ICLR 2026

Elastic Optimal Transport: Theory, Application, and Empirical Evaluation



Pei Yang[1] **Yuhang Zhuang**[1] **Qi Tan**[2]

yangpei@scut.edu.cn 202421044690@mail.scut.edu.cn tanqi@scnu.edu.cn

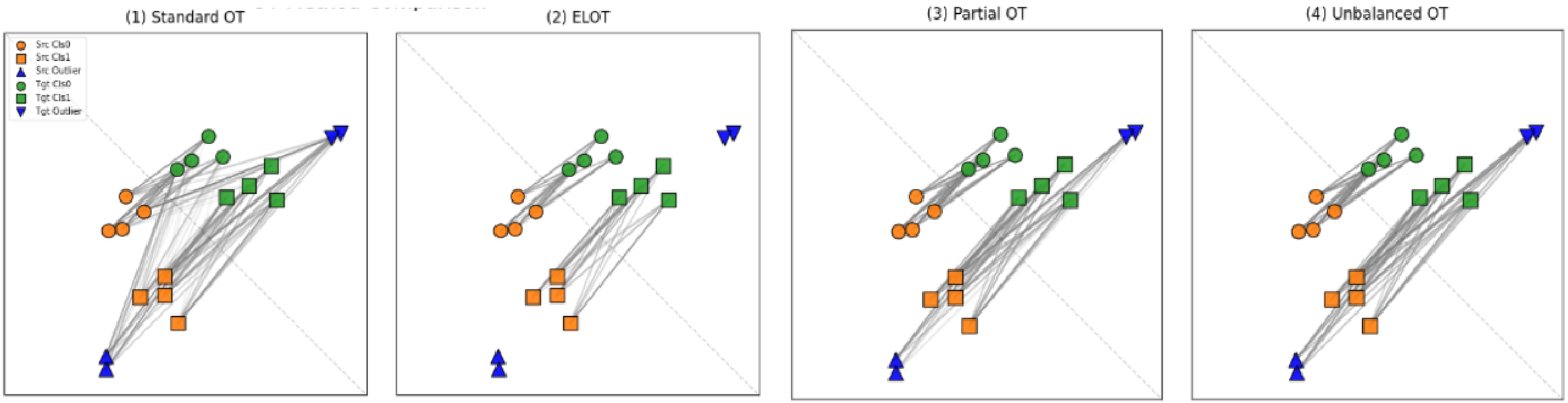
[1]South China University of Technology [2]South China Normal University

Outline

- Motivation
- Limitations of Classical Optimal Transport
- Elastic Optimal Transport
  - Theory
  - Applications (Domain Adaptation)
  - Evaluation
- Conclusion

Motivation

- A motivating example illustrating distinctive advantages of elastic optimal transport (ELOT) over classical optimal transport including Kantorovich OT, Partial OT, and Unbalanced OT.



Limitations of Classical Optimal Transport

- Kantorovich Optimal Transport
  - Full-mass constraint is likely to fit noise and outlier, or undesired pairs, and prevents partial matching
- Partial Optimal Transport
  - It is challenging to determine the fixed budget of mass
- Unbalanced Optimal Transport
  - It is usually unknown to what extent the ‘soft’ penalty should be imposed on the marginal constraints.

Elastic Optimal Transport: Formulation

- The distinctive characteristics of ELOT are twofold:
  - Adaptive-mass preserving
  - Mixed-sign ground cost matrix
- ELOT relies on the combination of mixed-sign cost and the marginal inequality constraints to achieve the goal of adaptive-mass preserving.

$$\mathcal{W}(\mu,\nu) = \min_{\gamma \in \Pi_e(\mu,\nu)} < \gamma, \mathcal{C} >$$

