

# Teachable MachineとScratchで作る画像認識によるIoTシステムの研究

## Research on IoT system based on image recognition made with Teachable Machine and Scratch

宮崎 英一<sup>1</sup> ・ 坂井 聡<sup>1</sup>

Miyazaki Eiichi, Sakai Satoshi

### 要旨

現在、小学校でのプログラミング教育が必須化され、多くの教育現場では、プログラミ的思考を育み、教科の学びをより確実にするためのプログラミング教育が行われている。一方、コンピュータを活用してより良い社会を育むといった点には、思考を形にするために、それを構成出来るプログラミング環境が必要であり、初学者のプログラミング教材として取り上げにくい。そこで本研究では、プログラミングのコードを書かずに、マウスのドラッグアンドドロップだけで実現出来るAIを用いた画像認識によるIoTシステムを提案する。この教材はAIで画像認識を行い、その結果を実空間のデバイスやインターネット上のプログラムと連携させる事が可能であり、自らが社会の問題点を考え、それを解決する事をプログラミングを通して学ぶものである。

キーワード：AI、IoT、プログラミング教育、Society 5.0

### 1. はじめに

現在、日本国政府はこれからの日本における成長戦略の1つとしてSociety 5.0に対応できる人材の育成を進めている。これに伴い、学校教育現場においても、小学校でのプログラミング教育の必修化、中学生での技術教育におけるコンピュータとの双方向性プログラミングの導入、プログラミングによる問題解決の導入、高校における情報の必修化、大学入試における教科「情報」の導入等、急速な勢いで小・中・高・大と連続的な形で教育現場の情報化が進んでいる。

今回のプログラミング教育におけるプログラミング教育で育成する資質・能力として図1<sup>1</sup>に示すような内容が挙げられている。多くの教育現場では、プログラミ的思考を育み、教科の学びをより確実にするためのプログラミング教育が行われている。一方、コンピュータを活用してより良い社会を育むといった点においては、思考を形にするために、それを構成出来るプログラミング環境が必要であり、初学者のプログラミング教材として取り上げにくい。

段階	知識・技能	思考力・判断力・表現力	学びに向かう力・人間性
小学校	身近な生活でコンピュータが活用されていることや、問題の解決には必要な手順があることに気付くこと。	発達の段階に即して、「 <b>プログラミング的思考</b> 」を育成すること。	発達の段階に即して、コンピュータの働きを、よりよい人生や社会づくりに生かそうとする態度を涵養すること。
中学校	社会におけるコンピュータの役割や影響を理解するとともに、簡単なプログラムを作成できるようにすること。		

図1 プログラミング教育で育成する資質・能力

そこで本研究では、プログラミングのコードを書かずに、マウスのドラッグアンドドロップだけでプログラミングが

1 香川大学教育学部

実現出来るAIを用いた画像認識によるIoTシステムを提案する。

現在、AI、機械学習、ディープラーニングといった単語は、我々の身の回りに広く氾濫し、以前のように一部の専門家だけが扱う対象では無い。更にスマートフォンやインターネットの普及に伴い、音声認識やインターネット上のレコメンドの提示等のAI処理は日常生活の一部になっている。このため、児童・生徒においてもAIは特別なものでなく、普通に利用できるツールとなっている。しかし残念ながらツールであるため、一般的なユーザからは完全なブラックボックスになっているため、自分で独自のカスタマイズ等は行えない。そこで何らかの方法で自分独自のAIシステムが実現出来れば、これを用いて現実世界の問題点を見つけて解決するという「学びにむかう力・人間性」の向上に寄与出来ると考えた。

ここでは本研究では、AIで画像認識を行い、その結果を実空間のデバイスやインターネット上のプログラムと連携させる事が可能なプログラミング教材システムを提案する。このプログラミング教材の利用により、児童・生徒自らが社会の問題点を考え、それを解決する事をプログラミングを通して学ぶ事が実現出来ると考えるものである。

## 2. 画像認識システム

本研究では、画像認識システムとしてGoogle Teachable Machine<sup>ii</sup>を用いた。これは、AIの機械学習ツールで、WEBアプリケーションとなっている。このため、アプリケーションを各自のコンピュータにインストールする必要がないので、独自のアプリケーションのインストールが困難な場合が多い学校教育現場においても導入が簡単である。更に複雑な計算はサーバ側で処理されるので、GIGAスクール端末のように非力なコンピュータでも問題なく利用できるという利点もある。しかし授業に際しては、インターネットを介してアクセスを行う事から、アクセス制限やインターネットの不測のトラブルがある事も留意すべきである。

