一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
IBISML2019-28(2020-01)

Siamese Networkによる筆記画像と動的情報を用いたハイブリッド認証モデル

林　　大介† 川又　泰介† 加納　　徹†† 赤倉　貴子††

† 東京理科大学大学院工学研究科　〒125–8585 東京都葛飾区新宿 6–3–1
†† 東京理科大学工学部　〒125–8585 東京都葛飾区新宿 6–3–1
E-mail: †{4418518,4417701}@ed.tus.ac.jp, ††{kano,akakura}@rs.tus.ac.jp

あらまし　現在の Web-testing における受験者認証法は，試験開始時の ID とパスワードのみであるため，試験中のなりすましが容易に行える．そこで，著者らは，受験者がタブレット PC に解答文字を記入した際の，筆圧やペン傾斜などの動的情報を基に受験者認証を行った．しかし，先行研究の認証精度は不十分であり，実用的に Web-testin g へ適用するのは困難であった．その理由として，動的情報のみを用いて認証していたことがあげられる．そこで本研究では，Siamese Network による筆記画像モデルと MLP を用いた動的情報モデルを結合したハイブリッド認証モデルを提案した．結果として，ハイブリッド認証モデルは，筆記画像モデルと動的情報モデルを単体で用いるよりも高い精度を得た．

キーワード　Siamese Network，MLP，筆記認証，画像処理，DTW

Hybrid authentication model using handwritten images by Siamese Network and dynamic information

Daisuke HAYASHI†, Taisuke KAWAMATA†, Toru KANO††, and Takako AKAKURA††

† Graduate School of Engineering, Tokyo University of Science　6–3–1 Niijuku, Katsushika-ku, Tokyo, 125–8585 Japan
†† Faculty of Engineering, Tokyo University of Science　6–3–1 Niijuku, Katsushika-ku, Tokyo, 125–8585 Japan
E-mail: †{4418518,4417701}@ed.tus.ac.jp, ††{kano,akakura}@rs.tus.ac.jp

Abstract　Since the common examinee authentication method in Web-testing is based only on the ID and password at the beginning of the test, it is easy to spoof during the test. To solve this problem, we proposed an examinee authentication method based on dynamic information such as pen pressure and pen tilt when the examinee wrote the answer characters on a tablet PC. However, the accuracy of our previous research was insufficient. Hence, it was practically difficult to apply the examinee authentication method of previous research to Web-testing. One reason for this is because authentication was performed using only dynamic information. In this research, we proposed a hybrid authentication model using handwritten images model by siamese network and dynamic information model by MLP. Therefore, our hybrid authentication model obtained better accuracy than using handwritten images model and dynamic information model alone.

Key words　Siamese Network，MLP，Writing authentication, Image processing，DTW

1. まえがき

情報化の進展を受けて，高等教育機関で e-Learning の導入が進んでおり，講義における時間的・空間的制約は緩和されつつある．しかし，試験における制約は依然として存在する．そのため，大学への通学が困難な社会人学生などは，単位認定試験などで大学に行く必要がある．

試験の制約を緩和するために，Web 上で試験を受験できる Web-testing が有効であると考えられる．しかし，高等教育機関でその導入は進んでいない．原因として，現状の認証方式は，

試験開始時の ID とパスワードのみであるため，試験中に「なりすまし」や「カンニング」といった不正行為が容易であることがあげられる．ゆえに，試験時間全体を通した継続的認証手法を確立する必要があるといえる．

継続的認証手法を確立するために，バイオメトリクスに着目する [1]．バイオメトリクスは，指紋や顔などの身体的特徴を用いたものと，音声や署名などの行動的特徴を用いたものに分けられる．近年では，指紋認証でログオンする PC や，顔認証を利用した入館ゲートなど，様々な場面でバイオメトリクスが用いられている．

ここで，Web-testing で受験者認証するには，受験者の負担にならないために，試験で自然に行われる行為を利用することが必要である．さらに，継続的に受験者認証用のデータを取得できなければならない．指紋認証に関しては，受験者が問題を解くごとに指紋取得の動作を行わせる必要がある．ゆえに，試験の妨げとなり，身体的負担を伴ってしまう．一方，顔認証に関しては，顔を録画しながらテストを行うため，受験者に精神的不快感を与えてしまう恐れがある．したがって，プライバシーの問題が存在するといえる．

そこで，著者らは図 1 で示すように，タブレット PC を用いて，試験中の解答文字より受験者認証を行う方法を提案した [2] [3]．受験者が解答を記入するごとに，解答文字による 5 種類の動的情報より受験者認証を行うものである．5 種類の動的情報とは，x 座標，y 座標，筆圧，x 傾斜，y 傾斜の時系列データを指す．動的情報を用いた理由は，筆記画像よりも他人が真似ることが困難であり，個人認証に適合しているためである [4]．解答文字より認証を行っているため，受験者の負担にならずに継続的な認証が可能である．

