

毛筆文字の印象の分析 パート1 : 人工知能による 定量化

著者	八柳 祐一, 杉? 哲子
雑誌名	静岡大学教育実践総合センター紀要
巻	32
ページ	1-6
発行年	2022-03-31
出版者	静岡大学教育学部附属教育実践総合センター
URL	http://doi.org/10.14945/00028684

毛筆文字の印象の分析 パート 1

～人工知能による定量化～

八柳祐一・杉崎哲子

(技術教育系列・国語教育系列)

Quantitative characterization of brush character using artificial intelligence

Yuichi YATSUYANAGI and Satoko SUGIZAKI

要旨

To estimate an impression on a shape of a classic Chinese calligraphy quantitatively, the authors develop a machine learning system based on a two-layer traditional neural network. The paper consists of two parts. In part 1, an implementation of the system is explained. The system is built on TensorFlow which is an optimized library for a machine learning system developed by Google. The number of classical calligraphy images used for training is 935. After repeating the learning 300 times, the correct answer rate reaches 95.4%. In a lecture for junior high and high school students held in Aug. 2nd 2021, impressions for the Chinese calligraphy written by them were conjectured by the system. In part 2, detailed background on this research work and the discussion on the results derived by the system will be explained.

キーワード： 機械学習 TensorFlow 書表現 印象 毛筆文字

はじめに

本論文で紹介する取組みの端緒は、共著者から寄せられた「書の制作において、自分の意図する表現と他者が感じる印象を定量的に比較することはできないか」という問題提起に遡る¹。それに対し、ニューラルネットワークを用いた機械学習システムで実現できる可能性がある、と提案したところ、実践する方向で話がまとまった。毛筆文字を自動認識するシステムは、古典の自動解読など枚挙に暇がない。Google で検索を行うと、いくらでも事例を見つけることができる。一方、本研究において人工知能を適用する対象は、文字認識ではなく文字から受ける印象の定量化であることを主張したい。

書の印象は一つに定まるものではないが、書道教育においては、書道史に対する理解を含めて書の古典の印象を鑑賞用語で捉え、それらの具象化を可能にするために用筆を学習する。例えば、九成宮醴泉銘には「厳正」、孔子廟堂碑には「温雅」という鑑賞用語が当てはめられている。したがって、書道学習を経験した者は、それを知った上で印象を捉えることが多い。一方、書道学習を経験していない者は、既定の鑑賞用語に縛られずに印象を捉えることが可能である。よって、本研究では、鑑賞用語を平易な言葉に直して文字と印象の語を紐付けし、その組合せを人工知能に学習させることによって、既定の鑑賞用語に縛られることなく文字の印象を決定づける字形の特徴を定量的に抽出するシステムを実現することを試みる。なお、書の

印象に関するより詳細な説明は、パート2の第4章を参照されたい。

具体的に、人工知能にできることを確認しよう。毛筆文字に関する印象には、二つの観点がある。一つは書の制作者が意図した印象、もう一つは制作された毛筆文字を鑑賞する者が受ける印象である。印象に関する評価が歴史的に定まっている書の古典と、その文字の印象を簡潔な現代語に翻訳した対応表を人工知能に学習させることにより、人工知能は書道の基礎的学習である書の古典(碑文や法帖)の印象を知識として得る。学習に用いたものと別の毛筆文字に対して印象を演繹させることは、人工知能が今度は鑑賞者としての立場で印象を述べることを意味する。ここに人間の意識は介入しないため、「定量的」評価を達成していると考ええる。さらに重要なのは、鑑賞者の立場に立つ人工知能が、毛筆文字の意味を理解しない点である。人間ならば「花」という毛筆文字に対して無意識のうちに「華やかな」や「綺麗」といった印象語を連想する。しかし、今回用いた人工知能には、文字の意味を理解する機構が実装されていないので、字形のみに依存した定量的な印象語の導出が可能となっている。また、時に、人工知能は書道学習の経験者が予想しなかった印象語を演繹することがあり、それが新たな発見に繋がることもあることを付け加える。これは、将棋界における藤井聡太氏が、これまでの常識とは全く異なる戦法を人工知能との対戦により学び、目覚ましい戦歴を築きつつあることから想像できるだろう。