また、このシステムはビジュアルプログラミングの形態をとっており、プログラムのコードを書かずにマウスの操作だけで機械学習による画像認識が実現されている。よってプログラムの文法等の知識的な学習の時間が省け、プログラミング教育本来の目的である問題解決に時間をかける事が可能である。更に本研究では、画像認識を用いたが、同ページではこれ以外に音声・ポーズ等の認識モデルも提供されている。これらの利用は画像認識と同じ手順で行えるので、認識を行うTeachable Machineの判別モデルの部分さえ変更すれば、そのまま、音声・ポーズ等認識が可能であり、幅広い展開が期待できる。

Google Teachable Machineのモデル生成には、機械学習の種類のひとつ、深層学習 (Deep Learning) が使われており、

複数の画像や音声から特徴点 (似た部分) を抽出し、学習<sup>iii</sup>を行っている。更に学習以外にも結果の混合行列、学習の繰り返し数による正確さのグラフ化等で結果を可視化してくれるので、機械学習の専門的な知識が無い学習者に対しても判定結果を直感的に理解させやすいという特徴がある。

### 2.1 画像認識テスト<sup>iv</sup>

ここでは上記で述べたTeachable Machineを用いて画像認識テストを行った。その結果を図2に示す。同図では、認識モデルとして著者 (miya)、人形 (DANBOARD<sup>v</sup>)、はさみ (scissor) の3つを対象とした。モデル名の後の括弧内は、認識された特に出力されるタグを示しており、学習者側で自由に変更可能である。

同図に示すように各モデルで500枚程度の画像を準備した。この画像の撮影においてはWEBカメラがそのまま使用できるので、ノートPCやタブレット端末等のカメラを持ったコンピュータならば特別なハードウェアの追加は必要無い。更に以前のTeachable Machineでは、写真を1枚1枚撮影するタイプであったが、最近のシステムでは、ビデオカメラの動画から静止画を自動で生成してくれるので、更に撮影の手間が省けるようになっている。

同図の枠線で囲んでいる部分が画像認識の判定結果である。ここでは撮影対象を「はさみ (scissor)」としているが、scissorの部分が100%で認識されている事が分かる。この認識について、大学の講義においてこのTeachable Machineを用いて画像認識を行ったが、ほとんどの場合100%近い判別が可能であり、学生からも「画像判別の正確さに驚いた」とい反応が多かった。

この学習結果は図3に示す。同図\_1は、学習モデルにおける混同行列<sup>vi</sup> (confusion matrix) を示しており、アルゴリズムの性能を可視化するための特有の表配置である。同図マトリックス (miya-miya : 78) (miya-DANBOARD : 0) (miya-scissor : 0) において、miyaを認識したデータ数78個、miyaをDANBOARDで認識したデータ数0、miyaをscissorとして認識したデータ数0個と極めて精度の高い判別が行われている事が示された。

同図\_2、3は、機械学習における各エポックごとの精度と損失を示している。同図横軸エポックは学習回数を示しており、右に行くほど学習回数が増えている事を示している。精度は1が最高値であり、学習の初期段階で早くも限りなく1に近づいている事が示されている。また損失は良い学習状態になるほど0に近づく。同図\_3でも早い学習段階で0に近づいている。このように効果の評価からも精度よい判別が実現出来ている事が示された。

このような画像認識システムを教育現場に持ち込む場合、画像判別精度そのものが大きく学習の意欲につながる。もし判別精度が低ければ、多く児童・生徒は授業そのものに興味を持ってくれないだろう。児童・生徒に興味を



