Unsupervised Foundation Model-Agnostic Slide-Level Representation Learning

Appendix

Hyperparameter	Value
Heads	8
Number of Mamba-2 layers	2
Embedding dimension	768
Input dimensions	768, 1024, 1280, 1536
Dropout	0.25
Attention hidden dimension	96
Teacher momentum	0.99
Contrastive loss temperature	0.2
Optimizer	AdamW [22]
Learning rate	5e-4
Warmup epochs	50
Weight decay	0.1
Epochs	2000
Batch size	1024
Tile embeddings per patient	768

Table 5. Hyperparameters for COBRA pretraining

A. Implementation details

FM pretraining The detailed pretraining settings for COBRA can be found in Tab. 5. We used 25% dropout in all MLPs.

A.1. Additional information on evaluation

A.1.1 MLP downstream classification

An MLP classifier is implemented using a two-layer architecture, with an input layer of 768 dimensions and a hidden layer of 256 dimensions. The hidden layer employs SiLU [14] activation, followed by a dropout layer (50%) for regularization. The output layer consists of a fully connected layer with the appropriate number of output classes. Cross-entropy loss with class weighting is applied to handle class imbalance. The classifier is trained using the AdamW [22] optimizer with a learning rate of 0.0001 and weight decay of 0.01, employing a one-cycle policy for 32 epochs. Training is conducted in a 5-fold cross-validation setup, with early stopping and best model checkpoints monitored by validation loss.

A.1.2 Linear probing

Linear probing is implemented using a logistic regression objective based on sklearn. We use the default sklearn L2 regularization (set to 1.0) with an lbfgs solver. We set the maximum iterations to 10,000 and apply balanced class weights. Training is conducted in a stratified sampling setting with 10 random runs, using 5, 10, and 25 cases per class in each run.

B. Inference modes

COBRA is designed to be flexible and versatile, supporting three primary inference modes: single-FM, multi-FM, and hybrid-FM. These modes allow for adaptability across various histology datasets and computational resources.

Single-FM inference mode In this mode, COBRA utilizes patch embeddings extracted from a single feature extractor, such as Virchow2. The framework generates attention weights for the patches based on these embeddings and aggregates them to produce the final patient-level embedding. This mode is computationally efficient and achieves state-of-the-art performance with minimal overhead, making it ideal for scenarios requiring simplicity and resource efficiency.

Multi-FM inference mode In multi-FM inference mode, COBRA integrates embeddings generated independently by multiple FMs. Each FM produces its own patch embeddings, which COBRA's embedding module then projects into a shared embedding space. These projected embeddings from all FMs are averaged for each corresponding patch, resulting in unified patch embeddings that integrate diverse morphological representations. COBRA processes these averaged embeddings through its encoding and attention modules to produce attention weights. Finally, these attention weights are applied to the original embeddings from a selected primary FM to obtain the final patient-level representation. While this mode might improve robustness by leveraging multiple FMs simultaneously, performance gains compared to the single-FM mode appear marginal.

Hybrid-FM inference mode The hybrid-FM inference mode allows COBRA to incorporate patch embeddings from previously unseen FMs without retraining the model. First, the patch embeddings (from one or more of COBRA's pretraining FMs) are mapped into COBRA's shared embedding space via the embedding module. Subsequently, the framework generates attention weights based on these encoded embeddings. Finally, these weights are applied to the original external patch embeddings of the previously unseen FM to generate a patient-level representation. This ability ensures that COBRA remains adaptable, allowing seamless integration and effective utilization of new FMs without requiring any retraining.

Handling different magnifications COBRA is equipped to process patch embeddings extracted at various magnifications, including 0.5 MPP (20×), 1.14 MPP (9×), and 2 MPP (5×). This flexibility ensures compatibility with a wide range of histology datasets, allowing for diverse applications without requiring adjustments to the core architecture.

C. Data

Overall, our study comprises a total of 4,652 WSIs from 3,292 patients, including the organs lung, stomach, breast and colon. We use 3,048 WSIs for pretraining COBRA and training the classifiers, and 1604 WSIs for external validation. The slides for TCGA are available at https://portal.gdc.cancer.gov. The slides for CPTAC are available at https://proteomics.cancer.gov/data-portal. The molecular data for TCGA and CPTAC are available at https://www.cbioportal.org [2].

TCGA BRCA (training) We collected N=1,041 primary cases from the TCGA Breast Invasive Carcinoma (BRCA) cohort. For each case, we downloaded the corresponding molecular status: ER (N=1041; 770 positive, 271 negative), PR (N=1041; 704 positive, 337 negative), HER2 (N=1041; 125 positive, 916 negative), and PIK3CA driver mutation (N=1023; 687 WT, 336 MUT). We defined ER positive, PR positive, HER2 positive and PIK3CA MUT as positive classes for AUPRC and F1 scores.

TCGA CRC (training) We collected N=558 primary cases from the TCGA Colorectal Carcinoma (CRC) cohort. For each case, we downloaded the corresponding molecular status: MSI status (N=429; 368 MSS, 61 MSI), Lymph Node status (N=556; 318 N0, 238 N+), CRC sidedness (N=398; 230 left, 168 right), BRAF (N=501; 450 WT, 51 MUT), KRAS (N=501; 296 WT, 205 MUT), and PIK3CA driver mutation (N=501; 377 WT, 124 MUT). We defined MSI high, N+, right-sided CRC, BRAF MUT, KRAS MUT and PIK3CA MUT as positive classes for AUPRC and F1 scores.

TCGA LUAD (training) We collected N=461 primary cases from the TCGA Lung Adenocarcinoma (LUAD) cohort. For each case, we downloaded the corresponding molecular status: STK11 (N=461; 394 WT, 67 MUT), EGFR (N=461; 411 WT, 50 MUT), KRAS (N=461; 317 WT, 144 MUT), and TP53 driver mutation (N=461; 239 MUT, 222 WT). We defined STK11 MUT, EGFR MUT, KRAS MUT and TP53 MUT as positive classes for AUPRC and F1 scores.

TCGA NSCLC (training) We collected N=462 primary cases from the TCGA Lung Squamous Cell Carcinoma (LUSC) cohort and the aforementioned N=461 primary cases from the TCGA LUAD cohort. We defined LUAD as the positive class for AUPRC and F1 scores.

TCGA STAD (training) We collected N=326 primary cases from the TCGA Stomach Adenocarcinoma (STAD) cohort. They were only used for the training of COBRA.

CPTAC BRCA (testing) We collected N=120 primary cases from the CPTAC Breast Invasive Carcinoma (BRCA) cohort. For each case, we downloaded the corresponding molecular status: ER (N=120; 79 positive, 41 negative), PR (N=120; 70 positive, 50 negative), HER2 (N=120; 14 positive, 106 negative), and PIK3CA driver mutation (N=120; 82 WT, 38 MUT).

CPTAC COAD (testing) We collected N=110 primary cases from the CPTAC Colon Adenocarcinoma (COAD) cohort. For each case, we downloaded the corresponding molecular status: MSI status (N=105; 81 MSS, 24 MSI), Lymph Node status (N=110; 56 N0, 54 N+), CRC sidedness (N=108; 51 left, 57 right), BRAF (N=106; 91 WT, 15 MUT), KRAS (N=106; 71 WT, 35 MUT), and PIK3CA driver mutation (N=106; 87 WT, 19 MUT).

CPTAC LUAD (testing) We collected N=106 primary cases from the CPTAC Lung Adenocarcinoma (LUAD) cohort. For each case, we downloaded the corresponding molecular status: STK11 (N=106; 88 WT, 18 MUT), EGFR (N=106; 72 WT, 34 MUT), KRAS (N=106; 74 WT, 32 MUT), and TP53 driver mutation (N=106; 55 MUT, 51 WT).

CPTAC LUSC (testing) We collected N=108 primary cases from the CPTAC Lung Squamous Cell Carcinoma (LUSC) cohort and the aforementioned N=106 primary cases from the CPTAC LUAD cohort.

D. Results

D.1. Full Classification

Here, we provide the complete classification results of our experiments for the metrics AUC, AUPRC, F1 score and balanced accuracy. Tabs. 6 to 9 compare all models at 20× including COBRA-ENC, which was computed using the encoded embeddings ($\boldsymbol{H}_S$) with Virchow2 patch embeddings as shown in Eq. (4). In line with Wang et al. [42], using the original patch embeddings ($\boldsymbol{H}^{f_{e_n}}$) is beneficial. Tabs. 10 and 11 show the complete AUC results at 5× and 9×.

D.2. Linear probing few-shot classification

Tabs. 12 to 23 show the complete results of our linear probing few-shot classification experiments for the metrics AUC, AUPRC, F1 score and balanced accuracy with k=5,10 and 25 samples per class.

E. Heatmaps

COBRA's approach to interpretability in WSI analysis is based on an aggregation method where each tile embedding is assigned a weight through a softmax-normalized attention score. These attention scores are used directly to compute a weighted average of the tile embeddings, yielding a slide-level representation that reflects the importance of each tile without requiring complex, non-linear transformations. Unlike GradCam[33]-based interpretability methods used with tile embedding MIL approaches, COBRA's attention scores are linearly applied to aggregate tile embeddings. This means that the attention scores correspond precisely to the actual weights used in generating the final slide embedding, allowing for direct interpretability without any intermediate non-linearities that might distort the contribution of each tile.

In Figs. 5 to 8, we provide interpretability heatmaps for slides from TCGA-CRC and in Figs. 9 and 10, we show interpretability heatmaps for slides from CPTAC-COAD. These heatmaps display the attention values across the slide, with tiles associated with higher attention scores consistently aligning with tumor regions. In contrast, non-tumorous areas and background regions receive lower attention values. This pattern demonstrates COBRA's capability to emphasize diagnostically relevant areas based solely on the unsupervised training with tile embeddings.

While this tile-based attention approach lacks the spatial precision of pixel-level methods, it offers a computationally efficient way to highlight regions of model focus. By operating directly on tile embeddings, COBRA can produce interpretable heatmaps that outline primary areas of interest, indicating its utility in scenarios where rapid, general interpretability is more practical than fine-grained spatial resolution.

Limitations

While COBRA has demonstrated promising results, several limitations exist that warrant further investigation. First, the pretraining process involves a limited number of tissue types, which may restrict its generalizability to other histopathological contexts. Second, the diversity of downstream tasks and evaluation datasets is currently narrow, potentially limiting the framework's applicability across varied clinical scenarios. Third, the self-supervised learning (SSL) strategy exclusively employs a contrastive loss function based on MoCo-v3, leaving room for exploration of alternative or complementary loss functions that could enhance representation quality. Finally, the resulting patient-level embedding is formulated as a linear combination of patch embeddings, which may not fully capture the complex, non-linear relationships inherent in histopathological data. Addressing these limitations will be a focus of future research to improve the robustness and versatility of the proposed framework.

Competing Interests

JNK declares consulting services for Bioptimus, France; Panakeia, UK; AstraZeneca, UK; and MultiplexDx, Slovakia. Furthermore, he holds shares in StratifAI, Germany, Synagen, Germany, Ignition Lab, Germany; has received an institutional research grant by GSK; and has received honoraria by AstraZeneca, Bayer, Daiichi Sankyo, Eisai, Janssen, Merck, MSD, BMS, Roche, Pfizer, and Fresenius. GW declares consulting services for Synagen. TL declares consulting services for StratifAI. The remaining authors have no competing interests to declare.

Table 6. **Classification performance comparison.** AUC score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC-20×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{CTransPath}}$ [41]	$87.2_{1.5}$	$62.8_{2.5}$	$59.3_{7.4}$	$70.1_{2.3}$	$52.4_{5.9}$	$68.1_{2.5}$	$66.5_{2.1}$	$48.6_{1.5}$	$56.3_{3.0}$	$76.1_{4.6}$	$59.8_{2.3}$	$59.8_{1.0}$	$55.9_{7.7}$	$52.5_{2.6}$	$56.3_{6.3}$	$62.1_{4.2}$
$\overline{\text{Virchow}}$ [39]	$89.4_{0.6}$	$76.5_{6.8}$	$60.3_{2.0}$	$70.7_{1.7}$	$54.3_{6.9}$	$66.9_{4.8}$	$60.5_{3.9}$	$51.3_{5.7}$	$63.5_{3.7}$	$62.1_{6.7}$	$65.0_{2.6}$	$58.2_{5.1}$	$53.9_{7.1}$	$52.3_{3.1}$	$52.4_{5.7}$	$62.5_{4.9}$
$\overline{\text{CONCH}}$ [24]	$96.5_{0.3}$	$66.0_{10.3}$	$62.0_{7.6}$	$74.6_{1.6}$	$59.0_{7.4}$	$85.3_{1.5}$	$80.3_{2.0}$	$58.8_{11.0}$	$63.2_{3.1}$	$79.2_{0.5}$	$57.5_{3.4}$	$\mathbf{67.3}_{2.0}$	$55.7_{8.6}$	$53.4_{2.4}$	$63.2_{5.3}$	$68.1_{5.7}$
$\overline{\text{UNI}}$ [4]	$95.8_{1.1}$	$69.4_{2.4}$	$70.1_{12.1}$	$73.9_{0.8}$	$50.7_{4.7}$	$87.4_{3.1}$	$74.9_{2.3}$	$64.0_{3.4}$	$62.3_{5.0}$	$89.0_{1.6}$	$73.0_{3.4}$	$62.0_{7.8}$	$63.5_{2.8}$	$59.4_{3.7}$	$63.5_{6.2}$	$70.6_{4.9}$
$\overline{\text{H-Optimus}}$ [32]	$97.2_{0.4}$	$78.5_{2.7}$	$78.2_{3.5}$	$71.3_{1.1}$	$58.1_{3.7}$	$85.2_{2.7}$	$74.9_{3.3}$	$51.4_{4.5}$	$59.5_{4.9}$	$94.7_{0.7}$	$77.1_{7.1}$	$55.5_{3.6}$	$59.2_{4.2}$	$\mathbf{62.2}_{3.4}$	$62.4_{9.0}$	$71.0_{4.3}$
$\overline{\text{GigaPath}}$ [44]	$96.6_{0.7}$	$71.3_{1.9}$	$75.7_{6.8}$	$75.4_{1.3}$	$56.9_{6.2}$	$85.7_{1.1}$	$75.9_{1.8}$	$64.5_{2.7}$	$62.2_{5.9}$	$93.3_{1.6}$	$77.5_{2.6}$	$61.7_{2.9}$	$56.1_{5.0}$	$60.0_{1.5}$	$59.4_{7.9}$	$71.5_{4.0}$
$\overline{\text{Ensemble Prediction}}$	$97.2_{0.3}$	$77.2_{3.7}$	$78.5_{4.1}$	$73.3_{0.6}$	$\underline{59.5}_{5.1}$	$87.6_{2.8}$	$77.2_{2.7}$	$65.6_{2.0}$	$63.3_{4.1}$	$94.7_{1.0}$	$78.9_{5.4}$	$62.5_{3.9}$	$\underline{64.1}_{3.0}$	$60.5_{2.1}$	$\underline{64.5}_{9.1}$	$73.6_{4.0}$
$\overline{\text{Virchow2}}$ [47]	$95.8_{0.7}$	$79.6_{5.5}$	$78.3_{4.6}$	$72.1_{0.7}$	$\mathbf{60.9}_{5.6}$	$89.2_{2.8}$	$79.3_{2.7}$	$71.3_{1.8}$	$63.2_{4.8}$	$94.9_{1.2}$	$81.6_{4.5}$	$63.0_{1.9}$	$59.3_{6.2}$	$56.3_{3.8}$	$62.7_{11.6}$	$73.8_{4.7}$
$\overline{\text{Concatenated}}$	$97.4_{0.4}$	$75.7_{3.0}$	$\underline{80.2}_{2.2}$	$72.5_{0.8}$	$57.6_{4.8}$	$\underline{89.6}_{1.4}$	$79.1_{3.4}$	$67.5_{3.9}$	$61.8_{4.0}$	$95.0_{1.1}$	$82.2_{4.3}$	$61.6_{2.5}$	$59.7_{5.8}$	$\underline{62.0}_{2.4}$	$\mathbf{70.2}_{4.1}$	$74.1_{3.3}$
COBRA-ENC	$93.1_{0.3}$	$65.6_{2.4}$	$68.7_{2.3}$	$72.0_{1.9}$	$53.8_{2.8}$	$71.1_{0.9}$	$68.1_{2.9}$	$62.9_{4.1}$	$62.3_{4.1}$	$54.9_{6.5}$	$60.8_{9.0}$	$50.0_{2.7}$	$45.6_{1.6}$	$45.6_{2.5}$	$52.1_{2.1}$	$61.8_{3.7}$
GigaPath-SE [44]	$90.9_{1.3}$	$67.0_{4.4}$	$65.4_{4.4}$	$73.7_{1.4}$	$57.1_{5.2}$	$72.9_{0.9}$	$71.9_{3.3}$	$55.4_{4.7}$	$60.5_{4.6}$	$66.2_{2.1}$	$56.7_{4.5}$	$54.6_{5.2}$	$51.3_{2.9}$	$45.8_{3.1}$	$53.2_{5.7}$	$62.8_{3.9}$
MADELEINE [18]	$94.0_{0.6}$	$72.2_{8.7}$	$64.0_{6.7}$	$72.0_{2.8}$	$51.9_{3.9}$	$80.1_{1.7}$	$73.7_{1.3}$	$66.7_{2.7}$	$64.9_{1.6}$	$68.6_{9.1}$	$54.2_{6.7}$	$60.3_{7.3}$	$58.9_{6.6}$	$50.5_{1.6}$	$59.5_{8.6}$	$66.1_{5.5}$
CHIEF [42]	$93.6_{0.8}$	$64.2_{10.7}$	$62.8_{10.9}$	$73.4_{1.5}$	$50.1_{5.0}$	$83.0_{0.5}$	$77.5_{0.3}$	$63.4_{2.3}$	$\underline{65.4}_{1.5}$	$75.1_{4.8}$	$63.6_{4.3}$	$58.0_{1.7}$	$58.4_{3.8}$	$48.2_{4.2}$	$56.6_{3.2}$	$66.2_{4.9}$
COBRA†-CTP	$95.9_{0.6}$	$68.1_{5.1}$	$69.2_{4.7}$	$75.1_{1.6}$	$46.7_{4.1}$	$77.9_{0.8}$	$71.3_{1.3}$	$59.3_{1.5}$	$59.2_{1.3}$	$80.3_{1.5}$	$73.5_{3.4}$	$60.6_{2.5}$	$55.2_{4.5}$	$48.3_{3.4}$	$54.3_{5.2}$	$66.3_{3.2}$
COBRA-CTP	$95.9_{0.6}$	$65.0_{10.5}$	$66.0_{5.5}$	$74.8_{1.8}$	$49.2_{4.4}$	$78.6_{0.7}$	$72.2_{0.6}$	$62.0_{1.4}$	$60.9_{3.6}$	$80.3_{2.2}$	$73.2_{3.0}$	$61.3_{1.6}$	$52.4_{4.2}$	$48.1_{2.3}$	$56.3_{4.3}$	$66.4_{4.0}$
PRISM [34]	$99.2_{0.1}$	$\mathbf{87.6}_{1.6}$	$70.7_{2.4}$	$78.2_{0.5}$	$52.9_{8.5}$	$\mathbf{92.2}_{0.7}$	$\mathbf{84.2}_{0.5}$	$64.5_{6.0}$	$\mathbf{69.4}_{2.1}$	$79.1_{1.5}$	$59.9_{1.4}$	$\underline{67.2}_{2.4}$	$54.6_{6.2}$	$52.2_{1.8}$	$52.1_{6.8}$	$70.9_{3.8}$
COBRA†-UNI	$99.1_{0.2}$	$79.1_{2.8}$	$76.2_{4.6}$	$\underline{80.2}_{0.7}$	$55.0_{5.7}$	$86.0_{1.7}$	$78.1_{3.2}$	$60.3_{4.9}$	$62.3_{3.1}$	$89.1_{0.7}$	$83.5_{1.5}$	$65.7_{2.1}$	$\mathbf{65.3}_{4.2}$	$57.2_{2.1}$	$61.9_{2.1}$	$73.3_{3.1}$
COBRA†-H0	$\underline{\mathbf{99.4}}_{0.1}$	$\underline{86.9}_{2.0}$	$\mathbf{80.9}_{3.4}$	$79.9_{1.8}$	$56.7_{3.6}$	$87.8_{1.2}$	$72.8_{3.4}$	$59.9_{2.1}$	$58.0_{0.9}$	$\underline{95.2}_{1.1}$	$84.9_{3.7}$	$58.1_{2.6}$	$59.7_{7.2}$	$58.5_{2.4}$	$61.2_{3.4}$	$73.3_{3.1}$
COBRA-UNI	$98.8_{0.3}$	$79.4_{2.5}$	$76.5_{5.3}$	$78.9_{1.2}$	$52.2_{4.2}$	$88.1_{1.7}$	$\underline{80.5}_{3.0}$	$65.1_{4.3}$	$63.9_{5.2}$	$89.1_{1.1}$	$82.8_{1.5}$	$64.6_{2.2}$	$59.0_{8.4}$	$57.4_{2.1}$	$64.3_{5.8}$	$73.4_{3.9}$
COBRA-H0	$\underline{\mathbf{99.4}}_{0.2}$	$86.5_{1.8}$	$79.9_{2.9}$	$80.1_{2.4}$	$54.3_{4.7}$	$87.1_{1.0}$	$74.0_{4.2}$	$64.2_{4.9}$	$55.7_{2.3}$	$\mathbf{96.0}_{0.6}$	$86.2_{3.3}$	$58.2_{2.5}$	$62.2_{4.4}$	$57.2_{1.9}$	$62.9_{4.7}$	$73.6_{3.2}$
COBRA†-GP	$98.9_{0.3}$	$81.5_{2.3}$	$78.7_{4.2}$	$\mathbf{80.9}_{1.1}$	$56.9_{5.3}$	$87.8_{1.2}$	$77.5_{1.1}$	$65.4_{1.3}$	$64.6_{3.8}$	$93.5_{1.2}$	$85.6_{2.0}$	$64.7_{2.4}$	$59.2_{6.6}$	$57.4_{2.1}$	$56.9_{9.4}$	$74.0_{3.8}$
COBRA-V2	$98.1_{0.2}$	$84.0_{2.9}$	$80.0_{2.4}$	$78.4_{2.9}$	$59.2_{6.2}$	$\underline{89.6}_{2.0}$	$79.2_{2.4}$	$\underline{71.6}_{2.2}$	$63.6_{6.2}$	$94.1_{0.5}$	$\underline{87.8}_{2.0}$	$65.7_{2.5}$	$62.1_{10.4}$	$58.3_{1.9}$	$57.6_{7.5}$	$\underline{75.3}_{4.4}$
COBRA†-V2	$98.4_{0.2}$	$84.6_{1.9}$	$78.9_{3.6}$	$78.4_{2.6}$	$55.9_{7.1}$	$\underline{89.6}_{1.7}$	$80.0_{2.3}$	$\mathbf{72.2}_{1.7}$	$65.1_{4.5}$	$94.2_{0.6}$	$\mathbf{88.7}_{1.6}$	$64.8_{2.3}$	$60.6_{5.7}$	$58.6_{1.7}$	$61.5_{4.1}$	$\mathbf{75.4}_{3.3}$