where $\mathcal{C} \in \mathbb{R}_{\pm}^{n \times m}$ and

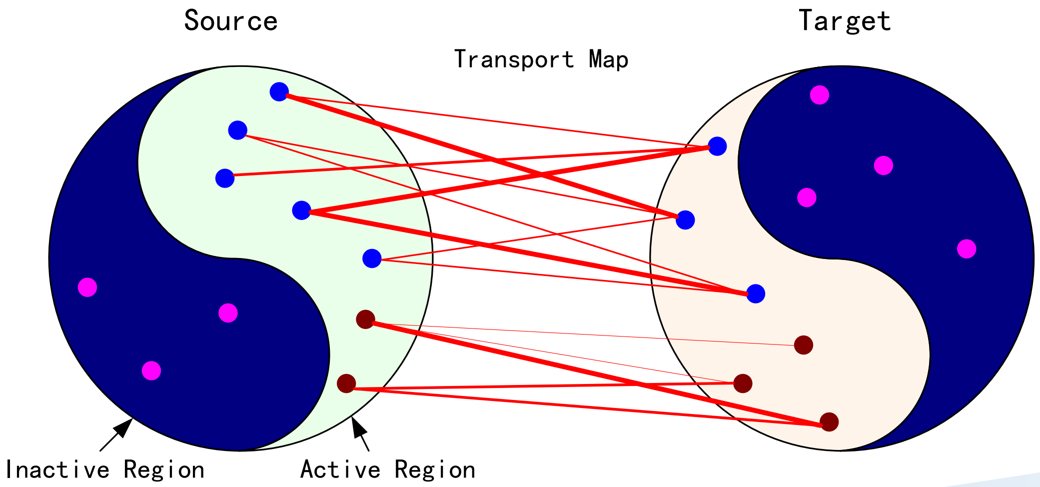
$$\Pi_e(\mu,\nu) = \{\gamma \in \mathbb{R}_{+}^{n \times m} \mid \gamma \mathbf{1}_m \leq \mu, \gamma^T \mathbf{1}_n \leq \nu\}.$$

ELOT: Novelty & Contribution

- A novel formulation of optimal transport, which is distinctive from POT/UOT in:
  - Adaptive-mass preserving
  - Self-determine mass according to task structure
  - Larger capacity by exploring the spectrum of mass
- The theoretical analysis sheds light on the mass transport mechanism of ELOT.
- A novel domain adaptation learning paradigm based on elastic optimal transport, corroborated on both unsupervised/partial domain adaptation tasks.

ELOT: Mass Transport Mechanism

- The source and target areas can be divided into active and inactive regions according the sign of ground cost.
- There is no mass transportation between the source and the target areas when their cost are positive.
- The mass allocation between the areas with negative cost acts like that of the constrained worst transport problem.



ELOT: Optimal Transport Solution

- Theorem 1 shows the equivalence between ELOT and its reformulation, which could be solved by the off-the-shelf optimal transport algorithms.

Theorem 1 (Optimal Transport Solution) *Let $\boldsymbol{\gamma}^*$ (or $\bar{\boldsymbol{\gamma}}^*$) and $\mathcal{W}(\mu, \nu)$ (or $\mathcal{W}(\bar{\mu}, \bar{\nu})$) be the optimal transport plan and the optimal transport distance for the elastic optimal transport problem (or the reformulation). Assume $\bar{\boldsymbol{\gamma}}_o^*$ is the matrix $\bar{\boldsymbol{\gamma}}^*$ with the last row and column removed.*
i) ***Optimal Transport Plan.*** *The optimal transport plan $\boldsymbol{\gamma}^*$ is equivalent to $\bar{\boldsymbol{\gamma}}_o^*$:*

