画像選択時の個人の好みによる CNN を用いた認証方式

2018SC003 浅野菜摘
指導教員：河野浩之

1　はじめに

現在，多くの web サイトにて文字パスワードが利用されている．その中でも 2020 年のトレンドマイクロの調査 [1] によると，複数のサービスで ID やパスワードを使いまわしている利用者が 85.7 %である．また，ID やパスワードを使いまわす理由として，異なるパスワードを設定すると忘れてしまうと答えた利用者が 71.4 %，異なるパスワードを考えるのが面倒と答えた利用者が 49.4 %である．

このように文字パスワードを使いまわすことについてのリスクを感じている利用者は多いものの，パスワード設定に関わる課題がある．この調査結果を踏まえ，特定の情報を用いて認証を解除する鍵として扱う記憶負担の少ない画像認証を提案する．

2　画像認証の先行研究

表 1 に画像を利用した本人認証の先行研究を示す．持田ら [4] の研究は，初めから認証鍵となる画像を登録せずにコンピュータが自動生成した画像から選択した画像の色や特徴量を抽出し，グラフや数値を用いた比較を行い本人を判別するものである．この手法は特定画像を保持しないため，不正アクセスのリスクを減らすことが可能である．

しかし，彼らの研究では選択画像から本人であると判別する特徴として利用した色の類似度は 8 割であったが特徴点として抽出したデータの類似度が 1〜2 割であり本人判定に利用できなかった．この特徴点が上手く利用できなかった原因として，色と特徴量を分けて類似度を計算しており，画像も線の図形で単調であったことがあげられる．

表 1　画像認証の先行研究の比較

著者	特徴	課題
森ら [2]	認証画像群でパス画像を一部分表示	認証を行うほどパス画像を特定が容易
Gho ら [3]	画像登録時に 3 つのルールで 12 枚の画像を登録	ルールをパス画像と共に覚えておく必要性有
持田ら [4]	画像登録時に特定の画像パスワードを持たない	単調な図形であり，特徴量を利用した類似度が低い

3　ランダム画像を用いた認証手法の提案

本研究の提案手法は，パスワードやパス画像を覚えておく負担を減らすことを目的とし，認証を解除する鍵として個人を判別するための特定の画像を利用せず認証を行うものである．本人固有の画像を利用しないという点では [2] らの研究と同様であるが，彼らの研究での課題であったランダムに作成する画像が単純であった点は画像を用いることによって複雑化を行い，画像の特徴の抽出法が色と特徴点を別物として扱っていたため，本手法ではこれらを一つにまとめて抽出することによって精度を上げを測る．

この認証手法を行う上で，個人の嗜好をコンピュータが学習するための登録と本人であるか判断を行う認証の 2 つに分けた操作を行う．

初めに，個人の嗜好をコンピュータが学習するために必要な画像のデータを取得するための登録フェーズのフローチャートを図 1 に示す．また，手順は以下の通りである．

(1)　ランダムにフィルタ加工を行った画像を複数枚提示．

(2)　被験者が表示画像群から好みの画像を 1 枚選択．

(3)　(2) での選択画像とそれ以外の非選択画像を異なるフォルダに保存．(2 つのフォルダ)

(4)　学習を行うために指定した枚数の画像が選択されるまで任意の回数 (n 回) 試行を行う．n は自然数

(5)　保存した選択画像 n 枚と非選択画像を 2 つのカテゴリに分け，それぞれのカテゴリの画像が持つ色や特徴をディープラーニングを用いて抽出し，個人データとして保存を行うことで登録フェーズは終了．

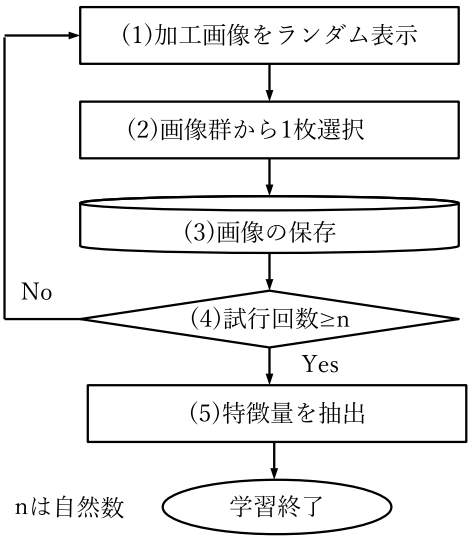

図 1　画像登録時のフローチャート

次に認証フェーズは，登録フェーズの (5) で抽出したデータと，登録時と同様にランダムに表示された画像群から新たに選択した画像を用いて選択画像の類似度を出力し，登録者本人であるか，否かの判別を行う．

4　実験結果

本研究で利用する画像は，kaggle（https://www.kaggle.com）に掲載されている，Fashion Product Images データセットの服の画像 50 枚である．服には色や柄の特徴が多くあり，画像の背景が白色で統一されている点からこのデータセットを利用する．

