.1\\ 20.0\\ 3.9\\ 3.8\\ 6.5\\ 5.0\\ 20.6\\ 3.8\\ 3.8\\ 21.9\\ 3.8\\ 6.5\\ 5.0\\ 0.7\\ 55.0\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0$	$\begin{array}{c} 83.0\\ 95.3\\ 4.1\\ 98.5\\ 100.0\\ 0.1\\ 93.7\\ 99.5\\ 1.0\\ 87.6\\ 99.1\\ 1.3\\ 92.1\\ 92.1\\ 92.1\\ 0.2\\ 56.4\\ 2.1\\ 0.2\\ 56.4\\ 0.5\\ 78.0\\ 52.4\\ 0.5\\ 78.0\\ 52.4\\ 0.5\\ 78.0\\ 52.4\\ 0.7\\ 55.5\\ 22.1\\ 3.8\\ 6.2\\ 0.7\\ 55.5\\ 22.1\\ 24.9\\ 54.2\\ 24.9\\ 54.2\\ 22.2\\ 0.7\\ 55.5\\ 22.1\\ 24.9\\ 54.2\\ 0.7\\ 55.5\\ 22.1\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 55.5\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 84.0\\ 95.9\\ 4.2\\ 100.0\\ 0.0\\ 91.6\\ 99.4\\ 1.2\\ 84.7\\ 99.3\\ 1.4\\ 91.6\\ 98.7\\ 2.1\\ 0.2\\ 56.2\\ 0.5\\ 76.9\\ 20.4\\ 3.8\\ 6.4\\ 21.0\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 3.8\\ 6.4\\ 0.7\\ 49.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 5.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 5.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 5.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 5.5\\ 0.7\\ 48.3\\ 20.8\\ 27.4\\ 58.5\\ 5.5\\ 27.4\\ 58.5\\ 5.5\\ 27.4\\ 58.5\\ 27.4\\ 58.5\\ 27.4\\ 58.5\\ 27.4\\ 58.5\\ 27.4\\ 58.5\\ 27.4\\ 58.5\\ 20.8\\ 20$	$\begin{array}{c} 822.5\\ 96.1\\ 4.4\\ 98.5\\ 100.0\\ 0.0\\ 99.0\\ 1.3\\ 83.3\\ 99.6\\ 1.3\\ 90.8\\ 99.0\\ 1.3\\ 83.3\\ 90.8\\ 98.5\\ 2.1\\ 0.2\\ 55.5\\ 0.2\\ 1.3\\ 90.8\\ 98.5\\ 5.3\\ 3.9\\ 0.2\\ 55.5\\ 0.5\\ 79.4$	$74.2 \\ 94.4 \\ 5.0 \\ 98.7 \\ 100.0 \\ 0.2 \\ 98.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.8 \\ 98.0 \\ 1.7 \\ 75.8 \\ 98.8 \\ 97.8 \\ 97.8 \\ 98.8 \\ 0.2 \\ 53.6 \\ 0.4 \\ 74.0 \\ 75.1 \\ 83.4 \\ 97.8 \\ 0.2 \\ 53.6 \\ 0.4 \\ 74.0 \\ 3.9 \\ 4.5 \\ 18.9 \\ 3.9 \\ 4.8 \\ 0.7 \\ 53.8 \\ 0.7 \\ 3.9 \\ 4.8 \\ 0.7 \\ 53.8 \\ 0.7 \\ 3.9 \\ 4.8 \\ 0.7 \\ 53.8 \\ 0.7 \\ 3.9 \\ 18.4 \\ 3.9 \\ 19.4 \\ 3.9 \\ 18.4 \\ 3.9 \\ 3.9 \\ 4.8 \\ 0.7 \\ 53.8 \\ 0.7 \\ 19.4 \\ 10.8 \\ $	$\begin{array}{l} 80.8\\ 95.4\\ 4.7\\ 96.4\\ 99.9\\ 0.1\\ 88.0\\ 98.0\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 1.5\\ 81.3\\ 98.0\\ 0.15\\ 81.3\\ 98.0\\ 0.2\\ 55.3\\ 0.2\\ 4.3\\ 93.5\\ 66.