

アモーダル補完を応用した文字型 CAPTCHA の生成手法の検討

上妻拓也^{1†} 梅澤猛² 大澤範高²

概要: Web サービスの悪用を防ぐために CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) を使ったボットと人間の判別が広く行われている。一般的に文字型 CAPTCHA が利用されているが、機械学習を応用した自動文字認識技術によってボットに突破されてしまう事例が増加している。本稿では、視覚の補完機能であるアモーダル補完を応用し、人間には負担が少なく、自動文字認識には攻撃コストと難度が高い CAPTCHA を生成する手法について検討を行った。対象とする CAPTCHA は、背景色の図形によって文字の一部を欠けさせた欠損画像と、欠損画像上の欠けた部分を覆い隠すことができる図形を描画したマスク画像の2枚を提示する形式をとる。欠損部にマスクがかかるように2枚の画像を重ね合わせると、アモーダル補完の効果により人間にとっては容易に文字認識ができる。画像の重ね合わせ操作は、ボットが攻撃に必要とするコストを増大させると期待できるが、欠損画像のみから文字認識されると効果がない。そこで、ノイズの付加、ノイズ集合の混入、文字幅・間隔の可変の3つの難読化手法を使った文字型 CAPTCHA について検討し、生成した欠損画像に対して Convolutional Neural Network (CNN) による自動文字認識率を調べた。また、画像の重ね合わせと操作が、ユーザに対してどの程度負荷となるか、正解率に影響があるかを調べるため、ユーザによる文字列読み取り実験も行った。

1. はじめに

Web サービス利用者が人間かボットかを判別する CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) において、画像中に含まれる文字列を解答する文字型ものは現在広く利用されているが、機械学習技術を応用した自動文字認識により、ボットによって突破することが可能であることが示されている。文字型 CAPTCHA 画像を複雑化させる対策が多く取られているが、深層学習を利用した文字認識に対しては効果が少なく、人間の負担が増すだけになるという問題がある。

本研究では、視覚の補完機能であるアモーダル補完を応用し、人間には負担が少なく、自動文字認識には攻撃コストと難度が高い CAPTCHA について検討を行った。

2. 関連研究

Web サービスをボットによる悪用から防ぐために CAPTCHA [1] が利用されている。CAPTCHA はサービス利用者にテストを出題し、解答の内容によって人間とボットの判別を行う。このテストは人間にとっては容易で、ボットに対しては困難となるように設計されている。CAPTCHA の例として次の4種類が挙げられる [2]。

(1) 文字型

文字列を含む画像を提示し、画像中の文字列を解答させる。

(2) 分類型

グループ分けした幾つかの画像群と、グループ分けされていない1枚の画像を表示する。画像群からグループ分けの法則を見つけ、グループ分けされていない1枚の画像がどのグループに属すかを解答させる。

(3) 画像型

大規模な画像データベースから画像をピックアップして提示し、画像中に写っている物体の名称を解答させる。

(4) 音声型

単語や何桁かの数字を発話している音声を歪ませたものを再生し、内容を解答させる。

本研究では、この中で最も広く利用されている文字型 CAPTCHA を対象とする。

2.1 文字型 CAPTCHA

文字型 CAPTCHA は文字列を含む画像を提示し、画像中の文字列を解答させることで人間とボットの判別を行う。文字型 CAPTCHA は画像に描画する文字、文字数、文字の並びなどを変化させることで様々な文字列画像を容易に生成することができるため、生成コストが小さい。また、世界中の多くの人が扱うことのできるアルファベットや数字を利用することで、グローバルな環境にも対応可能であることから、広く利用されている。

歪みやノイズを加えた画像内の文字を自動認識することは困難であったため、ボット対策として利用されてきたが、光学文字認識 (OCR) 技術の発展や機械学習の応用による自動文字認識精度の向上に伴い、既存の多くの文字型 CAPTCHA がボットに対して脆弱であることが示されている [3] [4] [5]。

2.2 文字型 CAPTCHA に対する攻撃

ボットによる文字型 CAPTCHA に対する攻撃は、前処理、領域分割、単一文字認識の3つのステップで構成される。前処理ではテキストエリアの切り出し、回転、ノイズの除去、二値化などを行い、領域分割と単一文字認識の精度を高める。領域分割では CAPTCHA 画像をそれぞれが

1 千葉大学大学院融合理工学府
Graduate School of Science and Engineering, Chiba University
2 千葉大学大学院工学研究院
Graduate School of Engineering, Chiba University

† 現在、(株) コナミデジタルエンタテインメント
Current affiliation is Konami Digital Entertainment Co., Ltd.