Table 7. **Classification performance comparison.** AUPRC score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUPRC-20×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{Virchow}}$ [39]	$90.3_{0.6}$	$36.3_{10.2}$	$43.1_{3.6}$	$68.2_{0.8}$	$40.1_{3.2}$	$78.4_{2.8}$	$68.4_{2.8}$	$17.8_{1.5}$	$46.1_{6.4}$	$35.7_{8.0}$	$25.8_{4.1}$	$57.8_{3.5}$	$40.2_{8.8}$	$57.4_{4.0}$	$21.4_{4.4}$	$48.5_{5.1}$
$\overline{\text{CTransPath}}$ [41]	$87.9_{1.5}$	$25.8_{2.0}$	$46.4_{8.3}$	$71.2_{2.3}$	$35.5_{2.5}$	$79.3_{1.6}$	$73.8_{1.6}$	$11.4_{0.3}$	$39.7_{3.3}$	$59.0_{8.7}$	$27.7_{1.2}$	$59.2_{1.3}$	$45.2_{8.4}$	$57.7_{2.3}$	$25.2_{4.0}$	$49.7_{4.3}$
$\overline{\text{CONCH}}$ [24]	$96.7_{0.3}$	$32.3_{10.0}$	$44.7_{8.3}$	$78.0_{1.0}$	$39.0_{4.5}$	$92.5_{0.7}$	$\underline{84.4}_{1.0}$	$18.6_{4.8}$	$47.9_{3.5}$	$62.5_{0.8}$	$28.2_{5.2}$	$\mathbf{70.6}_{3.0}$	$40.4_{7.5}$	$59.1_{2.1}$	$30.4_{7.3}$	$55.0_{5.0}$
$\overline{\text{UNI}}$ [4]	$96.1_{0.9}$	$26.9_{2.4}$	$58.9_{13.5}$	$72.6_{1.7}$	$36.9_{5.3}$	$93.1_{1.9}$	$79.8_{1.8}$	$20.2_{4.5}$	$40.0_{3.5}$	$75.5_{2.8}$	$35.5_{3.3}$	$63.1_{7.8}$	$50.4_{4.8}$	$64.0_{3.2}$	$35.4_{4.8}$	$56.6_{5.1}$
$\overline{\text{H-Optimus}}$ [32]	$97.3_{0.3}$	$38.3_{4.5}$	$68.2_{5.0}$	$68.5_{1.1}$	$40.4_{2.1}$	$91.2_{1.7}$	$81.2_{3.1}$	$13.6_{2.2}$	$38.3_{5.0}$	$89.2_{1.7}$	$38.7_{8.2}$	$53.2_{5.9}$	$46.6_{7.0}$	$64.2_{2.0}$	$30.7_{6.4}$	$57.3_{4.4}$
$\overline{\text{GigaPath}}$ [44]	$96.6_{0.7}$	$29.9_{3.1}$	$68.4_{7.6}$	$72.6_{2.3}$	$40.4_{2.9}$	$92.4_{0.6}$	$80.2_{1.1}$	$20.5_{5.1}$	$43.9_{7.5}$	$82.8_{2.2}$	$39.3_{2.5}$	$62.4_{4.5}$	$42.6_{6.8}$	$64.4_{1.3}$	$28.7_{9.5}$	$57.7_{4.7}$
$\overline{\text{Concatenated}}$	$97.4_{0.4}$	$34.2_{4.9}$	$69.7_{2.6}$	$69.6_{1.8}$	$42.1_{2.9}$	$94.4_{1.0}$	$83.1_{2.6}$	$21.0_{3.6}$	$38.7_{2.8}$	$86.5_{2.9}$	$40.9_{5.2}$	$59.7_{4.8}$	$47.4_{9.5}$	$\mathbf{65.2}_{0.8}$	$\mathbf{39.3}_{5.6}$	$59.3_{4.1}$
$\overline{\text{Ensemble Prediction}}$	$97.3_{0.3}$	$36.4_{5.6}$	$69.9_{5.1}$	$71.4_{1.4}$	$\underline{42.9}_{2.4}$	$93.0_{2.0}$	$82.4_{2.5}$	$20.0_{2.7}$	$40.0_{3.4}$	$85.6_{1.6}$	$38.3_{3.6}$	$61.4_{6.2}$	$\mathbf{50.7}_{6.5}$	$63.9_{1.4}$	$35.7_{6.0}$	$59.3_{3.9}$
$\overline{\text{Virchow2}}$ [47]	$95.9_{0.5}$	$40.4_{8.9}$	$\underline{70.2}_{5.4}$	$70.7_{1.5}$	$\mathbf{44.7}_{5.6}$	$94.6_{2.0}$	$83.0_{2.1}$	$26.2_{5.9}$	$40.5_{6.0}$	$84.9_{1.6}$	$41.5_{7.3}$	$65.7_{2.7}$	$47.3_{8.8}$	$62.0_{4.6}$	$32.7_{10.5}$	$60.0_{5.7}$
GigaPath-SE [44]	$91.3_{1.1}$	$28.8_{2.7}$	$51.4_{1.5}$	$69.9_{2.0}$	$40.0_{3.0}$	$80.5_{1.0}$	$73.0_{2.6}$	$17.2_{2.2}$	$42.6_{3.2}$	$39.9_{4.2}$	$18.4_{1.9}$	$55.3_{5.2}$	$38.3_{3.3}$	$51.8_{2.5}$	$22.7_{5.0}$	$48.1_{3.0}$
COBRA-ENC	$92.7_{0.5}$	$26.5_{3.4}$	$50.2_{3.4}$	$73.2_{3.0}$	$35.5_{2.3}$	$82.4_{0.8}$	$77.8_{2.5}$	$26.5_{2.1}$	$41.1_{2.9}$	$31.2_{6.8}$	$31.2_{9.4}$	$52.0_{3.5}$	$31.0_{1.2}$	$51.4_{2.5}$	$21.2_{3.1}$	$48.3_{3.8}$
CHIEF [42]	$94.7_{0.5}$	$26.5_{7.3}$	$52.5_{11.2}$	$73.2_{1.7}$	$32.7_{1.0}$	$90.0_{0.5}$	$82.1_{0.5}$	$17.7_{1.5}$	$\underline{51.4}_{2.9}$	$56.3_{6.7}$	$30.9_{3.4}$	$56.5_{2.3}$	$45.3_{5.8}$	$54.3_{4.3}$	$24.9_{3.8}$	$52.6_{4.6}$
COBRA-CTP	$96.4_{0.4}$	$27.1_{6.0}$	$55.7_{6.3}$	$74.2_{1.4}$	$33.0_{3.5}$	$87.6_{0.4}$	$78.9_{0.7}$	$17.9_{1.1}$	$46.2_{5.2}$	$64.2_{3.8}$	$37.1_{5.1}$	$60.6_{1.5}$	$41.0_{4.0}$	$54.0_{2.1}$	$23.4_{2.4}$	$53.2_{3.5}$
COBRA†-CTP	$96.6_{0.4}$	$27.7_{3.0}$	$57.2_{4.8}$	$74.2_{1.1}$	$34.0_{2.7}$	$86.7_{0.6}$	$78.3_{0.7}$	$18.4_{0.5}$	$44.4_{2.2}$	$66.0_{3.8}$	$40.4_{3.7}$	$61.0_{2.6}$	$43.9_{6.3}$	$55.2_{2.9}$	$23.8_{3.5}$	$53.9_{3.1}$
MADELEINE [18]	$94.6_{0.6}$	$46.3_{8.6}$	$48.7_{9.5}$	$74.3_{1.2}$	$34.3_{2.9}$	$88.7_{1.0}$	$81.0_{0.8}$	$26.0_{2.5}$	$\mathbf{52.1}_{1.3}$	$50.1_{13.5}$	$28.0_{8.3}$	$59.5_{7.4}$	$44.0_{5.6}$	$55.1_{3.3}$	$30.4_{6.5}$	$54.2_{6.2}$
PRISM [34]	$99.3_{0.0}$	$\mathbf{51.3}_{3.5}$	$61.0_{2.7}$	$70.8_{0.8}$	$36.5_{6.7}$	$\mathbf{95.3}_{0.4}$	$\mathbf{86.9}_{1.0}$	$19.1_{3.3}$	$47.2_{3.5}$	$58.0_{3.3}$	$29.3_{2.0}$	$65.5_{3.7}$	$39.5_{8.7}$	$60.4_{0.9}$	$25.1_{4.4}$	$56.3_{3.8}$
COBRA†-UNI	$99.1_{0.2}$	$35.3_{3.6}$	$65.6_{5.2}$	$\mathbf{80.7}_{1.2}$	$36.9_{0.9}$	$91.7_{1.5}$	$82.5_{2.5}$	$19.4_{2.8}$	$41.6_{2.5}$	$77.4_{1.0}$	$45.5_{4.3}$	$\underline{67.4}_{2.6}$	$50.4_{6.6}$	$61.6_{2.2}$	$32.1_{2.8}$	$59.1_{3.1}$
COBRA-UNI	$98.9_{0.3}$	$35.7_{3.3}$	$64.4_{5.5}$	$78.9_{1.9}$	$36.0_{3.6}$	$93.2_{1.3}$	$\underline{84.4}_{2.9}$	$21.2_{3.8}$	$44.8_{6.6}$	$77.0_{2.2}$	$44.8_{2.8}$	$66.4_{3.9}$	$44.1_{10.5}$	$63.0_{1.6}$	$\underline{36.4}_{4.1}$	$59.3_{4.3}$
COBRA†-H0	$\underline{99.4}_{0.1}$	$\underline{50.3}_{2.8}$	$\mathbf{70.8}_{3.3}$	$79.6_{2.4}$	$41.0_{2.5}$	$92.6_{0.7}$	$80.1_{2.2}$	$14.9_{0.7}$	$36.3_{1.0}$	$\underline{89.3}_{1.8}$	$46.4_{4.8}$	$56.9_{3.7}$	$47.0_{10.8}$	$62.0_{2.1}$	$27.9_{2.9}$	$59.6_{3.7}$
COBRA-H0	$\mathbf{99.5}_{0.2}$	$49.2_{2.2}$	$69.8_{3.0}$	$79.1_{3.1}$	$38.2_{1.7}$	$91.9_{1.0}$	$80.6_{3.3}$	$17.6_{2.6}$	$34.9_{2.8}$	$\mathbf{91.0}_{1.2}$	$48.9_{5.3}$	$57.7_{3.0}$	$\underline{50.6}_{7.0}$	$59.9_{1.1}$	$29.2_{5.6}$	$59.9_{3.4}$
COBRA†-GP	$98.9_{0.3}$	$40.2_{3.4}$	$70.1_{4.4}$	$\underline{79.8}_{0.7}$	$38.6_{4.2}$	$93.5_{0.6}$	$81.8_{0.6}$	$26.6_{4.2}$	$48.8_{4.8}$	$87.1_{2.2}$	$49.0_{3.7}$	$65.4_{2.8}$	$44.8_{10.9}$	$62.2_{2.3}$	$28.4_{7.1}$	$61.0_{4.4}$
COBRA-V2	$98.1_{0.2}$	$44.3_{6.1}$	$70.1_{1.9}$	$78.7_{2.0}$	$41.0_{5.0}$	$94.5_{1.2}$	$82.1_{2.1}$	$\underline{30.0}_{4.0}$	$46.4_{9.9}$	$85.2_{2.0}$	$\underline{54.2}_{6.2}$	$66.4_{2.9}$	$50.1_{12.9}$	$64.0_{1.1}$	$25.4_{6.7}$	$\underline{62.0}_{5.5}$
COBRA†-V2	$98.4_{0.2}$	$46.7_{4.8}$	$69.5_{2.7}$	$77.9_{1.7}$	$39.3_{4.5}$	$\underline{94.8}_{0.8}$	$83.0_{1.5}$	$\mathbf{30.7}_{5.5}$	$47.6_{8.5}$	$86.6_{1.4}$	$\mathbf{54.3}_{4.7}$	$66.8_{2.1}$	$\underline{50.6}_{10.4}$	$\underline{64.8}_{1.1}$	$32.2_{5.9}$	$\mathbf{62.9}_{4.7}$

Table 8. **Classification performance comparison.** F1 score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

