

座談会

画像認識応用における ディープラーニングのインパクト

情報通信研究機構 石寺永記氏,
富士通研究所 石原正樹氏, NEC 岩元浩太氏,
日立製作所 佐川浩彦氏, 日本IBM 中野宏毅氏
司会 佐藤敦 (NEC), 福島俊一 (JST)

近年、ディープラーニングが大きなブームとなり、画像認識の研究は大きな転機を迎えている。学習データとツールさえあれば、専門家でなくても、高精度な画像認識を実現できるようになってきた。このような「民主化」によって、画像認識の専門家は、今後何を狙ったらよいのだろうか。実問題の解決に向けたディープラーニングとの付き合い方に、議論は大いに盛り上がった。

佐藤 本日はお忙しい中お集まりいただきまして、ありがとうございます。今回の画像認識に関する特集では、現場目線の興味深い論文が集まったと思います。執筆者の皆様には感謝しております。

本日は招待論文の執筆者にお集まりいただき、「画像認識応用におけるディープラーニングのインパクト」をテーマにお話を伺いたいと思います。論文を読ませていただき、必ずしもディープラーニングを扱っていない方がおられるのは存じておりますが、このブームを皆様がどのように捉えているのか、企業の代表としてではなく、一研究者、一エンジニアとしての忌憚のないご意見をいただければと思います。

■ ディープラーニングの学習には大量データが必要

佐藤 20～30年前のニューロブームもすごいブームでしたが、結局、あまり使われませんでした。それに比べて、今回のブームは本物ではないかという意見が多く聞かれます。その一方で、技術的には以前とさほど変わらないという意見もあります。ディープラーニングを実際に使っておられる石寺さんや中野さんに、その辺の感触をまずお伺いしたいのですが。

石寺 何から何までニューラルネットでやってしまえばよいという考えは、少し危険だと思います。ディープラーニングは、良質な学習データが大量にないと良い性能が出ません。良質な学習データを大量に準備するには、非常にコストがかかるのですが、その議論がすっぱり抜けているように私は感じています。最近ではアノテーショ

ン付き画像DBがいろいろ公開されてますが、実際に応用システムを作るときは、それ専用の新しいDBを作る必要がでできます。そのDB構築にかかるコストの問題を解決しないと、ディープラーニングが広く使われるようになるには、まだ少し時間がかかるのではないかと思います。

佐藤 石寺さんの論文では、群衆の人数カウントでディープラーニングを使っていますね。学習に使う群衆の画像は、人物画像を切り貼りすることで人工的に作っている。そうやって作った数十万枚の大量データを学習させているから、性能が出ているのですね。

石寺 そのとおりです。碁のプログラムで、AI同士に碁の対決をさせて学習するのも、似たような話です。例題をどんどん人工的に与えて学習すれば、性能が良くなるのは事実ですが、本当に必要な例題がシミュレーションで生成できているかどうか、注意する必要があると思います。

中野 私は30年ほど前のブームのときにニューラルネットをちょっと触って、手書き数字認識をやっていたのですが、限界を感じていました。三層ぐらいだったら



石寺永記氏

1991年慶應義塾大学理工学研究科電気工学専攻修士課程修了。同年、日本電気(株)入社。以来、画像処理・パターン認識に関する研究開発に従事。2001年Essex大学研究員。2009年～2016年(株)NEC情報システムズマネージャを経て、現在、国立研究開発法人情報通信研究機構専門研究員。

よいのですが、層を増やすと誤差が伝搬してこないとか、いろいろ問題がありました。最近またブームになったということもあり、診断支援に使えないかということでやっていますが、30年前と比べてやはり格段の進歩があると感じています。

私が対象としているのはX線CTの画像でして、たとえば、肺腫瘍の診断支援ですと、アノテーション付きの画像は整備されているので、それを使ってネットワークをトレーニングすることができます。IBMとしても、大量のメディカルデータを保有しており、それにアノテーションを付けることにもかなり投資をしています。肺腫瘍は、素人では腫瘍かそうでないか分からないような画像なのですが、何十万画像というアノテーション付きの画像でトレーニングをすると、かなり高い精度が得られます。しかし、画像の99%は正常、1%弱が腫瘍というアンバランスな問題であり、工場の欠陥検査もそうですけれども、欠陥は非常に少ないわけです。そのようなアンバランスなデータからいかに高い認識率を得るかということは、まだまだ工夫の余地があるのではないかと考えています。

福島 アノテーションを付けるところに課題があるという話ですけれども、最近、クラウドソーシングとかでタグ付けするというのが結構ありますよね。私はもともと自然言語処理をやっていましたが、その手の作業を、昔はデータ作成者として専門性がある人に頼んでやってもらったのが、最近はクラウドソーシングを使っています。たくさんの人にやってもらって平均をとったり、統計的に処理して質の良いものを選んだりすることで結構使えるデータが作れて、コストは昔と比べるともう一桁、二桁、安くはできるという話を聞きます。画像系はそうでもないのでしょうか。