単一の文字を含む小さい領域へ分割する。単一文字認識では分割した単一文字画像から機械学習を利用して文字の認識を行う。

Convolutional Neural Network (CNN) による単一文字認識精度は人間の文字認識精度を上回っているため、文字型 CAPTCHA の堅牢性は領域分割の難度に依存している [3] [4]。そのため、ホロウスキーム [7]、Crowded Characters Together (CCT) [8]、2層構造 [9]などの様々な領域分割への対策が多くの CAPTCHA に使用されているが、それぞれの対策を無効化する前処理や、深層学習を組み合わせた領域分割手法が提案されており、高い精度で文字認識が行われてしまうことが明らかとなっている [5] [6]。

Yu らは、前処理と領域分割の画像処理を別々に行う従来の方法ではなく、CNN を用いて CAPTCHA 画像から文字認識を行う手法を提案している [10]。前処理と領域分割を行う従来の方法では、攻撃対象の CAPTCHA が使用している領域分割対策によって適用する画像処理を切り替える必要があり、未知の領域分割対策に対しては有効な画像処理やアルゴリズムを新たに開発する必要があった。Yu らの手法では CAPTCHA 画像を直接 CNN への入力とし、前処理と領域分割の過程を CNN に学習させることでこれを回避している。5文字からなる文字列を含む画像に、ノイズや歪みを加えて生成した CAPTCHA 画像データセットによる評価において、前処理と領域分割を行う従来の手法による文字認識精度 0.95 に対し、0.965 の精度であると報告している。

3. アモーダル補完を応用した CAPTCHA

本研究では、背景色の図形によって文字の一部を欠けさせた欠損画像と、欠損画像上の欠けた部分を覆い隠すことができる図形を描画したマスク画像の2枚を提示する手法を提案する。欠損画像、マスク画像の例をそれぞれ図 1 (a), (b) に示す。欠損画像は背景色の図形によって文字が欠けているため、画像中の背景色部分が背景か欠損部かの判別がつかない。これにより、どの断片同士が1文字を構成しているかの把握が困難になり、文字認識を行うことが難しくなっている。一方、マスク画像にはマスクのみが描画され、文字の情報を含まない。そのためマスク画像から文字認識を行うことも不可能である。



(a) 欠損画像 (b) マスク画像
図 1 欠損画像とマスク画像の例

この手法では、解答者が2枚の画像を適切な位置に重ね合わせ、合成画像を作成することで文字認識が可能となる。合成画像の例を図 2 に示す。人間はアモーダル補完の効果

により、マスクと文字の前後関係を把握し、マスクによって遮蔽された部分を自然な形で補完することで文字認識が容易になる。ボットもマスクによって背景と欠損部が明確になり、合成画像から文字認識を行うことが可能となるが、ボットは合成画像を作成するために画像の重ね合わせを総当たり的にやり、作成したすべての合成画像に対して文字認識を行う必要があると考えられるため、攻撃コストの増加が期待できる。



図 2 合成画像の例

4. CNN による認識実験

欠損画像から文字認識が行われてしまうと、ボットに画像の重ね合わせを要求することができない。そこで、欠損画像がボットに対して耐性を持つかを、CNN による文字認識で評価した。また、欠損画像に、ノイズ、ノイズ集合、文字サイズの可変化を加えることで耐性の向上を図った。

4.1 実験条件

(1) 文字画像

評価実験には縦横 40 px の文字画像を用いた。40×40 px の領域を縦横 5 等分した 5×5 マスを使用して (図 3) アルファベットの太文字 26 文字を用意した (図 4)。

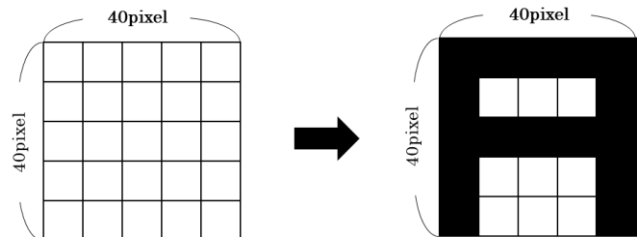


図 3 文字画像の作成例 (例: 'A')