また，本手法ではこれらの画像加工を HSV を用いて色彩変換することによって行う．HSV 色空間は OpenCV 上で，色相 (H)0〜179，彩度 (S) と明度 (V) は 0〜255 の範囲をとる．よってランダムに色相を 3 ずつ変化させた値，彩度は 15 ずつの値を利用しこの範囲で回転させ，明度は回転させると値によって色飛びが見られたため，元画像から 0.85，0.9，0.95，1 の 4 パターンの倍率を利用しランダムに組み合わせて画像の加工を行った．また，背景の白色はどの画像も共通の値であったため背景以外の部分の色彩変換を行った．

画像の登録や認証を行う際に表示する画像選択画面は，図 2 であり，GUI ライブラリ Tkinter を用いて作成した．表示される画像は 1 画面に 9 枚とし，用意した 50 枚の画像から 9 枚をランダムに選定し，HSV 色彩変換を行ったものである．

図 2　画像選択を行う GUI 画面と遷移の様子

本実験では被験者 6 名に対し実験を行った．各被験者に登録フェーズで画像を 50 回選択させ，選択画像 50 枚，非選択画像 400 枚を収集した．この収集した画像を選択画像，非選択画像の 2 つのカテゴリに分け CNN を用いて分類器を作成した．利用したパラメータは，出力時の活性化関数は出力値の合計が 1 となり，予測値としてそのまま利用できる Softmax，最適化手法は Adamax，学習率は 0.001，エポック数は 7，バッチサイズは 5 である．

また，認証フェーズで選択した画像 10 枚を 1 回の認証時に用いることとし，各被験者には 10 枚× 10 回の計 100 枚選択させた．この画像を先の分類器で分類予測を行い出力値が 0.5 を超えた方のカテゴリの画像であるとし，被験者本人が選択した画像のカテゴリと予測された数をカウントする．

登録フェーズで収集した画像 450 枚全てを利用した際の実験を実験 1 とし，被験者 6 名の全 60 試行の認証で本人が選んだ画像であると予測された画像の平均枚数と，最低枚数を表 2 に示した．本人が認証を行ったにも関わらず 1 枚も本人の選択画像であると予測されない試行があった．

この実験 1 での結果から，2 つのカテゴリの画像枚数が不均衡であったことから上手く予測が行えなかったと考え，非選択画像 400 枚をアンダーサンプリングした画像を利用した．方法としては，登録画像選択時に選択した画像の色相平均値に近い値を持つ画像を同画面に表示された非選択画像 8 枚から 2 枚のみ利用し，非選択画像を全部で 100 枚の全 150 枚で同様に CNN を用いて分類器を作成し認証実験を行ったところ平均，最低枚数を増加させることができた．この実験を実験 2 とし，表 2 に結果を示した．

また，被験者本人の認証だけでなく，コンピュータによるランダム攻撃を想定してランダムに選択した画像 10 枚 × 10 回で各被験者に対し同様の認証実験を行ったところ表 2 の結果となった．コンピュータの場合は登録者本人であると予測された画像の最高枚数についてである．

表 2　登録者本人とランダム攻撃による認証実験の結果

実験	認証者	平均枚数	最低枚数 (最大枚数)
1	登録者本人	4.02 枚	0 枚
2	登録者本人	7.68 枚	4 枚
1	ランダム攻撃	0.75 枚	3 枚
2	ランダム攻撃	1.82 枚	4 枚

5　まとめ

登録フェーズで収集した非選択画像をアンダーサンプリングすることで認証時に選択した画像が登録者本人が選択した画像であると予測される平均枚数を 3.66 枚増やすことができた．また，コンピュータでの攻撃を想定した認証での平均枚数も 1.07 枚増えたが登録時の画像選択画面の表示画像，収集方法に改良を加えれば学習精度を改善できる可能性がある．

参考文献

[1] トレンドマイクロ株式会社，“パスワードの利用実態調査 2020,” https://www.trendmicro.com/ja_jp/about/press-release/2020/pr-20200929-01.html, 参照 Jul.2021.

[2] 森 康洋, 高田 哲司, “秘密情報を変更せずに提供しうる安全性を 柔軟に変更可能な再認式画像認証の提案,” 情報処理学会論文誌, Vol.57, No.12, pp.2641-2653, Jun.2016.

[3] Goh Wen Bin, Sohail Safdar, Rehan Akbar, Suresh Subramanian, “Graphical Authentication Based on Anti-Shoulder Surfing Mechanism,” ICFNDS ‘18: Proc. of 2nd ICFNDS, No.20, pp.1–6, Jun.2018.

[4] 持田達範, 稲村勝樹, “個人の嗜好で識別を行う画像認証方式,” コンピュータセキュリティシンポジウム 2017 論文集, Vol.2017, No.2, pp.1500-1505, Oct.2017.