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 63.6\\ 4.1\\ 16.0\\ 3.9\\ 63.6\\ 16.5\\ 3.9\\ 17.3\\ 3.9\\ 4.3\\ 0.7\\ 61.9\\ 0.7\\ 61.9\\ 0.7\\ 5.6\\ 23.0\\ 0.7\\ 5.6\\ 23.0\\ 0.7\\ 5.6\\ 23.0\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5$	$\begin{array}{c} 74.3\\ 92.5\\ 5.1\\ 96.9\\ 0.2\\ 91.4\\ 83.2\\ 99.4\\ 83.2\\ 98.6\\ 56.0\\ \hline \\ 94.1\\ 97.5\\ 2.8\\ 0.6\\ 56.0\\ \hline \\ 94.1\\ 94.2\\ 64.0\\ \hline \\ 94.2\\ 64.0\\ \hline \\ 3.7\\ 16.1\\ 3.9\\ 94.2\\ 64.0\\ \hline \\ 3.7\\ 16.1\\ 3.9\\ 94.2\\ 64.0\\ \hline \\ 3.7\\ 16.1\\ 3.9\\ 4.1\\ 0.7\\ 50.0\\ 0.7\\ 55.0\\ 0$	$\begin{array}{c} 41.9\\ 41.9\\ 87.9\\ 26.8\\ 92.3\\ 1.1\\ 54.1\\ 94.7\\ 3.4\\ 28.4\\ 91.6\\ 3.0\\ 43.7\\ 89.3\\ 0.2\\ 43.7\\ 1.2\\ 89.3\\ 0.2\\ 45.0\\ 0.2\\ 45.0\\ 0.2\\ 45.0\\ 0.2\\ 4.5\\ 1.5\\ 1.5\\ 9\\ 4.0\\ 4.0\\ 4.0\\ 16.0\\ 4.0\\ 16.0\\ 4.0\\ 16.0\\ 16.0\\ 16.0\\ 18.2\\ 1.8\\ 10.4\\ 4.5\\ 1.8\\ 10.4\\ 4.5\\ 1.8\\ 10.4\\ 4.5\\ 1.8\\ 10.4\\ 4.5\\ 1.8\\ 10.4\\ 1.5\\ 1.5\\ 1.8\\ 10.4\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5\\ 1.5$	$\begin{array}{r} 46.7\\ 84.9\\ 7.6\\ 21.9\\ 94.5\\ 1.2\\ 67.0\\ 96.6\\ 2.8\\ 33.4\\ 93.6\\ 2.9\\ 46.9\\ 90.1\\ 4.5\\ 0.2\\ 46.0\\ 0.2\\ 67.6\\ 67.5\\ 4.0\\ 4.2\\ 45.8\\ 4.2\\ 15.5\\ 4.0\\ 15.3\\ 15.5\\ 15.3\\ 15.5\\ 15$	34.7 81.7 8.4 33.6 84.2 1.2 41.8 92.9 3.8 31.4 91.2 41.8 92.9 3.8 31.4 91.2 3.8 31.4 91.2 3.8 31.4 91.2 3.0 38.8 85.6 5.4 0.2 39.4 0.5 76.0 77.0 4.2 1.6 0.7 4.3.9 3.8 13.6 3.8 13.6 3.6 3.6 3.8	35.1 80.0 8.4 32.3 89.0 36.4 92.0 3.9 30.5 91.9 30.5 91.9 30.5 0.2 37.4 47.2 50.00 14.3 326.4 47.2	33.2 80.3 80.3 8.4 24.6 89.5 1.1 36.8 91.5 29.5 91.4 3.9 29.5 91.4 3.9 29.5 37.0 35.9 84.6 0.2 37.0 0.2 53.7 23.7 0.2 53.7 23.7 0.2 53.4 4.1 2.0 8.4 4.1 1.2 2.0 8.4 4.1 1.2 2.0 8.4 4.1 1.2 0.7 53.4 2.0 8.4 4.1 1.7 0.7 54.2 1.7 7.9 25.3 7.9 25.4 2.17 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 7.9 25.3 1.7 1.7 7.9 25.3 1.7 1	$\begin{array}{c} 31.5\\ 80.7\\ 80.7\\ 8.5\\ 26.4\\ 87.