佐藤 ImageNetという一般物体認識用の画像データがありますよね。数千万枚の大規模データですが、クラウドソーシングでアノテーションを付けています。公開されているアノテーション付き大規模データは、それほど多くはないと思います。企業でやっているかもしれないですが、公にはしないですよ。そのようなデータを持っていること自体が、その企業の強みになるので。

福島 持っていること自体は強みですが、テキスト系だと、評価に使うデータやトレーニングに使うデータの作り方を発表しているところも目に付きます。クラウドソーシングをして、生データの質が悪くても、その結果をもとにノイズとかをうまく削りながら、ちゃんと学習に使えるデータにするためにどんな工夫をしているかも、



中野宏毅氏

日本アイ・ビー・エム（株）東京
ラボラトリー・ワトソンヘルス開
発所属。1983年大阪大学基礎工学
部卒業。同年、日本アイ・ビー・
エム（株）入社。画像認識技術お
よびマルチモーダル・マイニング
の研究開発に従事。2015年より現
職。学術博士、電子情報通信学会
および精密工学会会員

ポイントになっていると思います。

石寺 研究のポイントが、学習のアルゴリズムとか特徴
量設計ではなくて、いかにデータを集めるかとか、アノ
テーションを付けるかとか、そのような課題にシフトす
るかも知れません。学習データさえ準備できれば、あと
はディープラーニングに突っ込めばいいみたいな、そう
いう世界になってきていて、研究者がやるべき研究もだ
んだん変わるのではないかというのが、私が感じている
ところです。

石原 私のお付き合いしている大学の先生も画像へのア
ノテーション付けにアマゾン・メカニカル・タークとい
うクラウドソーシングのサービスを活用されています。
海外の労働力を活用することで、低コストで大量に集め
られるけど、問題となるのはアノテーションの質だそう
です。石寺さんがおっしゃっていたように、この先生も学
習に使うデータの質をいかに上げていくかという研究テー
マへ移りつつあるようなことをおっしゃっていました。

福島 テキスト系だと、日本語のテキストは日本人に頼
みますが、日本人は几帳面なので、平均的には質がよい
そうです。画像だと国に関係なく頼める分、質がいろい
ろばらつくのでしょうか。

石原 はい。先ほどご紹介した先生が工夫されていたのは、
あるサンプル画像に対しては、こういったアノテーションを
付けてほしいという正解を依頼者側で持っていて、その正
解通りに正確なアノテーション付けができた人だけにフィル
タリングしています。データの質を担保するには、そういった
仕掛けを裏側に入れて、同じような見方をしてくれる人
だけを絞り込む工夫が必要なようです。

岩元 先ほど石寺さんから話がありましたが、そもそも
画像がない場合が結構多いような気がしています。私が
最近お付き合いしている、小売業界、物流、製造などの
お客さんが扱っている認識対象の画像を、現場で大量に
撮影するのは難しいですし、シミュレーションも結構難
しいので、まだディープラーニングは使えていません。
現実的には、実世界で良質なデータを集めるのは難しい

かなと感じています。

石原 そのデータに関してなのですけども、医療分野で少し特徴的なのは、結構データがたくさんあることです。日本はCTとMRIの設置台数でいうと世界一位でして、医用画像を世界で最も多く保有する医用画像大国でもあります。それこそ毎日大量に画像データが発生していて、病院内には大量に蓄積されているのですが、それがディープラーニングの学習データに使えるかというと、使えません。国内で公開している医用画像のデータベースは、実はわずかしかな。どちらかというと米国のほうが多く公開している状況です。国内では、データ漏えいのリスクだったり、倫理上の配慮から、なかなか簡単には施設外にデータを出せない。出せないからデータが使えない、あるのに使えないという状況が医療分野ではあるのかなと。

福島 それはやはり個人情報と紐づいたら大変だということですか。

石原 そうですね。医療情報というのは究極の個人情報なので、取り扱いには非常に慎重にならざるを得ないのが現実問題としてあります。

佐川 岩元さんの内容とも重なりますが、現場を監視するだけなら、それほど細かいところまで見える映像は必要ないかもしれませんが、画像から良否を判定したりとか、製品の状態を判定しようとすると、ある程度良いアングルから撮った画像が必要になります。そのような現場の画像は、非常に少ないのが現状だと思います。また、現場で撮影すると、画像に秘密情報が含まれる場合もあるので、画像を提供していただきがたい場合もあるかと思っています。そういう意味では、産業系もなかなかデータを集めるのが難しい部分があるのかなと推察しています。