しかし，実試験を想定した条件での認証精度は十分ではなかった．その理由として，動的情報のみを対象としており，筆記画像による画像処理を行っていなかったことが挙げられる．また，複数の特徴量を結合させるにあたり，式 (1) の線形回帰を用いていたことも理由として考えられる（X：複合距離，w：重み，x：距離, e：誤差，i：各特徴量の添字）．

$$X = \sum_{i=1}^{5} w_i x_i + e \tag{1}$$

図 2 と図 3 は，ある受験者が 3 と筆記した際の学習時と試験時の筆圧，x 傾斜，y 傾斜を表している．横軸の j，k，l は，筆記時間の要素数を意味する．ここで，y 傾斜に着目すると，学習時に比べて試験時は大きく低くなっていることがわかる．y 傾斜とは，筆記面の垂直方向を 0 度として，受験者側を正の角度，反対側を負の角度と設定したものである．

調査した結果，学習時と試験時の姿勢の違いが原因であると分かった．学習時は筆記数字が指定されていたため，画面上の指示文を注視する必要はなかった．しかし，試験時は画面に表示される設問と選択肢を注意深く読む必要があるため，姿勢が前かがみになっていた．ゆえに，ペンが受験者の反対側に傾きやすくなるため，y 傾斜が低い値となっていた．したがって，動的情報のみでは，正規の受験者となりすましを判別するのが

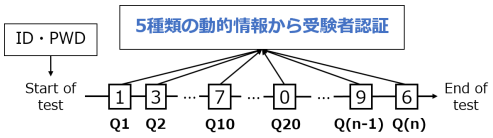

図 1　Web-testing の解答文字を利用した認証方式

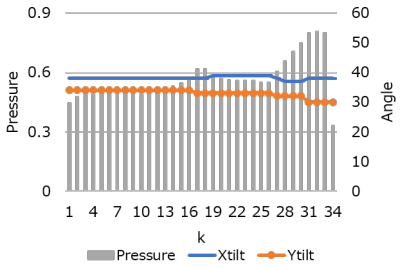

図 2　学習時のペン傾斜と筆圧

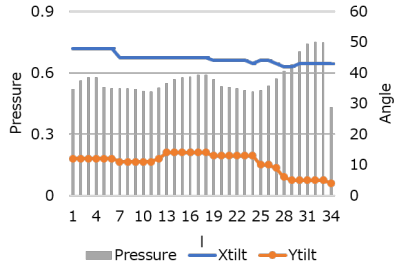

図 3　試験時のペン傾斜と筆圧

困難であった．

以上の背景より，本研究では Web-testing における「なりすまし」「カンニング」といった不正行為のうち，なりすましを防止することを目的とする．先行研究 [2] [3] より高精度の受験者認証を実現するために，Siamese Network [5] による筆記画像モデルと MLP (Multi-Layer Perceptron) [6] を用いた動的情報モデルを結合したハイブリッド認証モデルを構築する．

2 章で受験者認証の概要，3 章で MLP を用いた動的認証モデルについて述べる．4 章では Siamese Network による筆記画像モデル，5 章では筆記画像モデルと動的情報モデルを結合したハイブリッド認証モデルについて記す．

2.　受験者認証の概要

提案するアルゴリズムを図 4 に示す．提案アルゴリズムは，学習フェーズと試験フェーズからなる．

学習フェーズ：各受験者がタブレット PC で記入した「0〜9」の数字を登録する．1 回目に取得したものを学習データ 1，学習データ 1 の取得から 1 ヶ月後に取得したものを学習データ 2 とする．

動的情報においては，学習データ 1 と学習データ 2 を DTW（Dynamic Time Warping）[1] によって類似度計算を行い，特徴量ごとに距離ベクトルを算出する．MLP に各特徴量の距離ベクトルを適用することで，5 種類の特徴量を結合した距離ベクトルを出力する．