F1-20×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{Virchow}}$ [39]	$78.2_{2.0}$	$2.1_{4.2}$	$3.3_{4.4}$	$4.6_{5.9}$	$2.0_{4.0}$	$53.0_{9.6}$	$48.8_{6.4}$	$9.2_{8.4}$	$17.2_{20.4}$	$40.5_{1.3}$	$26.3_{2.2}$	$34.1_{24.1}$	$4.6_{4.9}$	$51.7_{22.8}$	$19.5_{11.7}$	$26.3_{11.5}$
$\overline{\text{CTransPath}}$ [41]	$75.9_{2.7}$	$3.5_{4.3}$	$9.0_{13.2}$	$32.9_{12.7}$	$9.9_{12.4}$	$41.0_{10.4}$	$42.8_{5.0}$	$0.0_{0.0}$	$4.7_{9.4}$	$39.7_{1.1}$	$25.5_{2.4}$	$52.5_{14.9}$	$12.5_{18.6}$	$\underline{54.4}_{27.3}$	$0.0_{0.0}$	$27.0_{11.7}$
$\overline{\text{H-Optimus}}$ [32]	$85.4_{1.8}$	$29.4_{7.0}$	$56.3_{5.3}$	$53.7_{15.0}$	$25.2_{7.5}$	$78.7_{10.1}$	$65.7_{14.2}$	$0.0_{0.0}$	$27.3_{22.2}$	$46.4_{7.7}$	$31.6_{6.3}$	$12.3_{22.8}$	$11.5_{10.8}$	$38.0_{24.3}$	$17.8_{19.6}$	$38.6_{13.8}$
$\overline{\text{CONCH}}$ [24]	$89.4_{0.8}$	$11.1_{14.9}$	$36.2_{16.5}$	$58.3_{8.6}$	$36.5_{7.6}$	$67.5_{7.8}$	$54.2_{11.5}$	$18.8_{11.9}$	$16.5_{10.4}$	$42.6_{3.1}$	$21.5_{11.3}$	$54.9_{8.2}$	$15.6_{19.8}$	$34.1_{14.8}$	$26.2_{13.3}$	$38.9_{11.7}$
$\overline{\text{Ensemble Prediction}}$	$81.2_{2.4}$	$18.6_{11.2}$	$56.9_{7.0}$	$59.5_{12.0}$	$19.0_{7.2}$	$83.7_{6.6}$	$66.0_{14.3}$	$1.5_{3.0}$	$24.1_{16.6}$	$52.1_{11.1}$	$36.1_{8.0}$	$20.2_{23.3}$	$15.8_{17.6}$	$37.0_{22.5}$	$14.4_{18.4}$	$39.1_{13.6}$
$\overline{\text{UNI}}$ [4]	$74.6_{3.0}$	$7.4_{9.4}$	$49.2_{10.6}$	$63.7_{11.8}$	$22.9_{16.4}$	$82.9_{5.1}$	$67.2_{10.8}$	$12.0_{7.8}$	$27.6_{11.4}$	$49.8_{9.0}$	$30.8_{2.5}$	$27.0_{20.4}$	$26.6_{19.6}$	$45.6_{15.0}$	$22.5_{14.2}$	$40.7_{12.3}$
$\overline{\text{Concatenated}}$	$78.0_{3.8}$	$20.7_{7.9}$	$43.4_{15.7}$	$62.0_{12.4}$	$27.7_{3.1}$	$\mathbf{86.8}_{1.8}$	$72.3_{12.5}$	$4.0_{4.9}$	$16.8_{18.1}$	$\mathbf{74.3}_{10.4}$	$36.7_{6.6}$	$19.2_{15.0}$	$21.4_{15.0}$	$45.4_{20.6}$	$15.2_{18.7}$	$41.6_{12.6}$
$\overline{\text{Virchow2}}$ [47]	$80.8_{1.9}$	$25.7_{14.0}$	$56.5_{6.3}$	$59.4_{10.7}$	$24.1_{7.5}$	$85.2_{6.1}$	$67.8_{11.4}$	$9.6_{11.7}$	$26.1_{15.8}$	$60.4_{13.8}$	$45.1_{8.0}$	$37.6_{30.9}$	$32.2_{20.7}$	$27.2_{18.4}$	$15.7_{14.6}$	$43.6_{14.5}$
$\overline{\text{GigaPath}}$ [44]	$84.3_{2.7}$	$16.7_{9.2}$	$56.1_{8.6}$	$45.3_{9.1}$	$23.1_{10.2}$	$83.7_{2.4}$	$\underline{75.0}_{6.1}$	$19.7_{9.1}$	$34.8_{17.4}$	$67.1_{8.5}$	$38.3_{6.8}$	$23.4_{19.8}$	$28.4_{16.2}$	$44.9_{18.7}$	$\mathbf{30.8}_{16.2}$	$44.8_{12.0}$
MADELEINE [18]	$84.9_{0.6}$	$11.1_{14.5}$	$1.1_{2.2}$	$62.3_{4.0}$	$15.2_{12.4}$	$53.4_{8.1}$	$45.5_{8.3}$	$22.7_{11.7}$	$5.4_{8.6}$	$22.0_{19.9}$	$17.7_{15.0}$	$23.5_{24.2}$	$5.0_{6.1}$	$26.7_{15.6}$	$26.0_{14.0}$	$28.2_{12.7}$
CHIEF [42]	$82.3_{1.0}$	$12.5_{11.3}$	$36.3_{18.5}$	$51.4_{7.8}$	$22.8_{4.9}$	$69.3_{7.0}$	$62.7_{4.5}$	$3.6_{7.3}$	$18.7_{19.4}$	$46.1_{7.8}$	$25.4_{2.2}$	$62.1_{4.6}$	$18.1_{20.6}$	$41.3_{33.7}$	$8.5_{12.8}$	$37.4_{13.8}$
GigaPath-SE [44]	$70.6_{2.4}$	$19.3_{2.7}$	$41.0_{9.1}$	$58.7_{7.4}$	$\mathbf{40.4}_{5.4}$	$66.6_{11.9}$	$65.0_{17.3}$	$9.4_{5.6}$	$\mathbf{42.3}_{6.9}$	$40.8_{2.8}$	$24.6_{3.2}$	$19.8_{20.9}$	$\underline{35.7}_{9.2}$	$54.3_{15.6}$	$18.0_{13.4}$	$40.4_{10.5}$
COBRA†-CTP	$85.4_{0.7}$	$18.8_{12.1}$	$51.9_{8.3}$	$55.5_{3.3}$	$15.4_{12.8}$	$77.4_{2.5}$	$74.1_{1.6}$	$12.4_{7.0}$	$37.8_{10.7}$	$45.1_{18.6}$	$31.4_{3.4}$	$56.4_{9.8}$	$30.8_{15.8}$	$52.8_{26.6}$	$1.9_{3.8}$	$43.1_{11.5}$
COBRA-CTP	$85.5_{0.9}$	$19.2_{10.5}$	$38.5_{20.1}$	$58.7_{3.7}$	$24.9_{13.4}$	$77.6_{1.2}$	$70.8_{2.8}$	$9.8_{5.1}$	$37.4_{18.9}$	$49.5_{3.1}$	$32.3_{6.8}$	$\underline{62.9}_{3.9}$	$25.2_{21.2}$	$\mathbf{66.5}_{3.8}$	$7.4_{11.3}$	$44.4_{10.8}$
COBRA†-H0	$94.2_{0.6}$	$\underline{55.3}_{4.3}$	$\mathbf{65.9}_{2.6}$	$64.8_{5.0}$	$\underline{37.7}_{7.4}$	$\underline{86.4}_{2.2}$	$65.4_{8.9}$	$0.0_{0.0}$	$14.1_{10.8}$	$67.1_{15.0}$	$37.9_{7.0}$	$19.0_{18.4}$	$21.8_{16.2}$	$34.4_{19.2}$	$7.9_{6.8}$	$44.8_{10.3}$
COBRA-ENC	$84.9_{0.7}$	$33.7_{5.7}$	$47.3_{6.8}$	$62.7_{1.9}$	$35.8_{5.0}$	$79.0_{1.4}$	$71.1_{2.3}$	$17.9_{3.3}$	$28.7_{11.0}$	$36.7_{1.6}$	$28.9_{5.6}$	$45.6_{5.8}$	$\mathbf{37.4}_{4.1}$	$49.6_{6.9}$	$\underline{29.5}_{6.1}$	$45.9_{5.3}$
COBRA-H0	$\underline{94.6}_{0.3}$	$52.2_{4.8}$	$65.1_{1.7}$	$68.2_{5.0}$	$35.4_{4.3}$	$81.5_{6.5}$	$70.1_{6.8}$	$6.1_{8.7}$	$16.5_{15.9}$	$72.5_{10.9}$	$40.6_{7.2}$	$20.7_{17.9}$	$20.4_{14.3}$	$32.0_{19.0}$	$17.6_{18.4}$	$46.2_{11.2}$
COBRA†-UNI	$90.5_{1.2}$	$34.9_{13.6}$	$56.8_{7.8}$	$\underline{73.1}_{5.8}$	$30.9_{16.5}$	$82.3_{3.6}$	$71.9_{6.8}$	$21.2_{4.6}$	$22.8_{21.1}$	$58.9_{6.4}$	$\underline{46.4}_{5.0}$	$35.1_{10.9}$	$30.7_{20.6}$	$33.3_{15.0}$	$15.9_{16.0}$	$47.0_{12.0}$
COBRA-UNI	$90.0_{3.0}$	$32.5_{17.3}$	$58.4_{7.6}$	$71.8_{2.1}$	$31.1_{12.4}$	$83.9_{2.6}$	$72.8_{6.2}$	$\underline{25.4}_{3.9}$	$22.5_{19.0}$	$62.7_{4.4}$	$41.0_{7.0}$	$37.6_{10.6}$	$33.9_{10.3}$	$37.6_{15.2}$	$16.5_{20.8}$	$47.8_{11.2}$
COBRA-V2	$89.9_{1.1}$	$46.9_{8.0}$	$63.3_{5.7}$	$69.1_{5.0}$	$29.3_{15.3}$	$83.7_{2.5}$	$67.8_{18.5}$	$23.7_{2.8}$	$35.3_{22.4}$	$\underline{73.0}_{6.2}$	$\mathbf{49.8}_{4.2}$	$50.3_{10.3}$	$24.7_{21.6}$	$22.3_{14.7}$	$0.0_{0.0}$	$48.6_{11.7}$
PRISM [34]	$\mathbf{96.2}_{0.3}$	$\mathbf{56.1}_{6.3}$	$52.7_{1.3}$	$\mathbf{75.0}_{1.8}$	$27.5_{12.5}$	$74.6_{2.7}$	$62.9_{13.8}$	$24.3_{4.4}$	$23.5_{11.0}$	$54.1_{2.9}$	$29.2_{3.2}$	$\mathbf{66.6}_{2.9}$	$20.0_{20.5}$	$46.0_{5.7}$	$27.9_{4.7}$	$49.1_{8.3}$
COBRA†-V2	$89.7_{0.6}$	$47.9_{4.6}$	$63.1_{5.3}$	$69.4_{3.4}$	$28.9_{14.7}$	$84.3_{2.0}$	$69.3_{10.0}$	$24.0_{2.1}$	$30.9_{21.3}$	$72.3_{7.4}$	$46.3_{6.8}$	$49.7_{18.2}$	$34.8_{18.6}$	$24.4_{17.5}$	$9.1_{11.6}$	$\underline{49.6}_{11.7}$
COBRA†-GP	$92.2_{0.5}$	$34.2_{6.2}$	$\underline{65.2}_{5.1}$	$67.7_{5.5}$	$34.1_{9.2}$	$84.2_{1.3}$	$\mathbf{77.7}_{1.9}$	$\mathbf{28.1}_{4.0}$	$\underline{42.0}_{10.5}$	$68.7_{7.8}$	$46.2_{3.3}$	$35.7_{16.0}$	$27.1_{23.1}$	$43.8_{13.9}$	$6.5_{12.9}$	$\mathbf{50.2}_{10.1}$

Table 9. **Classification performance comparison.** Balanced accuracy score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

Balanced Acc-20×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{Virchow}}$ [39]	$80.5_{1.4}$	$50.6_{1.1}$	$50.6_{1.1}$	$50.3_{0.9}$	$49.8_{0.4}$	$58.2_{4.0}$	$53.5_{3.2}$	$50.4_{3.5}$	$52.6_{4.3}$	$56.8_{1.9}$	$55.5_{1.9}$	$52.8_{4.7}$	$48.6_{1.0}$	$51.2_{0.8}$	$49.5_{4.2}$	$54.1_{2.7}$
$\overline{\text{CTransPath}}$ [41]	$77.3_{2.0}$	$50.2_{0.2}$	$52.1_{3.2}$	$57.8_{3.2}$	$48.1_{3.1}$	$57.6_{3.5}$	$59.1_{1.1}$	$48.4_{1.7}$	$50.7_{1.5}$	$55.6_{2.1}$	$53.3_{3.6}$	$54.7_{2.9}$	$52.6_{4.2}$	$50.4_{0.9}$	$50.0_{0.0}$	$54.5_{2.5}$
$\overline{\text{CONCH}}$ [24]	$89.7_{0.7}$	$52.2_{5.6}$	$55.6_{5.4}$	$67.4_{4.0}$	$54.8_{6.5}$	$73.5_{2.9}$	$65.6_{3.7}$	$56.8_{5.3}$	$53.4_{2.2}$	$60.4_{3.8}$	$54.1_{5.0}$	$\underline{60.6}_{4.8}$	$53.2_{4.3}$	$51.5_{1.6}$	$54.9_{4.4}$	$60.2_{4.3}$
$\overline{\text{H-Optimus}}$ [32]	$87.0_{1.3}$	$58.1_{3.0}$	$68.8_{3.2}$	$62.5_{4.8}$	$52.5_{1.8}$	$74.4_{3.1}$	$63.4_{3.0}$	$50.0_{0.0}$	$54.9_{5.9}$	$64.0_{9.8}$	$59.7_{5.0}$	$50.8_{1.6}$	$52.4_{3.1}$	$53.3_{3.5}$	$54.9_{8.8}$	$60.4_{4.6}$
$\overline{\text{UNI}}$ [4]	$79.6_{1.9}$	$51.0_{2.2}$	$62.7_{8.4}$	$65.2_{5.0}$	$50.4_{4.6}$	$75.4_{2.4}$	$66.1_{4.0}$	$51.6_{2.5}$	$54.3_{2.2}$	$67.8_{9.9}$	$61.6_{3.3}$	$54.8_{5.8}$	$55.8_{4.5}$	$54.4_{3.2}$	$55.7_{5.7}$	$60.4_{4.9}$
$\overline{\text{Ensemble Prediction}}$	$84.1_{1.7}$	$54.3_{4.1}$	$69.5_{4.4}$	$64.8_{4.1}$	$52.8_{1.2}$	$78.3_{3.3}$	$67.1_{4.3}$	$49.6_{0.8}$	$53.8_{4.4}$	$70.0_{11.4}$	$65.9_{4.8}$	$52.3_{3.2}$	$53.9_{4.4}$	$54.6_{3.6}$	$53.8_{6.2}$	$61.7_{4.8}$
$\overline{\text{GigaPath}}$ [44]	$86.1_{2.2}$	$52.6_{3.9}$	$69.2_{5.5}$	$60.7_{3.7}$	$52.5_{2.2}$	$74.0_{2.4}$	$67.2_{2.0}$	$56.9_{6.4}$	$54.3_{3.1}$	$82.9_{4.6}$	$68.2_{5.5}$	$54.2_{5.2}$	$54.1_{3.7}$	$56.0_{2.7}$	$\mathbf{58.3}_{6.8}$	$63.1_{4.3}$
$\overline{\text{Concatenated}}$	$81.9_{2.6}$	$54.4_{3.0}$	$64.1_{6.8}$	$64.9_{3.9}$	$54.4_{1.8}$	$78.1_{3.1}$	$69.5_{4.5}$	$50.3_{0.9}$	$53.2_{4.9}$	$\mathbf{85.5}_{6.0}$	$68.5_{6.1}$	$53.3_{4.1}$	$55.2_{4.3}$	$\mathbf{58.7}_{5.5}$	$55.2_{6.7}$	$63.1_{4.6}$
$\overline{\text{Virchow2}}$ [47]	$83.6_{1.3}$	$57.3_{5.5}$	$68.5_{4.1}$	$65.0_{3.5}$	$54.5_{2.3}$	$\mathbf{78.8}_{4.0}$	$69.8_{4.9}$	$52.6_{3.6}$	$53.8_{5.0}$	$77.4_{11.5}$	$73.3_{4.7}$	$54.4_{3.9}$	$53.7_{4.3}$	$53.7_{2.8}$	$52.6_{5.4}$	$63.3_{5.0}$
GigaPath-SE [44]	$76.5_{1.6}$	$53.7_{1.0}$	$61.1_{2.9}$	$63.6_{2.3}$	$\mathbf{56.1}_{2.8}$	$65.2_{4.0}$	$65.2_{7.4}$	$48.5_{5.3}$	$55.0_{3.0}$	$60.2_{2.5}$	$54.1_{2.2}$	$53.7_{6.2}$	$51.9_{2.1}$	$47.6_{1.5}$	$52.2_{4.5}$	$57.6_{3.7}$
MADELEINE [18]	$85.1_{0.8}$	$53.3_{4.4}$	$50.2_{0.3}$	$65.0_{3.2}$	$50.7_{1.5}$	$66.7_{3.1}$	$62.3_{2.1}$	$57.4_{3.8}$	$51.2_{2.1}$	$56.9_{6.5}$	$56.2_{5.8}$	$53.5_{3.1}$	$50.7_{1.1}$	$50.4_{1.1}$	$\underline{57.7}_{4.7}$	$57.8_{3.4}$
COBRA-ENC	$85.9_{0.7}$	$60.2_{5.0}$	$62.5_{3.5}$	$66.1_{1.4}$	$53.3_{2.1}$	$61.7_{2.8}$	$60.8_{4.7}$	$53.5_{2.1}$	$53.8_{4.4}$	$51.9_{1.7}$	$58.7_{7.9}$	$48.8_{2.6}$	$47.7_{2.9}$	$46.0_{3.2}$	$55.6_{2.8}$	$57.8_{3.6}$
CHIEF [42]	$84.2_{0.8}$	$52.9_{3.2}$	$59.8_{5.5}$	$62.5_{3.4}$	$47.7_{3.7}$	$72.4_{3.0}$	$67.8_{1.7}$	$50.5_{1.3}$	$54.6_{4.9}$	$63.2_{8.4}$	$52.8_{4.2}$	$52.5_{2.9}$	$53.1_{3.5}$	$50.0_{0.1}$	$51.7_{2.4}$	$58.4_{3.8}$
COBRA-CTP	$86.7_{0.7}$	$54.5_{2.7}$	$60.8_{6.6}$	$66.0_{2.4}$	$50.7_{3.3}$	$68.5_{1.9}$	$62.1_{2.1}$	$51.7_{1.2}$	$56.8_{4.9}$	$68.6_{3.3}$	$62.1_{6.4}$	$55.9_{3.5}$	$51.2_{1.6}$	$49.5_{2.7}$	$51.1_{2.3}$	$59.7_{3.5}$
COBRA†-CTP	$86.9_{0.6}$	$54.7_{3.5}$	$65.4_{5.6}$	$64.6_{2.2}$	$49.4_{1.8}$	$66.0_{0.7}$	$64.2_{1.3}$	$52.9_{2.0}$	$55.0_{2.4}$	$68.3_{8.1}$	$62.0_{4.4}$	$54.5_{3.5}$	$53.9_{1.6}$	$49.1_{2.3}$	$50.4_{0.8}$	$59.8_{3.4}$
PRISM [34]	$\mathbf{96.3}_{0.3}$	$\mathbf{75.5}_{5.1}$	$65.6_{0.7}$	$\mathbf{74.1}_{2.0}$	$52.0_{2.9}$	$\underline{78.4}_{1.6}$	$69.5_{5.0}$	$57.8_{5.6}$	$54.4_{2.5}$	$71.4_{2.7}$	$58.9_{1.9}$	$\mathbf{61.9}_{3.3}$	$52.0_{4.8}$	$50.5_{2.5}$	$53.7_{4.2}$	$64.8_{3.4}$
COBRA†-H0	$94.4_{0.6}$	$\underline{74.6}_{3.0}$	$\mathbf{75.0}_{2.2}$	$69.4_{2.2}$	$\underline{55.3}_{4.1}$	$78.3_{3.9}$	$64.4_{3.2}$	$48.4_{1.1}$	$50.1_{1.7}$	$81.5_{10.0}$	$67.7_{4.6}$	$53.8_{4.0}$	$55.2_{5.5}$	$53.8_{3.8}$	$51.1_{1.4}$	$64.9_{4.1}$
COBRA-H0	$\underline{94.8}_{0.3}$	$72.3_{3.0}$	$74.4_{1.3}$	$71.3_{3.1}$	$52.8_{1.8}$	$74.7_{6.1}$	$66.1_{4.6}$	$50.0_{3.1}$	$50.2_{2.2}$	$84.8_{5.6}$	$69.9_{4.2}$	$52.2_{3.6}$	$54.9_{4.0}$	$53.1_{3.0}$	$55.3_{6.1}$	$65.1_{3.8}$
COBRA†-UNI	$91.2_{0.9}$	$61.3_{8.1}$	$68.7_{5.2}$	$\underline{73.5}_{2.1}$	$51.8_{4.4}$	$74.4_{1.5}$	$70.2_{3.7}$	$56.4_{3.5}$	$53.9_{5.3}$	$76.2_{4.6}$	$72.9_{3.0}$	$58.1_{2.6}$	$\mathbf{57.9}_{6.0}$	$55.2_{2.8}$	$54.6_{4.8}$	$65.1_{4.3}$
COBRA-UNI	$90.9_{2.4}$	$61.5_{8.9}$	$69.9_{4.7}$	$70.7_{1.1}$	$51.6_{4.8}$	$77.1_{2.1}$	$\underline{71.3}_{3.2}$	$\underline{58.9}_{3.7}$	$53.8_{5.2}$	$79.1_{3.5}$	$69.8_{4.8}$	$58.0_{2.3}$	$53.1_{3.8}$	$56.8_{4.0}$	$55.8_{8.4}$	$65.2_{4.7}$
COBRA-V2	$90.6_{0.9}$	$68.2_{5.0}$	$73.3_{3.9}$	$70.2_{3.3}$	$54.4_{3.0}$	$67.3_{9.0}$	$70.3_{6.1}$	$56.6_{1.2}$	$\mathbf{59.1}_{6.6}$	$\underline{84.9}_{4.5}$	$\underline{75.0}_{6.5}$	$56.9_{2.6}$	$\underline{56.9}_{6.5}$	$53.6_{2.6}$	$49.4_{0.6}$	$\underline{65.8}_{4.8}$
COBRA†-GP	$92.6_{0.5}$	$60.3_{3.5}$	$\underline{74.5}_{3.5}$	$71.0_{3.8}$	$52.9_{6.3}$	$73.0_{5.1}$	$67.5_{2.4}$	$\mathbf{60.1}_{2.8}$	$\underline{57.7}_{3.1}$	$81.3_{4.1}$	$\mathbf{75.5}_{3.2}$	$57.6_{3.9}$	$55.3_{5.6}$	$\underline{56.9}_{1.8}$	$50.3_{2.8}$	$\underline{65.8}_{3.8}$
COBRA†-V2	$90.5_{0.5}$	$68.7_{3.0}$	$73.2_{3.6}$	$71.0_{3.0}$	$53.1_{2.1}$	$68.3_{6.3}$	$\mathbf{72.0}_{3.8}$	$57.2_{1.8}$	$57.5_{6.8}$	$\underline{84.9}_{4.3}$	$74.2_{8.7}$	$56.6_{4.0}$	$\underline{56.9}_{5.4}$	$55.1_{3.5}$	$51.6_{2.3}$	$\mathbf{66.1}_{4.4}$