佐藤 ディープラーニングを使う場合、学習に使うデータにやはり議論が集中するようですね。どうデータを集めるか、どうアノテーションを付けるか、コストの問題もあるし、法律や倫理的な問題もある。異常データに至っては、データそのものが少ないという問題がある、というところですね。

■ 出力結果が説明できないことに課題も

佐川 私はディープラーニングに直接かかわってはいないのでですけども、巨大なネットワークになるので、どのくらいの数のレイヤーを使えばよいとか、ノード数をどのくらいにすればよいとか、どのような特徴抽出をどんな形で入れていけばよいのか、といったネットワークの設計も苦労しているようなお話を聞きます。ネット

ワークの設計には職人技が必要とも聞いています。そういうところも、研究のネタの1つになっていくのかなという感じがしています。

岩元 結局、理論的になぜよいとか、どういう設計がよいとか何も分かっていないので、とりあえずいろいろやってみる。GPUがあるので、とりあえずパラメータを振って回して、ああ良い結果が出たというのが、今のやり方なのですよ（笑）。データをディープラーニングにただ突っ込んで、今までよりも良い結果が出ましたみたいな（笑）。でもそれは、どの部分が研究なのだろうとか、分からなくなってくる。

佐川 なぜできるのかよく分からないというのが、アルゴリズムを駆使して認識をやっている技術屋としては、不安に思うところがあるかなと思います。

石寺 そこはすごく大きいと思っています。僕も前回のニューロブームの生き残りですけども（笑）、前回のブームがなぜ終わったかという、そこに深く関係していると思います。前回のニューロブームでは、そこそこ性能が良かったのに導入が嫌われた理由がいくつかあります。1つは、未知データに対して時々変な挙動をすることです。それが予想できないから怖くて使えないという意見が当時もあり、導入が見送られたケースが多くあったように思っています。もう1つは、メンテナンス性の悪さです。僕もOCRをやっていたのですが、この文字が読めないから、読めるようにしてくれというお客さんがとても多い。それをニューラルネットで学習し直して、本当に読めるようになるかは、やってみないと分からない。現場からは、それではお客さんのところに持っていけない、という声がすごくあったのですよね。結局、それでニューラルネットがあまり広がらなかったとみています。佐川さんがおっしゃった話は当時からあって、それにどう答えていくかというのが、すごく大切なのではないかと思います。

佐藤 未知データに対して変な挙動をするのは、ディープラーニングでも指摘されていますね。パンダの画像に



石原正樹氏

（株）富士通研究所メディア処理研究所研究員。専門は画像検索。2013年より医用画像解析の研究に従事。上級医療情報技師（日本医療情報学会認定）、医用画像情報専門技師。



佐川浩彦氏

1991年東京大学工学系研究科情報工学専攻修士課程修了。同年、(株)日立製作所入社。同社研究開発グループにて、手話認識・生成、対話システム、AR等に関する研究開発に従事。博士(工学)。

ちょっとノイズを加えると、99%の信頼度でテナガザルと答えてしまうとか。人間が見ると、どうみても同じパンダの画像に見えるのだけれども、テナガザルと答えてしまう。未知データに対する挙動の問題は、まだ残っていますね。

岩元 そうすると、自動運転なんてそんなの使いたくないですね。変なノイズをうまく与えると、全然違う結果を出せるように、ハッキングできるような論文を見たことがある。そうだとすると、自動運転システムにノイズを投げて、そのシステムを破壊するみたいな怖いことができるかもしれない。

福島 ディープラーニングのブラックボックス問題はよくいわれますね。必ずしも画像ではありませんが、ビジネス系の予測問題とかでお客さんに、なぜこういう予測をしたかを説明しないと受け付けてもらえないので、ディープラーニングではなく、ちゃんと理由の説明できる機械学習技術が導入してもらえという話です。それに対して画像認識は見れば合っているかどうか分かるので、ディープラーニングを受け入れやすいと思うのですが、そんなことはないということなのですかね。

佐藤 大体は良い結果を出すけれども、たまに突然変な答えを出すことがある。人を介さずに自動判断させるのは、ちょっと怖いですね。

中野 医療系でもそういう危険性がありますので、コンピュータが勝手に確定診断をつけるのは到底無理で、あくまでも診断支援という位置付けです。コンピュータが出す確信度を見て、読影医が最終的に判断するという使い方になります。実際に診療の現場で使っていただくためには、日本では薬事法、アメリカですとFDAを通さなければいけないのですが、確信度だけ提示するというのでは到底通らないような気がしています(笑)。こういう理由でこういう判断をしたのですというエビデンスに基づいた説明が必要になるのですが、そこはまだ解けていないと思っています。