図 4 作成した文字画像

(2) 欠損画像

文字画像に欠損箇所をつくる背景色図形は、対象となる文字画像内 (図 5 (a)) で無作為に選択した座標 (図 5 (b)) を中心とする 16×16 px の正方形 (図 5 (c)) とした。確実に欠損箇所が生じるよう、文字色の座標のみを選択対象とした。また、欠損箇所が偏ることを避けるため、背景色図

形は互いに 6 px 以上の間隔を空けることとした。

文字色部分の少ない文字 (例: 'I') と多い文字 (例: 'W') の欠損が同程度となるように、描画する背景色図形の数はいくつ (文字の黒画素数/256) の範囲から無作為に決定した。

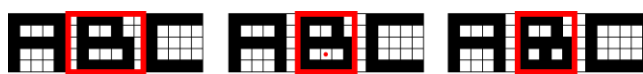


(a) 選択可能範囲 (b) 座標の選択 (c) 欠損の生成

図 5 欠損画像の作成例

(3) ノイズ付加

対象となる文字画像内で無作為に選択した座標を含む 8 × 8 px のマスについて、文字色と背景色を反転させることでノイズの付加を行った (図 6)。



(a) 選択可能範囲 (b) 座標の選択 (c) ノイズの付加

図 6 ノイズの付加例

(4) ノイズ集合

CNN を用いた単一文字認識は、欠損やノイズを含む文字であっても非常に高い精度を示す。そこで、文字色正方形のノイズで作成した一文字分の集合を混入することで認識精度に与える影響を調べた。文字画像の作成時の 5 × 5 マスから 3 ~ 7 マスをランダムに選択して文字色とすることでノイズ集合を作成した (図 7)。ただし、人間が文字の一部だと誤認識してしまうのを避けるため、同じ列・行で文字色とするのは 3 マス未満という制約を設けた。文字列中に含まれる 0 ~ 1 文字を無作為に選択しノイズ集合と入れ替えた。

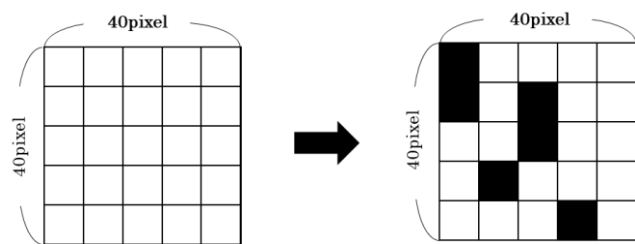


図 7 ノイズ集合の作成例

(5) 文字幅・間隔の可変性

文字列サイズが均等であると学習モデルによる文字分割が容易であると考えられるため、文字幅や文字の間隔を可変とすることで認識精度を低下させられる可能性について調べた。

文字画像の作成時の 5 × 5 マスを単位とし、列ごとに {2, 4, 6, 8, 10, 12} px の中から無作為に幅を変更した (図 8)。ただし、文字のバランスが大きく崩れるのを避けるため、対象とする列は文字の横線が描かれた列と文字間のみとし

た。

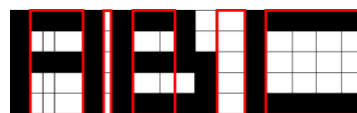


図 8 文字幅・間隔の変更例

(6) データセット

アルファベット大文字 26 文字の中から重複を許す 3 ~ 5 文字を無作為に選択して文字列を作成し、学習用セット 10,000 件、バリデーションセット 2,000 件、テストセット 1,000 件を合わせてデータセットを構成した。

(7) CNN 構造

CNN の実装には、深層学習フレームワークである TensorFlow [12] をバックエンドとしたライブラリである Keras [13] を用いた。CNN のモデルは、Yu らによって提案された文字型 CAPTCHA 認識用モデルを使用した。これは中間層に畳み込み層 10 層、プーリング層 4 層、全結合層 3 層を含む全 17 層からなるディープニューラルネットワークである。4 回目のプーリング層の後、モデルは 5 つに分岐し、各出力がそれぞれ文字列中の 1 文字に対応している。このモデルは従来の文字型 CAPTCHA に対して 0.965 の認識精度を示している。