本論文では、端緒となった問題提起に対する解である毛筆文字から受ける印象を定量的に評価するための機械学習システムの紹介、および中高生を対象にした「ひらめき☆ときめきサイエンス」向けに収集した毛筆文字を機械学習システムで処理した結果例についてパート1として述べる。毛筆文字に関する導入説明、印象を定量化する必要性、人工知能から得られた結果の分析については、パート2として稿を改めて述べた²。

「人工知能」という言葉が初めて導入されたのは、1956年にDartmouth大で開催された国際学会であると言われており、その提言者であるJohn McCarthyによると、人工知能とは、広く「知的な機械、特に、知的なコンピュータプログラムを作る科学と技術」と定義している³。

人工知能は、これまでに3回のブームを経験している。第1次は1950年代後半～1960年代で、コンピュータによる推論や探索が可能となり、特定の問題に対して解を提示できるようになった。第2次は1980年代で、コンピュータが推論するのに必要な様々な情報を与えることにより、専門家のような提案を行うことができるエキスパートシステムが実用化された。現在は2000年代に始まった第3次のブームと分類される。最大の特徴は「機械学習」が可能になったことだろう。第2次では人間が、情報の取捨選択を行ったり、推論機構のアルゴリズムを作ったりしなければいけなかったのに対して、今回のブームで実現された機械学習では、大量のデータを与えることにより、人工知能が自ら知識を獲得することが可能となった。Googleの子会社であるDeepMind社が開発した囲碁対戦プログラムであるAlphaGoでは、過去の棋譜から学習を行い、さらにAlphaGo同士を戦わせて、あらたな戦術を開拓した。その結果、2016年3月に囲碁におけるトップ棋士である李9段(韓国)に4勝1敗で勝った。第3次ブームの成果は、画像認識(人や物体の識別)、音声認識(会話からの文字おこし)や、文章の他言語への自動翻訳など、非常に多岐に渡る分野に応用され、大きな成果を挙げていると言えよう。

第3次のブームでは、さらに個人で機械学習を体験できるフレームワークの整備が進んだことも大きな特徴となっている。特に、Googleは精力的にフレームワークの開発と公開を行っており、Pythonというプログラミング言語の知識があれば、300行程度のコーディングで画像分類を行える人工知能システムを構築可能である。そこで、今回の取組みでは、Googleから提供されているフレームワーク(ライブラリ)であるTensorFlowを用いて機械学習システムを構築した⁴。その成果物を用いて、中高生を対象にした「ひらめき☆ときめきサイエンス」での実践も行ったので、それらの様子も含めて報告を行いたい。

論文の構成は、以下のとおりである。第1章ではニューラルネットワークのメカニズムに関して概略を述べる。第2章では、構築したシステムの構成を説明するとともに、印象の定量化を行うために必要となる学習(トレーニング)の内容を説明する。また、「ひらめき☆ときめきサイエンス」向けに参加者から収集した毛筆文字を人工知能で処理した結果を紹介する。最後に第3章でまとめる。

1. ニューラルネットワークのしくみ

まず初めに、ニューラルネットワークを用いた機械学習システムについて説明をする。

ニューラルネットワークを構成する最小単位は、パーセプトロンと呼ばれる構造である。最も基本的なパーセプトロンは、二つの入力と一つの出力をもつ。パーセプトロンは脳のメカニズムを模しており、○が神経細胞(ニューロン)、神経細胞同士を接続する線が神経繊維である。