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 1.2\\ 22.3\\ 87.5\\ 89.1\\ 32.2\\ 32.2\\ 32.2\\ 36.7\\ 36.2\\ 36.2\\ 36.2\\ 36.4\\ 36.2\\ 9.2\\ 4.1\\ 1.9\\ 8.9\\ 4.1\\ 2.2\\ 9.2\\ 4.1\\ 1.9\\ 8.7\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 9.4\\ 4.1\\ 1.9\\ 6.1\\ 17.8\\ 14.6\\ 6.1\\ 17.8\\ 14.6\\ \end{array}$	$\begin{array}{c} 42.59\\ 85.9\\ 97.7\\ 52.8\\ 97.7\\ 94.0\\ 0.9\\ 48.1\\ 94.0\\ 0.9\\ 48.1\\ 92.3\\ 2.8\\ 91.0\\ 5.0\\ 0.2\\ 40.6\\ 0.3\\ 60.3\\ 60.3\\ 60.3\\ 60.3\\ 60.3\\ 39.0\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 40.6\\ 0.2\\ 10.2\\ 0.2\\ 10.2\\ 0.7\\ 53.2\\ 0.7\\ 53.2\\ 0.7\\ 55.2\\ 215.2\\ 15.2\\ 10.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 0.7\\ 55.4\\ 23.6\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 48.4\\ 88.1\\ 7.2\\ 73.8\\ 97.6\\ 0.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 92.5\\ 41.3\\ 93.4\\ 93.4\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 0.2\\ 40.8\\ 1.9\\ 9.1\\ 4.1\\ 2.0\\ 0.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 1.9\\ 9.5\\ 4.1\\ 2.0\\ 0.7\\ 4.1\\ 2.0\\ 0.7\\ 51.8\\ 14.6\\ 0.7\\ 42.4\\ 2.0\\ 0.7\\ 51.8\\ 14.6\\ 0.7\\ 52.8\\ 1.0\\ 0.7\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0\\ 1.0$	$\begin{array}{r} 42.4\\ 84.2\\ 84.2\\ 84.2\\ 8.0\\ 34.8\\ 92.0\\ 1.1\\ 36.6\\ 90.3\\ 4.2\\ 25.8\\ 90.5\\ 25.8\\ 37.6\\ 90.3\\ 37.7\\ 0.3\\ 37.6\\ 0.2\\ 37.7\\ 0.3\\ 50.8\\ 51.4\\ 24.3\\ 31.7\\ 1.8\\ 8.5\\ 4.2\\ 8.5\\ 4.2\\ 1.8\\ 0.7\\ 50.1\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 0.7\\ 1.8\\ 8.9\\ 4.2\\ 1.8\\ 0.7\\ 50.1\\ 0.7\\ 3.2\\ 4.9\\ 2.2\\ 9.5\\ 0.7\\ 3.2\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 3.2\\ 0.7\\ 0.7\\ 3.2\\ 0.7\\ 0.7\\ 3.2\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7\\ 0.7$	$\begin{array}{c} 35.2\\ 82.6\\ 8.2\\ 11.5\\ 78.7\\ 1.2\\ 26.8\\ 88.2\\ 32.5\\ 88.5\\ 32.5\\ 88.5\\ 32.5\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.3\\ 32.6\\ 83.4\\ 23.5\\ 32.4\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 8.3\\ 4.2\\ 7.9\\ 4.2\\ 1.6\\ 0.7\\ 50.4\\ 0.7\\ 52.1\\ 14.1\\ 4.0\\ 15.6\\ 3.6\\ 3.6\\ 3.6\\ 3.6\\ 3.6\\ 3.6\\ 3.6\\ 3$