石原 おっしゃるように、まだ医療機器プログラム認定

のハードルが非常に高いですね。2016年3月に厚生労働省から機械学習を取り込んだ医療機器プログラム認定に関する指針が出ましたが、さまざまな問題提起がされています。審査時に、臨床での有効性を示すことができれば、承認を得られることが示されていますが、承認後に再学習をした場合には性能が変わってしまうので再審査が必要とされています。この再審査時に再び有効性と安全性を示す必要があるのですが、果たして再学習のたびに、また数カ月～数年かかる臨床試験をやるのかというと、できませんよね。なので、臨床で運用しながら再学習で性能を徐々に上げていくことが実質できません。必然的に、医療機器ベンダ各社は、事前に想定されるデータをいかに幅広く大量に学習させて、承認前にどれだけ高い性能を獲得できるかが勝負になってくるといわれています。

佐藤 ニューラルネットというのは、基本的には入力と出力を対応づける非線形関数であって、パラメータを減らすために、畳み込み層を多く使った深いネットワークにしているけど、それでもパラメータ数は多いので、それなりの大量なデータを用意しないと、過学習になって変な振る舞いをする。しかも、再学習すると出力の傾向が変わってしまう。もちろん使える場合もあるけど、使い方が難しい場合もある、ということですね。

■ 特徴抽出が自動設計できるようになったが

佐藤 ちょっと話を変えると、今のニューラルネットが本物といわれる1つの理由は、特徴が学習できるようになったことにあると思います。従来は識別器だけを機械学習で設計して、特徴は一生懸命人間が考えて設計するやり方だった。そこに研究的な要素があったり、コンパクトで良い特徴を設計するのがやりがいだったりするのですけど、ディープラーニングによって特徴抽出も学習できるようになってきた。今回の論文を読ませていただき、特徴設計に関する内容が結構あったのですが、それがディープラーニングに置き換わるのか、置き換わるのは無理なのか、何かご意見はありますか。

石原 医用画像というのは、腫瘍など見ても素人目にはよく分からない、もやもやした特徴を捉える必要があります。工業製品と違って、かっちりとした理論で作りたいのかなと。そこでディープラーニングである程度特徴を設計できるのは大きなメリットというか、うまく価値がはまっているのかなと思っています。ただ、医師にその特徴を分かるように説明をしないといけなくて、その作り出した特徴量を我々が解釈して、ちゃんとかいいう場合には問題なく動きますという説明ができるよう

にすることが一番重要なこと。

中野 おっしゃるように、細胞とか生体は特徴量が作りにくいので、ディープラーニングによって特徴抽出が作れることで助かっているところはあります。ではどこが研究テーマになるかという、1つはやはり極端にアンバランスなデータから、いかに良いネットワークを作るかだと思いますし、そこをやっているところです。

岩元 私はこれまで特徴量の設計を研究してきて、それが面白かったのですが、ディープラーニングの方が性能的に良い結果が出てきているので、特徴抽出はもう全部学習でやってしまえみたいな流れになっているのかなと感じます。

先日の学会で聞いた話ですが、今まで画像認識は特徴の設計が研究のポイントだったけど、ディープラーニングがはやってきて、画像認識の研究者ではない素人でも、そのツールを使えばすごいシステムを簡単に作れるようになった。それを「画像認識の民主化」と言っていました(笑)。今はもうだれでも使えますと。そうすると、いかにそのツールを使って、良いネットワークアーキテクチャを作り出すかという、アーキテクチャ設計がポイントになって、特徴設計は勝負の土俵ではなくなってきているというのは感じますね。自分の仕事がなくなってしまっても(笑)。

石原 私もこの前びっくりしたのは、医学系の学会に行ったら、お医者さん自身がディープラーニングを使って、自分の病院で持っているデータを使って試してみましたという発表があったんです。MRI画像の撮影部位をディープラーニングで認識するという内容でしたが、もうそこまで時代が変わってしまったのかと衝撃を受けました(笑)。こうなってくると、もはや病院からデータを出す必要性もないし、アノテーションも医師自ら付けられるので、これは鬼に金棒といわざるを得ない。先ほどの岩元さんの民主化の話と同じく、今は誰でも簡単に使えるツールも出てきているので、ディープラーニングのコモディティ化も進んでいる印象です。

福島 でも逆にその現場の技術者が、認識技術を民主化された道具、コモディティ化した道具として使って、より良いものを作るためには、アーキテクチャや段取り設計が重要になっているということですね。そこに対して何かアドバイスがありますか。

石寺 特徴抽出に関しては、僕は騒ぎすぎではないかと思っています。前回のニューロブームのときに、特徴抽出はどうあるべきかという議論は結構されていて、完全系だったらなんだって良いという議論もあったほどです。