一方，筆記画像においては，学習データ 1 と学習データ 2 を

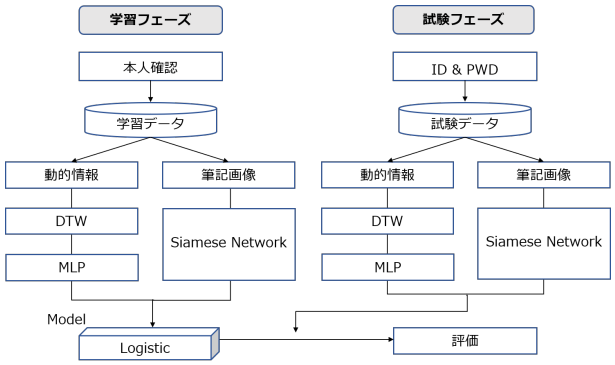

図 4　提案アルゴリズム

表 1　実験データ

データ	人数	取得時期	記述内容	データ数
学習データ 1	10	試験の 1 ヶ月半前	「0〜9」を 5 セット	500 文字
学習データ 2	10	試験の 2 週間前	「0〜9」を 5 セット	500 文字
試験データ	10		英語テスト 40 問 (四肢択一式)	400 文字

Siamese Network に入力して，各々を特徴ベクトルに変換する．各々の特徴ベクトルのユークリッド距離を求めることで，距離ベクトルを算出する．筆記画像と動的情報の距離ベクトルをロジスティック回帰に適用することで，ハイブリッド認証モデルを構築する．

試験フェーズ：Web-testing 中に得られた解答文字を試験データとして扱う．筆記画像と動的情報を学習フェーズで構築したハイブリッド認証モデルに入力し，推定確率を算出する．その値を閾値処理することで，正規の受験者となりすましの判定を行う．

2.1　実験データ

本研究で利用するデータを表 1 に示す．受験者 10 名から 1 ヶ月の間隔をあけて，学習データ 1，学習データ 2 を取得する．試験データは，学習データ 2 の取得から 2 週間後に，同じ受験者から取得する．

受験者が解答文字を記入するたびに，x 座標，y 座標，筆圧，x 傾斜，y 傾斜の動的情報を抽出する．これら 5 種類の動的情報は，時系列データとして保存される．

ある受験者が 3 と筆記した際の，動的情報の例を図 5 に示す．表の行が時系列データの数，列が各特徴量を表している．本研究では，ペンダウン（筆圧 >0）の時系列データを利用する．ペンダウンの行数は 92 行から 172 行までの計 81 行あり，特徴量は 5 種類ある．したがって，この例では，$81 \times 5 = 405$ 個の動的情報を取得できる．

また，抽出した動的情報の前処理は，以下の 3 段階により行う．

（1） ペンアップ時間（筆圧＝0）の時系列データを除外する．

（2） 左利きの受験者には，x 傾斜の正負を反転することで，右利きの受験者同様に対応する．

（3） 筆記時間の初期値を 0 に統一する．

2.2　出題画面の例

図 6 に本実験で利用した Web-testing システムを示す．受験者は，画面中央の枠内にタブレット PC を用いて解答番号を記入する．記入後に送信ボタンを押すと，x 座標，y 座標，筆圧，x 傾斜，y 傾斜の時系列データがサーバに送信される．

行数	x座標	y座標	筆圧	x傾斜	y傾斜
1	85	28	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮
92	52	59	0.3340	46	32
93	52	59	0.5078	46	32
94	52	59	0.5830	46	32
⋮	⋮	⋮	⋮	⋮	⋮
170	54	105	0.7783	39	17
171	54	105	0.7715	39	17
172	54	104	0.5410	39	17
⋮	⋮	⋮	⋮	⋮	⋮
179	60	148	0	0	0

図 5　動的情報の例

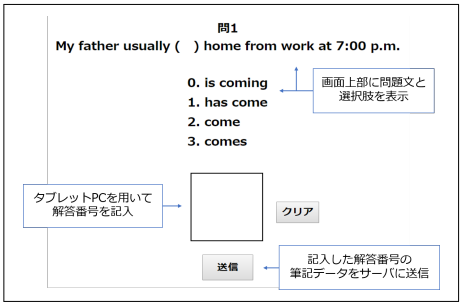

図 6　Web-testing システム

出題形式は，「0〜9」の数字を均等に出題するために四択問題の解答選択肢を，問 1〜8:「0, 1, 2, 3」，問 9〜16:「4, 5, 6, 7」，問 17〜24:「8, 9, 0, 1」，問 25〜32:「2, 3, 4, 5」，問 33〜40:「6, 7, 8, 9」と設定した．このように問題設定することで「0〜9」の数字をすべて選択肢に含めた場合でも，条件を満たす認証精度を算出できることを意図とした．

2.3　評価指標

2.3.1　AUC

本研究では，AUC（Area Under the ROC Curve：ROC 曲線下面積）を第一に優先する評価指標として用いる．ROC 曲線とは，横軸に FPR（False Positive Rate），縦軸に TPR（True Positive Rate）を置いて，閾値を変化させた際の推定確率をプロットしたものである．ROC 曲線下の面積値が AUC となる．AUC，FPR，TPR は，それぞれ以下の式で算出される．