Kinetics Acc@1	48.8	44.2	51.4	43.7	40.3	44.6	33.2	36.4	25.8	2.7	1.0	0.2	0.3	0.4	2.0	1.6	1.0	0.5	0.3	0.3	0.3
Kinetics Acc@5	75.5	70.9	77.9	70.7	67.6	71.7	59.9	63.0	51.8	9.7	4.3	1.3	1.4	1.7	7.0	6.5	3.5	2.2	1.3	1.3	1.3
Kinetics Loss	2.4	2.6	2.1	2.5	2.7	2.5	3.2	3.0	3.5	5.5	6.1	6.1	6.1	6.1	5.7	5.8	6.0	6.1	6.1	6.1	6.1
UCF-101 Acc@1	84.4	75.1	69.9	63.2	75.0	63.4	58.8	66.6	48.7	19.7	11.1	2.8	0.8	2.1	15.2	13.3	6.6	8.7	6.5	7.0	2.7
UCF-101 Acc@5	95.4	92.5	89.1	82.3	91.6	86.2	81.7	86.3	75.3	42.2	28.9	8.5	5.0	8.2	35.5	33.8	17.9	25.2	23.1	20.2	11.2
UCF-101 Loss	0.6	1.0	1.3	1.7	1.0	1.5	1.7	1.4	2.3	4.3	5.0	4.8	4.7	4.6	3.7	3.8	4.5	4.0	4.2	4.2	4.5
Task Mean	73.0	65.6	66.6	58.8	63.7	57.5	53.3	57.5	49.8	17.2	14.3	4.2	3.3	5.0	15.7	14.7	8.8	10.9	10.3	10.2	5.8
Video Reg																					
IWildCam MAE Score	1.3	1.4	1.3	1.4	1.4	1.6	1.4	1.5	1.6	2.0	1.9	1.9	2.6	2.1	1.8	1.8	1.9	1.8	2.2	1.8	2.1
IWildCam MSE Loss	3.7	4.4	4.0	4.0	4.1	5.4	4.3	5.0	5.9	7.1	6.5	6.2	12.5	8.5	5.1	6.3	6.0	6.2	8.6	6.4	8.4
Task Mean	1.3	1.4	1.3	1.4	1.4	1.6	1.4	1.5	1.6	2.0	1.9	1.9	2.6	2.1	1.8	1.8	1.9	1.8	2.2	1.8	2.1
GATE																					
Full GATE Mean	69.0	66.8	66.8	64.6	64.3	63.4	62.1	62.2	58.5	56.3	54.4	48.4	42.8	39.6	37.5	37.2	36.2	36.9	35.0	34.9	31.8
Big GATE Mean	76.6	74.5	74.4	72.8	72.0	71.9	70.6	70.0	66.8	66.7	64.8	58.5	53.1	46.8	43.8	43.4	41.9	41.5	40.9	39.8	37.1
Base GATE Mean	68.3	65.6	65.7	62.6	63.7	60.7	60.2	60.7	58.6	55.1	53.5	48.2	42.8	38.0	36.5	36.3	35.4	36.6	34.8	34.8	30.4
Small GATE Mean	77.7	74.9	74.6	73.3	72.4	71.2	68.9	69.1	65.3	65.7	61.7	58.5	49.3	40.5	35.7	35.4	35.9	35.3	34.1	34.4	30.4
Full GATE Rank	1.0	3.0	2.0	4.0	5.0	6.0	8.0	7.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0	18.0	17.0	19.0	20.0	21.0
Big GATE Rank	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0	21.0
Base GATE Rank	1.0	3.0	2.0	5.0	4.0	7.0	8.0	6.0	9.0	10.0	11.0	12.0	13.0	14.0	16.0	17.0	18.0	15.0	20.0	19.0	21.0
Small GATE Rank	1.0	2.0	3.0	4.0	5.0	6.0	8.0	7.0	10.0	9.0	11.0	12.0	13.0	14.0	16.0	17.0	15.0	18.0	20.0	19.0	21.0