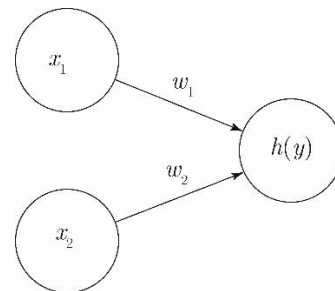


図1.1 パーセプトロンの例

神経細胞には電位 x_i が記録されており、神経繊維を通して他の神経細胞に伝達される。個々の神経繊維の太さは異なっており、その太さに応じて電位の伝わりやすさが変化する。その伝わりやすさを w_i と表記する。右側の神経細胞には、

$$y = w_1x_1 + w_2x_2$$

が伝達される。この y を重み付き総和と呼ぶことにする。さらに、右側の神経細胞内では、伝達された y の値があるしきい値 b を超えているか否かの判定を行い、超えている場合1、それ以外は0を出力とする活性化関数 $h(y)$ に入力される。

$$h(y) = \begin{cases} 0 & (w_1x_1 + w_2x_2 \leq b) \\ 1 & (w_1x_1 + w_2x_2 > b) \end{cases}$$

最終的に右側の神経細胞の電位は、活性化関数の出力である1か0となる。

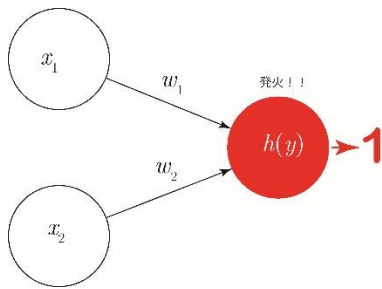


図 1.2 発火した場合の例

上記のパーセプトロンを多数、多段階に積み上げた構造がニューラルネットワークである。ニューラルネットワークの例を図 1.3 に示す。

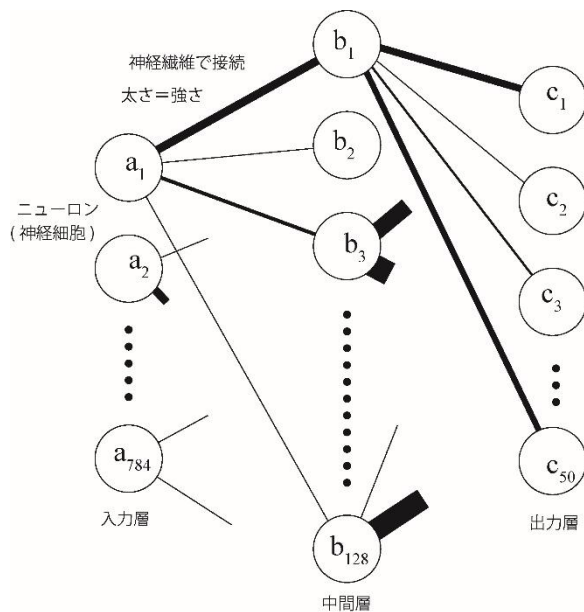


図 1.3 ニューラルネットワーク

図 1.3 のようなニューラルネットワークを 2 層ネットワークと呼ぶ。層の数は神経繊維の層の個数で表現するのが一般的なので、3 層ではなく 2 層となる。画像の分類問題を扱う場合、入力層には画像データを設定する。ただし、画像データのサイズが大きいとニューラルネットワークでの学習に膨大な時間が必要となるので、簡易的な問題の場合には 28×28 ドット程度の画像を入力とすることが多い。中間層のニューロン数は、学習に用いる画像数や出力層の数に依存し、結果を見ながら調整をする。出力層は分類結果の個数に一致する個数のニューロンを用意する。出力層のニューロン一つが一つの解答に対応し、各解答が正解であろう確率が出力層に求められる。すなわち、出力層の値の総和は、1 である。

機械学習では、入力の画像(学習画像)と、正解(出力層のある一つのニューロンの値のみが 1 で他はすべてゼロ)の組みを 1000~数万セット用意する。学習では、

1. 学習画像を入力として、出力層の値を計算
2. 正解と照合し、出力層の値が正解に近づくよう、誤差逆伝搬法などのアルゴリズムを用いて、神経繊維の重みの値を調整

を繰り返す。全ての学習画像と正解のセットについて 1 回学習するだけでは、多くの場合、勉強不足なので、これを多数回繰り返す。

上記の学習によって鍛えられたニューラルネットワークに、学習画像とは別の画像を与えると、出力層に解答(推論結果)が得られる。一般的に、ニューラルネットワークを用いた機械学習では、誤差逆伝搬法などによるトレーニングに時間がかかるが、推論は非常に高速に行うことができる。