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}(\theta)\mathrm{FRR}'(\theta)d\theta \tag{2}$$

$$\mathrm{FPR}(\theta) = \int_0^\theta \mathrm{P}(s|E=0)ds \tag{3}$$

$$\mathrm{TPR}(\theta) = \int_0^\theta \mathrm{P}(s|E=1)ds \tag{4}$$

$\mathrm{P}(s|E)$ は，受験者 E が本人（$E=1$）もしくは他人（$E=0$）であったときに，類似度が s であった割合である．$\mathrm{FPR}(\theta)$ は，類似度の閾値を θ と定めたときに，なりすましを誤って正規の

受験者と判定する確率である．一方，TPR(θ) は，正規の受験者を正しく正規の受験者と判定する確率である．

AUC の推定指標は，0.9 以上で High accuracy, 0.7 以上 0.9 未満で Moderate accuracy，0.5 以上 0.7 未満で Low accuracy となる [7]．したがって，本研究では High accuracy となる 0.9 以上を目指す．

2.3.2 EER

第二に優先する評価指標として，EER（Equal Error Rate：等価エラー率）を用いる [8]．EER とは，FRR（False Rejection Rate）と FAR（False Acceptance Rate）が等しくなった時の誤り率である．FRR と FAR は，以下の式で求める．

$$\mathrm{FRR}(\theta) = 1 - \int_0^{\theta} \mathrm{P}(s|E=1)ds \tag{5}$$

$$\mathrm{FAR}(\theta) = \int_0^{\theta} \mathrm{P}(s|E=0)ds \tag{6}$$

FRR(θ) は，正規の受験者を誤ってなりすましと判定する確率である．一方，FAR(θ) は，なりすましを誤って正規の受験者と判定する確率である．FRR と FAR は，閾値 θ の変化に対してトレードオフの関係にある．ゆえに，FRR と FAR が等しくなる最適閾値 θ の誤り率 EER を評価指標に用いる．

$$\theta_{EER} = \arg\min_{\theta} \frac{\mathrm{FRR}(\theta) + \mathrm{FAR}(\theta)}{2} \tag{7}$$

$$EER = \min \frac{\mathrm{FRR}(\theta_{EER}) + \mathrm{FAR}(\theta_{EER})}{2} \tag{8}$$

3. 動的情報モデル

3.1 類似度計算

データを分析するにあたり，筆記データの類似度（距離）を数値化する必要がある．しかし，筆記データは，字画の長さや筆記速度の違いにより，同じ文字を記入しても長さ（要素数）が異なる場合がある．そのため，単純な差分では距離を求めることはできない．

そこで，本研究では，異なる長さの時系列データ間の距離計算に，動的時間伸縮法である DTW を用いる．DTW は，二つの時系列データの各点を総当たりで計算して累積最短距離を算出する手法である．二つの時系列データが類似しているほど累積最短距離が小さくなる．ゆえに，この距離を基準に正規の受験者かなりすましかを判定する．DTW による距離計算の手順を以下に述べる．

比較する二つの時系列データ R，Q を式 (9) と式 (10) で表す．I と J は時系列データ R, Q の要素数である．r_i と q_j は時系列データの i 番目と j 番目のデータであり，式 (11) で表す．

$$R = r_1, r_2, \ldots, r_i, \ldots, r_I \tag{9}$$

$$Q = q_1, q_2, \ldots, q_j, \ldots, q_J \tag{10}$$

$$r_i, q_j \in (Cx_{i,j}, Cy_{i,j}, P_{i,j}, Sx_{i,j}, Sy_{i,j}) \tag{11}$$

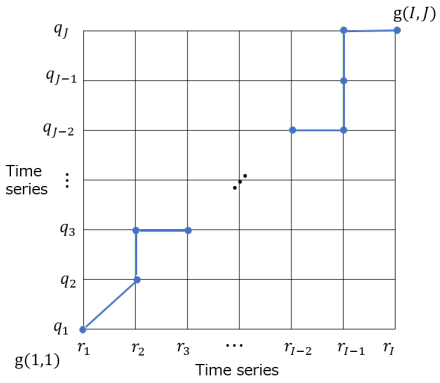

図 7 DTW

累積最短距離を算出する過程を図 7 に示す．初めに，累積最短距離 $g(i,j)$ の初期値 $g(1,1)$ を式 (12) とする．比較する時系列データの局所距離 $d(i,j)$ を，式 (13) のように差分の絶対値を求める．式 (14) で $g(i,j)$ の値を $g(I,J)$ まで変化させながら，累積最短距離を逐次計算していく．