Table 3: Full experiments table: Black/Bold best model, Green second best, Blue third best, and red the worst performing model. Models prefixed with 's' refer to 'from scratch' trained models, rather than pretrained. This table showcases the full set of data we use to evolve GATE using EEVEE.



(a) Best k=13 discovered combina- (b) Best k=14 discovered combination tion









(d) Best k=16 discovered combina- (e) Best k=17 discovered combina- (f) Best k=18 discovered combination tion













(j) Best k=22 discovered combina- (k) Best k=23 discovered combina- (l) Best k=24 discovered combination tion

Figure 7: Degradation of predictive power when a given benchmark is removed and the meta-model trained from scratch, for different best combinations in varying k.





(a) Best k=24 discovered combina- (b) Best k=25 discovered combina- (c) Best k=26 discovered combination tion



tion

(d) Best k=27 discovered combination

Figure 8: Degradation of predictive power when a given benchmark is removed and the meta-model trained from scratch, for different best combinations in varying k.

0.02 0.03



Figure 9: Ranking Heatmap for bigGATE We show how the various models on the y-axis rank on the metrics on the x-axis, where brighter is higher/better rank. From left to right we apply a spearman correlation sorting to capture tasks more similar to imagenet1k more towards the leftmost side, and, dissimilar ones towards the rightmost side. From top to bottom we rank models based on average rank.

#### Model vs Dataset Rankings



Figure 10: Ranking Heatmap for baseGATE: We show how the various models on the y-axis rank on the metrics on the x-axis, where brighter is higher/better rank. From left to right we apply a spearman correlation sorting to capture tasks more similar to imagenet1k more towards the leftmost side, and, dissimilar ones towards the rightmost side. From top to bottom we rank models based on average rank.

# Model vs Dataset Rankings



Figure 11: Ranking Heatmap for smallGATE: We show how the various models on the y-axis rank on the metrics on the x-axis, where brighter is higher/better rank. From left to right we apply a spearman correlation sorting to capture tasks more similar to imagenet1k more towards the leftmost side, and, dissimilar ones towards the rightmost side. From top to bottom we rank models based on average rank.



Figure 12: Architecture Variation: Results of keeping the pretraining method the same as ImageNet1k classification and varying the architecture across various key task domains.



Figure 13: Pretraining Scheme Variation: Results of varying the pretraining method and keeping the architecture as ViT B16 across various key task domains.



- whisper ----- mpnet

– bart – bert –

Figure 14: Modality Variation: Results of attempting modality shifting from audio and text to vision tasks.



Figure 15: Modality Variation: Results of attempting modality shifting from audio and text to vision tasks.

# **NeurIPS Paper Checklist**

860	1.	Claims
861 862		Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?
863		Answer: [Yes]
864 865		Justification: All the claims made are substantiated with rigorous empirical results and communicated via tables and figures.
866		Guidelines:
867 868		• The answer NA means that the abstract and introduction do not include the claims made in the paper.
869 870 871 872 873 874 875		<ul> <li>The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.</li> <li>The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.</li> <li>It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.</li> </ul>
876	2.	Limitations
877		Question: Does the paper discuss the limitations of the work performed by the authors?
878		Answer: [Yes]
879		Justification: We have an explicit limitations section.
880		Guidelines:
881 882		• The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
883		• The authors are encouraged to create a separate "Limitations" section in their paper.
884 885 886 887 888		• The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
889 890 891		• The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
892 893 894 895 896		• The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
897 898		• The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
899 900		• If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
901 902 903 904 905 906		• While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.
907	3.	Theory Assumptions and Proofs
908 909		Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