でもなぜ特徴抽出をみんな研究していたかという、処理を軽くしたかったからだと思うのです。組み込み用の小さなプロセッサでもそこそこうまくいく特徴を一生懸命作って、実用化させたという流れだったと僕は理解しています。そう考えると、ディープラーニングの特徴は、すごいラグジュアリーな特徴ですよ。ガボールのような特徴や色の組合せがたくさん出てきて、それが複雑に絡まった中間状態がどんどん出てきて、計算リソースが十分だったら確かによいのだけれども、GPUが必要なんて言ったら、そんな高いものは使えないというお客さんもいます。

佐川 最近、カメラを含めたセンサを現場にばらまいてデータを集めるIoTがはやっていますが、画像をそのままサーバに送るとすごいトラフィックが発生するので、現場である程度処理してからサーバに送るような話があります。そういうところは、非常にチープなものしか置けないと思いますので、従来のアルゴリズム的な方法で特徴抽出してから送るようになるのかなと思います。それで集めたデータを大量に処理するところで、リッチな環境でディープラーニングを使うという住み分けというか、使う状況で使う手法を変えていくのかなという気もしています。

岩元 私が今やっている研究はスマートフォンで画像認識することが多いので、いかに低コストで実現するかとか、クラウドと連携してどう効率的に処理するかとかが多いです。学会レベルでは、ディープラーニングが華々しく語られるから(笑)、お客さんも自分たちでツール使ったりして簡単にできるんでしょ、と思われることも多いですね。でも実際やってみると、リアルタイムで処理するのは、そんなに簡単ではない。そこが研究というか、腕の見せ所なのかなという気はします。

中野 そうですね。商品の識別などには実はディープラーニングは使えないのではないかと思っています。たとえば、本の表紙は非常に似通ったデザインが多いのですけれども、キーポイント法では識別できても、ディ-



岩元浩太氏

2003年、早稲田大学大学院理工学研究科電子・情報通信学専攻(修士課程)修了。同年、日本電気(株)に入社。現在、同社データサイエンス研究所主任研究員。画像・映像認識、画像・映像検索技術の研究開発および事業化、またISO/IEC 15938(MPEG-7)標準化のプロジェクトエディタとして国際標準化活動に従事。2012年、文部科学大臣表彰若手科学者賞、情報処理学会情報処理規格調査会標準化貢献賞を受賞。2013年、先端技術大賞経済産業大臣賞を受賞。2014年、精密工学会高城賞を受賞。

プラーニングで100万冊の本を学習できるとは思えない。用途によってはディープラーニングが使えない分野があるかなと思っています。

福島 ちょっと話が戻ってしまうのですが、先ほどのクラウドとエッジの話で、大量の学習データからディープラーニングで学習してモデルを作るところはクラウド側でやれて、今度はそのモデルによる識別はエッジで現場のデータでもやるみたいな形だとうまく分かりますが、そんな形で簡単にいくのですか。現場でまたどんどん適応的にやっていくと難しくなってきますか。要するに、クラウドとエッジとでのアーキテクチャ設計が重要だといったときに、クラウドとエッジでディープラーニングを使いこなすための作り方はどうなりますか。

石原 興味深い話ですね。以前見かけた事例ですが、監視カメラみたいなプライバシーを含んだような情報をクラウドに上げるとやはり抵抗があるので、エッジ側で、ディープラーニングの何層かあるうちの、中間の3層目ぐらいまで通しておく。そうすると、モザイクがかかったような画像になり、プライバシー情報はだいぶ落ちて、人間が見てもよく分からない(笑)。それをクラウドにアップして、残りの層をクラウドで処理して認識するという研究があった。それを聞いて、ああ確かになるほどなと思った。プライバシー保護のために、そういうアーキテクチャが使えるかなと。

佐藤 データ量が減るから、トラフィックの問題とかも解決されますよね。

石原 そうですね。たとえば病院内にはCT画像とか、レントゲン画像を院内に保管しておくPACSというシステムがあるのでありますが、今、問題になっているのが、データが大量にありすぎて困っていることです。果たして画像をそのまま取っておくのが良いことなのか、ちょっと今疑問に思い始めています。もちろん、法的な制約があったり、人が見るから一応そのまま生データを取っておくのですが、ディープラーニングが活躍する世界になってくると、もはや画像ではなく特徴量の形で十分なのかなと。特徴量になれば、データ量はだいぶ減っているので、システムはさくさく動くし、管理コストも下がる。そういう面では少しポジティブな見方を持っています。

石寺 そういう意味だと、今のディープラーニングと昔のニューラルネットの違いというのは、中間層の結果が使い回せるところかもしれないですね。たかだか3層や4層だと、せっかく学習したニューラルネットを、ちょっと違うものに適用しようとする、学習を全部や

り直すのが前提だったけど、ディープラーニングだとネットワークの前半は、とりあえず使い回してしまっても、そこそこいけますよなんていう報告がちらほら出ている、という意味では新しいのかもしれないですね。