Table 10. **Classification performance comparison.** AUC score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC-5×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{Virchow}}$ [39]	$84.0_{1.9}$	$69.0_{2.9}$	$65.7_{2.9}$	$61.2_{2.8}$	$52.6_{10.7}$	$76.1_{2.8}$	$74.5_{1.1}$	$54.2_{2.2}$	$64.9_{8.2}$	$81.7_{0.3}$	$63.9_{2.2}$	$46.9_{2.4}$	$61.6_{7.3}$	$47.8_{1.6}$	$52.5_{6.8}$	$63.8_{4.7}$
$\overline{\text{CTransPath}}$ [41]	$91.5_{0.8}$	$67.9_{5.4}$	$62.6_{6.1}$	$68.3_{2.5}$	$50.0_{5.2}$	$79.1_{2.1}$	$75.5_{2.9}$	$52.4_{1.1}$	$57.2_{6.6}$	$79.1_{2.0}$	$62.0_{1.5}$	$59.4_{3.4}$	$55.8_{6.3}$	$50.5_{1.5}$	$51.6_{4.8}$	$64.2_{4.0}$
$\overline{\text{UNI}}$ [4]	$92.2_{1.1}$	$63.5_{7.8}$	$71.7_{4.5}$	$68.0_{1.7}$	$53.0_{3.7}$	$79.7_{1.3}$	$73.9_{0.6}$	$50.7_{2.9}$	$63.0_{6.6}$	$82.0_{2.3}$	$68.2_{2.6}$	$55.4_{2.7}$	$53.0_{7.0}$	$52.4_{3.2}$	$50.5_{8.2}$	$65.1_{4.5}$
$\overline{\text{GigaPath}}$ [44]	$96.2_{0.8}$	$62.9_{14.5}$	$71.3_{4.6}$	$70.5_{2.2}$	$\underline{57.8}_{5.7}$	$79.8_{1.4}$	$76.0_{1.8}$	$54.1_{3.6}$	$56.7_{5.2}$	$81.9_{2.7}$	$58.4_{8.9}$	$57.0_{3.2}$	$54.9_{5.7}$	$52.0_{3.1}$	$53.6_{1.8}$	$65.5_{5.5}$
$\overline{\text{H-Optimus}}$ [32]	$92.9_{0.6}$	$72.1_{4.3}$	$70.7_{2.5}$	$65.0_{1.8}$	$53.6_{4.9}$	$78.8_{1.5}$	$73.1_{1.1}$	$52.2_{1.2}$	$60.9_{6.1}$	$84.9_{1.9}$	$65.4_{2.6}$	$61.1_{2.8}$	$54.6_{9.6}$	$53.2_{3.3}$	$52.3_{7.5}$	$66.1_{4.3}$
$\overline{\text{CONCH}}$ [24]	$97.8_{0.1}$	$74.4_{2.7}$	$67.9_{5.7}$	$74.2_{0.5}$	$57.5_{8.0}$	$81.6_{0.7}$	$79.0_{1.5}$	$62.5_{2.2}$	$55.4_{10.7}$	$80.4_{2.2}$	$63.9_{9.2}$	$60.5_{1.1}$	$64.3_{7.8}$	$56.1_{3.4}$	$58.8_{1.7}$	$69.0_{5.1}$
$\overline{\text{Virchow2}}$ [47]	$98.3_{0.3}$	$70.8_{17.1}$	$\underline{74.8}_{4.8}$	$74.9_{1.5}$	$57.3_{6.6}$	$\mathbf{90.9}_{0.8}$	$\underline{80.1}_{1.3}$	$\mathbf{70.2}_{1.7}$	$66.0_{5.5}$	$\mathbf{94.7}_{0.5}$	$81.5_{2.5}$	$\underline{61.3}_{2.6}$	$65.2_{8.5}$	$58.9_{2.7}$	$62.7_{12.2}$	$\underline{73.8}_{6.5}$
GigaPath-SE [44]	$90.2_{1.2}$	$63.7_{4.4}$	$66.7_{7.2}$	$74.0_{2.0}$	$49.9_{6.2}$	$75.3_{5.4}$	$73.8_{2.9}$	$51.8_{6.9}$	$65.7_{5.0}$	$75.5_{7.1}$	$61.1_{15.2}$	$51.9_{3.5}$	$57.5_{7.4}$	$47.9_{3.5}$	$55.3_{4.7}$	$64.0_{6.4}$
PRISM [34]	$91.6_{1.2}$	$64.0_{4.9}$	$62.5_{5.7}$	$71.9_{0.9}$	$53.2_{5.8}$	$75.1_{1.5}$	$71.4_{2.7}$	$61.2_{2.5}$	$64.3_{6.5}$	$79.1_{1.0}$	$66.3_{1.1}$	$56.4_{4.6}$	$62.8_{4.1}$	$49.2_{3.2}$	$\underline{64.0}_{4.4}$	$66.2_{3.8}$
MADELEINE [18]	$95.4_{0.4}$	$71.7_{2.0}$	$70.2_{2.3}$	$72.1_{1.8}$	$\mathbf{59.2}_{5.6}$	$79.0_{0.7}$	$77.7_{1.0}$	$62.6_{3.8}$	$59.8_{5.1}$	$75.6_{4.4}$	$60.3_{2.8}$	$57.4_{3.2}$	$\underline{66.6}_{8.4}$	$51.2_{0.4}$	$53.8_{6.3}$	$67.5_{3.9}$
COBRA-UNI	$97.1_{0.4}$	$66.3_{15.4}$	$71.9_{3.8}$	$74.6_{1.0}$	$54.7_{3.6}$	$84.2_{0.9}$	$73.6_{1.1}$	$59.6_{5.0}$	$65.7_{3.3}$	$79.2_{1.9}$	$68.7_{1.1}$	$56.1_{3.8}$	$51.2_{5.9}$	$53.5_{3.7}$	$59.7_{4.2}$	$67.7_{5.1}$
COBRA-H0	$97.2_{0.3}$	$78.2_{5.9}$	$71.9_{3.6}$	$72.4_{1.7}$	$51.6_{3.6}$	$82.5_{1.6}$	$75.1_{1.1}$	$56.4_{3.7}$	$59.8_{3.0}$	$83.7_{2.1}$	$71.0_{1.6}$	$58.3_{4.5}$	$51.9_{5.8}$	$50.6_{4.2}$	$54.5_{8.7}$	$67.7_{4.0}$
COBRA†-CTP	$96.6_{0.4}$	$70.5_{3.9}$	$70.5_{2.1}$	$74.3_{1.0}$	$53.2_{2.4}$	$82.2_{0.9}$	$77.1_{0.8}$	$65.7_{2.9}$	$66.0_{2.6}$	$79.3_{1.3}$	$67.9_{2.8}$	$60.6_{3.2}$	$53.0_{8.8}$	$47.2_{2.4}$	$51.6_{2.3}$	$67.7_{3.2}$
COBRA-CTP	$96.5_{0.3}$	$71.9_{2.3}$	$70.0_{1.7}$	$74.7_{1.2}$	$51.9_{3.1}$	$82.8_{0.8}$	$77.7_{0.8}$	$61.1_{7.3}$	$65.2_{2.2}$	$79.9_{1.6}$	$69.1_{2.4}$	$58.5_{4.0}$	$58.2_{6.2}$	$48.8_{2.7}$	$51.9_{3.7}$	$67.9_{3.3}$
COBRA†-H0	$97.6_{0.4}$	$78.5_{4.0}$	$71.3_{4.0}$	$73.1_{1.3}$	$51.8_{5.2}$	$81.8_{1.1}$	$74.9_{1.8}$	$56.7_{2.0}$	$63.1_{4.1}$	$82.3_{1.6}$	$71.4_{1.8}$	$59.6_{3.1}$	$56.0_{6.6}$	$50.6_{2.4}$	$55.7_{5.3}$	$68.3_{3.5}$
COBRA†-UNI	$97.1_{0.4}$	$77.1_{1.6}$	$71.5_{2.5}$	$75.2_{1.1}$	$57.3_{2.0}$	$82.7_{0.7}$	$74.3_{1.5}$	$57.1_{3.2}$	$\underline{68.2}_{4.2}$	$78.8_{1.9}$	$70.5_{2.2}$	$55.9_{4.3}$	$54.6_{7.0}$	$49.3_{6.8}$	$58.4_{4.4}$	$68.5_{3.5}$
CHIEF [42]	$95.8_{0.3}$	$77.3_{2.4}$	$68.1_{2.9}$	$72.7_{1.3}$	$51.9_{7.6}$	$84.3_{0.5}$	$\mathbf{81.0}_{0.4}$	$68.7_{3.0}$	$\mathbf{70.4}_{1.9}$	$78.0_{0.7}$	$67.5_{2.5}$	$59.0_{8.0}$	$58.2_{8.7}$	$49.6_{1.6}$	$52.8_{2.3}$	$69.0_{4.0}$
COBRA-V2	$\underline{98.9}_{0.3}$	$\mathbf{82.6}_{0.9}$	$74.6_{2.6}$	$\mathbf{80.1}_{2.7}$	$55.9_{3.3}$	$88.8_{1.0}$	$78.2_{1.3}$	$\underline{69.5}_{2.7}$	$66.3_{3.4}$	$\underline{94.3}_{1.2}$	$\mathbf{83.5}_{1.5}$	$\mathbf{61.5}_{1.9}$	$59.8_{12.3}$	$\mathbf{60.0}_{1.8}$	$53.5_{6.8}$	$\underline{73.8}_{4.1}$
COBRA†-V2	$\mathbf{99.0}_{0.2}$	$\underline{81.6}_{1.5}$	$\mathbf{75.5}_{2.7}$	$\underline{79.9}_{1.8}$	$51.4_{7.8}$	$\underline{89.0}_{1.3}$	$79.0_{1.2}$	$67.2_{2.9}$	$62.1_{4.7}$	$94.1_{0.7}$	$\underline{82.9}_{2.7}$	$\underline{61.3}_{1.5}$	$\mathbf{68.1}_{8.5}$	$\underline{59.4}_{3.1}$	$\mathbf{69.6}_{3.4}$	$\mathbf{74.7}_{3.7}$

Table 11. **Classification performance comparison.** AUC score of models trained on TCGA deployed on CPTAC datasets. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). For the other COBRA entries, we used the inference mode from (Eq. (6)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping, CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC-9×[%]	NSCLC	LUAD				BRCA				COAD						Average
Model	ST	STK11	EGFR	TP53	KRAS	ESR1	PGR	ERBB2	PIK3CA	MSI	BRAF	LN	KRAS	Side	PIK3CA	
$\overline{\text{CTransPath}}$ [41]	$90.1_{0.9}$	$64.8_{10.4}$	$63.4_{7.7}$	$70.9_{2.3}$	$56.5_{5.5}$	$77.6_{2.7}$	$72.8_{2.9}$	$53.4_{2.0}$	$59.4_{4.3}$	$71.6_{18.8}$	$63.1_{2.7}$	$60.0_{1.5}$	$59.1_{4.1}$	$52.3_{1.1}$	$53.8_{8.0}$	$64.6_{6.8}$
$\overline{\text{Virchow}}$ [39]	$90.9_{2.1}$	$70.9_{5.2}$	$71.0_{2.7}$	$71.7_{1.2}$	$52.6_{6.1}$	$76.3_{1.7}$	$72.5_{1.8}$	$48.3_{3.0}$	$60.6_{6.0}$	$82.9_{1.1}$	$62.2_{3.2}$	$56.6_{2.2}$	$50.3_{18.8}$	$52.7_{1.9}$	$58.2_{0.9}$	$65.2_{5.8}$
$\overline{\text{CONCH}}$ [24]	$97.8_{0.3}$	$66.7_{20.5}$	$63.5_{10.4}$	$74.8_{1.4}$	$\mathbf{60.2}_{7.0}$	$81.6_{1.7}$	$78.8_{0.8}$	$59.0_{9.5}$	$58.9_{11.3}$	$81.5_{0.9}$	$57.2_{12.7}$	$\underline{\mathbf{64.2}}_{4.5}$	$63.2_{4.7}$	$58.3_{1.5}$	$57.8_{4.1}$	$68.2_{8.3}$
$\overline{\text{GigaPath}}$ [44]	$97.5_{0.3}$	$74.9_{5.9}$	$75.9_{5.0}$	$73.1_{2.3}$	$\underline{57.6}_{4.3}$	$84.9_{1.3}$	$77.3_{2.2}$	$57.1_{2.3}$	$65.1_{6.6}$	$84.7_{11.5}$	$72.8_{8.4}$	$60.3_{2.4}$	$61.4_{5.9}$	$56.1_{3.4}$	$54.0_{7.9}$	$70.2_{5.5}$
$\overline{\text{H-Optimus}}$ [32]	$97.1_{0.6}$	$81.2_{4.2}$	$69.3_{5.8}$	$73.5_{1.0}$	$49.9_{8.2}$	$83.9_{2.0}$	$76.8_{1.8}$	$58.5_{2.1}$	$55.4_{9.0}$	$90.8_{1.5}$	$76.5_{6.0}$	$63.0_{2.8}$	$62.6_{5.1}$	$59.4_{2.3}$	$\underline{63.3}_{2.9}$	$70.7_{4.5}$
$\overline{\text{UNI}}$ [4]	$96.4_{0.6}$	$68.6_{11.1}$	$75.7_{5.3}$	$73.4_{1.5}$	$56.6_{5.2}$	$85.4_{1.6}$	$79.3_{0.6}$	$62.1_{3.6}$	$57.5_{16.2}$	$89.3_{1.3}$	$74.6_{3.7}$	$61.8_{2.5}$	$59.3_{9.7}$	$59.8_{2.0}$	$60.2_{7.6}$	$70.7_{6.5}$
$\overline{\text{Virchow2}}$ [47]	$97.2_{0.8}$	$82.6_{2.2}$	$74.7_{3.2}$	$73.3_{2.0}$	$52.9_{5.1}$	$\mathbf{92.1}_{1.7}$	$\underline{80.0}_{2.6}$	$69.6_{2.2}$	$64.5_{5.1}$	$93.0_{1.4}$	$75.8_{5.7}$	$60.5_{2.0}$	$\mathbf{64.2}_{4.7}$	$60.0_{2.0}$	$\mathbf{63.5}_{5.9}$	$73.6_{3.5}$
GigaPath-SE [44]	$90.5_{0.8}$	$60.4_{8.9}$	$68.0_{11.8}$	$74.7_{1.5}$	$49.4_{4.4}$	$75.1_{2.6}$	$68.7_{3.6}$	$58.5_{4.1}$	$60.9_{2.5}$	$80.0_{2.4}$	$58.4_{7.0}$	$56.5_{3.4}$	$58.1_{6.4}$	$46.9_{1.7}$	$50.8_{6.4}$	$63.8_{5.4}$
MADELEINE [18]	$95.6_{0.4}$	$69.8_{5.9}$	$72.0_{2.3}$	$73.7_{2.0}$	$44.6_{7.2}$	$79.8_{1.7}$	$76.4_{1.2}$	$64.7_{2.0}$	$60.9_{6.1}$	$73.9_{2.5}$	$60.0_{2.2}$	$63.1_{3.1}$	$49.7_{8.4}$	$53.1_{0.4}$	$59.6_{2.5}$	$66.5_{4.0}$
COBRA†-CTP	$96.4_{0.3}$	$75.5_{3.5}$	$69.4_{6.7}$	$74.4_{1.4}$	$52.1_{3.8}$	$81.6_{0.6}$	$76.2_{0.2}$	$65.7_{1.5}$	$62.0_{2.1}$	$81.3_{1.0}$	$72.1_{2.5}$	$59.9_{2.2}$	$52.6_{8.1}$	$49.4_{4.4}$	$55.4_{8.0}$	$68.3_{4.0}$
COBRA-CTP	$96.4_{0.4}$	$75.9_{3.4}$	$71.4_{2.5}$	$74.8_{1.5}$	$51.2_{3.3}$	$83.2_{0.5}$	$78.0_{0.5}$	$63.5_{4.9}$	$63.6_{2.6}$	$81.9_{1.4}$	$74.0_{3.6}$	$58.4_{5.9}$	$56.2_{4.5}$	$51.0_{2.2}$	$52.0_{7.4}$	$68.8_{3.6}$
CHIEF [42]	$95.4_{0.4}$	$74.5_{3.0}$	$68.8_{4.4}$	$73.6_{1.2}$	$55.1_{5.3}$	$85.7_{0.9}$	$\mathbf{81.0}_{0.1}$	$68.1_{3.0}$	$66.5_{1.9}$	$76.2_{8.5}$	$70.8_{1.9}$	$62.2_{1.0}$	$58.3_{10.4}$	$49.5_{2.0}$	$50.5_{3.9}$	$69.1_{4.3}$
PRISM [34]	$97.9_{0.4}$	$80.0_{3.1}$	$71.8_{2.0}$	$74.4_{1.5}$	$56.8_{5.6}$	$84.8_{0.5}$	$77.3_{0.9}$	$65.4_{1.6}$	$\mathbf{68.5}_{2.3}$	$80.7_{2.0}$	$61.2_{5.4}$	$58.0_{3.3}$	$50.5_{4.0}$	$54.5_{5.6}$	$57.3_{3.9}$	$69.3_{3.3}$
COBRA-H0	$\mathbf{99.4}_{0.2}$	$\mathbf{84.4}_{1.6}$	$72.8_{3.5}$	$79.2_{1.7}$	$51.9_{5.4}$	$84.5_{1.8}$	$78.3_{2.2}$	$63.7_{1.6}$	$58.1_{3.6}$	$93.2_{0.8}$	$81.5_{2.7}$	$\underline{\mathbf{64.2}}_{1.8}$	$57.7_{6.4}$	$56.5_{5.7}$	$57.8_{7.8}$	$72.2_{3.8}$
COBRA†-H0	$\underline{99.3}_{0.2}$	$83.4_{2.3}$	$73.6_{3.3}$	$78.7_{2.1}$	$52.4_{5.5}$	$84.1_{2.0}$	$77.2_{0.9}$	$66.7_{1.7}$	$62.3_{3.6}$	$91.4_{0.7}$	$82.5_{3.3}$	$63.8_{2.7}$	$56.2_{4.5}$	$57.3_{2.8}$	$58.3_{2.5}$	$72.5_{2.9}$
COBRA†-UNI	$98.9_{0.3}$	$71.6_{16.2}$	$74.8_{3.4}$	$\mathbf{80.5}_{1.7}$	$56.3_{4.1}$	$87.2_{0.7}$	$79.1_{0.9}$	$65.5_{2.6}$	$66.0_{3.8}$	$89.5_{1.4}$	$\mathbf{85.2}_{1.9}$	$59.9_{4.9}$	$62.0_{8.7}$	$58.4_{3.9}$	$56.5_{5.7}$	$72.8_{5.6}$
COBRA-UNI	$98.8_{0.2}$	$79.3_{1.8}$	$\underline{76.5}_{4.0}$	$79.9_{1.7}$	$56.0_{2.8}$	$88.2_{0.6}$	$79.7_{0.6}$	$64.8_{2.7}$	$66.1_{3.4}$	$88.9_{1.0}$	$\underline{84.7}_{2.3}$	$61.2_{2.9}$	$62.8_{4.4}$	$56.7_{7.4}$	$57.7_{6.8}$	$73.4_{3.5}$
COBRA-V2	$98.8_{0.1}$	$\underline{83.6}_{1.1}$	$75.8_{2.7}$	$79.7_{2.4}$	$54.9_{6.9}$	$88.8_{1.8}$	$79.0_{0.8}$	$\mathbf{70.9}_{3.0}$	$\underline{66.6}_{3.8}$	$\underline{94.7}_{1.3}$	$83.7_{1.9}$	$62.4_{0.6}$	$\underline{63.5}_{10.6}$	$\underline{61.0}_{2.8}$	$51.8_{12.9}$	$\underline{74.3}_{5.0}$
COBRA†-V2	$98.9_{0.2}$	$\underline{83.6}_{1.7}$	$\mathbf{76.7}_{3.9}$	$\underline{80.0}_{1.8}$	$53.0_{4.4}$	$\underline{89.6}_{1.6}$	$79.5_{1.2}$	$\underline{70.6}_{2.6}$	$65.8_{4.8}$	$\mathbf{95.1}_{0.9}$	$82.5_{2.5}$	$61.7_{0.5}$	$58.4_{12.8}$	$\mathbf{61.9}_{2.9}$	$61.2_{3.2}$	$\mathbf{74.6}_{4.2}$