図2 Teachable Machine学習結果

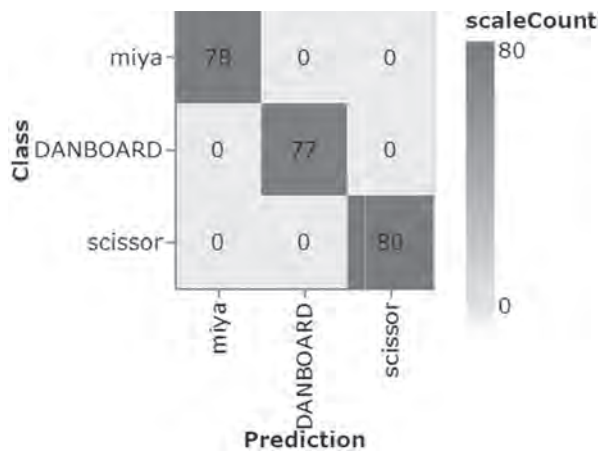


図3\_1 混同行列

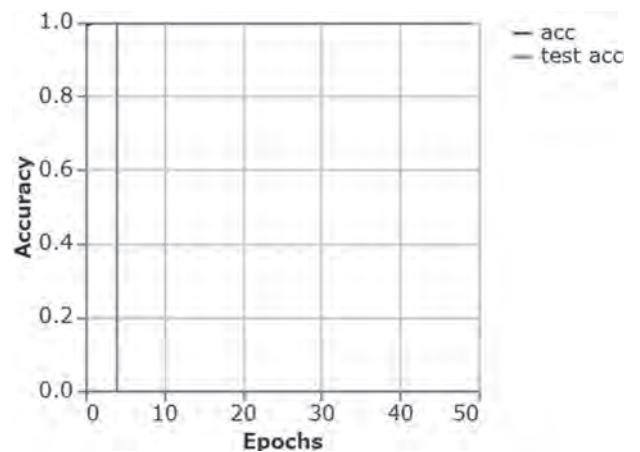


図3\_2 エポックごとの精度

持ってもらうためには、皆が驚くくらいの判別精度が必要になる。そのためにも、本システムで用いたTeachable Machineは大きな学習効果を持つと言える。

更にTeachable Machineの大きな特徴として画像判別に用いる学習モデルをネットワークで共有できるという点である。この学習モデルエクスポートの画面を図4に示す。ここではモデルをローカル（各自のコンピュータ）保存、もしくはネットワーク保存（クラウドモデル）の両方が選択出来る。授業においてクラウドで保存する場合の利点としては、何らかの理由で、学習モデルの撮影が困難だった場合でも、クラウドモデルのURLを知っていれば、直ちにそのモデルが利用可能になるという点である。つまり授業の対象年齢によってWEBカメラを用いた撮影や学習（それ

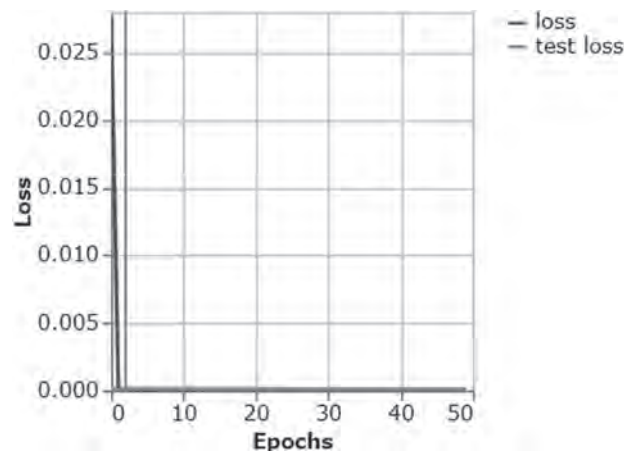


図3\_3 エポックごとの損失



でもボタンを押すだけであるが) が困難だった場合、先生・教師が事前に学習モデルを準備しておき、授業開始時に学習モデルのURLを提示する事で、クラス全員が同じ学習モデルを用いた画像認識が実現出来るという利点がある。

### 3. Scratchとの連携

上記で述べたTeachable Machineを用いれば、簡単に画像認識が実現出来る事が示された。ただ、このシステム単独

では、画像認識は実現できるが、その判定結果を直接外部に受け渡す事が出来ない。しかし、上記で述べたように学習モデルがクラウドモデルとしてエクスポートされるので、これを外部プログラムから使用すれば、外部プログラムから画像認識が行える事が可能になる。

本研究では、この外部プログラムとして小学校のプログラミング教育によく用いられているScratchを使用した。ただし、正式版のScratchではこの機能が実装されておらず、独自の拡張版であるScratch 拡張機能 (TM2Scratch™) を利用させて頂いた。この拡張版が有効なカスタマイズされたScratchにはTeachable Machine以外にも後で説明するIFTTT等の様々な拡張が施されている。

