# Task-Adaptive Neural Network Search with Meta-Contrastive Learning

Pyatkin Stanislav
MIPT

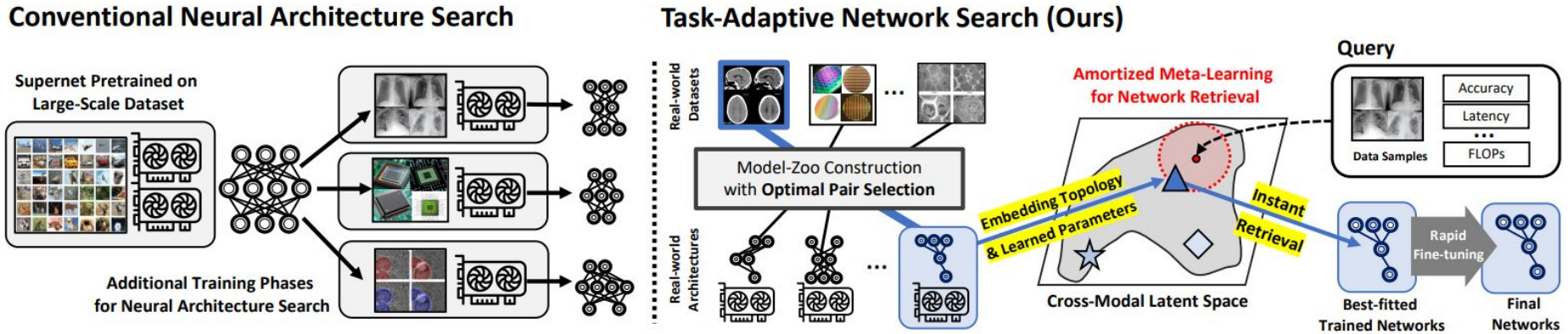



Figure 1: **Comparison between conventional NAS and our method**: Conventional supernet-based NAS approaches (Left) sample subnets from a fixed supernet trained on a single dataset. TANS (Right) can dynamically select the best-fitted neural networks that are trained on diverse datasets, adaptively for each query dataset.

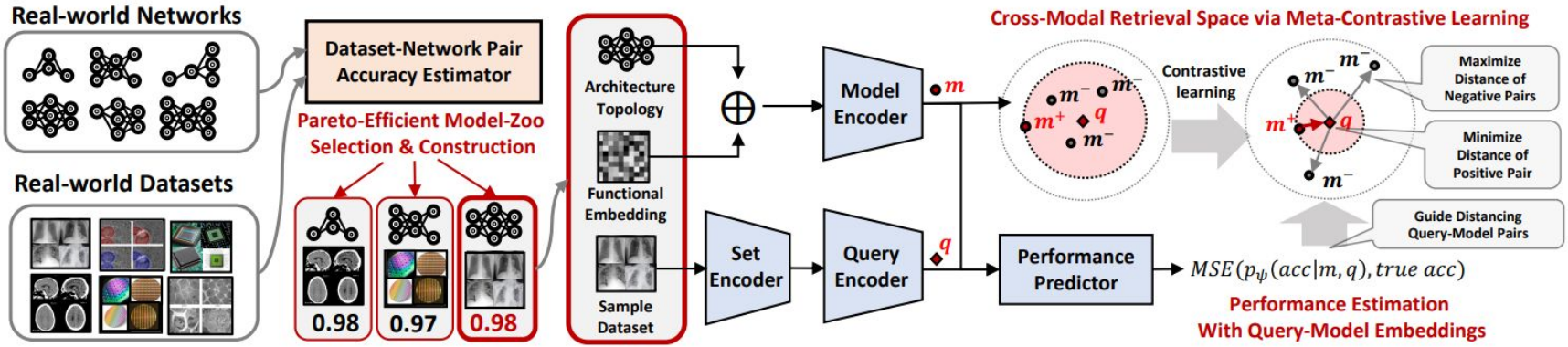



Figure 3: **Illustration for overall framework of our proposed method (TANS)**: We first construct our model-zoo with pareto-optimal pairs of dataset and network, rather than exhaustively train all possible pairs. We then embed a model and a dataset with a graph-functional model and a set encoder. After that, we meta-learn the cross-modal retrieval network over multiple model-query pairs, guided by our performance predictor.

# Problem statement

$$\tau = \{D^\tau, M^\tau, s^\tau\}$$

$$\boldsymbol{q} = E_Q(D^\tau; \boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{m} = E_M(M^\tau; \boldsymbol{\phi}),$$

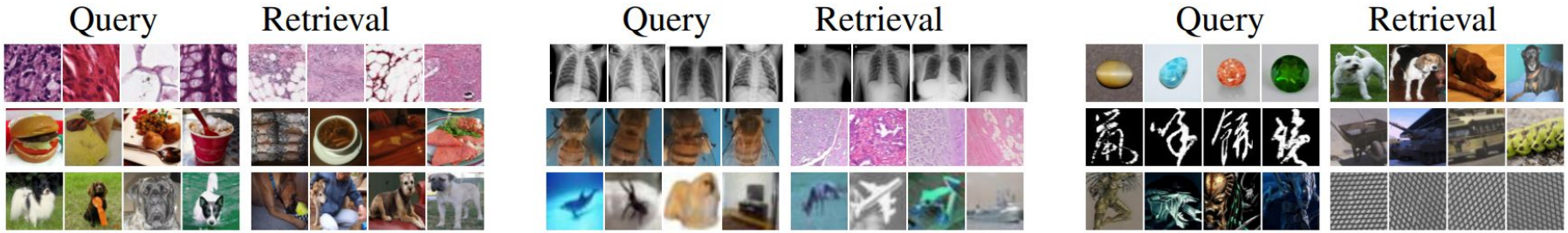
$$\mathcal{L}_m(\tau; \boldsymbol{\theta}, \boldsymbol{\phi}) = \max\left(0, \alpha - \log \frac{\exp(f_{sim}(\boldsymbol{q}, \boldsymbol{m}^+))}{\exp\left(\sum_{\gamma \in p(\tau), \gamma \neq \tau} f_{sim}(\boldsymbol{q}, \boldsymbol{m}^-)\right)}\right)$$