Table 12. **Few shot performance comparison.** AUC score of models on CPTAC datasets with k=5 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC[%]-k=5	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$72.3_{15.4}$	$60.8_{11.7}$	$56.5_{7.7}$	$51.1_{10.5}$	$53.9_{7.6}$	$49.1_{9.8}$	$48.3_{5.7}$	$52.8_{5.0}$	$50.3_{6.9}$	$50.2_{4.1}$	$54.5_{9.1}$
$\overline{\text{CTransPath}}$ [41]	$64.1_{13.5}$	$55.6_{8.5}$	$56.6_{10.5}$	$50.6_{10.2}$	$58.2_{9.0}$	$58.9_{5.4}$	$49.3_{4.7}$	$59.4_{5.1}$	$47.8_{12.8}$	$49.7_{5.1}$	$55.0_{9.0}$
$\overline{\text{H-Optimus}}$ [32]	$68.6_{17.1}$	$63.9_{10.9}$	$63.3_{10.6}$	$51.3_{6.1}$	$62.1_{9.7}$	$51.1_{8.4}$	$48.1_{5.2}$	$71.9_{8.2}$	$55.9_{15.9}$	$51.6_{4.4}$	$58.8_{10.4}$
$\overline{\text{UNI}}$ [4]	$67.8_{15.5}$	$60.8_{6.8}$	$60.3_{12.2}$	$53.3_{7.7}$	$61.9_{11.0}$	$59.8_{10.0}$	$53.6_{7.2}$	$67.4_{6.4}$	$53.8_{9.6}$	$50.7_{10.4}$	$58.9_{10.0}$
$\overline{\text{GigaPath}}$ [44]	$71.6_{13.9}$	$58.1_{10.2}$	$62.9_{11.2}$	$54.1_{8.0}$	$63.3_{11.0}$	$58.5_{6.8}$	$53.2_{7.1}$	$69.9_{9.4}$	$56.8_{10.1}$	$\underline{52.6}_{6.6}$	$60.1_{9.7}$
$\overline{\text{CONCH}}$ [24]	$83.1_{8.8}$	$61.8_{8.9}$	$56.8_{8.7}$	$54.8_{9.2}$	$60.5_{12.7}$	$64.8_{9.3}$	$51.8_{9.0}$	$66.2_{7.1}$	$55.0_{5.8}$	$52.3_{5.8}$	$60.7_{8.7}$
$\overline{\text{Virchow2}}$ [47]	$72.4_{13.6}$	$61.1_{7.3}$	$62.6_{8.9}$	$52.6_{9.8}$	$65.6_{12.1}$	$62.2_{7.2}$	$56.9_{7.1}$	$78.0_{6.9}$	$59.7_{6.6}$	$\mathbf{53.3}_{4.2}$	$62.4_{8.8}$
GigaPath-SE [44]	$65.2_{9.4}$	$57.2_{7.4}$	$58.3_{4.7}$	$52.5_{7.7}$	$58.4_{8.0}$	$54.0_{6.4}$	$53.7_{11.0}$	$54.4_{11.4}$	$51.1_{11.4}$	$47.2_{8.8}$	$55.2_{8.9}$
CHIEF [42]	$73.5_{13.1}$	$60.9_{7.7}$	$58.7_{7.9}$	$54.6_{8.5}$	$63.1_{7.9}$	$\mathbf{66.6}_{4.5}$	$53.7_{7.1}$	$64.2_{8.1}$	$49.8_{12.8}$	$48.7_{5.6}$	$59.4_{8.7}$
COBRA†-CTP	$77.5_{11.3}$	$62.0_{7.4}$	$59.9_{9.6}$	$60.6_{7.0}$	$61.7_{6.6}$	$60.2_{5.2}$	$51.8_{5.0}$	$61.7_{7.0}$	$53.2_{14.2}$	$47.7_{4.9}$	$59.6_{8.3}$
MADELEINE [18]	$87.8_{5.8}$	$63.2_{7.4}$	$59.5_{7.5}$	$54.7_{8.6}$	$62.6_{8.5}$	$62.5_{11.0}$	$\underline{59.3}_{7.6}$	$68.3_{4.2}$	$56.4_{7.1}$	$52.2_{4.5}$	$62.6_{7.5}$
COBRA†-UNI	$86.5_{8.4}$	$\underline{71.8}_{6.2}$	$62.8_{10.5}$	$\underline{60.8}_{7.7}$	$66.4_{10.2}$	$61.7_{9.1}$	$57.5_{11.6}$	$71.7_{8.1}$	$61.5_{10.5}$	$49.1_{8.8}$	$65.0_{9.2}$
COBRA†-H0	$\underline{88.6}_{7.6}$	$\mathbf{74.0}_{11.5}$	$\mathbf{68.4}_{8.5}$	$60.5_{7.5}$	$64.9_{10.8}$	$54.3_{7.9}$	$52.8_{8.3}$	$\underline{78.8}_{7.5}$	$\underline{61.7}_{14.3}$	$51.3_{5.7}$	$65.5_{9.3}$
PRISM [34]	$\mathbf{96.9}_{1.7}$	$70.2_{9.6}$	$59.0_{9.5}$	$\mathbf{65.6}_{8.6}$	$\mathbf{73.0}_{10.3}$	$\underline{66.3}_{7.7}$	$57.1_{9.7}$	$71.2_{5.0}$	$58.6_{5.3}$	$52.1_{2.7}$	$\underline{67.0}_{7.6}$
COBRA†-V2	$86.7_{6.8}$	$66.9_{9.0}$	$\underline{63.4}_{7.6}$	$59.4_{9.4}$	$\underline{71.7}_{10.4}$	$64.9_{6.3}$	$\mathbf{59.8}_{9.7}$	$\mathbf{82.2}_{8.5}$	$\mathbf{66.6}_{9.9}$	$51.0_{3.5}$	$\mathbf{67.3}_{8.4}$

Table 13. **Few shot performance comparison.** AUC score of models on CPTAC datasets with k=10 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC[%]-k=10	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{CTransPath}}$ [41]	$63.8_{13.8}$	$55.9_{7.1}$	$56.1_{10.1}$	$53.9_{8.4}$	$66.0_{7.1}$	$60.9_{5.6}$	$49.7_{6.8}$	$66.6_{9.3}$	$54.6_{13.5}$	$50.4_{4.7}$	$57.8_{9.1}$
$\overline{\text{Virchow}}$ [39]	$75.9_{8.8}$	$62.7_{11.7}$	$56.7_{8.6}$	$56.4_{10.6}$	$59.6_{12.4}$	$58.9_{8.4}$	$48.6_{5.7}$	$62.6_{8.1}$	$53.5_{4.0}$	$50.5_{2.4}$	$58.5_{8.6}$
$\overline{\text{H-Optimus}}$ [32]	$74.6_{13.1}$	$66.8_{7.2}$	$63.5_{6.4}$	$57.9_{9.0}$	$71.6_{8.2}$	$59.8_{7.7}$	$51.1_{7.2}$	$77.7_{8.9}$	$62.5_{8.4}$	$53.8_{5.0}$	$63.9_{8.4}$
$\overline{\text{CONCH}}$ [24]	$85.8_{7.4}$	$60.9_{10.9}$	$59.1_{6.2}$	$60.6_{7.0}$	$73.9_{7.9}$	$66.9_{9.5}$	$53.6_{7.5}$	$68.1_{6.0}$	$61.8_{5.8}$	$\mathbf{55.4}_{5.8}$	$64.6_{7.6}$
$\overline{\text{UNI}}$ [4]	$73.4_{14.8}$	$63.3_{9.2}$	$62.9_{9.0}$	$60.9_{8.3}$	$70.6_{10.8}$	$66.4_{5.6}$	$55.6_{8.7}$	$76.3_{7.4}$	$63.5_{8.3}$	$54.2_{7.3}$	$64.7_{9.2}$
$\overline{\text{GigaPath}}$ [44]	$78.7_{10.1}$	$59.6_{8.3}$	$65.2_{8.9}$	$59.7_{8.1}$	$72.1_{9.6}$	$62.6_{5.1}$	$56.6_{9.0}$	$78.1_{9.2}$	$65.2_{9.6}$	$53.7_{4.5}$	$65.2_{8.4}$
$\overline{\text{Virchow2}}$ [47]	$76.5_{7.1}$	$60.2_{6.8}$	$64.1_{7.1}$	$59.2_{9.1}$	$76.0_{8.3}$	$67.4_{6.1}$	$59.4_{5.1}$	$\underline{82.6}_{8.5}$	$70.0_{7.0}$	$\underline{54.6}_{3.9}$	$67.0_{7.1}$
GigaPath-SE [44]	$71.4_{7.8}$	$60.0_{6.6}$	$61.0_{7.1}$	$57.6_{4.5}$	$62.1_{3.1}$	$57.7_{8.0}$	$53.8_{10.3}$	$55.2_{10.1}$	$56.8_{9.4}$	$48.9_{6.7}$	$58.5_{7.7}$
CHIEF [42]	$76.2_{12.0}$	$65.0_{4.4}$	$60.2_{8.6}$	$58.2_{8.0}$	$70.8_{8.1}$	$\underline{68.9}_{6.1}$	$56.6_{8.8}$	$71.8_{10.6}$	$57.9_{13.5}$	$50.1_{4.1}$	$63.6_{8.9}$
COBRA†-CTP	$82.1_{9.7}$	$67.0_{4.2}$	$60.9_{7.9}$	$64.1_{6.2}$	$67.2_{6.0}$	$61.0_{6.2}$	$54.3_{6.0}$	$71.3_{10.7}$	$61.7_{12.7}$	$47.7_{3.1}$	$63.7_{7.8}$
MADELEINE [18]	$90.0_{5.4}$	$64.9_{7.5}$	$60.9_{5.8}$	$61.2_{7.7}$	$74.5_{6.8}$	$64.7_{10.2}$	$\mathbf{63.0}_{6.2}$	$71.0_{6.7}$	$60.2_{4.7}$	$54.0_{3.1}$	$66.4_{6.7}$
PRISM [34]	$\mathbf{97.8}_{0.7}$	$\underline{74.9}_{9.3}$	$63.0_{7.2}$	$\mathbf{70.9}_{6.8}$	$77.0_{7.7}$	$\mathbf{72.5}_{6.6}$	$58.7_{7.9}$	$74.4_{3.8}$	$62.0_{8.1}$	$51.5_{3.8}$	$70.3_{6.7}$
COBRA†-H0	$\underline{92.7}_{4.3}$	$\mathbf{78.8}_{3.9}$	$\mathbf{72.6}_{4.2}$	$67.5_{6.5}$	$75.5_{5.9}$	$59.4_{9.1}$	$54.0_{8.8}$	$\underline{82.6}_{7.4}$	$67.5_{8.7}$	$52.4_{5.9}$	$70.3_{6.7}$
COBRA†-UNI	$91.0_{5.7}$	$73.5_{6.2}$	$\underline{69.7}_{5.3}$	$\underline{69.4}_{7.1}$	$\underline{77.1}_{6.3}$	$63.6_{6.6}$	$58.2_{8.4}$	$78.9_{5.9}$	$\underline{70.6}_{6.7}$	$51.9_{5.9}$	$\underline{70.4}_{6.5}$
COBRA†-V2	$90.7_{4.0}$	$71.4_{3.8}$	$69.3_{6.2}$	$68.8_{6.1}$	$\mathbf{78.2}_{6.1}$	$64.4_{7.8}$	$\underline{62.7}_{7.2}$	$\mathbf{85.3}_{5.5}$	$\mathbf{76.6}_{7.7}$	$53.2_{4.2}$	$\mathbf{72.1}_{6.0}$

Table 14. **Few shot performance comparison.** AUC score of models on CPTAC datasets with k=25 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUC[%]-k=25	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{CTransPath}}$ [41]	$71.8_{14.1}$	$61.1_{6.7}$	$60.6_{5.6}$	$60.3_{8.2}$	$67.6_{6.5}$	$62.3_{7.2}$	$50.3_{5.8}$	$76.2_{6.2}$	$63.9_{12.0}$	$53.0_{4.0}$	$62.7_{8.2}$
$\overline{\text{Virchow}}$ [39]	$79.9_{9.3}$	$68.0_{7.8}$	$64.2_{8.0}$	$60.4_{9.8}$	$61.6_{13.1}$	$58.1_{9.9}$	$53.8_{6.9}$	$72.4_{10.8}$	$62.2_{5.8}$	$52.6_{5.0}$	$63.3_{8.9}$
$\overline{\text{UNI}}$ [4]	$80.1_{10.6}$	$65.3_{9.8}$	$69.0_{6.8}$	$65.1_{7.2}$	$73.4_{11.2}$	$63.9_{6.7}$	$56.8_{6.8}$	$83.5_{5.0}$	$67.3_{8.3}$	$57.7_{3.3}$	$68.2_{7.9}$
$\overline{\text{H-Optimus}}$ [32]	$82.4_{9.1}$	$70.9_{9.4}$	$72.1_{3.3}$	$61.5_{6.4}$	$74.5_{6.8}$	$58.6_{10.8}$	$49.1_{5.3}$	$85.3_{7.5}$	$70.5_{8.6}$	$\mathbf{59.7}_{4.6}$	$68.5_{7.5}$
$\overline{\text{GigaPath}}$ [44]	$82.2_{10.0}$	$66.2_{7.8}$	$73.7_{4.0}$	$63.9_{8.2}$	$75.2_{8.2}$	$62.6_{8.5}$	$59.5_{7.1}$	$82.5_{9.2}$	$68.8_{7.8}$	$56.4_{2.7}$	$69.1_{7.7}$
$\overline{\text{CONCH}}$ [24]	$91.3_{5.3}$	$70.1_{8.8}$	$66.0_{7.3}$	$64.8_{6.8}$	$76.4_{7.3}$	$66.7_{10.9}$	$56.4_{6.2}$	$77.8_{5.3}$	$68.1_{6.8}$	$58.4_{5.8}$	$69.6_{7.2}$
$\overline{\text{Virchow2}}$ [47]	$83.9_{7.7}$	$69.5_{7.6}$	$71.0_{4.2}$	$63.0_{8.2}$	$79.0_{5.5}$	$66.3_{10.2}$	$63.9_{6.2}$	$89.1_{4.6}$	$74.5_{6.1}$	$\underline{58.8}_{3.4}$	$71.9_{6.7}$
GigaPath-SE [44]	$77.8_{9.0}$	$64.0_{6.4}$	$63.0_{4.0}$	$61.4_{6.8}$	$56.9_{7.7}$	$59.4_{8.9}$	$55.8_{8.8}$	$61.9_{6.3}$	$61.6_{9.3}$	$49.5_{5.1}$	$61.1_{7.4}$
CHIEF [42]	$84.3_{11.0}$	$69.4_{6.2}$	$67.1_{6.0}$	$65.6_{7.1}$	$74.3_{5.9}$	$\mathbf{70.6}_{5.7}$	$55.5_{7.5}$	$78.0_{7.3}$	$65.0_{13.6}$	$50.9_{3.5}$	$68.1_{7.9}$
Cobra†-CTP	$88.6_{7.6}$	$70.2_{6.1}$	$69.0_{5.3}$	$70.3_{6.0}$	$72.5_{4.4}$	$63.7_{7.4}$	$51.9_{6.6}$	$80.1_{6.5}$	$68.6_{9.6}$	$49.8_{3.4}$	$68.5_{6.5}$
MADELEINE [18]	$93.4_{4.4}$	$70.8_{6.9}$	$67.3_{6.0}$	$66.7_{6.4}$	$77.7_{6.5}$	$65.2_{9.6}$	$\mathbf{66.3}_{3.2}$	$77.1_{4.0}$	$60.5_{3.4}$	$56.1_{4.0}$	$70.1_{5.8}$
PRISM [34]	$\mathbf{98.1}_{0.6}$	$\mathbf{82.6}_{5.0}$	$73.2_{5.3}$	$\underline{72.2}_{4.4}$	$\underline{79.1}_{6.8}$	$\underline{70.5}_{4.2}$	$59.7_{7.4}$	$78.2_{3.5}$	$62.9_{6.0}$	$51.1_{3.5}$	$72.8_{5.0}$
Cobra†-UNI	$94.2_{3.7}$	$73.1_{5.6}$	$74.2_{6.2}$	$\mathbf{73.3}_{5.0}$	$77.6_{8.6}$	$66.6_{8.2}$	$57.3_{7.9}$	$84.5_{5.2}$	$75.1_{7.8}$	$55.6_{5.4}$	$73.2_{6.6}$
Cobra†-H0	$\underline{95.5}_{3.3}$	$\underline{79.0}_{4.7}$	$\mathbf{78.1}_{3.8}$	$70.7_{5.8}$	$75.7_{7.3}$	$60.0_{11.4}$	$51.8_{5.8}$	$\underline{89.6}_{4.3}$	$\underline{76.1}_{7.6}$	$56.5_{6.6}$	$\underline{73.3}_{6.5}$
Cobra†-V2	$93.4_{4.9}$	$73.9_{4.7}$	$\underline{75.3}_{5.2}$	$71.7_{7.4}$	$\mathbf{81.6}_{4.9}$	$65.7_{10.8}$	$\underline{64.8}_{5.7}$	$\mathbf{90.3}_{4.1}$	$\mathbf{82.2}_{5.2}$	$58.3_{4.8}$	$\mathbf{75.7}_{6.1}$