このカスタマイズされたScratchを用い試作したプログラムを図5に示す(以下の図においてネットワーク上の情報となる部分についてはセキュリティ面から塗りつぶしている事に留意されたい)。Scratchは多くの小学校でプログラミング教材としての利用実績があり、現場への導入も簡単である。ここでは、画像認識によって判定された結果により「宮崎」と猫がしゃべっている事が示している。本プログラムは同図に示すようにA、B-1、B-2、B-3、Cのグループで構成されている。Aは上記でクラウドに保存した学習モデルのURLであり、このブロックを実行する事で、Scratchから画像認識結果を利用できるようになる。B-1、2、3はそれぞれ判別された画像に対してどのような処理を行うかを定義している。CのグループはIFTTT™を制御す



図4 学習結果のクラウドモデル

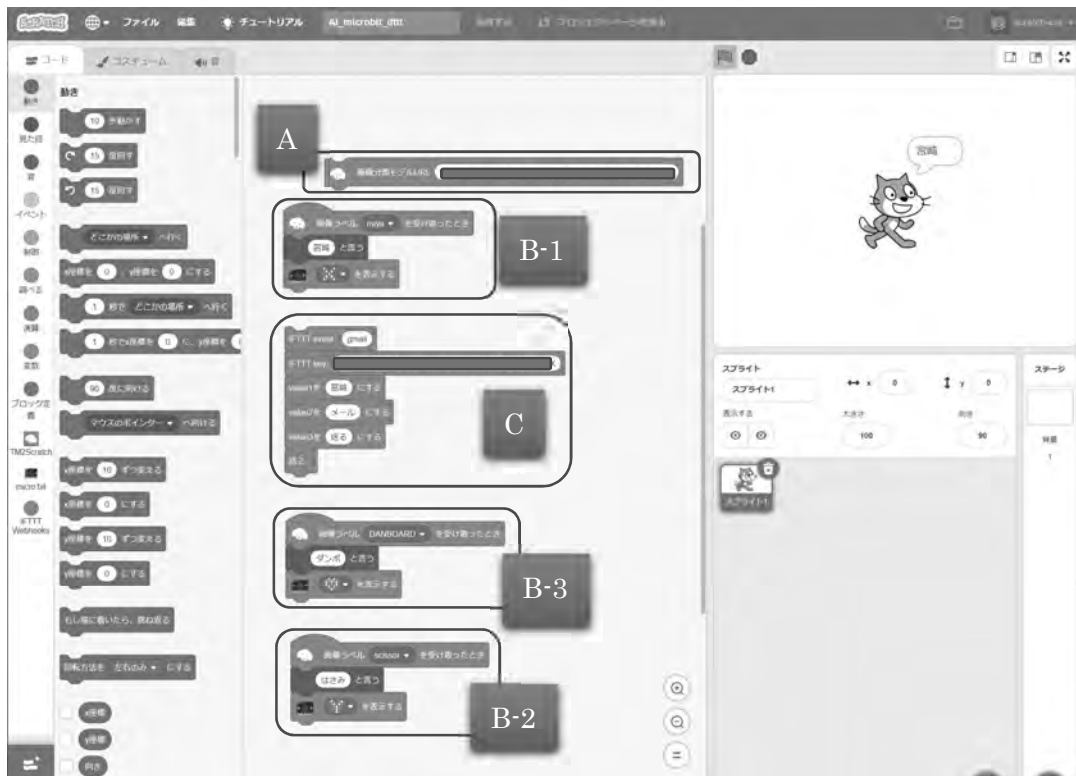


図5 カスタマイズされたScratchの画像認識実行画面

るものである。IFTTTは、IF This Then Thatの略であり、「イフト」と呼ばれている。これは、アプリやWebサービス、ネットワークに接続できるIoT製品を連携するプラットフォームであり、ノンプログラミングでTwitterやFacebook、Gmail、InstagramなどのさまざまなWebサービスを連携して1つのサービスにする事が出来る。

このプログラムで画像認識後に実行されるプログラムのブロックを図6に示す。ここでは

- 1) 「画像ラベル (miya) を受け取った時」の1ブロックで画像認識の結果を得る
- 2) 画面上に (宮崎) という文字を表示する「(宮崎) と言う」ブロック
- 3) スピーカーから音声で (宮崎) と喋る「(宮崎) としゃべる」ブロック
- 4) 外部のデバイス (マイクロビット) のLED表示を「X」と表示する「(X) を表示する」ブロック

から構成されている。このように児童・生徒が思考するプログラミング部分は極めて単純なブロックから構成されているが、実世界とインターネットを結びつけることが体感できるシステムとなっている。