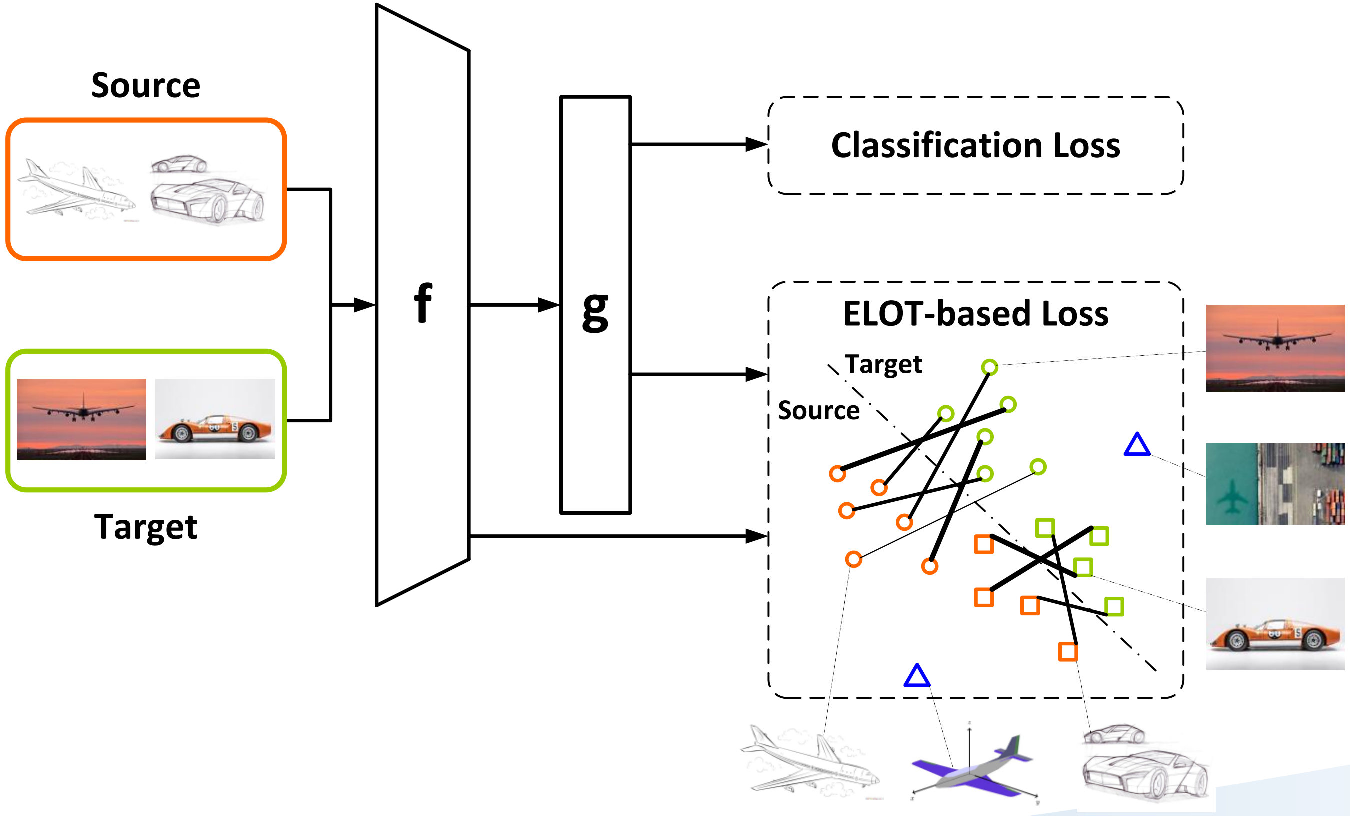
$$\bar{\boldsymbol{\gamma}}_o^* = \underset{\gamma \in \Pi_e(\mu,\nu)}{\arg\min} < \gamma, \mathcal{C} >= \boldsymbol{\gamma}^*. \tag{11}$$

ii) ***Optimal Transport Distance.*** *The difference of $\mathcal{W}(\mu, \nu)$ and $\mathcal{W}(\bar{\mu}, \bar{\nu})$ is a constant term which can be ignored in optimization (i.e., the optimization is not affected by the specific value of σ):*

$$\mathcal{W}(\bar{\mu}, \bar{\nu}) - \mathcal{W}(\mu, \nu) = \sigma\Big(\|\mu\|_1 + \|\nu\|_1\Big). \tag{12}$$

ELOT Application: Domain Adaptation

- The objective of ELOT is to minimize the elastic optimal transport distance between the source domain and target domain, as well as the empirical classification loss on the source domain.



Evaluation: Domain Adaptation

- We compare ELOT with various domain adaptation algorithms including OT-based methods such as ROT, DeepJDOT, JUMBOT, and m-POT on benchmark datasets such as VisDA (below), Office-Home, and Office-31.

Method	Accuracy
DANN$^{(*)}$	67.63±0.34
CDAN$^{(\#)}$	70.10
ALDA$^{(*)}$	71.22±0.12
ROT$^{(\#)}$	66.30
DeepJDOT$^{(\#)}$	68.00
JUMBOT$^{(\#)}$	72.50
m-POT$^{(*)}$	$\underline{73.59}$±0.15
ELOT (ours)	**76.32**±0.28