Table 15. **Few shot performance comparison.** AUPRC score of models on CPTAC datasets with k=5 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and underline indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUPRC[%]-k=5	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$72.1_{14.8}$	$25.3_{10.0}$	$39.3_{6.7}$	$55.0_{6.9}$	$67.6_{3.9}$	$59.8_{7.2}$	$13.1_{2.1}$	$25.8_{3.3}$	$17.5_{5.3}$	$54.4_{3.7}$	$43.0_{7.3}$
$\overline{\text{CTransPath}}$ [41]	$63.7_{12.5}$	$22.7_{5.9}$	$40.8_{10.2}$	$55.9_{9.0}$	$73.0_{5.8}$	$68.7_{4.8}$	$13.6_{1.9}$	$36.2_{6.2}$	$17.9_{7.9}$	$55.6_{4.1}$	$44.8_{7.4}$
$\overline{\text{UNI}}$ [4]	$70.5_{13.0}$	$26.7_{7.3}$	$46.0_{11.9}$	$56.1_{6.1}$	$75.7_{8.6}$	$67.9_{6.9}$	$16.4_{5.1}$	$43.3_{6.3}$	$22.1_{9.0}$	$55.2_{7.2}$	$48.0_{8.5}$
$\overline{\text{H-Optimus}}$ [32]	$70.7_{15.0}$	$31.8_{13.0}$	$\underline{48.5}_{11.4}$	$54.6_{4.3}$	$75.1_{7.2}$	$62.4_{6.4}$	$12.7_{1.8}$	$47.3_{11.7}$	$25.4_{11.4}$	$56.6_{3.7}$	$48.5_{9.6}$
$\overline{\text{GigaPath}}$ [44]	$71.0_{13.9}$	$25.0_{7.8}$	$48.1_{12.3}$	$57.3_{6.4}$	$77.5_{8.0}$	$67.2_{5.7}$	$14.2_{2.5}$	$46.8_{10.4}$	$21.5_{6.3}$	$\mathbf{57.7}_{5.5}$	$48.6_{8.5}$
$\overline{\text{CONCH}}$ [24]	$83.7_{10.2}$	$27.2_{10.8}$	$39.0_{6.9}$	$58.1_{7.9}$	$75.0_{9.2}$	$72.3_{6.3}$	$14.5_{4.4}$	$44.9_{7.2}$	$21.1_{4.4}$	$56.8_{3.9}$	$49.3_{7.5}$
$\overline{\text{Virchow2}}$ [47]	$73.3_{13.5}$	$24.7_{5.8}$	$48.2_{10.7}$	$56.5_{8.2}$	$78.0_{8.4}$	$69.7_{5.8}$	$16.1_{3.3}$	$54.6_{9.6}$	$\underline{26.5}_{10.4}$	$\underline{57.6}_{3.6}$	$50.5_{8.5}$
GigaPath-SE [44]	$65.9_{9.6}$	$25.2_{8.2}$	$40.7_{5.0}$	$57.2_{8.0}$	$72.8_{5.7}$	$63.1_{6.5}$	$15.8_{5.8}$	$31.7_{7.8}$	$18.1_{7.8}$	$52.2_{6.2}$	$44.3_{7.2}$
CHIEF [42]	$73.2_{14.4}$	$25.1_{5.5}$	$42.4_{8.0}$	$58.8_{8.3}$	$75.4_{4.9}$	$\mathbf{73.5}_{3.2}$	$14.5_{2.5}$	$40.8_{9.0}$	$19.7_{8.5}$	$54.7_{4.6}$	$47.8_{7.7}$
Cobra†-CTP	$78.2_{13.1}$	$26.6_{6.3}$	$43.6_{8.7}$	$63.1_{7.3}$	$75.2_{3.8}$	$69.3_{4.3}$	$15.1_{2.6}$	$38.7_{9.5}$	$21.5_{9.9}$	$54.3_{4.1}$	$48.6_{7.6}$
MADELEINE [18]	$88.6_{6.1}$	$28.3_{9.0}$	$41.7_{8.1}$	$58.8_{8.5}$	$77.5_{6.3}$	$70.9_{9.2}$	$\mathbf{20.5}_{6.6}$	$50.6_{4.5}$	$24.2_{6.8}$	$56.1_{2.7}$	$51.7_{7.1}$
Cobra†-UNI	$87.0_{11.1}$	$\underline{33.4}_{9.2}$	$47.1_{11.6}$	$\underline{63.4}_{7.8}$	$79.1_{7.2}$	$69.8_{7.3}$	$17.7_{6.4}$	$49.0_{8.9}$	$25.7_{9.3}$	$55.2_{6.8}$	$52.7_{8.7}$
PRISM [34]	$\mathbf{96.8}_{2.1}$	$31.9_{7.6}$	$40.0_{9.1}$	$\mathbf{64.9}_{5.9}$	$\mathbf{83.3}_{7.1}$	$\underline{72.4}_{7.4}$	$18.1_{4.7}$	$47.6_{4.5}$	$24.7_{5.3}$	$56.9_{3.2}$	$53.7_{6.0}$
Cobra†-H0	$\underline{88.9}_{9.0}$	$\mathbf{40.5}_{17.1}$	$\mathbf{53.4}_{10.2}$	$62.9_{6.8}$	$76.9_{7.6}$	$63.7_{6.5}$	$16.3_{4.5}$	$\underline{59.6}_{12.5}$	$25.9_{10.6}$	$55.6_{4.9}$	$\underline{54.4}_{9.7}$
Cobra†-V2	$87.5_{7.7}$	$29.5_{7.1}$	$46.7_{7.8}$	$61.8_{8.2}$	$\underline{82.3}_{7.0}$	$71.5_{5.7}$	$\underline{19.7}_{5.0}$	$\mathbf{64.0}_{10.3}$	$\mathbf{31.4}_{9.2}$	$56.9_{3.7}$	$\mathbf{55.1}_{7.4}$

Table 16. **Few shot performance comparison.** AUPRC score of models on CPTAC datasets with k=10 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUPRC[%]-k=10	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$77.4_{8.1}$	$26.3_{9.9}$	$40.3_{9.5}$	$59.3_{8.2}$	$74.4_{8.0}$	$66.6_{5.6}$	$13.3_{2.8}$	$35.9_{9.7}$	$19.5_{3.9}$	$55.5_{2.7}$	$46.9_{7.4}$
$\overline{\text{CTransPath}}$ [41]	$64.1_{12.9}$	$21.2_{4.9}$	$40.6_{9.4}$	$57.7_{7.4}$	$78.9_{5.2}$	$70.1_{4.4}$	$13.7_{2.8}$	$43.1_{9.2}$	$25.9_{11.6}$	$57.1_{4.1}$	$47.2_{7.9}$
$\overline{\text{UNI}}$ [4]	$76.2_{12.3}$	$25.6_{6.3}$	$48.7_{10.3}$	$62.9_{6.9}$	$82.5_{7.1}$	$73.1_{3.8}$	$15.4_{4.4}$	$55.7_{12.4}$	$28.6_{9.5}$	$\underline{59.2}_{5.6}$	$52.8_{8.4}$
$\overline{\text{CONCH}}$ [24]	$86.6_{7.4}$	$27.3_{9.3}$	$40.9_{7.5}$	$62.8_{7.6}$	$84.9_{5.1}$	$73.9_{6.7}$	$15.7_{3.7}$	$44.7_{8.0}$	$31.1_{6.8}$	$\mathbf{59.6}_{3.4}$	$52.8_{6.8}$
$\overline{\text{H-Optimus}}$ [32]	$76.9_{11.2}$	$32.7_{12.5}$	$48.6_{7.4}$	$61.2_{6.8}$	$82.7_{5.4}$	$68.3_{6.7}$	$14.8_{3.9}$	$59.6_{14.4}$	$31.0_{9.1}$	$58.5_{4.1}$	$53.4_{8.8}$
$\overline{\text{GigaPath}}$ [44]	$80.0_{9.1}$	$24.7_{8.2}$	$51.2_{10.7}$	$62.0_{6.3}$	$83.5_{6.5}$	$69.7_{3.8}$	$18.0_{5.4}$	$59.1_{11.8}$	$31.9_{9.7}$	$58.9_{5.0}$	$53.9_{8.1}$
$\overline{\text{Virchow2}}$ [47]	$77.6_{6.8}$	$25.0_{4.5}$	$48.9_{9.9}$	$62.4_{7.9}$	$\underline{86.5}_{5.0}$	$74.0_{4.2}$	$18.2_{3.2}$	$62.1_{14.2}$	$\underline{34.5}_{10.5}$	$58.0_{3.3}$	$54.7_{7.8}$
GigaPath-SE [44]	$72.6_{8.0}$	$23.6_{4.9}$	$43.6_{7.0}$	$59.7_{4.6}$	$74.7_{2.7}$	$64.9_{6.0}$	$15.4_{5.5}$	$33.6_{8.3}$	$19.9_{5.4}$	$52.7_{4.1}$	$46.1_{5.9}$
CHIEF [42]	$76.4_{13.5}$	$26.6_{5.3}$	$45.0_{9.0}$	$61.6_{7.4}$	$81.6_{4.7}$	$\underline{75.1}_{5.9}$	$17.6_{4.6}$	$49.1_{14.3}$	$26.1_{11.6}$	$56.5_{3.9}$	$51.6_{8.8}$
COBRA†-CTP	$83.3_{10.9}$	$28.1_{5.7}$	$46.0_{7.0}$	$66.2_{5.9}$	$79.1_{4.1}$	$68.4_{5.1}$	$16.2_{4.0}$	$50.5_{14.8}$	$29.0_{12.1}$	$54.2_{3.2}$	$52.1_{8.2}$
MADELEINE [18]	$90.7_{5.3}$	$26.9_{6.3}$	$44.5_{8.1}$	$63.6_{6.4}$	$85.5_{3.8}$	$72.0_{8.1}$	$\underline{19.9}_{4.5}$	$49.6_{6.9}$	$30.7_{5.5}$	$58.8_{2.1}$	$54.2_{6.0}$
PRISM [34]	$\mathbf{98.0}_{0.8}$	$\underline{35.8}_{11.3}$	$45.4_{8.3}$	$69.2_{4.6}$	$85.8_{4.9}$	$\mathbf{76.1}_{5.4}$	$\underline{19.9}_{5.9}$	$50.8_{5.8}$	$25.5_{3.9}$	$56.8_{2.5}$	$56.3_{6.0}$
COBRA†-UNI	$92.5_{4.3}$	$32.6_{6.1}$	$\underline{53.6}_{7.4}$	$\mathbf{71.4}_{6.3}$	$85.8_{4.5}$	$71.1_{5.2}$	$17.1_{4.4}$	$60.5_{11.0}$	$\underline{34.5}_{9.1}$	$57.2_{4.4}$	$57.6_{6.6}$
COBRA†-H0	$\underline{93.6}_{3.7}$	$\mathbf{42.3}_{6.6}$	$\mathbf{56.9}_{5.9}$	$69.1_{6.1}$	$84.1_{4.5}$	$68.5_{7.9}$	$16.6_{4.9}$	$\underline{65.3}_{12.5}$	$31.3_{9.5}$	$56.3_{5.6}$	$\underline{58.4}_{7.2}$
COBRA†-V2	$92.1_{3.2}$	$30.3_{3.0}$	$53.4_{7.2}$	$\underline{70.9}_{5.5}$	$\mathbf{87.3}_{3.8}$	$72.1_{6.2}$	$\mathbf{21.0}_{5.1}$	$\mathbf{67.7}_{11.4}$	$\mathbf{42.5}_{11.4}$	$57.3_{3.3}$	$\mathbf{59.5}_{6.7}$

Table 17. **Few shot performance comparison.** AUPRC score of models on CPTAC datasets with k=25 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

AUPRC[%]-k=25	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$80.6_{8.2}$	$30.2_{9.6}$	$47.8_{8.3}$	$62.5_{8.3}$	$75.8_{8.2}$	$65.9_{7.1}$	$16.6_{3.9}$	$48.7_{15.4}$	$24.7_{5.7}$	$57.0_{4.0}$	$51.0_{8.5}$
$\overline{\text{CTransPath}}$ [41]	$72.9_{13.8}$	$25.4_{4.7}$	$44.5_{7.2}$	$63.1_{7.3}$	$79.6_{4.8}$	$70.5_{6.1}$	$13.3_{2.4}$	$57.8_{9.5}$	$31.2_{9.5}$	$57.1_{3.1}$	$51.5_{7.6}$
$\overline{\text{UNI}}$ [4]	$81.1_{9.1}$	$26.4_{6.6}$	$53.5_{9.6}$	$66.1_{5.9}$	$83.6_{8.0}$	$70.2_{5.1}$	$18.0_{6.4}$	$64.2_{7.8}$	$30.0_{9.6}$	$60.9_{3.8}$	$55.4_{7.4}$
$\overline{\text{CONCH}}$ [24]	$92.0_{4.6}$	$35.6_{10.5}$	$47.5_{10.7}$	$67.0_{7.3}$	$85.7_{4.7}$	$72.7_{8.2}$	$17.1_{4.7}$	$56.2_{6.8}$	$30.8_{8.5}$	$60.9_{4.7}$	$56.5_{7.4}$
$\overline{\text{GigaPath}}$ [44]	$83.7_{8.8}$	$27.5_{6.3}$	$\underline{59.3}_{6.8}$	$64.0_{5.9}$	$85.3_{5.5}$	$69.2_{5.7}$	$18.9_{6.0}$	$66.0_{12.3}$	$33.7_{11.8}$	$59.9_{3.1}$	$56.8_{7.7}$
$\overline{\text{H-Optimus}}$ [32]	$83.0_{8.6}$	$34.4_{9.8}$	$57.0_{5.0}$	$63.5_{4.1}$	$85.4_{4.1}$	$66.2_{7.3}$	$16.1_{4.0}$	$70.9_{10.8}$	$34.4_{9.4}$	$\mathbf{62.0}_{3.9}$	$57.3_{7.2}$
$\overline{\text{Virchow2}}$ [47]	$84.9_{6.5}$	$30.2_{7.3}$	$54.3_{7.5}$	$64.6_{6.8}$	$\underline{88.5}_{2.9}$	$72.5_{7.9}$	$22.3_{5.3}$	$72.8_{10.2}$	$32.8_{10.9}$	$61.2_{3.7}$	$58.4_{7.3}$
GigaPath-SE [44]	$80.1_{7.1}$	$25.4_{5.2}$	$46.2_{6.5}$	$63.2_{5.5}$	$70.9_{5.8}$	$66.6_{6.4}$	$17.9_{5.4}$	$38.2_{7.3}$	$23.0_{8.0}$	$54.0_{3.8}$	$48.5_{6.2}$
CHIEF [42]	$85.2_{11.0}$	$30.9_{4.5}$	$52.2_{7.0}$	$67.3_{7.2}$	$83.8_{4.2}$	$\mathbf{76.3}_{5.1}$	$16.6_{5.0}$	$58.6_{12.7}$	$29.9_{9.2}$	$54.9_{3.2}$	$55.6_{7.5}$
COBRA†-CTP	$90.2_{6.5}$	$31.6_{5.2}$	$53.1_{3.5}$	$70.7_{6.1}$	$83.2_{2.2}$	$71.1_{6.2}$	$16.7_{5.7}$	$64.5_{10.6}$	$34.9_{10.6}$	$54.0_{2.4}$	$57.0_{6.5}$
MADELEINE [18]	$94.1_{3.8}$	$34.1_{9.7}$	$50.7_{9.4}$	$69.2_{5.3}$	$86.7_{3.6}$	$72.6_{6.8}$	$\mathbf{24.1}_{6.7}$	$57.5_{4.0}$	$29.5_{4.0}$	$\underline{61.4}_{3.3}$	$58.0_{6.1}$
PRISM [34]	$\mathbf{98.3}_{0.8}$	$\mathbf{45.9}_{8.9}$	$54.0_{7.4}$	$69.6_{4.4}$	$86.9_{4.3}$	$\underline{75.1}_{3.6}$	$17.9_{3.6}$	$53.0_{6.3}$	$25.8_{6.3}$	$57.4_{2.9}$	$58.4_{5.3}$
COBRA†-UNI	$95.0_{2.9}$	$32.5_{7.2}$	$57.8_{6.2}$	$\mathbf{74.5}_{5.0}$	$86.2_{6.1}$	$72.6_{7.1}$	$18.9_{6.1}$	$67.8_{7.5}$	$36.3_{9.2}$	$58.8_{4.4}$	$60.0_{6.4}$
COBRA†-H0	$\underline{96.0}_{2.7}$	$\underline{41.8}_{7.4}$	$\mathbf{63.8}_{4.2}$	$71.5_{6.0}$	$85.9_{5.1}$	$68.8_{10.0}$	$17.1_{4.7}$	$\mathbf{79.0}_{6.9}$	$\underline{38.9}_{7.1}$	$59.7_{4.7}$	$\underline{62.2}_{6.2}$
COBRA†-V2	$94.4_{3.7}$	$34.5_{5.8}$	$59.2_{5.7}$	$\underline{72.6}_{7.2}$	$\mathbf{89.7}_{2.4}$	$72.2_{8.9}$	$\underline{24.0}_{7.8}$	$\underline{75.0}_{8.8}$	$\mathbf{43.2}_{8.4}$	$59.6_{3.6}$	$\mathbf{62.4}_{6.6}$

Table 18. **Few shot performance comparison.** F1 score of models on CPTAC datasets with k=5 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

F1[%]-k=5	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$50.4_{29.7}$	$16.9_{12.5}$	$33.0_{19.5}$	$26.5_{27.4}$	$50.0_{27.8}$	$39.9_{25.3}$	$9.9_{8.7}$	$36.6_{2.3}$	$\underline{24.0}_{3.2}$	$41.1_{25.0}$	$32.8_{20.7}$
$\overline{\text{UNI}}$ [4]	$48.3_{21.2}$	$23.1_{9.5}$	$33.5_{16.3}$	$30.4_{20.0}$	$58.4_{22.9}$	$53.2_{21.0}$	$16.4_{8.5}$	$36.5_{13.2}$	$23.8_{1.9}$	$41.6_{28.5}$	$36.5_{18.0}$
$\overline{\text{H-Optimus}}$ [32]	$58.4_{19.2}$	$27.2_{10.4}$	$41.1_{11.1}$	$30.9_{22.5}$	$54.4_{34.1}$	$49.5_{29.0}$	$10.5_{9.4}$	$31.9_{16.7}$	$22.5_{8.1}$	$42.7_{28.0}$	$36.9_{20.8}$
$\overline{\text{CTransPath}}$ [41]	$58.1_{18.8}$	$25.0_{9.7}$	$42.7_{12.3}$	$30.8_{23.1}$	$53.7_{21.7}$	$53.8_{17.4}$	$13.1_{8.9}$	$33.8_{11.4}$	$22.2_{5.0}$	$40.5_{27.7}$	$37.4_{17.1}$
$\overline{\text{GigaPath}}$ [44]	$54.3_{18.1}$	$24.7_{8.7}$	$42.8_{12.9}$	$39.7_{18.1}$	$55.3_{35.3}$	$45.1_{28.4}$	$13.1_{9.7}$	$30.6_{14.5}$	$22.7_{6.0}$	$48.5_{24.3}$	$37.7_{19.7}$
$\overline{\text{Virchow2}}$ [47]	$59.4_{18.6}$	$23.5_{11.3}$	$46.2_{9.2}$	$38.3_{21.9}$	$57.6_{22.0}$	$45.6_{27.5}$	$15.7_{9.4}$	$35.8_{12.6}$	$\mathbf{26.6}_{2.7}$	$42.4_{28.6}$	$39.1_{18.3}$
$\overline{\text{CONCH}}$ [24]	$70.4_{17.5}$	$32.0_{11.6}$	$37.8_{13.9}$	$44.2_{21.4}$	$45.5_{26.1}$	$\mathbf{55.3}_{23.9}$	$15.5_{9.1}$	$38.2_{6.1}$	$21.6_{7.8}$	$46.9_{24.1}$	$40.7_{17.6}$
GigaPath-SE [44]	$50.5_{14.3}$	$27.5_{6.0}$	$39.9_{8.9}$	$41.0_{20.9}$	$55.9_{25.4}$	$40.4_{18.3}$	$17.0_{8.2}$	$28.7_{13.1}$	$23.7_{7.1}$	$\mathbf{54.3}_{19.1}$	$37.9_{15.5}$
CHIEF [42]	$63.1_{17.8}$	$27.5_{7.0}$	$45.7_{6.0}$	$35.4_{16.4}$	$62.7_{20.5}$	$48.4_{26.7}$	$14.2_{10.0}$	$34.1_{11.4}$	$22.8_{3.8}$	$36.2_{28.2}$	$39.0_{16.9}$
COBRA†-UNI	$66.3_{20.6}$	$32.1_{12.0}$	$41.4_{10.4}$	$37.5_{16.0}$	$60.7_{29.4}$	$51.9_{21.8}$	$13.8_{12.2}$	$28.7_{16.2}$	$23.8_{4.5}$	$44.1_{28.1}$	$40.0_{18.7}$
COBRA†-CTP	$65.2_{19.0}$	$30.3_{5.4}$	$48.3_{7.7}$	$39.5_{16.9}$	$\underline{62.9}_{21.6}$	$\underline{54.9}_{18.8}$	$11.4_{8.5}$	$34.0_{12.7}$	$21.8_{8.8}$	$39.9_{25.5}$	$40.8_{15.9}$
COBRA†-H0	$74.3_{14.4}$	$\mathbf{39.2}_{8.9}$	$\mathbf{49.7}_{8.2}$	$41.2_{17.9}$	$49.4_{35.3}$	$48.3_{29.9}$	$9.7_{9.9}$	$32.9_{19.0}$	$23.3_{9.9}$	$41.8_{27.5}$	$41.0_{20.3}$
COBRA†-V2	$74.8_{10.1}$	$36.0_{6.2}$	$47.6_{7.5}$	$\underline{45.5}_{10.8}$	$\mathbf{64.3}_{21.8}$	$43.4_{30.4}$	$\underline{19.5}_{11.0}$	$32.7_{22.4}$	$23.9_{7.6}$	$42.8_{26.0}$	$43.0_{17.5}$
MADELEINE [18]	$\underline{76.0}_{9.8}$	$32.3_{6.0}$	$45.3_{7.7}$	$43.7_{14.3}$	$55.0_{19.5}$	$45.4_{24.5}$	$\mathbf{21.2}_{3.3}$	$\underline{38.9}_{4.3}$	$22.6_{8.6}$	$\underline{50.5}_{20.7}$	$\underline{43.1}_{13.8}$
PRISM [34]	$\mathbf{91.7}_{3.4}$	$\underline{37.1}_{8.9}$	$\underline{49.5}_{5.8}$	$\mathbf{54.8}_{18.8}$	$59.0_{27.1}$	$41.2_{30.5}$	$12.4_{10.3}$	$\mathbf{41.9}_{12.9}$	$23.4_{6.9}$	$32.0_{22.5}$	$\mathbf{44.3}_{17.2}$