佐藤 転移学習ですね。特徴抽出を大規模データで作っておいて、その特徴を使って違うタスクを認識するものですね。人間が同じ特徴抽出器でいろんなものを認識しているという、アナロジーかもしれないけれども(笑)。

石寺 前回のニューロブームでも、ガボールのような特徴が出てくることは、あったではないですか。今回のブームでは、グーグルの猫のように構造を持った模様や、よく分からない変な模様が中間層に出てきてますが、いろんな画像に対してある程度普遍性を持っているようだと、とても面白いと思います。そこまで圧縮できるなら、相当圧縮効率が良さそうですね。

石原 画像診断の世界だと、医師が腫瘍を見て、これは良性か悪性かと判断されてレポートを書くのですけれども、それを見るときいきなり良性とは書かずに、いろいろ理由を書くのですよね。たとえば、気管支に拡張が見られて、腫瘍の形状は扁平で中心部に空洞があり、これら所見を総合的に勘案すると、悪性腫瘍と疑われるなどと書かれています。このように医師は、細かい特徴をいくつか拾い上げて、それで最終判断をしている。もし同じように、途中で何を特徴として最終判断に至ったのか、そのプロセスを人が分かるように出力できると、ディープラーニングの使い勝手もかなり上がると思うのですけどね。

佐藤 ネットワークの中間層を可視化する研究はありますが、人間が理解するのはなかなか難しいですね。どういう特徴を捉えているのか、見てもよく分からなくて(笑)。それを説明できるようになれば良いのですが、認識結果が良ければ、理由は説明できなくてもいいという人もいます。人間だって、なぜそう認識できるのか、説明しにくいことも多いからでしょう。

岩元 先日、ある学会で、ディープラーニングと人間が考えるロジックみたいなものを融合させるのが、今後重要になるという議論がありました。ディープラーニングは良い結果を出すけれども、完全にブラックボックスなので、結果が説明できないと活用が難しい面もある。ディープラーニングとロジックを融合させる研究が、今後増えていくのではないかと。

佐藤 そうですね、私もそれはよく思っているところです。結局、ディープラーニングというのは、ボトムアップで信号が伝わって最後に出力されるだけですね

(笑). 人間には、よく分からないときはもう一度確認して、詳しく分析するみたいなトップダウン処理があるではないですか. そういう処理がディープラーニングにはまったく入っていないので、認識結果の良し悪しが判断できない. そのような処理が入ってくると、より人間らしく判断できるようになるかもしれないですね. まだまだやることはありますね (笑).

■ ディープラーニングで課題は解決できるのか

佐川 言葉というか、ディープラーニングという名前だけが独り歩きしている感じも結構ありますね. 使っている人を見ていると、本当に使い物になるのかどうか、判断しかねるところもあるように思います.

佐藤 昔もそうでした、ロボットが出たときに職を失うと言われながら、やはりロボットは使われている. 私が郵便区分機の開発に携わったときには、手作業でやっていた人が職を失うと言われたけど、やはり自動化は進んでいく. 何か変化があると必ずネガティブな反応をする人がいるけど、それで効率化が進むなら使われていくのでしょう.

石寺 最近まで現場に近いところで開発をしていたのですけれども、困ることもあった (笑). 「ディープラーニングって凄いらしいじゃない」「なんで使わないの」って、多くのユーザが言うのだけれども、ディープラーニングの利用がとても難しい応用に対しても期待が膨らんでしまっただけで、たとえば欠陥検査では、ユーザに「欠陥というのはどんなパターンがありますか」と聞いても、「分からない」という答えが返ってくるわけです. そこをなんとかするために一生懸命アルゴリズムを考えるわけだけれど、ディープラーニング急進派の人々は「それは100万枚の欠陥画像を集めてくれれば全部解決しますよ」と、いうわけですね. そんなに画像を集めることができないからスクラッチでアルゴリズムを考えているんだということを、いちいち説明しなくてはならないわけです. それが、結構大変でした. ただ、ディープラーニングを否

定したいのではないのです. そもそもできるかどうか分からないことをまずディープにやらせて、なんか良さそうな結果が出たら、そこから真面目に考えるとかね (笑), そういった取っ掛かりになると思うのですよ.

岩元 そうすることが現場では結構多いです. お客様の課題に対してとりあえずディープラーニングを使ってみて、いけると思ったら実際には全然うまくいかないことが多い. それで結局、どうにかしてくれみたいな話が、我々専門家のところに落ちてきて、そこから解決しなければならない課題が逆に見えてくる. そういう意味では、チャンスではありますね. これまで眠っていたニーズをたたき起こすというか、こんなこともできるんじゃないかとお客様に思ってもらう意味で、ディープラーニングのインパクトはあるかもしれないですね.

佐藤 逆にディープラーニングがうまくいっている実際の応用というのは、何があるのでしょうか.