学習時間短縮のため、CNN への入力の際には画像の縦横を半分にリサイズし、コントラスト正規化を行った。バッチサイズは 128、エポック数は 300 とした。学習係数の最適化には SGD (Stochastic Gradient Descent) を使用し、学習係数は 0.001 とした。なお、ドロップアウトは全ての全結合層後に行い、パラメータは 0.5 とした。

4.2 実験結果

ノイズ付加、ノイズ集合、文字幅・間隔の可変性を組み合わせた際の、300 エポックの学習を行った CNN モデルによるテストセット 1,000 件に対する正解率を表 1 に示す。

表 1 条件の組み合わせに対する CNN の正解率

ノイズ付加	ノイズ集合	文字幅・間隔	正解率
—	—	—	0.946
—	—	✓	0.917
✓	—	—	0.914
✓	—	✓	0.788
—	✓	—	0.959
—	✓	✓	0.912
✓	✓	—	0.913
✓	✓	✓	0.830

最も低い正解率を示したのは、ノイズの追加と文字幅・間隔の可変性を行ったテストセットの 0.788 であった。一方、最も高い正解率を示したのは、ノイズ集合のみを加えたテストセットの 0.959 であった。

ノイズの追加、文字サイズ可変化は個別に行うことで正解率を低減できていることから、有効な手法であることが示唆された。ノイズ集合については追加することで正解率が高くなっている。学習モデルはノイズ集合を高い精度で認識可能であり、ノイズ集合を追加することは、分類の簡単なクラスを1つ追加することと同じになってしまったためであると考えられる。

5. ユーザによる認識実験

提案手法は、画像の重ね合わせをユーザに要求するため、操作負荷や誤解答の増加が懸念される。そこで、ユーザによる CAPTCHA 読み取り実験を行った。

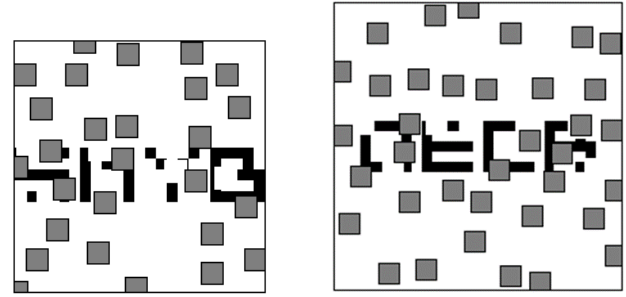
5.1 実験条件

欠損画像とマスク画像を提示し、画像の重ね合わせを行うことができる実験用アプリケーションを PC 上に作成した。ユーザは、マウスまたはキーボードの方向キーによって画像の重ね合わせ操作を行うことができる。操作方法の違いによる影響を調べるため、操作条件を変えたタスクを用意した。実験参加者は大学学部生および大学院生 6 名であった。

ユーザにはまず今回使用する文字画像の一覧を提示し、各アルファベットがどの文字画像に対応するかを確認させた。

欠損画像の生成手法が正答率と操作負担へ与える影響を調べるため、A) 欠損のみ、B) ノイズを付加、C) 文字幅・間隔を可変とする、D) ノイズを加え、文字幅・間隔を可変とする 4 種類の欠損画像を問題セットとして用意した。

各問題セットは、次のタスク 1~4 から構成した。1) 欠損画像のみを提示し、マスク画像を使わずに読み取りを行う、2) 欠損画像に対して、あらかじめ適切な位置にマスク画像を重ねた状態で読み取りを行う、3) 欠損画像の上下移動とマスク画像の回転で画像を重ね合わせて読み取りを行う、4) 欠損画像の上下左右移動で画像を重ね合わせて読み取りを行う。タスク 3 においては、欠損画像の横幅を一辺の長さとする正方形のマスク画像を用い、マウスによるドラッグおよびホイールスクロールと、キーボードの上下キーを使って欠損画像の上下移動を、マウスの右クリックと、キーボードの左右キーを使って 90 ずつでのマスク画像回転を行うものとした。タスク 4 においては、欠損画像の横幅+40 px を一辺の長さとする正方形のマスク画像を用い、マウスによるドラッグおよびホイールスクロールと、キーボードの上下左右キーを使って欠損画像の上下左右移動を行うものとした。



