



松原 仁 Matsubara Hitoshi 人工知能研究者

京都橘大学工学部情報工学科教授。公立はこだて未来大学特命教授。元人工知能学会会長、前情報処理学会副会長。著書に『AIに心は宿るのか』（集英社インターナショナル、2018年）『やさしくわかる！文系のための東大の先生が教えるChatGPT』（ニュートンプレス、2024年）など

認識するAI(1)

認識するAIが世の中で広く使われるようになりつつあります。それは便利である一方で懸念すべき面もあります。今回は画像認識や音声認識を中心に考えてみましょう。

画像の認識

AIによる画像認識の能力は、2010年代以降劇的に進歩しました。ディープラーニングが最初に有効だと分かったのが画像認識でした。AIによる人間の顔の認識精度は我々人間を超えています。人間は知っている人の顔を95～96%程度の精度で認識するといわれています（知り合いの顔でも100%は難しいのです）。それに対してAIは98%程度の精度です。少しの差と思うかもしれませんが、たくさんの顔を対象とするときはこの差がかなり効いてきます。

AIによる顔認証の技術は、多くの国で出入国の本人チェックに使われています。以前はパスポートを人間の係官に見せて、その係官が目の前の顔を見てパスポートの写真と見比べて本人かどうかをチェックしていました（そこで怪しいとなると別室に連れていかれるのでしょ）。今はパスポートをAIに見せて、AIはカメラで目の前の顔を見て本人かどうかチェックしています（怪しいとなると係官のいる別室に連れていかれるはず）。以前の人間より今のAIのチェックのほうが精度が高く、しかもチェックが速くできるということなのです。建物や部屋に入るときに（カードなどではなく）顔をカメラに見せてAIが確認すると戸が開けられるというしくみも普及しつつあります。

AIで食事の内容も認識できるようになりました。食事を取る前に写真に撮ってAIに送ると、何をどれくらい食べるのかをAIが認識して栄養素

（塩分が多過ぎるとか）やバランス（野菜が少ないとか）、カロリー（食べ過ぎとか）などのチェックをしてくれます。

AIによる画像認識は外観検査にも広く使われるようになっていきます。道路、線路、建物などの古くなったインフラ施設は、定期的に外観を検査して怪しいようであれば本格的に内部を調べる必要があります。従来は人間の検査員が現場に行って人間の目で外観を検査していましたが、インフラ施設はたくさんあって場所もいろいろあるので現場に行くのが大変でした。また検査は誰でもできるわけではなく、訓練を受けた専門の人間である必要があります。そのために検査の回数を増やすことが難しく、検査と検査の間に重大な破損が起きてしまうことがあります。訓練を受けた専門技師でも人間なので見落としも生じます。AIによる外観検査であれば、専門技師以外の人間がその装置を操作すれば済みす。施設によっては人間が近づくのが危険な場所もありますが、ドローンを飛ばして画像を取得すれば人間が近づかずに済みす。さらには見落としも減らすことができます。

医療では画像診断でAIが活躍しています。筆者が最初に知ったのは肺がんの画像診断でした。アメリカのAIベンチャーの会社がたくさんの肺のレントゲン画像をAIに学習させました。画像と一緒にその肺がどういう状態か（正常か肺がんか、肺がんであればどのステージか）という人間の医者診断結果も学習させています。そうしてできた肺がん画像診断AIはとても優秀で、人間の専門医の精度を超えたと報告されました。人間の専門医はどうしてもある確率で誤診をしてしまいますが、AIの誤診の確率は人間より低く抑えられています。人間の専門医だとまず見過ごしてしまうごく初期の肺がんもAIは検出できたとのこと。その後は大腸、胃、脳、目（眼底）などさまざまな部位の画像診断をAIが

できるようになりつつあります。人間の医者として、AIが画像診断をしてその結果を参考に、人間の医者がどういう治療を行うかを決定するということになるでしょう。

AIは画像の中から文字を見つけてそれをファイルに書き出したり翻訳したりすることもできます。大学の授業は以前は黒板あるいはホワイトボードに先生が書いて学生がノートにそれを書き写すというのが普通でしたが、今は先生が書き終わった後に学生がスマートフォンでそれを写真に撮るということがよくあります。その写真をAIのアプリに入力すると、AIが書かれている文字を認識してワードなどのファイルに書き出してくれます。書き写すことにある教育効果は期待できませんが、便利ではあります(筆者の書いた汚い字を正確に認識してくれるか怖くて確認していませんが)。

最近のAIは静止画像だけでなく動画も認識できます。例えば、人間がどういう動作をしているのか(歩いている、食事をしている、本を読んでいる、スマートフォンで話をしているなど)をかなり正確に認識することができます。無人のコンビニもできています。客が棚から何を幾つ取り出したかを認識して、レジでは(改めてバーコードをスキャンすることなく)代金を表示します。怪しい行動もチェックできるので万引きの防止にもなります。万引きをした場合は顔を含めた情報をAIが把握しているので犯人の特定をしやすいでしょう。顔が写っていないなくても歩行のようすだけからもAIは精度よく個人認証ができます。人間の歩き方には癖があって一人一人違います。歩いている下半身だけが写っている動画からその人が誰かが分かるのです。

音声認識

音声を認識するAIについても少し触れておきましょう。

家庭や職場にAIスピーカーがあるという読者もいらっしゃるでしょう。人間が話をするとそれを聞き取って希望する音楽を流したり電気をオンオフしてくれたりします。台所では料理を

作るときにAIスピーカーに「5分経ったら教えて」と言えば5分後に教えてくれます。ときどきこちらの言うことが聞き取れずに言い直させられることもありますが、便利な道具として使われています。

客からの問合せや苦情を電話で受け付けるコールセンターでは人間のオペレーターではなくAIが対応するところが増えています。扱っている商品に関する語彙をあらかじめ教えておくことによって、客の話をかなりの精度で聞き取ることができます。定型的な問合せにはAIが返事をして、難しい問合せだと人間のオペレーターにつなぐという運用が多いようです。

また、会議のようすを録音して後で人間が文字起こしをして議事録にまとめるというのが普通でしたが、今はAIが自動的に文字起こしをするというサービスが広がっています。オンライン会議では(ほぼ)リアルタイムで発話をテロップの文章として画面に表示できます。耳が不自由な人向けの文字起こしAIも活躍しています。

音声認識AIを用いれば外国語の会話をそのまま日本語に翻訳できる(もちろん日本語の会話も外国語にできます)ので、スマートフォンで日本語で外国人と話ができるようになります。

懸念すべき点

AIの画像認識や音声認識の能力は格段に進歩しました。それはとても便利なのですが、懸念事項もあります。それはプライバシーが損なわれるおそれがあるということです。施設や道に多くのカメラが備え付けられています。そのことによって安全が守られているという側面はありますが、本人の意思に反して行動履歴が他の人に分かってしまうという側面もあります。音声認識も使いようによっては話者が誰か分かりません(例えば、悪口を言ったのが誰か分かっちゃうのです)。国によってはAIを用いて国民の行動を管理しているところもあります。

進歩したAIをどう使うかは、みんなで考えていく必要があるでしょう。