石寺 グーグル翻訳はディープラーニングを使っていて、評判が良いらしいじゃないですか. たぶんクラウドに飛ばしてしまうから、リソースはジャブジャブあるだろうし、使い方としては良いと思いますね.

福島 自然言語処理の分野は、もともとと辞書とルールでがりがりやっていたのが、徐々に機械学習ベースへ移行しています. だけれど、ディープラーニングで出る性能というのは従来のものに追いついてきたとか、ちょっと越し始めたぐらいで、画像認識や音声認識みたいに、大きく差をつけたというほどにはなっていないと聞きました. これからもっと進むとは思いますが.

逆に、グーグルやフェイスブックとか、ネット系のサービスにはデータがたくさん集まっているので、どんどん使えそうな場面が広がっているように見えますね. でも、現場のデータでいろいろやらなければいけないところは、そうはいかないですね.

石原 医療分野に応用している例を言えば、米国のEnlitic社というベンチャーですかね. 2年前ぐらいに、約2万件分の肺がん検診のCT画像データを使って、画像診断の専門医を上回る性能の肺がん検出システムを開発したことで話題となった. 検診といった大量の画像データが使えるところで、検出系のタスクがかなりうまくいっている印象です.

佐川 産業系の現場でもデータは少ないと思います. たとえば、音のわずかな変化を感知して判断するような場合です. ベテランの作業員の方が聞くと分かるのですが、我々が聞いてもまったく分からない. そういうものはデータとしてまったく残っていない



左：福島俊一，右：佐藤 敦

のです。作業員の経験に依存するし、異常が発生する頻度も少ない（笑）。そういう場合に、機械学習に適したデータをどのように集めるかが課題だと思います。

石原 なぜ米国ではそんなに多くのデータを公開できたのかなと、個人的にずっと考えていたのですが、データがたくさんあるとこんな良いことがあるというイメージが、医療従事者とうまく共有できたのかなと思います。医療の世界では、データを出すことはリスクという意識がどうしてもあるので、そのリスクに勝るようなメリットを我々の方で示してあげることが、今後我々がやるべき事かなと思っています。

福島 ディープラーニングは、こういうところでうまくいっているとか、ここが良いとかありませんか。もっとデータを出してもらえれば、適用が進むみたいなの。

石原 画像で解決できることには、ディープラーニングで挑む認識系の世界と、人手作業の補助のような、結構、地味な世界もあります。現場では意外と後者の作業系の課題が多くあって、「毎回、画像の位置を合わせる作業がすごく大変なんだよね」という話を聞いていました。そういうところでは、我々画像の専門家が話を聞くだけでアルゴリズムを設計してしまった方が早いし、現場も受け入れやすい。認識系でディープラーニングを使って性能を上げていく話と、従来通りハンドクラフトで設計した技術で日々の問題を解決していく話、両方を並行して進めていく必要があると思っています。単純にある日突然、ハンドクラフトの設計がまったく必要なくなるという訳ではなく、地道にやっていくべきかと。

福島 お客様の問題解決というと、認識というところはまだ問題解決の一部分ですからね。お客様にとっては、認識技術によって異常らしいとか言われても、それを人間がきちんと解釈して、その異常をなくすような手を打つところまでいって問題解決です。認識のところにディープラーニングを使うとしても、問題解決までの全体をトータルで見えていくことが、現場の人間には重要だと思います。

石原 そうなんです。先ほど申し上げた認識系の課題ばかりが今スポットライトを浴びていますけど、それ以外の作業系の課題などでメリットを示して、お客さんからデータをたくさん出してもらって、そのデータを使って、並行してディープラーニングに活用する。そんな両輪で攻めて行けばうまくいくのかなと思います。

岩元 ディープラーニングは非常に強力なツールなのですが、それだけでは実際問題は解決できないと思います。今おっしゃったような既存の技術とか従来型の

研究をうまく組み合わせると、全体システムとして非常によいものができる、そのような強力なツールが出てきたと捉えれば良いのかなと。

中野 一般論として、サンプルはたくさんあってクラスが有限だと、ディープラーニングが使えるのですけれども、先ほど言ったような、100万種類の書籍があって、100万種類の表紙画像があるというケースでは、100万クラスの識別をディープラーニングを使って行うのは無理かと思っています。

佐藤 今回のブームは本物だという意見は、私もよく分かります。確かに前回のブームとは違い、全然使えないわけではなくて、条件が揃えば性能がちゃんと出る。ただ、それをうまく使っていくためには、その特性をよく理解する必要があります。たとえば、トラフィックが重要なら、ディープラーニングよりもコンパクトな特徴をハンドメイドで設計しなくてはならないだろうし。専門家としてディープラーニングをうまく使っていくというのが、我々がやらなくてはいけないことかなと思います。

