

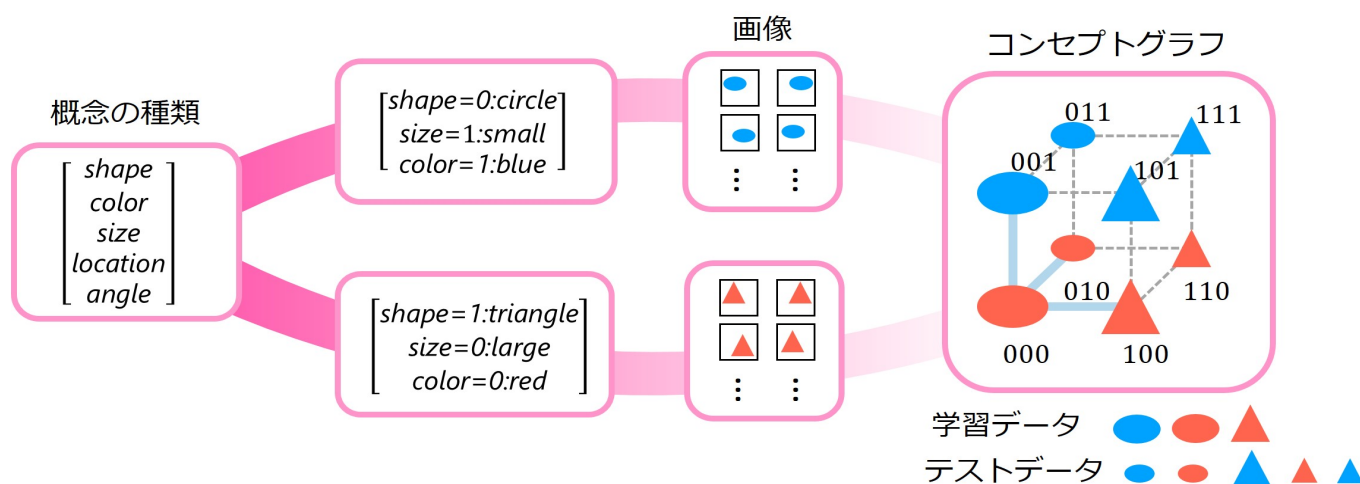


背景

生成系AIの性能が向上し、適用範囲が広がっています。しかし、その性能の詳しい理解はまだ不十分で、安心・安全な使用に課題が残っています。本研究では、生成系AIの代表例である拡散モデルを対象に、その性能を科学的に分析し、より深く理解することをめざします。

成果の概要

実問題の構造を反映させたシンプルな学習タスクを設計し、拡散モデルが「学習データ中の概念をどう学習し」、「学習した概念をどう組み合わせているか」、明らかにしました。こうした知見を活かしモデルの性能を制御することで、生成モデルの安全な活用を可能にします。



技術のポイント

- 画像生成における「想像力」を、学習データから概念を学び、その学んだ概念を組み合わせて新しい画像を生成する能力として定義
- 拡散モデルの「想像力」を評価するタスクを設計するために、「コンセプトグラフ」という理論フレームワークを提案
- コンセプトグラフに基づいて生成されたシンプルな学習タスクを拡散モデルに解かせることで、モデルの「想像力」を評価

この研究がもたらす未来

拡散モデルの性能を理解しその振る舞いを予測可能にすることで、より安全・安心な深層学習モデルの社会実装を可能にします。

コラボレーションパートナー

ハーバード大学、ミシガン大学

出展企業

NTT Research, Inc.

問い合わせ先

tech-promotion@ntt-research.com