2. 行ったことの概要

今回行ったことは、以下のとおりである。

1. 毛筆文字の画像認識を行うことができるニューラルネットワークシステムを構築する。
2. 高等学校芸術科書道の教材から採取した毛筆文字 985 文字と、その文字の印象を表す語 50 語の対応関係を学習データとして用意する。一つの毛筆文字に対して、一つの印象を表す語を関連づけ、対応表を作成し、人工知能のトレーニングを行う。なお、教材から採取した毛筆文字データについて、学習を行う際に背景にノイズが含まれると誤認識する現象を確認したので、機械学習にかけるまえに背景画素の平滑化を行った。
3. 中高生を対象にした「ひらめき☆ときめきサイエンス」向けに、参加者から収集した毛筆文字をデジタルデータ化し、それを人工知能で処理することにより、各文字に対応した印象を表す語を推論する。

なお、はじめに述べたように、上記 2 で人工知能が学習する文字は、印象に関する評価が歴史的に定まったものであることを強調したい。現代語に翻訳された印象語との組みを学習することにより、人工知能は「書の古典」の制作者と同等の視点を獲得することができる。

2.1 作成したニューラルネットワークシステム

ニューラルネットワークシステムをフルスクラッチでプログラミングすることは可能であるが、学習過程の高速化が困難なため、一般的には、ニューラルネットワーク構築用に公開されているフレームワーク/ライブラリを用いる。今回は、Google が無料で公開している TensorFlow を用いた。

また、作成したニューラルネットワークについてのパラメタを以下に示す。

表 2.1 ニューラルネットワークのパラメタ

層の数	2
入力層の数 (入力画像の 画素数)	28×28=784
中間層の数	128
出力層の数 (印象を表す 語の総数)	50
活性化関数	Relu
最適化法	Adam
損失関数	交差エントロピー誤差

2.2 ニューラルネットワークの学習

学習に用いる毛筆文字は、主に高等学校芸術科書道の教材から採取した。採取した文字数は 985 文字である。最初はイメージスキャナにより 300dpi 程度でページごとにスキャンし、その後 1 文字単位で 28×28 ドットの BMP 画像に変換した。学習に用いた毛筆文字の一部を図 2.1 に示す⁵。



図 2.1 学習に用いた毛筆文字(一部)

印象を表す語は、以下の 50 語を選定した。これらを選定した基準等はパート 2 で述べる。各語には、先頭から順に 0~49 の番号が振られており、その番号をプログラムでは使用する。

0：力強い，1：繊細，2：穏やか，3：温かい，4：抑揚がある，5：緩急のある，6：平坦，7：面白い，8：重厚，9：明るい，10：冷静，11：軽快，12：鈍い，13：鋭い，14：不安定，15：尊厳，16：勢いのある，17：奥深い，18：伝統的な，19：独特な，20：質素な，21：乾いた，22：しっとり，23：寂しい，24：大胆，25：斬新，26：妖艶，27：直線的，28：伸びやか，

29：緊密，30：濃い，31：薄い，32：曲線的，33：かっこいい，34：暗い，35：可愛い，36：不気味，37：かたい，38：やわらかい，39：沈んだ，40：粘り気のある，41：計算された，42：突発的な，43：幼稚な，44：すっきりした，45：暑苦しい，46：懐の広い，47：自由気ままな，48：美しい，49：忙しい

初めに全データを 10 回、繰り返し学習させた。学習に伴う正答率の変化の様子を図 2.2 に示す。ここでの「正答」とは、文字と、その文字に紐付けられた印象を表す語の対応表を人工知能が正しく記憶しているか否かを表している点に注意して欲しい。

