

石川 博

早稲田大学理工学術院  
教授

認識の数理モデルと高階・多層確率場による高次元実データ解析

## §1. 研究成果の概要

深層学習の性能にネットワークの構造が大きな役割を果たすことが、今年度発表した画像復元ネットワーク自動設計の研究によって、一層明確になった<sup>1</sup>。提案方法は、畳み込みオートエンコーダ各層の詳細設計のパラメータを進化計算で自動最適化する方法であり、構造自動設計を物体認識以外の問題に適用した世界で最初の例である。この研究により、基本構造こそ標準的なネットワークであっても、詳細設計を自動最適化すれば、既存のいかなるネットワーク構造や手法よりも上回る高い性能が達成できることが明らかになった。

また、画像 1 枚に写るシーンの意味を理解する「画像理解」の方法実現へ向けて、画像と言語という異なるモダリティの表現を効果的に融合する「Dense Co-attention (密で双方向な注意) ネットワーク」を発表した<sup>2</sup>。1 枚の画像と、そこに映るシーンに関する自然言語の質問文(英語)のペアが与えられ、その質問に適切に答えるタスク(visual question answering)に適用し、最もメジャーなデータセットである VQA および VQA 2.0 において世界最高の精度を達成した。比較的単純な構造でありながら、基本構造を積み重ね多層構造を作ることが可能であり、将来的に、画像と言語の多様なマルチモーダルタスクへの発展が可能である。

応用面では、深層学習を応用して対話的にラフスケッチのペン入れができるツール、スマートインカー(Smart Inker)を提案した<sup>3</sup>。スマートインカーは、途切れた線を自然につなぎ、不要な線を効率的に消すことが可能な“スマート”ツール機能を持ち、自動出力された線画を効果的に修正することができる。実際に提案ツールを用いてラフスケッチにペン入れをするユーザテストを行った結果、商用のイラスト制作ソフトに比べ、提案ツールは簡単かつ短時間で線画作成が可能となり、イラスト作成経験がほとんどないユーザでもきれいな線画を作成できることが確かめられた。



### 【代表的な原著論文】

1. M. Suganuma, M. Ozay, T. Okatani, “Exploiting the Potential of Standard Convolutional Autoencoders for Image Restoration by Evolutionary Search,” ICML2018, PMLR 80:4771-4780.
2. D-K. Nguyen, T. Okatani, “Improved Fusion of Visual and Language Representations by Dense Symmetric Co-Attention for Visual Question Answering,” CVPR2018, pp.6087-6096.
3. E. Simo-Serra, S. Iizuka, H. Ishikawa, “Real-Time Data-Driven Interactive Rough Sketch Inking,” ACM Trans. on Graphics (Proc. of SIGGRAPH), 37, 4, 98:1-98:14, 2018.

## §2. 研究実施体制

### (1) 石川グループ

- ① 研究代表者: 石川 博 (早稲田大学理工学術院 教授)
- ② 研究項目
  - ・ トップダウンモデルと確率場生成
  - ・ 高階確率場の医用画像への応用
  - ・ ディープラーニングの応用

### (2) 岡谷グループ

- ① 主たる共同研究者: 岡谷 貴之 (東北大学大学院情報科学研究科 教授)
- ② 研究項目
  - ・ 高効率学習最適化法構築
  - ・ 動画認識への応用
  - ・ 多層モデル理論構築