910 Answer: [NA]

911	Justification: No theories were derived.
912	Guidelines:
913	• The answer NA means that the paper does not include theoretical results.
914	• All the theorems formulas and proofs in the paper should be numbered and cross-
915	referenced.
916	• All assumptions should be clearly stated or referenced in the statement of any theorems
017	• The proofs can either appear in the main paper or the supplemental material, but if
917	they appear in the supplemental material the authors are encouraged to provide a short
919	proof sketch to provide intuition.
920	• Inversely any informal proof provided in the core of the paper should be complemented
921	hy formal proofs provided in appendix or supplemental material
922	Theorems and Lemmas that the proof relies upon should be properly referenced
522	Fire arity and Dearly Dearly dearly life
923	4. Experimental Result Reproducibility
924	Question: Does the paper fully disclose all the information needed to reproduce the main ex-
925	perimental results of the paper to the extent that it affects the main claims and/or conclusions
926	of the paper (regardless of whether the code and data are provided or not)?
927	Answer: [Yes]
928	Justification: We do so both in the main paper, and in more detail in the appendix, in addition
929	to offering the codebase that reproduces all results.
930	Guidelines
	• The answer NA means that the namer days not include synapsiments
931	• The answer INA means that the paper does not include experiments.
932	• If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers. Making the paper reproducible is important regardless of
933	whether the code and data are provided or not
934	• If the contribution is a dataset and/or model, the outhors should describe the stars taken
935	• If the contribution is a dataset and/of model, the authors should describe the steps taken to make their results reproducible or verifiable
936	<ul> <li>Depending on the contribution remedualitility can be accomplished in various wave</li> </ul>
937	• Depending on the contribution, reproductoring can be accomplished in various ways. For example, if the contribution is a noval architecture, describing the architecture fully.
930	might suffice or if the contribution is a specific model and empirical evaluation it may
940	be necessary to either make it possible for others to replicate the model with the same
941	dataset, or provide access to the model. In general, releasing code and data is often
942	one good way to accomplish this, but reproducibility can also be provided via detailed
943	instructions for how to replicate the results, access to a hosted model (e.g., in the case
944	of a large language model), releasing of a model checkpoint, or other means that are
945	appropriate to the research performed.
946	• While NeurIPS does not require releasing code, the conference does require all submis-
947	sions to provide some reasonable avenue for reproducibility, which may depend on the
948	nature of the contribution. For example
949	(a) If the contribution is primarily a new algorithm, the paper should make it clear how
950	to reproduce that algorithm.
951	(b) If the contribution is primarily a new model architecture, the paper should describe
952	the architecture clearly and fully.
953	(c) If the contribution is a new model (e.g., a large language model), then there should
954	either be a way to access this model for reproducing the results or a way to reproduce
955	the detect)
956	(d) We recognize that reproducibility may be trially in some access in which access
957	(u) we recognize that reproducionity may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility
900 959	In the case of closed-source models, it may be that access to the model is limited in
960	some way (e.g., to registered users), but it should be possible for other researchers
961	to have some path to reproducing or verifying the results.
962	5 Open access to data and code
302	Overtient Dars the neuronalistic energy second to the late of the life of the fill of the second sec
963	Question: Does the paper provide open access to the data and code, with sufficient instruc- tions to faithfully raproduce the main experimental results, as described in surglemental
964	nons to rannung reproduce the main experimental results, as described in supplemental material?
900	111411141

966	Answer: [Yes]	
967	Justification: Full code and data are available and shared on github and huggingface.	
968	Guidelines:	
969	• The answer NA means that paper does not include experiments requiring code.	
970	• Please see the NeurIPS code and data submission guidelines (https://nips.cc,	/
971	public/guides/CodeSubmissionPolicy) for more details.	
972	• While we encourage the release of code and data, we understand that this might not be	e
973	possible, so "No" is an acceptable answer. Papers cannot be rejected simply for no	t
974	including code, unless this is central to the contribution (e.g., for a new open-source benchmark)	)
975	• The instructions should contain the exact command and environment needed to run to	
976 977	reproduce the results. See the NeurIPS code and data submission guidelines (https	, :
978	//nips.cc/public/guides/CodeSubmissionPolicy) for more details.	
979	• The authors should provide instructions on data access and preparation, including how	V
980	to access the raw data, preprocessed data, intermediate data, and generated data, etc.	
981	• The authors should provide scripts to reproduce all experimental results for the new	V
982	proposed method and baselines. If only a subset of experiments are reproducible, the	ý
983	should state which ones are omitted from the script and why.	1
984 985	• At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).	1
986	• Providing as much information as possible in supplemental material (appended to the	9
987	paper) is recommended, but including URLs to data and code is permitted.	
988	6. Experimental Setting/Details	
989	Question: Does the paper specify all the training and test details (e.g., data splits, hyper	;-
990	parameters, how they were chosen, type of optimizer, etc.) necessary to understand the	2
991	lesuits?	
992	Answer: [Yes]	
993 994	Justification: We describe these in the experiments section in summary, and in the appendix in detail.	< C
995	Guidelines:	
996	• The answer NA means that the paper does not include experiments.	
997	• The experimental setting should be presented in the core of the paper to a level of detai	1
998	that is necessary to appreciate the results and make sense of them.	
999	• The full details can be provided either with the code, in appendix, or as supplementa	1
1000	material.	
1001	7. Experiment Statistical Significance	
1002 1003	Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?	e
1004	Answer: [Yes]	
1005	Justification: Where relevant our results include error bars.	
1006	Guidelines:	
1007	• The answer NA means that the paper does not include experiments.	
1008	• The authors should answer "Yes" if the results are accompanied by error bars, confi	
1009	dence intervals, or statistical significance tests, at least for the experiments that suppor	t
1010	the main claims of the paper.	
1011	• The factors of variability that the error bars are capturing should be clearly stated (fo	r 1
1012	example, train/test spin, initialization, random drawing of some parameter, or overal	I
1013	<ul> <li>The method for calculating the error bars should be explained (closed form formula)</li> </ul>	
1015	call to a library function, bootstrap, etc.)	••
1016	• The assumptions made should be given (e.g., Normally distributed errors).	