Table 19. **Few shot performance comparison.** F1 score of models on CPTAC datasets with k=10 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

F1[%]-k=10	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{CTransPath}}$ [41]	$58.0_{15.7}$	$24.9_{9.2}$	$41.1_{12.9}$	$24.9_{18.3}$	$54.4_{22.5}$	$53.9_{15.2}$	$12.3_{9.2}$	$40.6_{4.3}$	$24.5_{3.5}$	$39.9_{24.2}$	$37.5_{15.1}$
$\overline{\text{Virchow}}$ [39]	$58.5_{21.0}$	$22.2_{11.8}$	$32.4_{14.4}$	$28.6_{28.3}$	$53.9_{23.2}$	$57.9_{21.2}$	$13.8_{8.0}$	$39.6_{1.7}$	$23.6_{6.0}$	$48.4_{22.1}$	$37.9_{17.8}$
$\overline{\text{H-Optimus}}$ [32]	$57.7_{21.8}$	$28.7_{11.6}$	$39.4_{15.6}$	$28.9_{21.2}$	$\underline{70.4}_{23.3}$	$49.6_{30.2}$	$10.7_{10.1}$	$40.4_{9.3}$	$21.9_{7.4}$	$41.6_{29.2}$	$38.9_{19.6}$
$\overline{\text{UNI}}$ [4]	$55.8_{21.8}$	$23.0_{9.9}$	$42.7_{14.4}$	$33.1_{22.3}$	$62.3_{22.4}$	$\mathbf{58.4}_{18.8}$	$13.2_{12.3}$	$42.1_{6.6}$	$22.8_{7.6}$	$40.3_{24.2}$	$39.4_{17.2}$
$\overline{\text{GigaPath}}$ [44]	$59.3_{15.9}$	$19.9_{13.7}$	$47.7_{8.2}$	$38.4_{14.4}$	$65.0_{22.3}$	$54.5_{26.7}$	$13.9_{10.6}$	$40.2_{12.2}$	$25.9_{7.0}$	$\underline{53.9}_{22.9}$	$41.9_{16.6}$
$\overline{\text{Virchow2}}$ [47]	$62.4_{15.1}$	$25.9_{10.4}$	$45.6_{10.5}$	$42.0_{14.7}$	$\mathbf{72.9}_{14.0}$	$52.2_{23.5}$	$18.4_{7.1}$	$41.4_{8.6}$	$22.3_{11.9}$	$37.9_{27.2}$	$42.1_{15.5}$
$\overline{\text{CONCH}}$ [24]	$72.8_{15.3}$	$31.3_{4.0}$	$42.5_{8.9}$	$44.1_{16.0}$	$51.7_{26.9}$	$49.0_{23.5}$	$17.8_{7.4}$	$40.0_{4.3}$	$26.6_{2.5}$	$48.4_{24.1}$	$42.4_{15.9}$
GigaPath-SE [44]	$60.8_{9.1}$	$25.4_{9.1}$	$44.2_{9.3}$	$46.1_{15.2}$	$57.6_{22.6}$	$47.7_{20.8}$	$15.5_{7.9}$	$34.0_{10.9}$	$21.7_{8.1}$	$\mathbf{57.3}_{10.0}$	$41.0_{13.3}$
CHIEF [42]	$67.8_{13.1}$	$32.7_{8.8}$	$44.3_{12.4}$	$35.7_{16.6}$	$64.5_{18.5}$	$48.7_{23.4}$	$14.9_{13.6}$	$41.5_{6.7}$	$24.3_{5.2}$	$40.1_{23.5}$	$41.5_{15.4}$
MADELEINE [18]	$\underline{80.2}_{5.2}$	$31.9_{7.4}$	$44.2_{9.4}$	$46.5_{13.9}$	$56.9_{24.9}$	$29.9_{22.7}$	$\mathbf{22.7}_{2.5}$	$40.3_{3.6}$	$24.7_{2.5}$	$47.7_{21.5}$	$42.5_{14.1}$
COBRA†-H0	$78.9_{8.3}$	$\mathbf{42.9}_{7.0}$	$48.6_{8.6}$	$39.3_{20.3}$	$54.8_{32.3}$	$49.4_{26.9}$	$11.8_{10.3}$	$41.7_{14.9}$	$22.8_{7.7}$	$45.0_{25.4}$	$43.5_{18.4}$
COBRA†-CTP	$72.5_{13.3}$	$34.0_{5.4}$	$47.4_{11.3}$	$38.8_{18.9}$	$67.0_{16.2}$	$\underline{58.1}_{22.6}$	$13.0_{11.2}$	$39.3_{15.3}$	$25.3_{6.8}$	$40.4_{21.2}$	$43.6_{15.2}$
COBRA†-UNI	$75.5_{16.7}$	$38.5_{9.8}$	$49.9_{8.7}$	$40.2_{15.0}$	$59.6_{29.2}$	$54.1_{24.6}$	$12.8_{10.7}$	$42.1_{16.5}$	$\mathbf{28.1}_{5.5}$	$41.1_{19.9}$	$44.2_{17.2}$
COBRA†-V2	$80.0_{7.9}$	$37.2_{8.7}$	$\underline{50.3}_{11.1}$	$\underline{48.2}_{11.9}$	$67.1_{18.0}$	$45.4_{29.2}$	$\underline{19.8}_{8.7}$	$\underline{44.8}_{17.4}$	$\underline{26.8}_{13.7}$	$41.3_{25.9}$	$\underline{46.1}_{16.8}$
PRISM [34]	$\mathbf{92.8}_{3.4}$	$\underline{40.1}_{8.4}$	$\mathbf{52.4}_{4.7}$	$\mathbf{56.3}_{9.6}$	$69.6_{17.5}$	$52.6_{23.1}$	$17.2_{10.3}$	$\mathbf{48.0}_{6.6}$	$24.3_{6.9}$	$41.6_{19.7}$	$\mathbf{49.5}_{12.7}$

Table 20. **Few shot performance comparison.** F1 score of models on CPTAC datasets with k=25 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

F1[%]-k=25	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$62.8_{16.5}$	$26.7_{13.0}$	$46.7_{13.4}$	$38.2_{27.2}$	$40.4_{25.5}$	$50.2_{24.6}$	$17.0_{5.0}$	$44.1_{7.8}$	$26.0_{3.0}$	$41.8_{22.3}$	$39.4_{17.9}$
$\overline{\text{CTransPath}}$ [41]	$64.1_{12.3}$	$25.5_{11.6}$	$47.3_{9.5}$	$34.7_{18.7}$	$60.0_{21.2}$	$49.1_{21.0}$	$15.6_{8.8}$	$41.2_{5.7}$	$25.7_{1.4}$	$42.3_{25.9}$	$40.5_{15.5}$
$\overline{\text{H-Optimus}}$ [32]	$65.4_{14.7}$	$27.4_{14.3}$	$46.6_{17.7}$	$42.6_{23.9}$	$66.2_{26.6}$	$49.3_{29.6}$	$13.5_{9.6}$	$43.7_{10.8}$	$27.1_{3.0}$	$25.1_{28.7}$	$40.7_{19.8}$
$\overline{\text{GigaPath}}$ [44]	$64.4_{15.0}$	$22.1_{14.3}$	$54.1_{7.6}$	$36.4_{20.0}$	$67.9_{22.9}$	$49.5_{29.6}$	$19.5_{8.2}$	$39.9_{11.6}$	$27.0_{4.7}$	$37.6_{23.0}$	$41.8_{17.4}$
$\overline{\text{CONCH}}$ [24]	$81.8_{8.5}$	$37.9_{9.4}$	$44.7_{11.6}$	$42.6_{15.4}$	$50.3_{27.6}$	$52.4_{19.5}$	$18.9_{7.2}$	$49.9_{9.1}$	$29.0_{10.0}$	$\underline{46.8}_{24.9}$	$45.4_{15.9}$
$\overline{\text{UNI}}$ [4]	$64.5_{15.4}$	$27.7_{10.1}$	$49.2_{10.4}$	$45.6_{25.1}$	$68.3_{22.9}$	$\underline{62.0}_{22.3}$	$18.0_{6.7}$	$47.4_{10.4}$	$26.9_{2.2}$	$45.2_{18.9}$	$45.5_{16.2}$
$\overline{\text{Virchow2}}$ [47]	$71.4_{9.2}$	$30.7_{11.5}$	$52.6_{5.5}$	$46.1_{18.0}$	$71.6_{17.8}$	$51.5_{22.8}$	$16.9_{9.5}$	$\underline{54.5}_{19.0}$	$29.3_{6.5}$	$30.6_{25.4}$	$45.5_{16.0}$
GigaPath-SE [44]	$64.4_{9.7}$	$21.1_{11.8}$	$42.4_{9.9}$	$44.5_{16.7}$	$48.9_{25.2}$	$43.1_{24.4}$	$20.2_{7.6}$	$36.3_{6.9}$	$25.0_{5.3}$	$\mathbf{50.9}_{16.3}$	$39.7_{15.0}$
CHIEF [42]	$76.1_{9.8}$	$33.9_{6.1}$	$52.8_{7.4}$	$45.1_{17.0}$	$68.2_{18.4}$	$53.5_{27.2}$	$18.1_{8.1}$	$42.7_{10.2}$	$24.5_{1.7}$	$34.4_{25.0}$	$44.9_{15.3}$
MADELEINE [18]	$85.1_{5.2}$	$37.9_{7.3}$	$45.0_{10.7}$	$50.3_{13.5}$	$57.2_{24.2}$	$40.0_{21.2}$	$\mathbf{23.2}_{1.5}$	$47.5_{7.4}$	$24.8_{2.7}$	$42.5_{22.8}$	$45.4_{14.1}$
COBRA†-CTP	$79.6_{7.3}$	$33.9_{6.7}$	$54.2_{4.6}$	$48.4_{15.3}$	$\mathbf{75.1}_{8.1}$	$58.4_{23.0}$	$14.1_{6.6}$	$47.9_{13.4}$	$26.0_{2.3}$	$33.7_{26.9}$	$47.1_{13.8}$
COBRA†-H0	$\underline{85.2}_{6.7}$	$\underline{42.9}_{6.0}$	$\mathbf{58.4}_{13.1}$	$50.2_{13.9}$	$65.8_{30.6}$	$49.8_{28.4}$	$13.2_{9.9}$	$52.5_{12.5}$	$31.0_{5.4}$	$30.9_{26.0}$	$48.0_{17.7}$
COBRA†-UNI	$83.9_{6.6}$	$40.4_{7.5}$	$56.9_{7.9}$	$53.5_{12.9}$	$\underline{74.2}_{13.5}$	$\mathbf{62.2}_{23.8}$	$18.0_{9.0}$	$50.2_{17.1}$	$\underline{33.3}_{6.7}$	$24.5_{20.9}$	$49.7_{13.9}$
COBRA†-V2	$84.9_{6.0}$	$40.7_{5.3}$	$56.9_{3.6}$	$\underline{53.8}_{14.5}$	$73.2_{15.1}$	$55.7_{22.7}$	$18.5_{7.7}$	$\mathbf{58.0}_{24.2}$	$\mathbf{38.6}_{8.5}$	$24.4_{21.0}$	$\underline{50.5}_{14.8}$
PRISM [34]	$\mathbf{93.4}_{1.8}$	$\mathbf{49.0}_{4.9}$	$\underline{57.6}_{2.9}$	$\mathbf{63.8}_{6.6}$	$66.0_{17.8}$	$55.7_{22.8}$	$\underline{21.1}_{6.3}$	$49.2_{4.1}$	$29.5_{4.4}$	$45.6_{17.7}$	$\mathbf{53.1}_{11.4}$

Table 21. **Few shot performance comparison.** Balanced accuracy score of models on CPTAC datasets with k=5 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

Balanced Acc.[%]-k=5	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$55.2_{8.3}$	$49.5_{5.3}$	$51.7_{2.1}$	$49.3_{3.5}$	$53.5_{6.3}$	$48.6_{3.8}$	$46.5_{3.7}$	$52.6_{2.1}$	$51.0_{3.7}$	$50.2_{2.3}$	$50.8_{4.5}$
$\overline{\text{CTransPath}}$ [41]	$58.7_{9.3}$	$52.5_{4.3}$	$54.1_{6.8}$	$50.6_{4.9}$	$53.5_{7.7}$	$54.3_{3.9}$	$51.0_{1.8}$	$51.3_{2.9}$	$48.4_{2.6}$	$\underline{51.3}_{1.5}$	$52.6_{5.2}$
$\overline{\text{UNI}}$ [4]	$58.9_{7.7}$	$53.5_{3.4}$	$52.8_{7.4}$	$52.1_{5.3}$	$53.4_{6.1}$	$53.3_{4.1}$	$51.9_{2.1}$	$54.9_{7.7}$	$48.9_{4.0}$	$51.1_{2.6}$	$53.1_{5.4}$
$\overline{\text{GigaPath}}$ [44]	$62.2_{9.7}$	$52.7_{5.3}$	$57.8_{7.0}$	$52.8_{6.7}$	$51.5_{4.0}$	$51.8_{3.6}$	$50.9_{4.7}$	$53.7_{4.9}$	$51.0_{6.0}$	$51.2_{3.0}$	$53.6_{5.8}$
$\overline{\text{H-Optimus}}$ [32]	$63.0_{9.9}$	$56.0_{5.7}$	$56.8_{5.5}$	$50.9_{5.4}$	$54.2_{5.5}$	$49.7_{3.6}$	$50.2_{3.2}$	$54.8_{7.6}$	$50.7_{5.8}$	$\underline{51.3}_{2.6}$	$53.8_{5.8}$
$\overline{\text{Virchow2}}$ [47]	$64.3_{9.9}$	$53.7_{4.4}$	$58.8_{7.3}$	$51.6_{6.9}$	$56.0_{8.0}$	$54.7_{3.3}$	$51.2_{4.4}$	$55.1_{6.5}$	$\underline{54.3}_{4.6}$	$50.5_{0.9}$	$55.0_{6.1}$
$\overline{\text{CONCH}}$ [24]	$74.9_{10.9}$	$58.3_{8.2}$	$52.9_{3.6}$	$53.4_{6.2}$	$54.3_{7.0}$	$\mathbf{58.0}_{7.3}$	$50.2_{6.7}$	$54.3_{6.4}$	$50.7_{4.6}$	$50.0_{3.2}$	$55.7_{6.8}$
GigaPath-SE [44]	$59.2_{4.3}$	$54.2_{5.4}$	$54.9_{3.8}$	$52.0_{6.8}$	$53.7_{5.2}$	$51.4_{4.9}$	$50.7_{7.7}$	$50.0_{4.7}$	$52.0_{6.3}$	$49.2_{4.3}$	$52.7_{5.5}$
CHIEF [42]	$65.4_{10.7}$	$54.4_{5.0}$	$55.8_{4.7}$	$51.7_{5.5}$	$54.5_{6.6}$	$57.1_{5.4}$	$51.7_{3.9}$	$52.1_{2.5}$	$48.9_{4.2}$	$49.5_{2.7}$	$54.1_{5.6}$
COBRA†-CTP	$67.7_{10.4}$	$56.6_{5.7}$	$57.8_{6.5}$	$55.4_{4.0}$	$55.1_{5.4}$	$53.5_{4.2}$	$50.9_{2.3}$	$54.7_{5.4}$	$50.4_{6.2}$	$48.5_{2.2}$	$55.1_{5.7}$
COBRA†-UNI	$73.1_{10.6}$	$61.6_{7.2}$	$56.3_{6.0}$	$54.8_{4.7}$	$54.2_{7.7}$	$56.4_{5.7}$	$\underline{53.0}_{7.0}$	$52.6_{4.9}$	$51.8_{4.8}$	$50.8_{4.9}$	$56.5_{6.6}$
MADELEINE [18]	$\underline{78.1}_{6.9}$	$58.2_{6.6}$	$57.2_{6.1}$	$52.7_{5.2}$	$57.4_{5.9}$	$55.9_{5.9}$	$52.2_{3.7}$	$54.5_{6.1}$	$51.8_{7.4}$	$\mathbf{51.8}_{2.6}$	$57.0_{5.8}$
COBRA†-H0	$77.2_{9.2}$	$\mathbf{65.3}_{7.6}$	$\mathbf{62.2}_{6.1}$	$\underline{56.4}_{5.8}$	$54.6_{6.9}$	$52.7_{4.4}$	$51.4_{4.4}$	$54.5_{8.3}$	$52.1_{8.2}$	$50.3_{2.0}$	$57.7_{6.6}$
COBRA†-V2	$77.2_{7.5}$	$62.6_{6.8}$	$\underline{59.3}_{5.4}$	$55.5_{4.2}$	$\underline{58.0}_{6.4}$	$55.4_{4.7}$	$\mathbf{56.0}_{5.4}$	$\underline{56.6}_{11.8}$	$54.2_{6.7}$	$50.0_{1.4}$	$\underline{58.5}_{6.5}$
PRISM [34]	$\mathbf{91.7}_{3.5}$	$\underline{64.9}_{8.5}$	$57.2_{7.5}$	$\mathbf{60.3}_{6.2}$	$\mathbf{60.8}_{7.2}$	$\underline{57.7}_{7.2}$	$50.5_{3.8}$	$\mathbf{62.0}_{7.7}$	$\mathbf{55.5}_{3.6}$	$50.2_{2.4}$	$\mathbf{61.1}_{6.1}$

Table 22. **Few shot performance comparison.** Balanced accuracy score of models on CPTAC datasets with k=10 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