$$\mathrm{g}(1,1) = |\mathrm{r}_1 - \mathrm{q}_1| \tag{12}$$

$$\mathrm{d}(\mathrm{i},\mathrm{j}) = |\mathrm{r}_\mathrm{i} - \mathrm{q}_\mathrm{j}| \tag{13}$$

$$\mathrm{g}(\mathrm{i},\mathrm{j}) = \mathrm{d}(\mathrm{i},\mathrm{j}) + \min \left\{ \begin{array}{c} \mathrm{g}(\mathrm{i},\mathrm{j}-1) \\ \mathrm{g}(\mathrm{i}-1,\mathrm{j}-1) \\ \mathrm{g}(\mathrm{i}-1,\mathrm{j}) \end{array} \right\} \tag{14}$$

逐次的に求めた累積最短距離 $g(I,J)$ を二つの時系列データの要素数の合計値で割って，式 (15) により正規化する．正規化された距離 x を 2 つの時系列データ R と Q の距離とする．

$$x = \frac{\mathrm{g}(\mathrm{I},\mathrm{J})}{I+J} \tag{15}$$

3.2 正 規 化

DTW によって算出した各特徴量における距離を，式 (16) によって最小値 0，最大値 1 に正規化する [9]．$\boldsymbol{x}^{(i)}$ は特定のサンプルであり，$\boldsymbol{x}_{\min}$ は各特徴量における距離の最小値，$\boldsymbol{x}_{\max}$ は最大値を表す．$\boldsymbol{x}^{(i)}_{\mathrm{norm}}$ は，$\boldsymbol{x}^{(i)}$ の正規化後のサンプルである．

正規化は，個人内距離と個人間距離を対象にして行う．個人内距離は，受験者本人を対象にして距離計算したものである．一方，個人間距離は，受験者本人をほかの受験者全員と比較して距離計算したものである．MLP へ適用するにあたり，個人内距離を正例，個人間距離を負例にラベル付けした．正例ラベル数は 2,500，負例ラベル数は 22,500 である．

$$x^{(i)}_{\mathrm{norm}} = \frac{x^{(i)} - x_{\min}}{x_{\max} - x_{\min}} \tag{16}$$

3.3 MLP

本研究では，距離結合に MLP を利用する．構築した MLP の構造を図 8 に示す．ネットワーク構造は，入力層，隠れ層，出力層の 3 層からなる多層パーセプトロンである．

入力層には，5 次元の距離ベクトルを入力する．隠れ層の直後に，活性化関数 ReLU（Rectified Linear Unit）[10] と

Dropout [11] を用いる．Dropout とは，ネットワークを学習する際に，ある更新で層の中のいくつかのノードを無効にして学習を行い，次の更新では別のノードを無効にして学習を繰り返すことを意味する．これにより，学習時にネットワークの自由度を強制的に小さくして汎化性能を上げ，過学習を抑制する．本研究では，Dropout の確率を p=0.5 に設定した．

出力層は式 (17) の Sigmoid 関数を用いる．$\phi(z)$ は活性化関数，x は距離ベクトル，w は重みベクトル，w_0 はバイアスユニットを表している．z は式 (18) で表される総入力である．$\phi(z)$ は z が大きいほど 1 に近づき，小さいほど 0 に近づく関数である．

$$\phi(z)= \frac{1}{1+e^{-z}} \tag{17}$$

$$z \ = \boldsymbol{w}^{\mathrm{T}} \cdot \boldsymbol{x} = w_0 x_0 + w_1 x_1 + \ldots + w_6 x_6 \tag{18}$$

MLP の出力を計算する順伝搬法は，以下の 3 つのプロセスから成り立つ [9]．複数の手順でこれらの手順を繰り返し，MLP の重みを学習した後，順伝播法を使ってネットワークの出力を計算する．

（1） 入力層を出発点として，学習データのパターンをネットワーク経由で順方向に伝播させ，出力を生成する．

（2） ネットワークの出力に基づき，活性化関数を使って誤差を計算する．この誤差を最小化することが目的となる．

（3） 誤差を逆方向に伝播させることで，ネットワーク内の各重みに対する偏導関数を求め，モデルを更新する．

学習の最適化手法には，Adam [12] を α=0.001，β_1=0.9，β_2=0.999 に設定して用いる．ミニバッチサイズは 256，エポック数は 50 に設定した．MLP の実装は，Python のライブラリである Keras [13] を用いる．

3.4 実験概要

本実験では，Web-testing 中のなりすましを防止するため，先行研究 [2] [3] よりも高精度な受験者認証を実現できるかを検証する．実際の Web-testing の環境下で受験者 10 名から，タブレット PC により四肢択一式の英語テスト 40 問を解答してもらう．試験の実施は，学習データ 2 の取得から 2 週間後に行った．試験より取得したデータを照合データとして扱う．照合データを認証モデルに入力することで，認証精度を出力する．