(a) タスク 3 (b) タスク 4

図 9 欠損画像とマスク画像の例

各問題セットにおいて、タスク 1~3 は 20 問、タスク 4 は 10 問で構成され、各タスクにおいて半数は全ユーザ共通の問題、残り半数はユーザ毎に異なる問題とした。また、問題セットの提示は各ユーザで異なる順番で行った。各タスクが終了する毎にユーザに対し「本タスクで提示された CAPTCHA の解答負荷は重い」という質問を行い、1 (強く不同意) ~5 (強く同意) の 5 段階で回答を得た。

5.2 実験結果

図 10 に、各問題セットにおける各タスクについての平均解答時間と、各ユーザの平均解答時間中の最大値と最小値をエラーバーで示す。タスク 2 に対する平均解答時間が最も短くなり、ユーザ間でも差があまり生じていないことがわかる。対して、画像の重ね合わせを行うタスク 3 とタスク 4 ではタスク 2 の倍以上の平均解答時間となり、ユーザ間で大きく差が生じていることがわかる。

図 11 に、各問題セットにおける各タスクについての平均正解率と、各ユーザの平均正解率の中での最大値と最小値をエラーバーで示す。平均解答時間と同様に、タスク 2 に対する平均正解率が最も高くなり、ユーザ間で差があまり生じていない。対して、画像の重ね合わせを行うタスク 3 とタスク 4 では平均正解率が低くなるものの、タスク 1 より高い値を示した。

図 12 に、各問題セットにおける各タスクについてのアンケート結果を示す。画像の重ね合わせを行うタスク 3 とタスク 4 に対しては解答負荷が重いと答えるユーザが多いことがわかる。

6. 考察

CNN による認識実験においては、ノイズの付加、ノイズ集合の混入、文字幅・間隔の可変化の 3 つの手法をそれぞれ組み合わせることで、文字認識精度の低下が可能か検証を行った。この結果、全ての場合において、ノイズの追加を行うことで正解率を低下できていることがわかる。文字幅・間隔の可変化についても同様である。ノイズの付加と文字幅・間隔の可変化の手法を組み合わせることで正解率は 0.788 となり、それぞれの手法を個別で行うよりも正

