



Vol. 96

CONTENTS

- 【コラム】説明できる人工知能：自分の知能の働きを説明できるのか？… 佐藤 真一
【解説】グローバル PBL —イノベーション創出と持続可能な開発のための人材育成— … 井上 雅裕
【解説】Computational Thinking は大学入試を変え得るか… 萩原 兼一

COLUMN

説明できる人工知能： 自分の知能の働きを説明できるのか？



顔の検出は、画像解析の重要な課題として古くから研究されている。1970年代、計算機自体の黎明期にすでに検討が開始されている。当時のアプローチは、我々人間はどうやってほかのものから顔を見分けているか、を考え、それに基づいて顔検出プログラムを作成していた。たとえば、まず目を見つけ、次いでその位置関係から鼻と口の位置を推定し、などとプログラムを作成していた。しかし、このアプローチでは高い性能は達成できなかった。すぐに想定に合わない事例（顔）が見つかり、それに合うようにプログラムを修正すると、ほかの場合にうまく動かない、というようなことになってしまったのである。考えてみると、そもそも我々自身どうやって顔を見分けているかうまく説明できるかという、実は疑問である（ぜひ試してみてください）。

1990年代後半になり、機械学習を用いた顔検出技術が提案され、きわめて実用的な性能を収め、大成功する（今スマホなどに搭載されている顔検出技術の基本的な形ができあがった）。どのように顔を検出するかは人間（開発者）は考えず、事例の顔画像をひたすら集めて機械学習システムに与え、具体的な詳細な手法（パラメータなど）は機械学習アルゴリズム、すなわち計算機に自動的に決めさせた。説明をあきらめることで、実用的な性能を達成したのである。

この学習データに基づく機械学習を利用したアプローチは、深層学習の後押しもあり、現在さまざまな分野で大成功を収めている。将棋や囲碁ではもはや人間は計算機に太刀打ちできないし、画像認識（与えられた画像を1,000クラスのカテゴリに分類するタスクなど：興味のある人はImageNet Large Scale Visual Recognition Challenge (ILSVRC) を調べてみてください）でも人間以上の精度を達成し、自動運転技術にも応用が進んでいる。その一方、こうした手法は、予期しない状況で予期しない振舞いを行うことがある点が問題になってきた。将棋プログラムが人間では思いもよらなかった好手を考えだし（これはいい点）、Google フォトがユーザの写真に不適切なラベル付けをしたり、自動運転車が人間では考えられないような誤りにより事故を起こしたりするのがその例である。

こうした問題に対応するため、現在、説明できる人工知能技術の研究が盛んに行われている。特に想定とは異なる振舞いをした場合に、なぜそのように判断したかを説明させ、適切に対応しようという考え方である。考えてみると、少なくとも画像認識については、大体人間がどのように画像を認識しているか説明できないし、機械も説明をあきらめることで高性能を手に入れたという経緯がある。そもそも説明できないのが知能の本質かもしれない。挑戦は続いている。

佐藤 真一(国立情報学研究所)

LOGOTYPE DESIGN...Megumi Nakata, ILLUSTRATION&PAGE LAYOUT DESIGN...Miyu Kuno