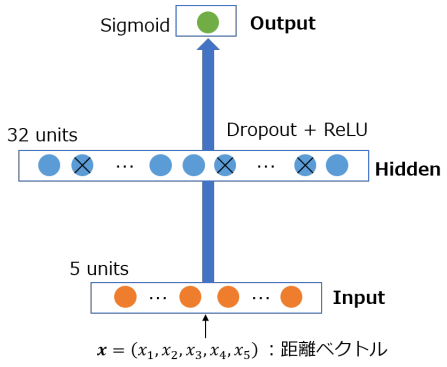

図 8 MLP の構造

表 2 動的情報モデルの結果

	AUC	EER[%]
MLP	**0.9338**	**14.99**
NB	0.9280	15.93
SVM	0.9220	16.55
RF	0.8709	18.74
LR [2] [3]	0.8480	15.19
KNN	0.8382	20.18

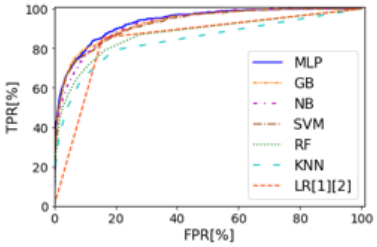

図 9 動的情報モデルにおける分類手法ごとの ROC 曲線

4 章の筆記画像モデルと 5 章のハイブリッド認証モデルに関しても，同様の概要で実験を行う．

3.5 動的情報モデルの結果

動的情報モデルの結果を表 2 に，ROC 曲線を図 9 に示す．試験データにおいて，正例ラベル数：708 に対し，負例ラベル数：14,618 となっている．MLP の有効性を確認するために，先行研究 [2] [3] の LR（Linear regression：線形回帰）に加えて，GB（Gradient boosting）[14]，NB（Naive Bayes）[15]，SVM（Support vector machine）[16]，RF（Random forest）[17]，KNN（K-nearest neighbor）[18] とも比較した．

結果より，MLP は AUC，EER ともに分類手法の中で最も良い精度を示していることがわかる．MLP の AUC は 0.9338 であり，LR [2] [3] の AUC：0.8480 よりも 0.0858 優れた精度となっている．EER に関しては，MLP は 14.99% となっており，LR [2] [3] の EER：15.19% よりも 0.20% 良い精度を示している．

一方，ROC 曲線を見ると，LR [2] [3] の線は折れ線になっていることがわかる．ゆえに，閾値を変化させていく際に，正規の受験者かなりすましかを判定するのは困難であるといえる．

4. 筆記画像モデル

4.1 筆記画像の生成法

本研究では，動的情報から筆記画像を生成する．受験者が「0〜9」の解答文字を筆記した際の静止画像を単に利用する方法では，精度が不十分であった．ゆえに，動的情報を基にして筆記画像を生成する方法を用いることにする．

筆記画像の生成法を図 10 に示す．横軸に x 座標（C_x），縦軸に y 座標（C_y）を取って，筆圧（P），x 傾斜（T_x），y 傾斜（T_y）の値をプロットする．x 座標，y 座標の最大値は 150 から 100 に縮小する．筆圧，x 傾斜，y 傾斜の値は，画素値として扱うために，最小値 0，最大値 255 になるようにスケール変換する．プロットした筆圧，x 傾斜，y 傾斜の画像を組み合わせるこ

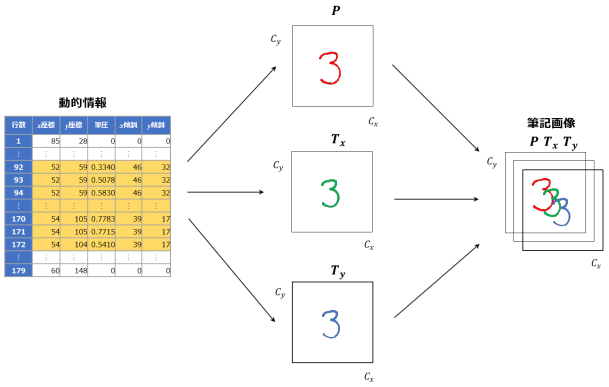

図 10　筆記画像の生成法

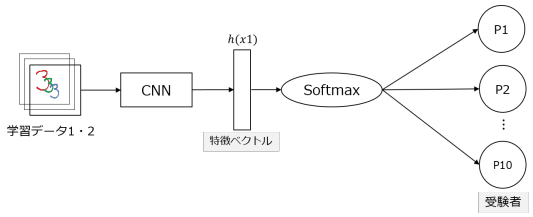

図 11　既存法の CNN

とで，3 次元の筆記画像を出力する．出力した 3 次元の筆記画像を Siamese Network に適用する際の学習データとして扱う．

4.2　既存法の CNN [19]