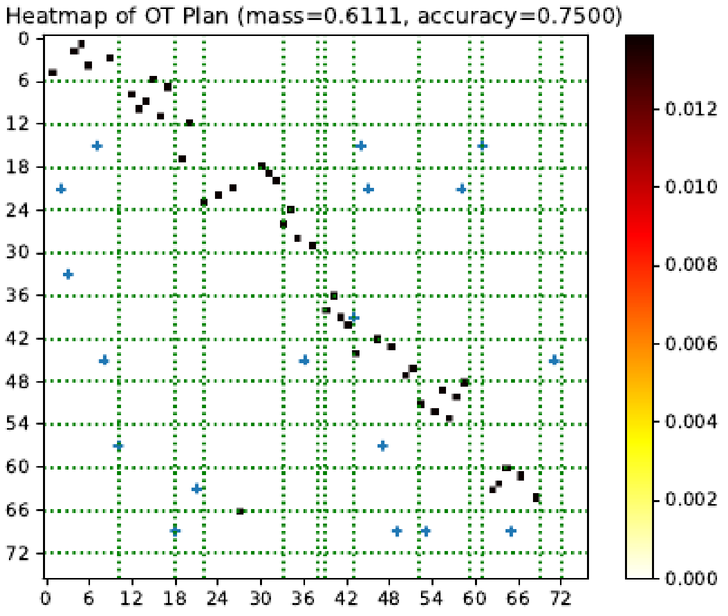
Evaluation: Partial Domain Adaptation

- Partial domain adaptation provides an ideal test bed for evaluating the robustness of the methods because the private classes in source domain could cause negative transfer.
- ELOT offers a more robust solution for partial domain alignment on benchmarks such as DomainNet (below) and Office-Home.

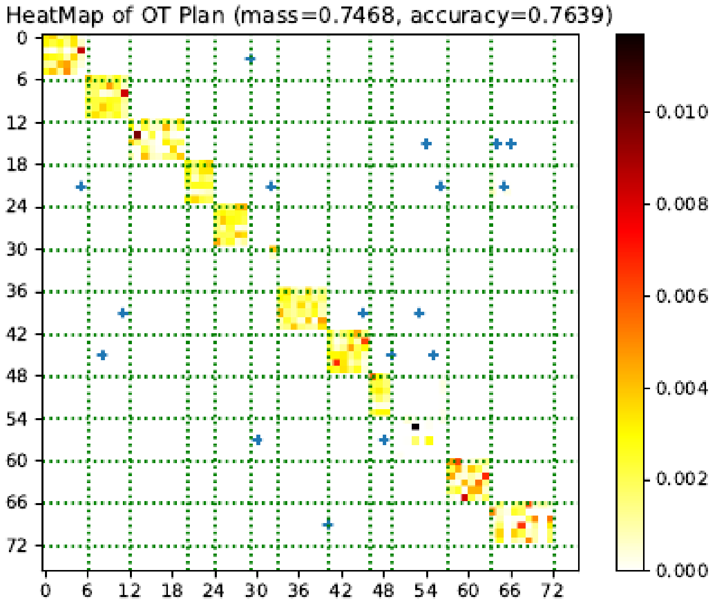
Method	C→P	C→R	C→S	S→C	S→R	S→P	Avg
ResNet-50	41.2	60.0	42.1	45.4	39.3	49.8	46.3
DANN	27.8	36.6	29.9	25.8	29.5	32.7	30.4
CDAN	37.5	48.3	46.6	35.4	38.5	43.6	41.7
PADA	22.5	32.9	30.0	17.5	23.9	26.9	25.6
BA3US	42.9	54.7	53.8	50.4	42.7	49.7	49.0
STCPDA	65.1	69.6	**69.6**	<u>64.4</u>	60.7	67.8	66.2
ARPM	<u>67.9</u>	**79.8**	66.3	62.5	<u>64.8</u>	<u>71.7</u>	<u>68.8</u>
ELOT (Ours)	**68.8**	<u>77.7</u>	<u>68.3</u>	**68.0**	**69.9**	**75.9**	**71.4**

Evaluation: Heatmap of Transport Plan

- The heatmap intuitively verified that ELOT transports masses between domains in an adaptive and class-aware way.