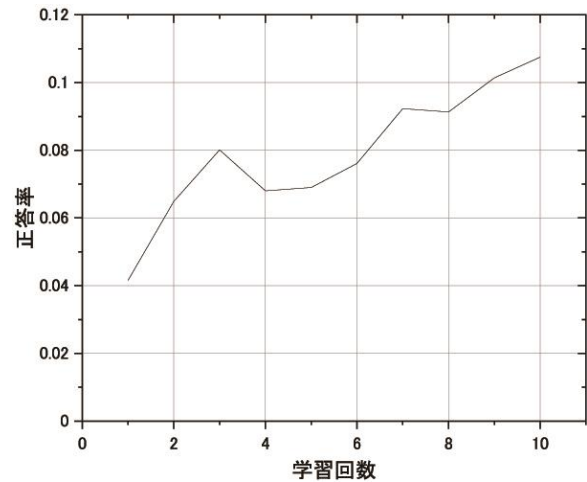


図 2.2 正答率の変化(10 回学習)

縦軸の値に注目すると、最大でも 1 割程度の正答率、すなわち、100 点満点のテストに換算して 10 点台である。実際にどの程度の成績がでるのか確かめるため、学習に用いたデータを用いて、正しく演繹できるかチェックを行った。結果の例を図 2.3, 2.4 に示す。各図において左側の文字が予想を行った毛筆文字、右側の縦棒グラフが、50 の印象を表す語について、その語が合致する確率をプロットした図である。縦棒グラフの左端から順に 0(力強い)~49(忙しい)の各語に対応し、確率が一番高い語を予想結果としている。

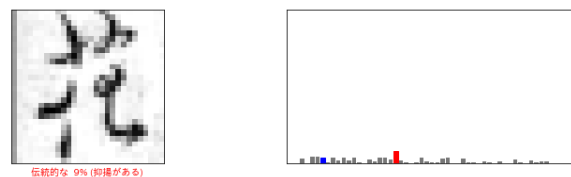


図 2.3 結果の例。正解は「抑揚がある」だが、推論した語は「伝統的な」となっている。

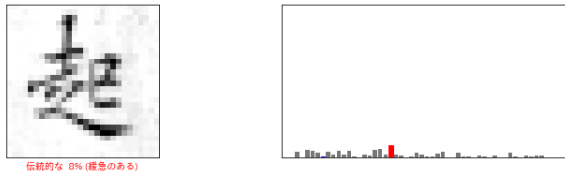


図 2.4 結果の例。正解は「緩急がある」だが、推論した語は「伝統的な」となっている。

図の予想結果を表す縦棒グラフをみると、色々な語に確率値が分散しており、推論結果に確信が持てない様子が見える。すなわち、人工知能が、まだ十分な学習を行えていないことを表す。

次に、10回に加えて、さらに290回(合計300回)の学習を行った。学習に伴う正答率の変化の様子を図2.5に示す。

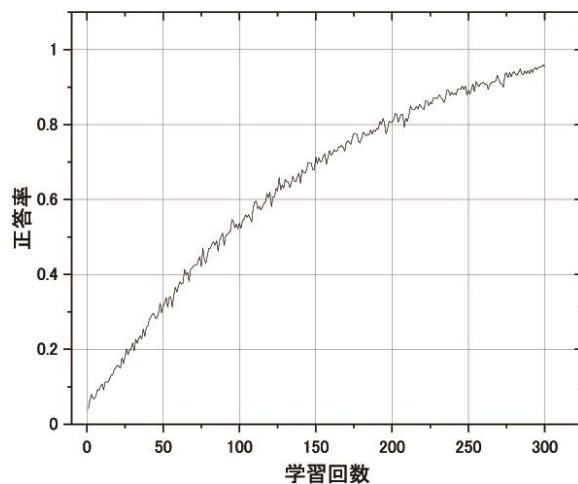


図 2.5 正答率の変化(合計300回学習)