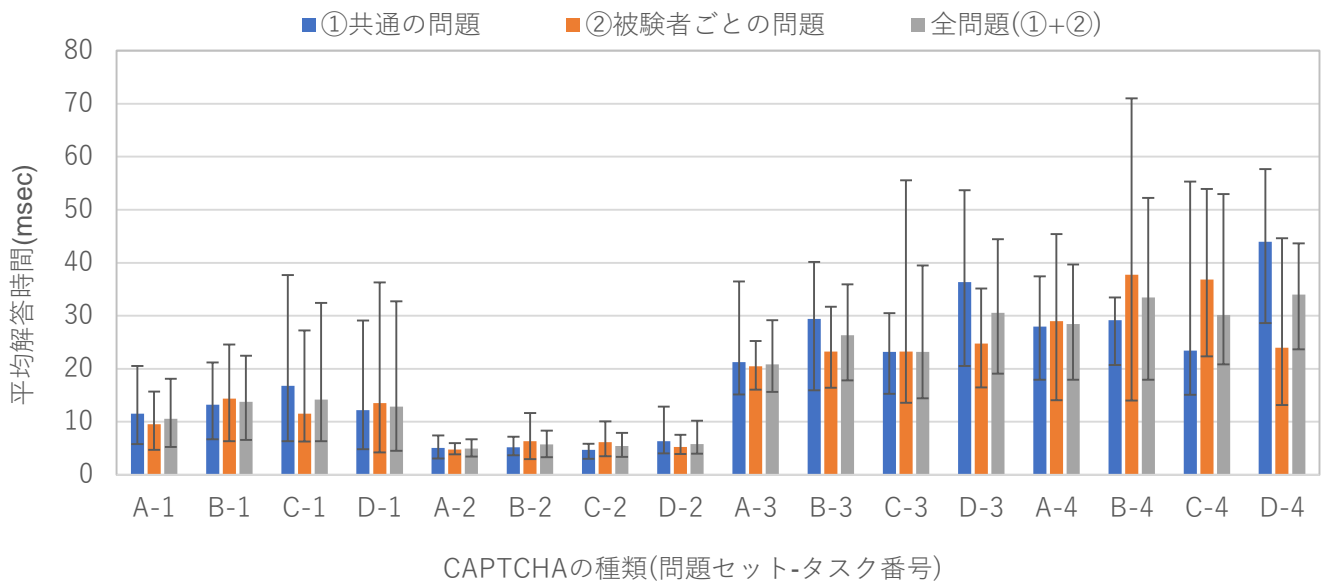


図 10 ユーザによる文字列読み取り時間

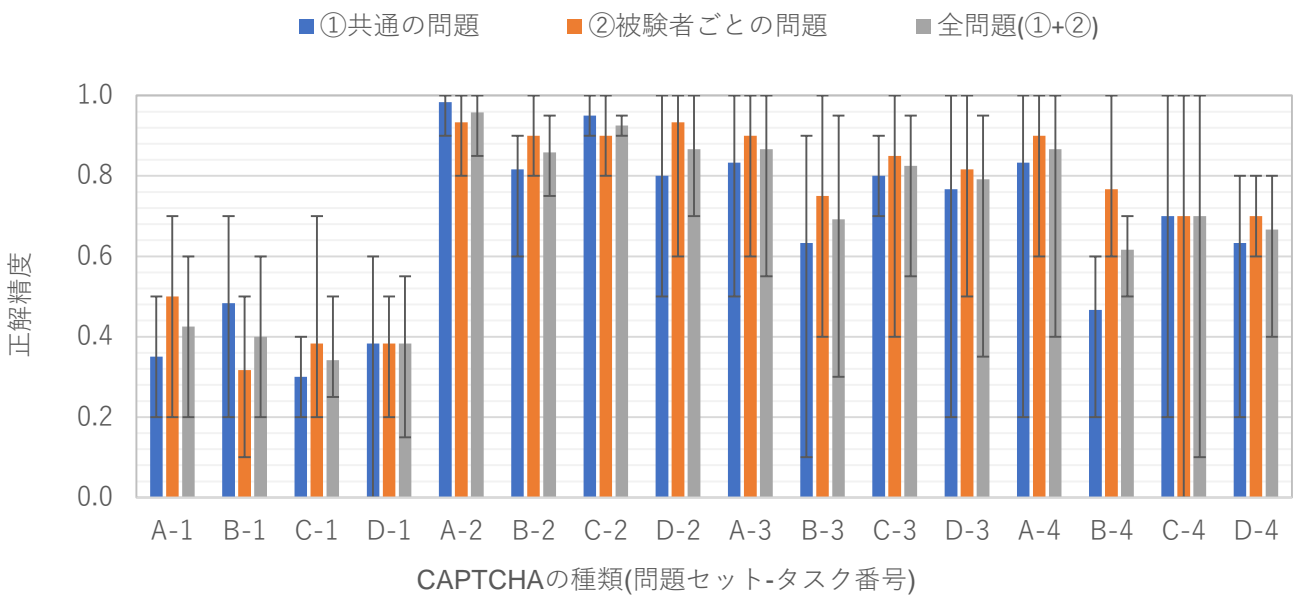


図 11 ユーザによる文字列読み取り正解精度

解率を低下させることができることを確認した。一方で、3つの手法全てを適用した場合には、正解率が0.830となったことから、ノイズ集合は他の手法と組み合わせても正解率の低下に対して効果がないと考えられる。

Yuらは、文字型CAPTCHAに対してCNNを用いて0.965の文字認識精度を示している。それに対し、本実験ではノイズの付加、文字幅・間隔の可変化を行った場合の欠損画像に対して、同じCNNモデルを使用し0.788の文字認識精度を示した。このことから、欠損画像は文字認識に対する耐性は従来の文字型CAPTCHAよりも高く、提案手法がボットに対して堅牢性を示していることが示唆さ

れる。しかし、実用的なCAPTCHAではボットの突破率が0.1%未満であることが要求される [6] ことから考えると、欠損画像に対する自動認識精度をさらに低下させることが必要である。