既存法の CNN (Convolutional Neural Network）での構造を図 11 に示す．出力層に Softmax 関数を用いて，受験者ごとに認証モデルを構築する必要がある．しかし，この構造で発生する問題点として二つあげられる．

一つ目は，受験者数分の異なるデータが大量に必要となることである．受験者を出力として分類するためには，各々の受験者のデータが膨大に必要である．二つ目は，受験者が新たに追加もしくは抜けるごとに，モデルを再調整する必要があることである．正しく受験者を分類できるために，CNN のモデル構造を調整しなければならない．これらより，膨大な受験者がいる大規模環境を想定すると非実用的であるといえる．

4.3　Siamese Network

本研究では，筆記画像の距離ベクトルを算出するために Siamese Network を用いる．Siamese Network は，二つの画像を入力にして，その画像同士が類似しているかの距離を算出するネットワークである．類似度を算出できるため．少量の画像により受験者を分類することが可能である．

ネットワークの構造を図 12 に示す．学習データ 1 と学習データ 2 の筆記画像を CNN に入力する．筆記画像は，各ピクセルの画素値を 255 で割ることで [0, 1] の範囲に正規化されている．各々のデータは 100×100 の同じサイズを入力とする．二つの CNN は同一構造である．

CNN に入力することで，各々の特徴ベクトルを出力できる．それらの特徴ベクトルのユークリッド距離を求めることで，距離ベクトルを算出する．これにより，特徴空間での非線形な変換を学習することが可能になり，より高い表現能力を獲得できる．

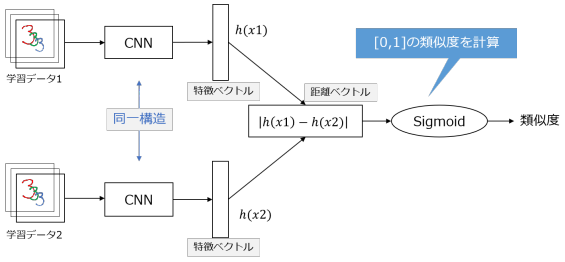

図 12　Siamese Network

距離ベクトルを Sigmoid 関数に適用することで類似度を求める．類似度は最小値 0，最大値 1 になっており，1 に近いほど正規の受験者であるといえる．最終的な出力が類似度になっているため，受験者の増減に応じたモデルの再調整が不要となっている．

学習の最適化手法は MLP と同様に，Adam [12] を α=0.001, β_1=0.9, β_2=0.999 に設定して用いる．ミニバッチサイズは 256，エポック数は 50 に設定した．Siamese Network の実装は，Python のライブラリである Keras [13] を用いる．

4.4　Siamese Network の CNN 構造

Siamese Network で用いる CNN の構造を図 13 に示す．CNN のネットワークは，畳み込み層，プーリング層，全結合層の 3 種類を組合わせて構成した．

畳み込み層（Conv）は，入力画像のピクセル範囲に対してフィルタを適用して，画像に畳み込む処理を行う．カーネルサイズはすべて 2×2 とした．処理後の出力マップには ReLU を適用する．

プーリング層（Pool）は，設定された正方領域の最大値や平均値に置き換えることで，平行移動に対しての不変性を向上させる．本研究では，すべて最大値と置き換える Max Pooling を用いる．プーリングサイズは 2×2 とした．

全結合層（Fc）は，隣接層のすべてのユニットを結合する．最終層の出力層では，シグモイド関数を適用し，最小値 0，最大値 1 とすることで出力を分類確率としている．

4.5　筆記画像モデルの結果

Siamese Network の畳み込み層のレイヤ数を変えて，認証精度を出力した結果を表 3 に示す．結果より，畳み込み層を 5 層用いた際の，AUC と EER が最も良い精度を示していることがわかる．ゆえに，本研究では 5 層の畳み込み層を用いることにする．

つぎに，5 層の畳み込み層に，Dropout 層，L2 ノルム，正規化層 [20] を組み合わせた結果を表 4 に示す．Dropout の確率は $p = 0.5$ として，各 ReLU 関数の直後に適用した．L2 ノルムのパラメータ λ は，各畳み込み層には $\lambda = 0.0002$，結合層には $\lambda = 0.001$ と設定した．正規化層は，平均を 0，標準偏差を 1 に近づける変換を行うものであり，各畳み込み層の直後に適用した．