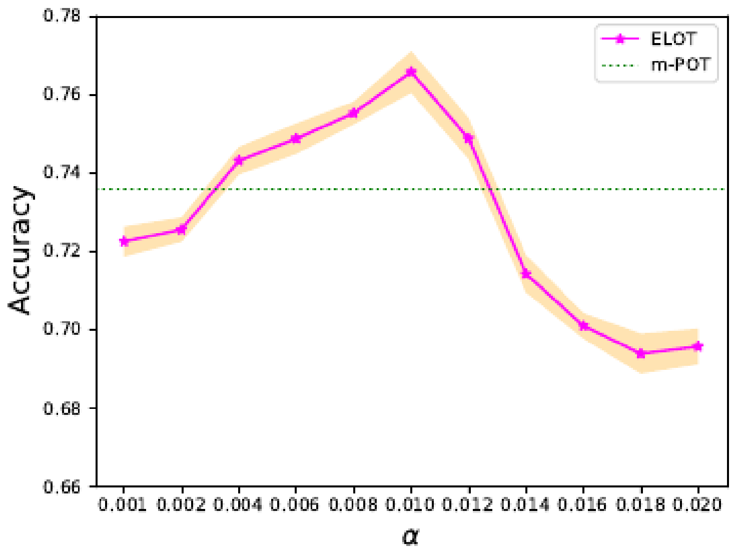
(a) $\epsilon = 0$

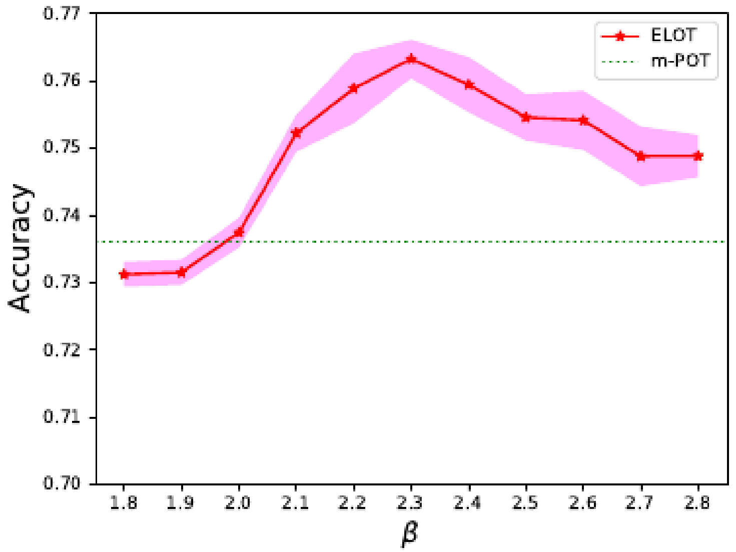


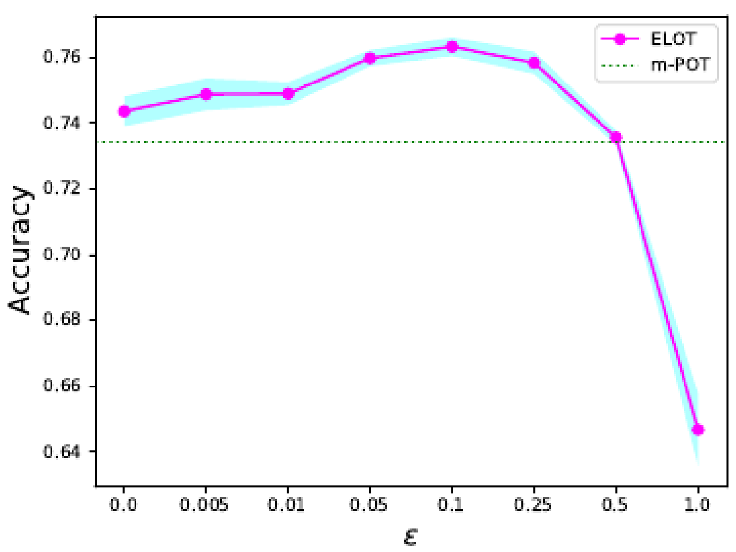
(b) $\epsilon = 0.1$

Evaluation: Parameter Sensitivity

- α adjusts the weight of the feature-wise cost.
- β adjusts the weight of the label-wise cost.
- ε is the entropy-regularized coefficient.







Conclusion

- A novel formulation of optimal transport along with the solid theoretical analysis to remedy the limitation of classical optimal transport.
- Elastic optimal transport leverages the geometry structure of the problem itself to transport adaptive-mass rather than full or fixed masses.
- It opens the pathway to unlock the problems of adaptive distribution matching in various areas of artificial intelligence.