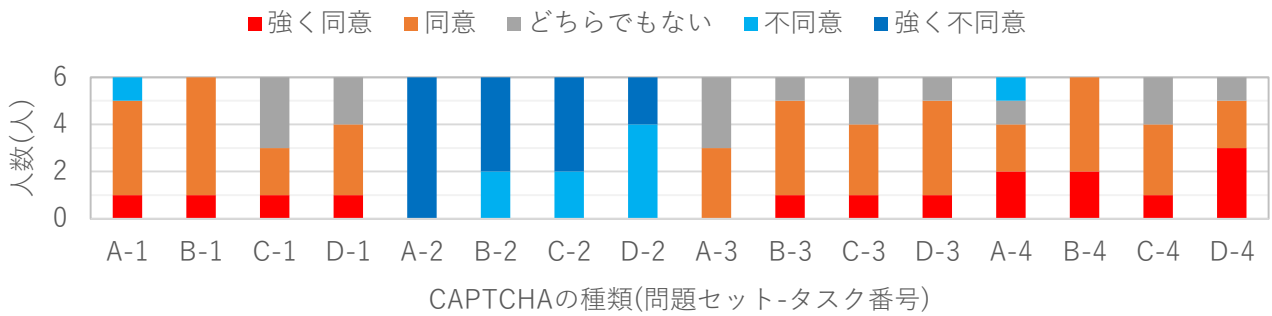


図 12 操作負荷についての解答 (Q: 本タスクで提示された CAPTCHA の解答負荷は重い)

ユーザによる認識実験においては、欠損画像のみを表示するタスク 1 の平均正解率が 0.50 以下なのに対し、合成画像を表示するタスク 2 の正解率が 0.85 以上を示している。また、タスク 2 に対する平均解答時間は 10 秒未満であり、他のタスクの半分以下であることがわかる。さらに、タスク 1 を解答負荷が重いと感じたユーザが多かったのに対し、タスク 2 に対して解答負荷を感じたユーザはいなかった。以上のことから人間は欠損部分をマスクで隠すことによって起こるアモダル補完の効果によって、合成画像からであれば、文字を認識しやすいということがわかった。しかし、タスク 3、タスク 4 ではタスク 2 よりも正解率が低くなっており、平均解答時間の点でも、タスク 3、タスク 4 は 20 秒以上の解答時間を要し、ユーザによっては 1 間に 1 分以上かかる例も見られた。タスク 3、タスク 4 において、画像の重ね合わせを適切に行うことができれば、タスク 2 の合成画像と同様の画像を作成することができるため、タスク 2 とタスク 3、タスク 4 の間に正解率と解答時間の差が生じたのは、画像の重ね合わせを適切に行うことが難しく、合成画像を作成することができなかったからであると考えられる。解答負荷の調査でも、タスク 3 とタスク 4 に対して負荷が重いと回答したユーザが多いため、本実験で行った画像の重ね合わせ方法はユーザに対して難易度が高く、解答負荷が重くなってしまっていることがわかった。

問題セットの違いによる影響については、タスク 3 とタスク 4 における平均解答時間に注目すると、A, C よりも B, D のほうが、解答時間が長くなっている。B, D はノイズを含んでいるため、ユーザはノイズと文字の判別がつかず、ノイズ上にマスクを重ねることを避けてしまうことが影響していると考えられる。同様の現象により、正解率に関しても A, C に比べて B, D のほうが低くなったと考えられる。

本実験の結果から、タスク 3、タスク 4 のような画像の重ね合わせ方法はユーザにとって解答負荷が重く、解答時間が長くなり、正解率にも影響がでてしまうことが示された。画像の重ね合わせの方法をより簡単化するなど、ユーザに対する負担を低下させる方法について更に検討を行う

必要がある。ただし、画像の重ね合わせパターンが減少すると、ボットの攻撃コストが低下する懸念もある。一方、タスク 2 に対しては解答時間が非常に短く、正解率も 0.85 以上であり、アンケート結果からもユーザは解答負荷を感じていない。そのため合成画像を作成することができれば、アモダル補完の効果によって文字を認識しやすくなるということは確認できた。

7. まとめ

本研究では、人間の持つ視覚の補完機能であるアモダル補完を応用し、人間には負担が少なく、自動文字認識には攻撃コストと難度が高い CAPTCHA を生成する手法について検討した。背景色の図形によって文字の一部を欠けさせた欠損画像と、欠損画像上の欠けた部分を覆い隠すことができる図形を描画したマスク画像の 2 つを提示し、欠損部にマスクがかかるように 2 つの画像を重ね合わせることで、アモダル補完の効果により人間には文字認識が容易になる。