最終的な正答率は95.4%に達し、画像認識用の機械学習システムとしては、まずまずの正答率を実現できることが確認できた。先ほどは誤答だった結果について、合計300回学習した後の結果を以下に示す。

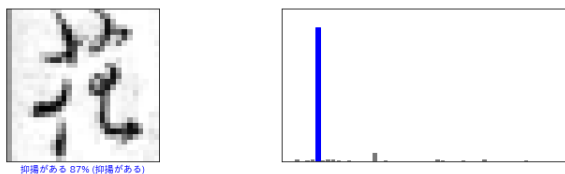


図 2.6 結果の例。正解である「抑揚がある」を演繹している。

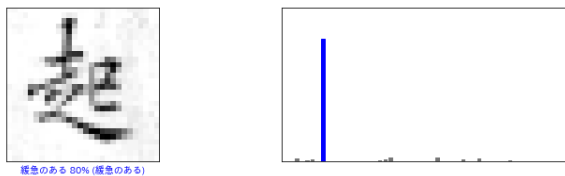


図 2.7 結果の例。正解である「緩急がある」を演繹してい

る。

今回は、予想結果の確率値が他の確率値よりも明確に大きな値となっており、ニューラルネットワークが、以前よりも自信をもって予想している様子が見える。これは、人工知能が学習に用いた毛筆文字の作者が意図した印象について十分、学習したことを意味する。

2.2 印象を表す語の鑑賞者としての演繹

「ひらめき☆ときめきサイエンス」向けに毛筆文字を提供してくれた参加者のプロフィールは以下のとおりである。

表 2.2 参加者のプロフィール

静岡東高校 3年	1名
科学技術高校 3年	1名
附属静岡中学校 2年	7名
高松中学校 1年	1名
高松中学校 2年	3名
高松中学校 3年	2名
南中学校 3年	1名
袖師小学校 6年	1名
常葉大学附属小学校 4年	1名
附属静岡中学校 3年「探究」受講生	19名
城南静岡中学校 1年	19名
城南静岡中学校 2年	16名
城南静岡中学校 3年	24名