$$\min_{\boldsymbol{\phi},\boldsymbol{\theta},\boldsymbol{\psi}} \sum_{\tau \in p(\tau)} \mathcal{L}_m(\tau;\boldsymbol{\theta},\boldsymbol{\phi}) + \mathcal{L}_q(\tau;\boldsymbol{\theta},\boldsymbol{\phi}) + \lambda \cdot \mathcal{L}_s(\tau;\boldsymbol{\psi}),$$

$$\mathcal{L}_s(\tau;\boldsymbol{\psi}) = (s^{\tau}_{\text{acc}} - S(\tau;\boldsymbol{\psi}))^2$$

# Inference

$$M^* = \max_{M^\tau} \{ f_{sim}(\tilde{\boldsymbol{q}}, \boldsymbol{m}^\tau) \mid \tau \in p(\tau) \},$$



# Results

| Target Dataset | Method | # Epochs | FLOPs (M) | Params (M) | Search Time (GPU sec) | Training Time (GPU sec) | Speed Up | Accuracy (%) |
|---|---|---|---|---|---|---|---|---|
| **Averaged Performance** | MobileNetV3 [26] | 50 | 132.94 | 4.00 | - | 257.78±09.77 | 1.00× | 94.20±0.70 |
| | PC-DARTS [65] | 500 | 566.55 | **3.54** | 1100.37±22.20 | 5721.13±793.71 | 0.04× | 79.22±1.69 |
| | DrNAS [10] | 500 | 623.43 | 4.12 | 1501.75±43.92 | 5659.77±403.62 | 0.04× | 84.06±0.97 |
| | FBNet-A [60] | 50 | 246.69 | 4.3 | - | 293.42±57.45 | 0.88× | 93.00±1.95 |
| | OFA [8] | 50 | 148.76 | 6.74 | 121.90±0.00 | 226.58±03.13 | 0.74× | 93.89±0.84 |
| | MetaD2A [31] | 50 | 512.67 | 6.56 | 2.59±0.13 | 345.39±28.36 | 0.74× | 95.24±1.14 |
| | **TANS (Ours)** | 10 | 181.74 | 5.51 | 0.002±0.00 | 40.19±03.06 | - | 95.17±2.20 |
| | **TANS (Ours)** | 50 | 181.74 | 5.51 | **0.002**±0.00 | **200.93**±11.01 | **1.28**× | **96.28**±0.30 |
| Colorectal Histology Dataset (Easy) | MobileNetV3 [26] | 50 | 132.94 | **4.00** | - | 577.18±04.15 | 1.00× | 96.23±0.07 |
| | PC-DARTS [65] | 500 | 534.64 | 4.02 | 2062.42±49.14 | 12124.18±1051.16 | 0.04× | 96.17±0.68 |
| | DrNAS [10] | 500 | 614.23 | 4.12 | 4183.20±188.60 | 11355.18±1352.62 | 0.04× | 97.51±0.13 |
| | FBNet-A [60] | 50 | 215.45 | 4.3 | - | 696.00±295.19 | 0.83× | 95.43±0.57 |
| | OFA [8] | 50 | 134.85 | 6.74 | 121.90±0.00 | 537.61±03.52 | 0.88× | 96.40±0.52 |
| | MetaD2A [31] | 50 | 506.88 | 5.93 | 2.58±0.12 | 784.45±79.32 | 0.73× | 96.57±0.56 |
| | **TANS (Ours)** | 10 | 171.74 | 4.95 | 0.001±0.00 | 98.56±04.24 | - | 96.87±0.21 |
| | **TANS (Ours)** | 50 | 171.74 | 4.95 | **0.001**±0.00 | **492.81**±21.19 | **1.17**× | **97.67**±0.05 |
| Food Classification Dataset (Hard) | MobileNetV3 [26] | 50 | 132.94 | 4.00 | - | 235.57±07.57 | 1.00× | 87.52±0.78 |
| | PC-DARTS [65] | 500 | 567.85 | **3.62** | 1018.49±6.31 | 6323.40±938.83 | 0.03× | 55.42±2.46 |
| | DrNAS [10] | 500 | 632.67 | 4.12 | 1276.38±0.00 | 5079.89±161.05 | 0.04× | 61.45±0.68 |
| | FBNet-A [60] | 50 | 251.29 | 4.3 | - | 251.24±3.31 | 0.94× | 84.33±1.41 |
| | OFA [8] | 50 | 152.34 | 6.74 | 121.90±0.00 | **190.86**±03.48 | 0.75× | 87.43±0.59 |
| | MetaD2A [31] | 50 | 521.11 | 8.23 | 2.60±0.23 | 324.62±34.97 | 0.72× | 89.72±1.53 |
| | **TANS (Ours)** | 10 | 179.83 | 5.07 | 0.002±0.00 | 40.59±04.84 | - | 93.11±0.24 |
| | **TANS (Ours)** | 50 | 179.83 | 5.07 | **0.002**±0.00 | 202.93±24.21 | **1.16**× | **93.71**±0.24 |