■ まとめ

佐藤 最後に、皆様一言ずつお願いします。

佐川 私はディープラーニングを使って研究を行っているわけではないので、知らないことがたくさんあるのですが、世の中の期待感は非常に大きいと感じています。ただ、まだよく分からない部分も多いと思っていますので、その辺の見極めをちゃんと客観的にしながら、ブームに踊らされるのではなく活用できたらと思っています。

石寺 ディープラーニングに関しては、使えるところに使っていけば良いと思っています。これまでは、できるだけ軽いリソースで認識できるようにと考えて、匠の世界で真面目に特徴抽出や識別系を考えてごりごりやっていました。今後は、認識できるかできないか分からないようなものに、ディープラーニングをやってみようと思います。うまく説明できないけれども、なぜか人間は認識できるようなものとか。それで、どうも認識できるらしいとか、人間と同じようなことができるらしいと分かったら、もうちょっと真面目に取り組んでいくというような、そういう使い方をしてみたいなと私は思っています。ブームで、すごいすごいと言っている人たちに対しては、少し遠くから眺めさせてもらおうかと思いますが、使えるところは上手に使っていければ良いかなと思っています。

石原 ディープラーニングは確かに今すごいブームで、お医者さんとお話しをしていると、我々もそうなのです

けれども、仕事を奪われるのではないかというすごい危機感がある。ただ、ディープラーニングでそんなに簡単にはすべては置き換えられない面もある。画像技術の過剰な期待に対して、我々専門家が適度なブレーキ役となって、地道にお客様の問題解決をしてゆく必要性があります。一方で人間がやらなくても良いようなところはどんどんディープラーニングに置き換えてもらって、それで人々が楽になるというポジティブなメッセージも伝えていきたい。このブームに乗って、これまで注目されてこなかった地味な課題にもスポットが当たれば良いですね。それで、画像処理にかかわる分野全体をどんどん盛り上げていけたらという想いがあります。

中野 大量にアノテーション付きのサンプル画像があるケースではかなり良い成績が出るということは実証されてきているように思います。医療系で使うには、先ほど言いましたように、たとえば50万画像あったら、49万9千個は良性の組織だというケースが通常ですので、ディープラーニングを利用して良性だと判断される画像をどんどん削ぎ落していったら、最終的に腫瘍に限りなく近いものと、本当の腫瘍だけを選別して、それを読影医に提示するということはかなり労力が減ると考えています。あと皆さん画像の専門家なので、フィードフォワードネットワークの議論が中心だったと思うのですが、音声のような時系列データですりカレントネットワークが使える可能性があります。フィードフォワード以外のネットワークも研究事例がでてきていますので、今後、一層ディープラーニングの用途が広がっていくのではと思います。

岩元 私は今まで特徴抽出の開発をやってきて、ディープラーニングは使ってないのですが、あえてブームに乗ってみようかなという思いはあります。アルファ碁みたいに、画像認識以外にもいろんな使い方で成果も出ていますし、学会でも非常に面白い発想で成果を出しているものがある。成果をぱぱっと出すタイプの研究がいろ

いろ出ているけど、まだそのスピード感に乗り切れていないので。ディープラーニングだけでは解決できない部分もあるので、従来タイプの研究とうまく融合すれば、今までできなかったことができるようなシステムが作れるのではと思っていて、あえてブームに乗って使ってみると面白いかなと考えています。

福島 実際に自分でディープラーニングをやっていたわけではないですけども、話を聞いていて思ったことをまとめます。皆がおっしゃっているようにディープラーニングは認識のための強力なツールになったのですが、先ほども言ったように、お客さんの問題解決から見れば、認識するというのはその一部分です。モノや状況を識別・判別をしてくれたり、異常を検出・診断してくれたりというところまでです。問題解決の全体を見てアーキテクチャやプロセスを作ることが重要だと思います。データが問題であればどうやってデータを集めるかというところまで含めて。また、データを集めるだけでなく、クラウドとエッジでどうやるかとかという話や、検出・識別・診断したあと、問題解決のための最後のアクションを決めるところまでどうするかも問題になってきます。ますますそういう全体アーキテクチャをちゃんと見て作っていくというのが重要になるだろうと思いました。

佐藤 アルファ碁も今回のブームの火付け役かなと思うのですよね。でもあれば、ディープラーニングを使って囲碁に勝つための仕組みを考えた人がエライのであって、ネットワークを学習させたら勝手にアルファ碁が出来上がったわけではない。ディープラーニングはもちろんブームなのですが、それをどううまく使って、現実的な問題解決につなげていくかが重要なのだと思います。これは、一般の人にはできないことです。やはり専門家でないとアイデアが出せないのです。そういう使い方を目指したいですね。いろいろご意見をいただき、ありがとうございました。(拍手)

