IOWN
INTEGRAL
NTT R&D FORUM 2024

DEVELOPMENT
δ 01-02

Fast optimal transport between large-scale data

Improving the training efficacy of generative AI by fast computing relationships between large-scale data

#Improving customer experience #Improving operational efficiency

大規模データ間の高速な最適輸送技術

Fast Optimal Transport between Large-Scale Data

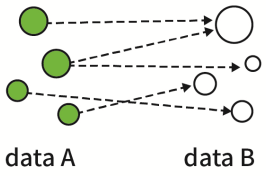
最適輸送技術とは？ What is Optimal Transport?

数値データ間の「最適な輸送経路」を計算
Calculate an optimal transportation plan between data

データ間の 類似度＋対応関係 が分かる
Obtain similarity and correspondence between data

計算コスト大 ⇒ 大規模データへ適用困難
Huge computational cost, so not for large-scale data

輸送経路
Transport plan



提案法 Proposed Method

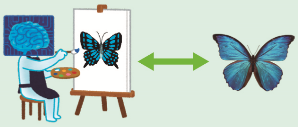
データに潜む巡回対称性を利用
Utilize cyclic symmetry hidden in data

計算コストの大幅削減を達成
Achieve to reduce the computational cost

巡回対称性を持つ画像データ
Image data with cyclic symmetry

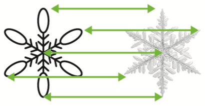
応用例 Applications

■ 生成AIの学習効率化 Training Generative AI



実世界データと
高速比較
Fast comparison
against real data

高速な
物体マッチング
Fast object matching



///Technical Issue

Existing optimal transport technologies compute relationships between data with high accuracy but require a huge computational cost.

///Research Goal

Reduces the computational cost of computing relationships between large-scale data required for training generative AI by 50%.

---Technology

In the optimal transport research field, we are the first to focus on cyclic symmetry hidden in data and to utilize it with various optimization method, which results in reducing the computational cost of optimal transport techniques.

---Applicable Business

In information and technology industry, our technology will improve the training of DL models including generative AI (Scheduled service launch: 4Q 2026).

---Novelty

- We are the first to propose a fast optimal transport technique using cyclic symmetry hidden in data.
- We are the first to show theoretically and experimentally that our technique can speed up to obtain the solution without reducing its quality compared to conventional techniques.

Exhibitors＝NIPPON TELEGRAPH AND TELEPHONE CORPORATION　Related URL＝[NTT Press Release [Development of technology that computes similarity and correspondence between large-scale data]]
[Contact URL]