

## 画像認識 AI を使った MS-Word 操作の習熟度分類の試み

### An Attempting to Classify the Proficiency of MS-Word Operation using Image Recognition AI

新井 正一, 吉岡 由希子, 小川 真里江  
Masakazu ARAI, Yukiko YOSHIOKA, Marie OGAWA  
目白大学  
Mejiro University  
Email: m.arai@mrjiro.ac.jp

あらまし：Microsoft Word で作成されたポスター画像を基にその操作スキルを自動的に分類し，多人数のクラス分け基礎資料とすることを目的に，IBM Watson Visual Recognition の活用を検討した。その結果，数少ないトレーニング枚数でも分類は可能であるが完全な分類の自動化には至らず，ある程度人手を介する必要があることが示された。

キーワード：画像分類，習熟度，Microsoft Word，Visual Recognition，アプリケーションスキル

#### 1. はじめに

本学では一般情報基礎科目として，1 年次の入学直後に学科単位で同一時間帯に『情報活用演習 1』が必修科目として開講されている。履修内容は，大学のネットワーク環境の理解から LMS の使い方に至るキャンパスリテラシーを学び，その後，情報の収集，整理，発信の一連のテーマを通して総合的に情報リテラシーを学ぶ。

近年，スマホの普及と共に自宅に家族共有の PC さえ持たない学生が散見され，PC 離れが一段と加速している。当然，PC に対する苦手意識がある学生も多くなり，一般社会で求められるオフィス系アプリケーションの操作能力も年と共に低下傾向にある。その一方で，高度な操作能力を持つ学生も見られ格差が広がっている。一般家庭に PC が普及し始めた頃，大学に入学し始めて PC を操作する学生と，毎日 PC に触れ操作に慣れている学生の格差が大きい時代の再来に近い現象が見られている。

この演習科目では 2012 年以來，多様化した学生に対応すべく，学習意欲，情報に関する知識および PC スキルなどの観点から，テストを実施し習熟度別のクラスを編成している<sup>(1)</sup>。運用を開始した当初は，操作スキルに対する格差は目立たなくなっていたが，学習意欲および情報に関する知識の格差が広がり，特に，意欲に対する配慮が必要になった。しかし，近年，再び PC 操作の習熟度の差が顕著になりつつあり，知識や意欲があっても実技が伴わず授業を進めるにあたって障壁となり，その対応を迫られている。

習熟度テストは，1 回目の授業で実施，翌週の第 2 回の授業までにその結果に応じてクラス分けをおこなっている。このテストを受験する学生は約 1000 名に及び，独自に開発された Web テストを使った知識および学習意欲に関する評価は迅速に処理されるが，実技の評価は人手を必要とし短時間での処理には困難を極める。このため，この数年，実技テスト

の結果はトップクラスに割り振る学生に限定し，操作スキルの確認のみに使われている。

ここでは，この実技テストの採点に画像認識 AI を導入することで，短時間に処理しすべての学生のスキルを把握し，クラス分けに反映する試みを報告する。

#### 2. 実技テスト

実技テストは，MS-Word を使って A4 サイズ 1 枚のポスターを作成する。作成にあたっては，使用するべき図形，表，グラフおよび文字フォントなどの作成要領と見本(図 1)が提示され，これと同じものを作る。



図 1 ポスター作成要領 (左) と見本 (右)

評価にあたってはクラス分けを目的としているため，大まかに分類する程度で詳細な点数化をおこなう必要はない。そこで，『指示どおりできている』から『ほとんどできていない』を Class A から Class D の 4 段階に分類することとした。分類のポイントは，ポスター全体のレイアウト，ポスターを構成する図形，表，グラフ，テキストボックスの要素に着目し次のように定めた。

- Class A:すべての要素が指示どおり使われている。  
 Class B:図形, 表, グラフの各要素が使われているが, グラフが正しく描かれていない。また, テキストボックスが使われていない。  
 Class C:図形, 表のみでグラフおよびテキストボックスが使われていない。  
 Class D:タイトル, 名前などテキストのみで, 表, グラフ, テキストボックスが使われていない。

### 3. 画像認識 AI を活用した分類

#### 3.1 画像認識 AI

評価にあたっては, 作品を画像化し IBM Watson Visual Recognition を使って分類した。この AI は, 一般画像に対して識別機能を持ち, 数千のクラスから適合するクラスを抽出, ラベルとその信頼度を表すスコアと共に提示される。この AI には単に識別するだけでなく, 転移学習を利用していると思われるカスタムクラスが用意され, ユーザが自ら分類クラスを定めユーザ自身が用意したデータを使ってトレーニングし分類できる機能が提供されている<sup>(2)</sup>。この機能を利用すると, 現在, Google からオープンソフトウェアライブラリとして提供されている TensorFlow などの機械学習フレームワークを活用するより, はるかに少ないトレーニングデータで分類精度を高めることができる。ここでは, このカスタムクラスの機能を使って分類を試みている。