本手法のボットに対する堅牢性は、欠損画像が持つ文字認識耐性に依存するため、欠損画像の文字認識耐性を CNN 学習モデルによって評価した。3 種類の欠損画像生成手法について検討を行った結果、ノイズの付加、文字幅・間隔の可変化によって、正解率を低減できることを確認した。さらにノイズの付加と文字幅・間隔の可変化を組み合わせることで、欠損画像の文字列全体の自動認識の正解率を 0.946 から 0.788 に低減可能であることを明らかにした。

提案手法がユーザに要求する画像の重ね合わせという動作によって、人間に対する誤解答の増加や操作負荷の増大が懸念されたため、ユーザによる文字列読み取り実験を行った。この結果、実験で行った画像の重ね合わせに対して操作負荷を感じるユーザが多く、解答時間の増加や正解率の低下が確認された。一方、画像の重ね合わせを正確に行うことさえできれば、アモダル補完の効果によって文字の認識が容易になることも確認できたため、画像の重ね合わせ方法を改良することで、重ね合わせ操作への負担感軽減や解答時間の短縮が期待できると示唆された。

参考文献

- [1] Luis von Ahn, Manuel Blum, Nicholas Hopper, and John Langford: “CAPTCHA: Using hard AI problems for security,” International Conference on the Theory and Applications of Cryptographic Techniques, Springer, Berlin, Heidelberg, pp.294-311, 2003.
- [2] Luis von Ahn, Manuel Blum, and John Langford: “Telling humans and computers apart (automatically) or how lazy cryptographers do AI,” Advanced in Cryptology. (Lecture Notes in Computer Science) pp.294-311, 2003.
- [3] Kumar Chellapilla, Kevin Larson, Patrice Simard, and Mary Czerwinski: “Computers beat humans at single character recognition in reading based human interaction proofs,” The Second Conference on Email and Anti-Spam, 2005.
- [4] Jeff Yan, and Ahmad Salah El Ahmad: “A Low-cost Attack on a Microsoft CAPTCHA,” Proceedings of the 15th ACM conference on Computer and communications security. pp.543-554, 2008.
- [5] Rafaqat Hussain, Hui Gao, and Riaz Ahmed Shaikh: “Segmentation of connected characters in text-based CAPTCHAs for intelligent character recognition,” Multimedia Tools and Applications 76.24 (2017): 25547-25561, 2017.
- [6] Mengyun Tan, Haichang Gao, Yang Zhang, Yi Liu, Ping Zhang, and Ping Wang: “Research on Deep Learning Techniques in Breaking Text-based CAPTCHAs and Designing Image-based CAPTCHA,” IEEE Transactions on Information Forensics and Security 13.10 (2018): 2522-2537, 2018.
- [7] Haichang Gao, Wei Wang, Jiao Qi, Xuqin Wang, Xiyang Liu, and Jeff Yan: “The robustness of hollow CAPTCHAs,” ACM SIGSAC Conference on Computer Communication pp. 1075-1086, 2013.
- [8] Haichang Gao, Wei Wang, Ye Fan, Jiao Qi, and Xiyang Liu: “The robustness of ‘connecting characters together’ CAPTCHAs,” J. Inf. Sci. Eng., vol. 30, pp. 347-369, 2014.
- [9] Haichang Gao, Mengyun Tang, Yi Liu, Ping Zhang, and Xiyang Liu: “Research on the security of Microsoft’s two-layer Captcha,” IEEE Transactions on Information Forensics and Security, vol.12, no.7, pp.1671-1685, 2017.
- [10] Yu Hu, Li Chen, and Jun Cheng: “A CAPTCHA recognition technology based on deep learning,” 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018.
- [11] 森拓真, 宇田隆哉, 菊池真之: “アモーダル補完を利用した動画 CAPTCHA の提案,” マルチメディア, 分散, 協調とモバイルシンポジウム 2011 論文集, pp. 1518-1525, 2011.
- [12] Google: “TensorFlow,” <<https://www.tensorflow.org>>, (参照 2019-05-12).
- [13] Keras Google group: “Keras Documentation,” <<https://keras.io/>>, (参照 2019-05-12).