

# 画像認識革命

## 編集にあたって

佐藤真一 (国立情報学研究所)

永野秀尚 (日本電信電話(株))

高速ネットワーク、動画共有サービス、ソーシャルネットワークサービス、スマートフォンなどの発展と普及に伴い、画像・映像などの視覚的メディアデータの量は爆発的に増加している。ただ、Webでの情報検索などが日常となっているテキストデータと比べ、視覚的メディアデータの場合、そもそも計算機でその内容を解析・認識することが難しく、大量のメディアデータを有効に活用できているとは言いがたかった。しかし、ここ数年、ディープラーニングに代表される手法により、画像認識の性能は飛躍的に向上し、まさに「画像認識革命」が起こっている。

ただ、画像認識の研究の歴史を振り返ってみると「革命」は今回が初めてではないことが分かる。1970年代にはじまった画像認識研究は当初考えられていたよりも困難な研究テーマであり、40年以

上を経過した現在も未解決な課題を多く抱えている。しかし、この40年以上の間、まったくブレークスルーがなかったわけではなく、その時々新しい技術が発明され、まさに「革命」を繰り返すことで画像認識技術は進歩してきた。先のディープラーニングを用いた画像認識も、1980年代に「革命」を起こした多層ニューラルネットワークを祖に持つ。その後、1990年代にサポートベクタマシンが「革命」を起こしニューラルネットワークは衰退する。その後も、局所画像特徴量 SIFT や Bag of Visual Words などの「革命」があり、画像認識技術は進歩してきた。そして、現在、ディープラーニングが、再び「革命」を起こしている。

本小特集では、画像の認識と理解について、まず1.で、これまでの研究開発の歴史を振り返るとともに明らかになってきた問題点を改めて捉え直し、それらに対する近年の有望な手法と展望について述べる。そして2.で近年の大きな革命であるディープラーニングを用いた画像認識について解説する。また、3.では「画像識別と画像復元」と題し、画像を入力としてその画像のカテゴリを計算機で予測する「画像識別」とその逆の手順をたどる「画像復



元」について解説する。これは、計算機による画像認識において、その内部で起こっていることを画像として復元し人間に理解させようという試みである。1.では「革命」の歴史を見ることができるとともに、当時は計算機パワーや関連技術の問題で結実しなかったが、現在では再び可能性を持つと考えられるアプローチなども紹介される。まさに第二のディープラーニングのように「革命」を起こす技術があるかもしれない。また、1.では、画像認識は人間の画像認識過程を計算機で実現することであるが、そもそも人間自身にも人間による画像認識がどのように行われているか分からないことが問題であることも述べられる。そこで、2.ではディープラーニングを用いた画像認識で用いられる畳込みニューラルネットワーク（CNN）と人間の脳との関係についても述べられる。CNNと人間の画像認識にかかわる視覚機能が非常に類似していることが述べられており、大変に興味深い。また、3.では、このようなCNNによる画像識別過程を可視化し、CNNの中でどのような処理が行われているかを見ることができ。ぜひ、1.から3.まで通して読んでいただきたい。

4.はソーシャルネットワークを用いた画像の認

識と理解に関する手法の解説である。ソーシャルネットワーク・ソーシャルメディアはまったく新しいメディアであり、そこで流通する大量の画像とそれらに付随するさまざまなデータ、そしてそれらの関係性を利用することで、1枚の画像それだけを対象としているときには決して得られなかったような情報が得られるようになる。極端な場合、画像そのものを見なくても、その画像の所有者の情報やその画像がどのように誰とシェアされているかなどの情報から、画像の内容を推察することも可能になり得る。まさに画像の認識と理解に「革命」を起こすかもしれない。

以上のように、本小特集では画像の認識と理解に関するこれまでとそしてこれまでとこれからの革命の一部を紹介する。いずれも興味深く、刺激的な内容になったかと思う。もちろん画像の認識と理解はさまざまな技術が必要であり、本小特集はその一部を紹介したに過ぎないが、この小特集を読むことで、このエキサイティングな分野に少しでも興味を持っていただけると幸いである。

(2015年5月13日)