更に図7に示すようにこのカスタマイズされたScratchではIFTTTが利用出来る。同図では、「宮崎」という認識が行われると自動的にGmailの制御に移り、送信メールに「宮崎」「メール」「送る」という本文で自動的に付記してメールを発信するものである。ここではわかりやすいようにGmailの制御を行ったが、IFTTTはIoTを目的としたサービスが多数準備されているので、スマート家電の制御も簡単に実行できる環境を誰もが構築可能である。

具体的な応用例とすれば、「画像認識を用いて個人を判別し、家の鍵を自動で開錠する」とか、「音声認識を用いて20度にエアコンを設定して」とかスマートホーム的な制御も簡単に実現出来る。

#### 4. 今後に向けて

本研究では、Teachable MachineとScratchを用いる事で、学習者がプログラミングのコードを書かずに実現出来るAIを用いた画像認識によるIoTシステムを提案した。この教材はAIで画像認識を行い、その結果を実空間のデバイスやインターネット上のプログラムと連携させる事が可能である。これにより、自分独自の画像認識AIシステムが実現された。しかし、この教材が「自らが社会の問題点を考え、それを解決する事をプログラミングを通して学ぶ」事が出来るかまでは、まだまだ検証出来ていない。

今後は、どのような授業構成を提案すれば、学びに向かう力を育む教材となるか、また児童・生徒からどのような問題点が提案されるか、その問題点を本システムで解決できるか等の解決すべき点が多々ある。今後は実際の教育現場において本システムの有効性を検証し、今後のプログラ



図6 画像認識後の処理

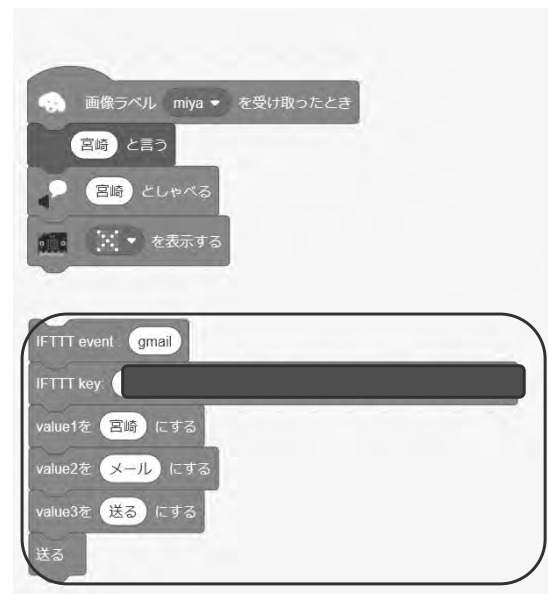


図7 IFTTTによるメール自動送信

ミング教育に役立つものとして改善を行う必要がある。

#### 5. 謝辞

本研究は、令和元年度「重度重複障害児のコミュニケーション獲得性に向けた支援者支援ディープラーニングシステムの試作（課題番号：19K11417）」の一部として行われたことを記して謝意を示す。

#### 6. 参考文献

- i プログラミング教育の在り方について（議論の取りまとめ）、[http://www.mext.go.jp/b\\_menu/shingi/chousa/shotou/122/attach/1372525.htm](http://www.mext.go.jp/b_menu/shingi/chousa/shotou/122/attach/1372525.htm)
- ii Teachable Machine <https://teachablemachine.withgoogle.com/>
- iii Google Teachable MachineでAIを体験してみよう、<https://mekurun.com/tips/teachablemachine/>
- iv SonyのNNCで人工知能を試してみる (02\_binary\_cnn\_sdcproj) <https://newtechnologylifestyle.net/sony%E3%81%AEn>

- nc% E3%81%A7%E4%BA%BA%E5%B7%A5%E7%9F%A5%E8%83%BD%E3%82%92%E8%A9%A6%E3%81%97%E3%81%A6%E3%81%BF%E3%82%8B02\_binary\_cnn-sdcproj/
- v よつばと！ <https://dengekidaioh.jp/product/yotsubato/>
- vi 混同行列 (Confusion Matrix) とは ～ 2 値分類の機械学習のクラス分類について、@TutomuNakamura、<https://qiita.com/TutomuNakamura/items/a1a6a02cb9bb0dcbb37f>
- vii TeachableMachine向けScratch拡張機能 (TM2Scratch) 使い方 (日本語版)、<https://mirapro.mext.go.jp/assets/tm2scratch.pdf>
- viii IFTTT、<https://ifttt.com/home>
- ix IFTTTとは、<https://www.otsuka-shokai.co.jp/words/ifttt.html>