1017	• It should be clear whether the error bar is the standard deviation or the standard error
1018	of the mean.
1019	• It is OK to report 1-sigma error bars, but one should state it. The authors should
1020	of Normality of errors is not verified
1021	• For asymmetric distributions, the authors should be coraful not to show in tables or
1022	• For asymmetric distributions, the authors should be called in the show in tables of figures symmetric error bars that would yield results that are out of range (e.g. negative
1023	error rates)
1025	• If error bars are reported in tables or plots. The authors should explain in the text how
1025	they were calculated and reference the corresponding figures or tables in the text.
1027	8. Experiments Compute Resources
1028	Question: For each experiment, does the paper provide sufficient information on the com-
1029	puter resources (type of compute workers, memory, time of execution) needed to reproduce
1030	the experiments?
1031	Answer: [TODO]
1032	Justification: [TODO]
1033	Guidelines:
1034	• The answer NA means that the paper does not include experiments.
1035	• The paper should indicate the type of compute workers CPU or GPU, internal cluster,
1036	or cloud provider, including relevant memory and storage.
1037	• The paper should provide the amount of compute required for each of the individual
1038	experimental runs as well as estimate the total compute.
1039	• The paper should disclose whether the full research project required more compute
1040	than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper)
1041	Cada Of Ethics
1042	9. Code Of Ethics
1043 1044	Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?
1045	Answer: [Yes]
1046	Justification: Yes it does abide by the code of ethics to our best of our understanding.
1047	Guidelines:
1048	• The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
1049	• If the authors answer No, they should explain the special circumstances that require a
1050	deviation from the Code of Ethics.
1051	• The authors should make sure to preserve anonymity (e.g., if there is a special consid-
1052	eration due to laws or regulations in their jurisdiction).
1053	10. Broader Impacts
1054	Question: Does the paper discuss both potential positive societal impacts and negative
1055	societal impacts of the work performed?
1056	Answer: [No]
1057	Justification: It's a method for finding optimal subsets of benchmarks from a large pool and
1058	a framework that automates model encoder evaluation. Societal impacts relate to improved
1059	research efficiency and hopefully compute usage, however this is too far from what one
1060	would consider strongly tied societal impacts.
1061	Guidelines:
1062	• The answer NA means that there is no societal impact of the work performed.
1063	• If the authors answer NA or No, they should explain why their work has no societal
1064	impact or why the paper does not address societal impact.
1065	• Examples of negative societal impacts include potential malicious or unintended uses
1065	(e.g., distinormation, generating take promes, surveinance), familiess considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific
1068	groups), privacy considerations, and security considerations.