この中から、附属静岡中学校2年生の7人が書いた「翔」について、人工知能がどのような判断を行ったのか、結果を示す。これらの結果に関する考察は、パート2で述べる。

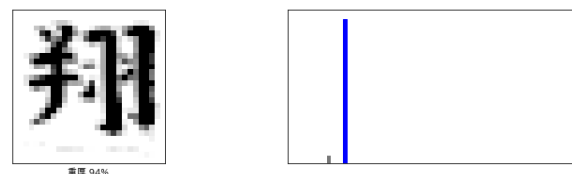


図 2.8 一人目の毛筆文字と鑑賞結果。「重厚」が演繹された。

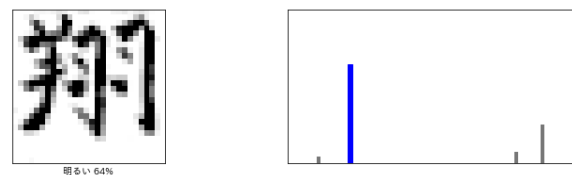


図 2.9 二人目の毛筆文字と鑑賞結果。「明るい」が演繹された。

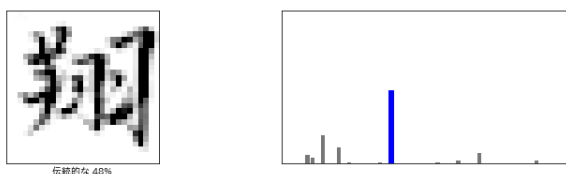


図 2.10 三人目の毛筆文字と鑑賞結果。「伝統的な」が演繹された。

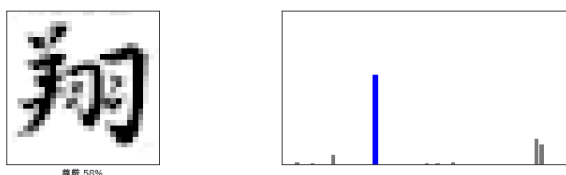


図 2.11 四人目の毛筆文字と鑑賞結果。「尊厳」が演繹された。

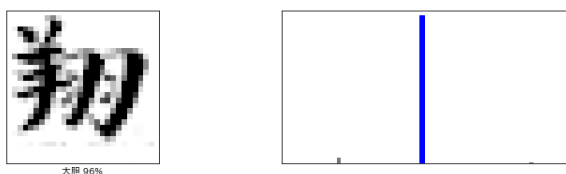


図 2.12 五人目の毛筆文字と鑑賞結果。「大胆」が演繹された。

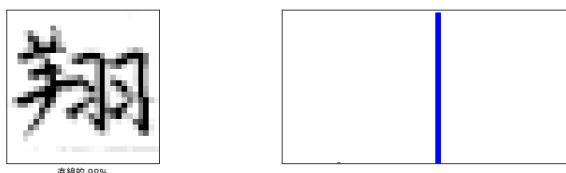


図 2.13 六人目の毛筆文字と鑑賞結果。「直線的」が演繹された。

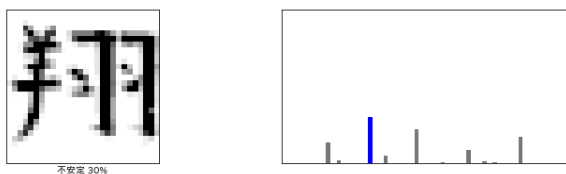


図 2.14 七人目の毛筆文字と鑑賞結果。「不安定」が演繹された。

References

¹ 杉崎 哲子, 八柳 祐一: 「書表現に活かす毛筆文字の印象: 古典の漢字を題材にした機械学習の活用に向けて」 静岡大学教育実践総合センター紀要, **31**, 96-106 (2021-03-25).

² 杉崎 哲子, 八柳 祐一: 「毛筆文字の印象の分析パート 2~人工知能による診断をふまえて~」 静岡大学教育実践総合センター紀要, **32**, 7-16 (2022-03-25).

「ひらめき☆ときめきサイエンス」での講演では、人工知能のメカニズムの解説や生徒が提供してくれた毛筆文字を用いた実演を行った。参加した生徒たちがこれらの結果を見たときに、お互い、どのような印象語と関連付けられているのか歓談し、どのような特徴がどのような印象を引き出したのか議論している姿が印象的だった。

3. まとめ

毛筆文字から受ける印象を定量的に評価するために、機械学習システムを開発した。学習用に準備した毛筆文字画像は985文字分であり、50種類の印象語を学習するにはデータが少なすぎるのでは、という懸念があったが、合計300回、繰り返して学習させることにより、学習に用いたデータから印象語を推定する正答率は95.4%に達した。このことから、少ないデータ数でも繰り返して学習を行うことにより実用に耐える精度を実現できることが確認できた。「ひらめき☆ときめきサイエンス」について、参加した生徒が人工知能というものを身近に感じることができたイベントになったのではないかと考えている。

今後は、本学習システムが、学習に用いた毛筆文字のどのような特徴を検知しているのか、検討を重ねる予定である。

【謝辞】 本取組みは、令和2(2020)年度科学研究費助成事業(研究成果公開促進費)(研究成果公开发表(B)・ひらめき☆ときめきサイエンス~ようこそ大学の研究室へ~KAKENHI 代表者: 杉崎 哲子 プログラム名: 書とプログラミング~「書のプログラミング」にチャレンジしよう~の支援を受けて行いました。

³ 一般社団法人人工知能学会: 「人工知能のFAQ」 <https://www.aigakkai.or.jp/whatsai/AIfaq.html> (2022年1月25日)

⁴ TensorFlow チュートリアル: <https://www.tensorflow.org/tutorials?hl=ja> (2022年1月25日)

⁵ 全国大学書道学会編: 「書の古典と理論」 光村図書 2020.