Balanced Acc.[%]-k=10	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$59.9_{7.2}$	$51.7_{3.9}$	$51.1_{3.8}$	$51.4_{5.2}$	$53.3_{7.2}$	$54.6_{4.6}$	$48.2_{3.1}$	$56.0_{3.5}$	$52.3_{4.6}$	$49.9_{3.2}$	$52.8_{4.8}$
$\overline{\text{CTransPath}}$ [41]	$58.5_{9.8}$	$53.1_{4.1}$	$53.0_{6.2}$	$49.9_{4.8}$	$56.7_{3.3}$	$56.4_{4.6}$	$49.6_{3.2}$	$56.8_{6.6}$	$50.3_{5.9}$	$50.1_{2.0}$	$53.4_{5.5}$
$\overline{\text{H-Optimus}}$ [32]	$64.7_{10.8}$	$58.2_{7.2}$	$55.2_{3.9}$	$52.2_{4.9}$	$57.4_{8.2}$	$52.8_{3.8}$	$51.1_{3.8}$	$57.0_{9.7}$	$50.3_{1.8}$	$51.2_{2.2}$	$55.0_{6.4}$
$\overline{\text{UNI}}$ [4]	$64.9_{9.9}$	$53.6_{4.7}$	$57.0_{6.3}$	$55.6_{4.8}$	$58.2_{7.4}$	$58.6_{6.2}$	$51.9_{8.4}$	$57.7_{8.1}$	$51.3_{2.6}$	$\mathbf{53.3}_{4.8}$	$56.2_{6.6}$
$\overline{\text{GigaPath}}$ [44]	$67.1_{8.1}$	$52.6_{4.7}$	$58.3_{5.2}$	$53.7_{7.0}$	$58.6_{8.5}$	$54.1_{4.8}$	$52.7_{4.9}$	$60.2_{8.8}$	$54.9_{4.9}$	$\underline{52.3}_{3.6}$	$56.5_{6.3}$
$\overline{\text{Virchow2}}$ [47]	$67.1_{7.7}$	$55.4_{5.6}$	$57.5_{5.3}$	$55.4_{6.2}$	$\underline{62.8}_{10.2}$	$57.2_{4.5}$	$53.3_{3.6}$	$60.0_{7.8}$	$53.7_{5.1}$	$51.4_{1.8}$	$57.4_{6.2}$
$\overline{\text{CONCH}}$ [24]	$75.9_{8.7}$	$56.0_{5.8}$	$54.1_{4.8}$	$56.0_{5.6}$	$61.3_{6.0}$	$58.9_{7.2}$	$52.3_{4.5}$	$57.1_{5.8}$	$54.3_{4.5}$	$49.9_{4.4}$	$57.6_{5.9}$
GigaPath-SE [44]	$64.6_{5.9}$	$54.6_{6.1}$	$57.8_{4.6}$	$55.2_{4.3}$	$56.8_{3.6}$	$53.1_{6.4}$	$50.2_{5.5}$	$52.5_{8.6}$	$51.9_{4.2}$	$51.4_{4.8}$	$54.8_{5.6}$
CHIEF [42]	$68.9_{10.9}$	$59.0_{6.9}$	$56.1_{6.1}$	$54.2_{5.2}$	$62.3_{5.3}$	$\underline{59.7}_{5.4}$	$52.3_{7.5}$	$58.1_{8.8}$	$50.4_{7.7}$	$50.2_{4.5}$	$57.1_{7.1}$
COBRA†-CTP	$74.3_{9.7}$	$59.9_{3.8}$	$58.4_{7.2}$	$56.1_{4.4}$	$58.2_{6.0}$	$55.4_{4.6}$	$52.2_{4.8}$	$58.4_{9.8}$	$52.0_{9.8}$	$49.5_{1.9}$	$57.4_{6.7}$
MADELEINE [18]	$80.6_{5.9}$	$59.2_{6.2}$	$57.4_{5.8}$	$57.1_{6.2}$	$60.2_{7.7}$	$55.2_{5.0}$	$54.3_{4.7}$	$56.6_{5.4}$	$50.6_{4.4}$	$51.0_{3.5}$	$58.2_{5.6}$
COBRA†-H0	$80.9_{5.6}$	$\mathbf{69.8}_{7.2}$	$\underline{61.5}_{5.1}$	$58.8_{6.1}$	$55.7_{6.1}$	$51.4_{3.3}$	$51.4_{4.1}$	$60.9_{10.1}$	$52.1_{2.3}$	$50.7_{2.4}$	$59.3_{5.7}$
COBRA†-UNI	$79.7_{8.7}$	$65.6_{7.7}$	$61.0_{6.5}$	$58.9_{4.4}$	$58.6_{6.2}$	$55.3_{3.8}$	$51.6_{6.5}$	$62.1_{10.9}$	$\underline{57.0}_{4.9}$	$51.1_{3.2}$	$60.1_{6.7}$
COBRA†-V2	$\underline{81.6}_{6.2}$	$64.0_{7.2}$	$\mathbf{61.9}_{8.2}$	$\underline{59.7}_{5.2}$	$59.2_{5.8}$	$56.5_{6.1}$	$\mathbf{56.2}_{4.5}$	$\underline{64.6}_{12.1}$	$\mathbf{59.2}_{9.2}$	$51.3_{2.4}$	$\underline{61.4}_{7.2}$
PRISM [34]	$\mathbf{92.9}_{3.3}$	$\underline{67.5}_{8.6}$	$60.4_{6.2}$	$\mathbf{62.0}_{4.4}$	$\mathbf{64.5}_{7.8}$	$\mathbf{62.8}_{6.8}$	$\underline{54.4}_{6.6}$	$\mathbf{65.7}_{6.9}$	$55.7_{4.6}$	$50.4_{2.9}$	$\mathbf{63.6}_{6.1}$

Table 23. **Few shot performance comparison.** Balanced accuracy score of models on CPTAC datasets with k=25 positive samples during training on TCGA. $\overline{\text{Overline}}$ indicates mean over patch embeddings, † indicates that embeddings of all four training FMs were used to generate the weighting vector (Eq. (8)). **Bold** indicates the best performance, and <u>underline</u> indicates the second-best performance. The abbreviations are as follows: ST: Subtyping CTP: CTransPath [41], H0: H-Optimus-0 [32], V2: Virchow-2 [47], GP: GigaPath [44], SE: Slide Encoder.

Balanced Acc.[%]-k=25	LUNG		LUAD			BRCA			COAD		Average
Model	ST	STK11	EGFR	TP53	ESR1	PGR	ERBB2	MSI	BRAF	Side	
$\overline{\text{Virchow}}$ [39]	$64.8_{7.0}$	$55.0_{8.0}$	$56.6_{6.6}$	$54.5_{6.9}$	$54.1_{7.4}$	$53.8_{5.2}$	$49.7_{5.1}$	$60.3_{7.9}$	$54.3_{4.3}$	$52.0_{2.9}$	$55.5_{6.3}$
$\overline{\text{CTransPath}}$ [41]	$64.2_{9.7}$	$54.6_{6.2}$	$56.1_{6.4}$	$56.6_{5.3}$	$59.9_{6.0}$	$55.9_{3.7}$	$50.4_{6.0}$	$57.9_{7.6}$	$52.2_{3.2}$	$50.0_{2.3}$	$55.8_{6.0}$
$\overline{\text{H-Optimus}}$ [32]	$71.4_{7.3}$	$59.5_{9.0}$	$62.4_{5.3}$	$57.7_{6.1}$	$54.0_{6.9}$	$51.1_{5.0}$	$52.2_{2.8}$	$61.8_{9.4}$	$54.9_{5.9}$	$50.5_{1.0}$	$57.5_{6.4}$
$\overline{\text{GigaPath}}$ [44]	$71.3_{9.7}$	$55.6_{7.0}$	$64.8_{4.7}$	$57.2_{6.7}$	$59.6_{7.7}$	$54.0_{4.2}$	$54.5_{6.3}$	$61.8_{8.3}$	$55.6_{7.4}$	$51.6_{1.8}$	$58.6_{6.7}$
$\overline{\text{UNI}}$ [4]	$70.4_{8.8}$	$56.6_{8.1}$	$60.6_{5.6}$	$58.9_{6.5}$	$61.1_{8.9}$	$58.1_{5.3}$	$51.9_{3.7}$	$65.1_{10.3}$	$54.9_{4.6}$	$\mathbf{53.6}_{2.0}$	$59.1_{6.8}$
$\overline{\text{Virchow2}}$ [47]	$73.9_{6.3}$	$60.1_{8.6}$	$63.4_{3.2}$	$58.5_{7.0}$	$63.0_{10.0}$	$55.6_{4.9}$	$53.1_{4.3}$	$\underline{72.3}_{12.5}$	$58.5_{8.0}$	$\underline{52.8}_{2.6}$	$61.1_{7.4}$
$\overline{\text{CONCH}}$ [24]	$83.1_{6.6}$	$62.8_{9.3}$	$57.4_{4.9}$	$58.1_{6.0}$	$62.3_{8.0}$	$58.3_{6.1}$	$52.5_{4.6}$	$67.0_{8.3}$	$58.6_{9.6}$	$52.0_{3.6}$	$61.2_{7.0}$
GigaPath-SE [44]	$68.6_{6.1}$	$53.5_{4.9}$	$56.9_{5.4}$	$55.5_{5.5}$	$54.4_{5.3}$	$53.6_{6.5}$	$53.6_{6.9}$	$57.3_{4.5}$	$52.7_{5.3}$	$49.8_{4.5}$	$55.6_{5.5}$
CHIEF [42]	$77.0_{8.9}$	$60.4_{5.5}$	$62.1_{5.7}$	$60.1_{6.6}$	$63.3_{6.8}$	$\underline{60.3}_{6.6}$	$53.1_{5.6}$	$61.5_{7.9}$	$50.0_{3.4}$	$49.5_{2.2}$	$59.7_{6.2}$
COBRA†-CTP	$80.7_{6.6}$	$60.5_{6.9}$	$63.5_{4.2}$	$61.6_{6.0}$	$62.8_{6.5}$	$56.4_{3.9}$	$50.0_{3.8}$	$67.1_{8.0}$	$53.2_{4.3}$	$49.4_{2.5}$	$60.5_{5.5}$
MADELEINE [18]	$85.6_{4.7}$	$64.1_{6.0}$	$59.3_{5.8}$	$61.4_{5.1}$	$64.7_{7.2}$	$57.1_{4.1}$	$\mathbf{55.7}_{3.2}$	$64.6_{6.9}$	$50.9_{4.8}$	$50.5_{2.7}$	$61.4_{5.2}$
COBRA†-H0	$\underline{86.5}_{4.7}$	$\underline{69.4}_{6.1}$	$\mathbf{70.2}_{6.1}$	$62.1_{6.1}$	$58.3_{7.6}$	$51.5_{5.5}$	$51.5_{4.8}$	$68.9_{10.9}$	$60.8_{8.6}$	$50.9_{2.3}$	$63.0_{6.7}$
COBRA†-UNI	$85.5_{5.3}$	$67.0_{7.1}$	$\underline{67.0}_{6.4}$	$\underline{63.8}_{5.0}$	$\underline{67.3}_{6.8}$	$58.7_{5.4}$	$53.9_{4.3}$	$69.6_{10.9}$	$\underline{64.6}_{8.7}$	$50.3_{3.1}$	$64.8_{6.7}$
COBRA†-V2	$86.0_{4.8}$	$67.5_{5.4}$	$66.8_{3.6}$	$63.6_{6.8}$	$64.1_{10.1}$	$58.6_{7.8}$	$54.8_{2.6}$	$\mathbf{76.3}_{13.2}$	$\mathbf{67.4}_{8.4}$	$50.7_{1.7}$	$\underline{65.6}_{7.3}$
PRISM [34]	$\mathbf{93.4}_{1.8}$	$\mathbf{74.7}_{4.9}$	$66.1_{3.5}$	$\mathbf{65.7}_{2.8}$	$\mathbf{68.0}_{8.2}$	$\mathbf{60.6}_{5.4}$	$\underline{55.0}_{6.2}$	$67.5_{3.1}$	$58.8_{4.6}$	$50.1_{3.0}$	$\mathbf{66.0}_{4.7}$

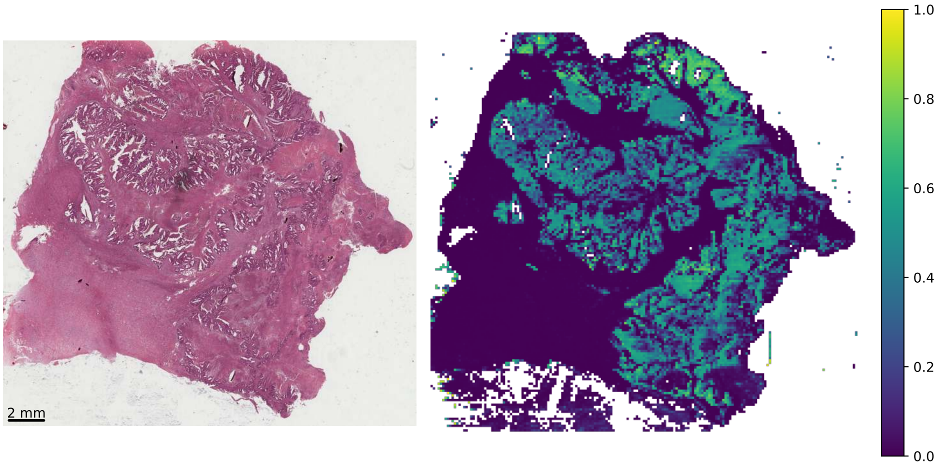

Figure 5. **COBRA Unsupervised Heatmap**. Patient: TCGA-CA-6715

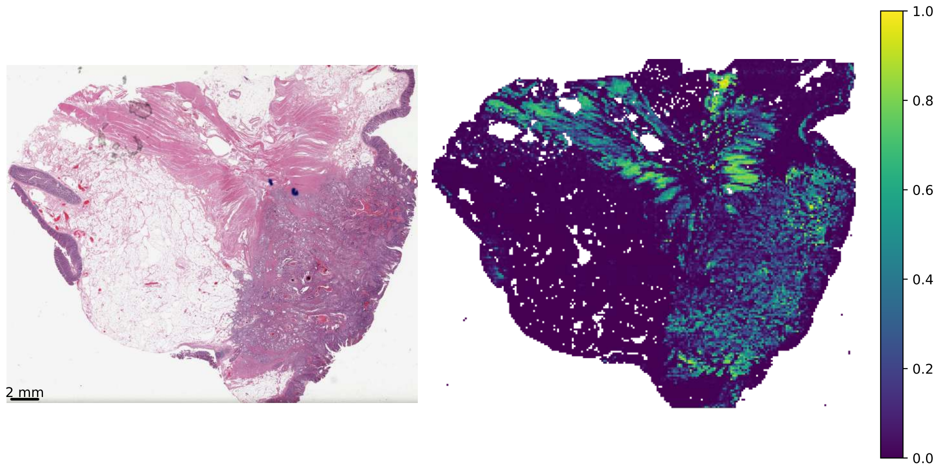

Figure 6. **COBRA Unsupervised Heatmap**. Patient: TCGA-CM-5349

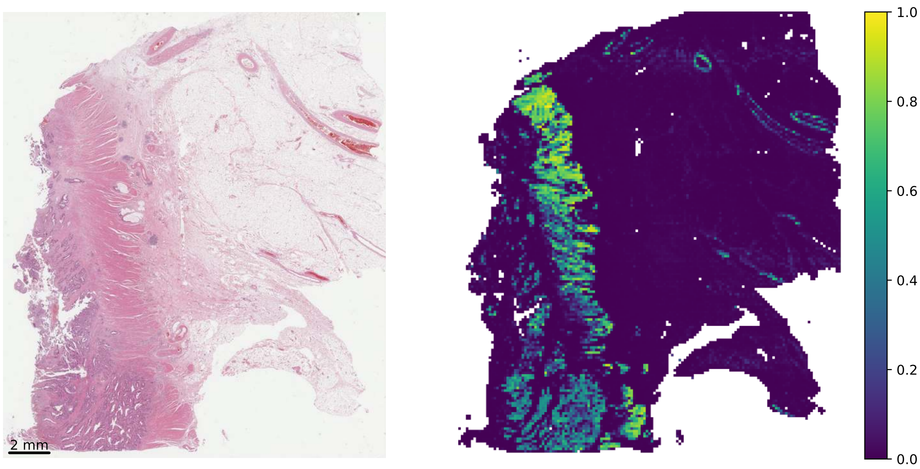

Figure 7. **COBRA Unsupervised Heatmap**. Patient: TCGA-EI-6508

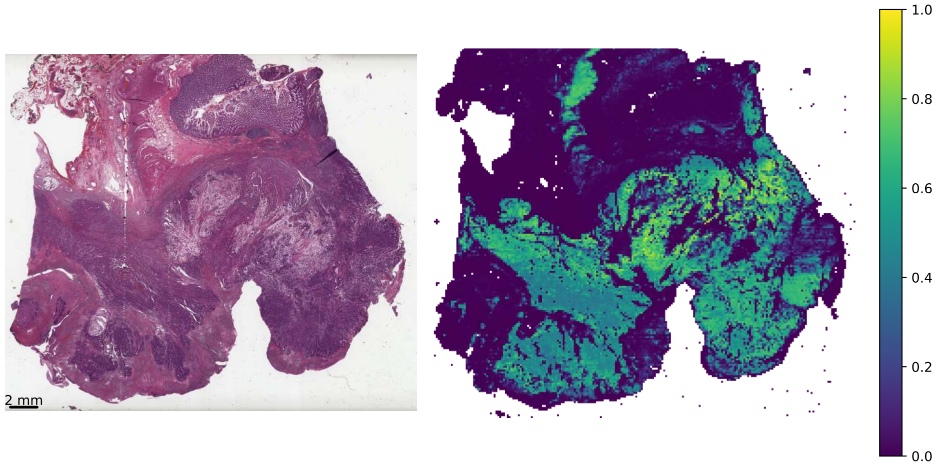

Figure 8. **COBRA Unsupervised Heatmap**. Patient: TCGA-CM-4743

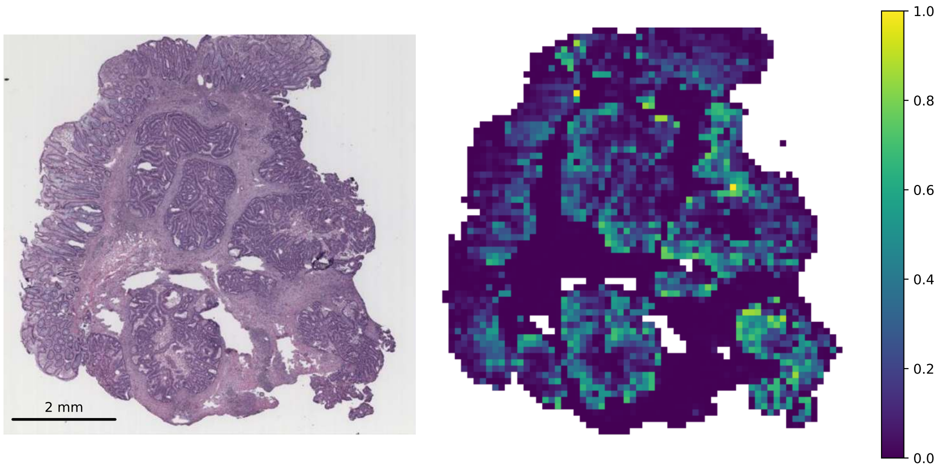

Figure 9. **COBRA Unsupervised Heatmap**. Patient: CPTAC-20CO007

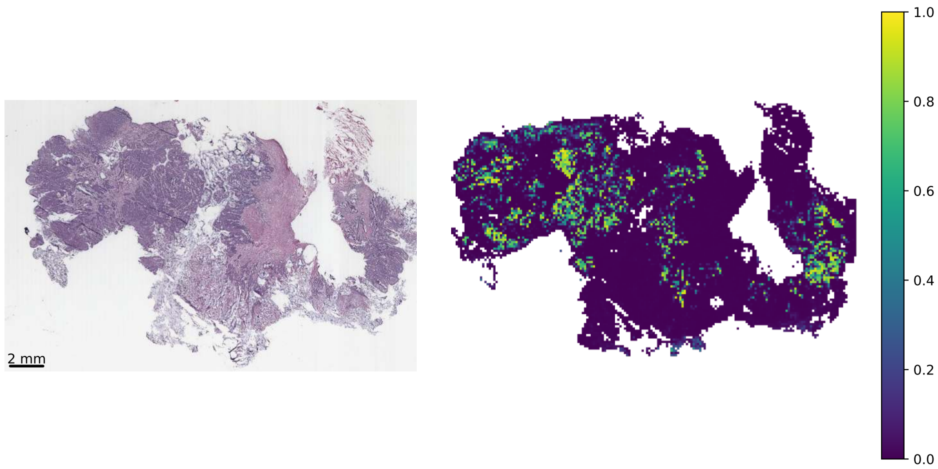

Figure 10. **COBRA Unsupervised Heatmap**. Patient: CPTAC-11CO062