1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083	<ul> <li>The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.</li> <li>The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.</li> <li>If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).</li> </ul>	
1084	11. Safeguards	
1085	Ouestion: Does the paper describe safeguards that have been put in place for responsible	
1086 1087	release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?	
1088	Answer: [NA]	
1089	Justification: It's a benchmark with datasets that are already public and previously published	
1090	in other papers.	
1091	Guidelines:	
1092	• The answer NA means that the paper poses no such risks.	
1093	• Released models that have a high risk for misuse or dual-use should be released with	
1094	necessary safeguards to allow for controlled use of the model, for example by requiring	
1095	that users adhere to usage guidelines or restrictions to access the model or implementing	
1096	safety filters.	
1097	• Datasets that have been scraped from the Internet could pose safety risks. The authors	
1098	• We recognize that providing effective sefective sefectives is chellenging, and many papers de	
1099	• we recognize that providing enective safeguards is chanenging, and many papers do not require this, but we encourage authors to take this into account and make a best	
1101	faith effort.	
1102	12. Licenses for existing assets	
1103	Question: Are the creators or original owners of assets (e.g., code, data, models), used in	
1104	the paper, properly credited and are the license and terms of use explicitly mentioned and	
1105	properly respected?	
1106	Answer: [Yes]	
1107	Justification: All datasets used have appropriate licenses, and the code packages used in	
1108	implementing our software framework have appropriate licenses as well.	
1109	Guidelines:	
1110	• The answer NA means that the paper does not use existing assets.	
1111	• The authors should cite the original paper that produced the code package or dataset.	
1112	• The authors should state which version of the asset is used and, if possible, include a	
1113	URL.	
1114	• The name of the license (e.g., CC-BY 4.0) should be included for each asset.	
1115	• For scraped data from a particular source (e.g., website), the copyright and terms of	
1116	service of that source should be provided.	
1117	• It assets are released, the license, copyright information, and terms of use in the	
1118	package snould be provided. For popular datasets, paperswithcode.com/datasets	
1120	license of a dataset.	
-		

1121 1122	• For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
1123	• If this information is not available online, the authors are encouraged to reach out to
1124	the asset's creators.
1125	13. New Assets
1126 1127	Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?
1128	Answer: [Ves]
1120	Instituction: Our codebase is fully documented
1129	Guidelines
1130	
1131	• The answer NA means that the paper does not release new assets.
1132 1133	• Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license,
1134	limitations, etc.
1135 1136	• The paper should discuss whether and now consent was obtained from people whose asset is used.
1137 1138	• At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.
1130	14 Crowdsourcing and Research with Human Subjects
1100	Question: For crowdsourcing experiments and research with human subjects does the paper
1140	include the full text of instructions given to participants and screenshots if applicable as
1141	well as details about compensation (if any)?
1143	Answer: [NA]
1144	Institution: No crowdsourcing with humans
1145	Guidelines
1145	The answer NA means that the same data not involve another size and some her it.
1146	• The answer NA means that the paper does not involve crowdsourcing nor research with human subjects
1147	<ul> <li>Including this information in the supplemental material is fine, but if the main contribu-</li> </ul>
1140	tion of the paper involves human subjects, then as much detail as possible should be
1150	included in the main paper.
1151	• According to the NeurIPS Code of Ethics, workers involved in data collection, curation,
1152	or other labor should be paid at least the minimum wage in the country of the data
1153	
1154	15. Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects
1155	
1156	Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Paview Board (IPR)
1159	approvals (or an equivalent approval/review based on the requirements of your country or
1159	institution) were obtained?
1160	Answer: [NA]
1161	Justification: Same as previous answer.
1162	Guidelines:
1163	• The answer NA means that the paper does not involve crowdsourcing nor research with
1164	human subjects.
1165	• Depending on the country in which research is conducted, IRB approval (or equivalent)
1166	may be required for any human subjects research. If you obtained IRB approval, you
1167	should clearly state this in the paper.
1168	• We recognize that the procedures for this may vary significantly between institutions
1169	and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the
1170	guidelines for their institution.
1171	• For initial submissions, do not include any information that would break anonymity (if
1172	applicable), such as the institution conducting the review.