#### 3.2 トレーニングおよび検証

トレーニング用および検証用データは, 本年度および昨年度受講した学生の作品を用いた。約 2000 枚の作品から, 前述の評価基準に従って A~D に該当する件数がそれぞれ 100 枚程度になるまで収集する。その後, それぞれのクラスについて 40 枚を 2 組, 20 枚を 1 組のデータセットを用意する。はじめに, トレーニングとして 40 枚のデータを学習させ, 次に, 20 枚のデータを用いて分類精度を測る。さらに, 残りの 40 枚を追加学習させ, 1 回目の 20 枚のデータと同じものを使って分類精度を再度確認する。

分類結果は, 対象となる一枚の画像に対して, A, B, C, D のそれぞれに分類される信頼度が 0~1 の数値として示される。この数値を基に分類結果を以下のような 4 つのカテゴリに分けた。

- T1:最大値を示すスコアが当該クラスと一致し, その値が 0.8 以上である。
- T2:最大値を示すスコアが当該クラスと一致するが, その値が 0.8 未満である。
- T3: 最大値を示すスコアが当該クラスと一致しないが, その値が 0.8 未満で, 当該クラスのスコアが 2 番目に大きい。
- T4: : 最大値を示すスコアが当該クラスと一致せず, その最大値が 0.8 以上である。

図 2 は, それぞれのクラスについて, 横軸にカテゴリー, 縦軸に 20 枚の検証用データの中でそのカテ

ゴリーに属する枚数を表している。図中の 40 および 80 は, トレーニングデータの枚数を意味している。カテゴリー分類の閾値 0.8 は経験的に定めたもので, この値より大きなスコアを記録するクラスが一つあると, その他のクラスのスコアは極めて小さくなる。これは, AI の判断に揺らぎがないと推測される。

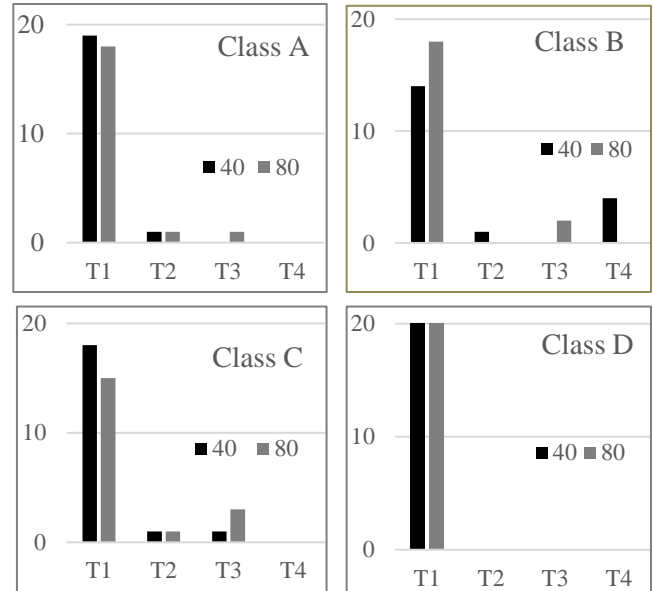


図 2 検証結果

図 2 から次のことが言える。

- ・ Class A および D は 40 枚のトレーニング枚数で分類可能であるが, B, C の分類には不十分である。特に, B ではカテゴリー T4 に 20 枚中 3 枚が誤認識されている。
- ・ しかし, トレーニングデータ数を増加させる事によって改善され T4 に属するケースはどのクラスにも認められず, 誤認識は解消されている。

#### 4. おわりに

このシステムを活用した実技テスト採点の完全自動化は難しく, カテゴリー T2, T3 に該当するデータについては, 人の目による分類が必要となる。しかし, カテゴリー T4 の誤認識を除去するトレーニングをおこなうことによって, 従来の採点の労力軽減にはつながるもの考えられる。ただし, 毎年問題を変更するとなると, その都度トレーニングが必要となり, その枚数を考えると実用性に疑問が残る。

#### 参考文献

- (1) 新井正一・小川真里江・吉岡由希子: 文系大学にみる新入生のパソコン離れ - 一般情報教育での 7 年間のアンケート調査から -, 2018 PC カンファレンス, CIEC
- (2) 立花 隆輝: Watson とロボットの音声対話機能, 日本ロボット学会誌, 第 35 巻, 第 3 号, p. 199-202(2017)