結果より，Dropout，L2 ノルム，正規化層，L2 ノルムと Dropout を組み合わせたものにおいて，いずれも 5 層の畳み込

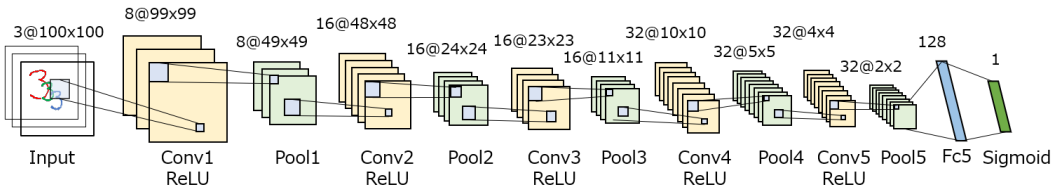

図 13　Siamese Network の CNN 構造

表 3　Siamese Network において畳み込み層のレイヤ数を変えた場合の結果

Number of Convolution layers	ACU	EER[%]
1 layer	0.6967	36.02
2 layers	0.6973	35.70
3 layers	0.7234	33.77
4 layers	0.8129	26.89
5 layers	**0.8468**	**23.83**
6 layers	0.7983	28.79

表 4　畳み込み層を 5 層用いた際に各関数を適用した結果

5 layers of convolution	ACU	EER[%]
With Dropout	0.8291	26.13
With L2-Norm	0.8213	29.93
With Batch-Norm	0.7832	29.38
With L2-Norm Dropout	0.7992	28.40

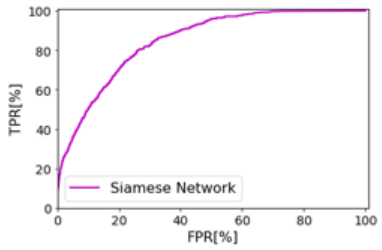

図 14　Siamese Network の ROC 曲線

み層を用いたものよりも AUC，EER ともに低い精度になっていることがわかる．したがって，本研究では，入力層，5 層の畳み込み層，5 層のプーリング層，結合層，出力層の構造で CNN を構築する．この CNN 構造で構築した Siamese Network の ROC 曲線は，図 14 に示す通りである．

5.　ハイブリッド認証モデル

筆記画像と動的情報の距離ベクトルをロジスティック回帰に適用することで，ハイブリッド認証モデルを構築する．

5.1　学 習 結 果

モデルの学習結果を表 5 に示す．Accuracy の学習結果は図 15，Loss の学習結果は図 16 に示す通りである．

結果より，Siamese Network の Accuracy は 100% となっており，ハイブリッド認証モデルと MLP よりも優れた精度になっていることがわかる．また，Loss : 0.0025 も同様に最も良い値となっている．本研究では，Accuracy が下がる直前のエポックを最適エポックとして，モデル構築を行った．

表 5　モデルの学習結果

	Hybrid	Siamese Network	MLP
Accuracy[%]	94.12	100	92.13
Loss	0.1498	0.0025	0.2098

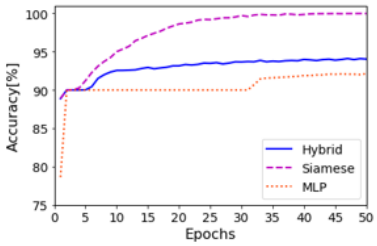

図 15　accuracy の学習結果

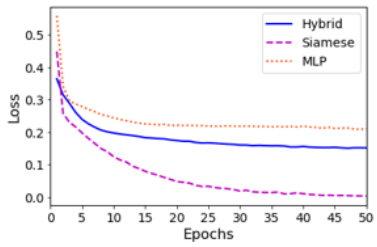

図 16　Loss の学習結果

表 6　ハイブリッド認証モデルの結果

	Hybrid	Siamese Network	MLP
AUC	**0.9355**	0.8468	0.9338
EER[%]	**14.01**	23.83	14.99

5.2　ハイブリッド認証モデルの結果

ハイブリッド認証モデルの結果を表 6 に，ROC 曲線を図 17 に示す．ハイブリッド認証モデルの AUC は 0.9355，EER は 14.01% となっている．AUC に関しては，High accuracy を意味する 0.9 以上の値を得た．また，ハイブリッド認証モデルは，筆記画像モデルと動的情報モデルを単体で用いるよりも優れた精度になっていることがわかる．したがって，ハイブリッド認証モデルの有効性を確認できた．

5.3　考　　察

ハイブリッド認証モデルを構築する際に，筆記画像モデルと動的情報モデルの結合が功を奏した理由を考察する．理由として，図 18 で示したように Siamese Network の出力を距離ベクトルにしたことがあげられる．結合する際に，距離ベクトルと

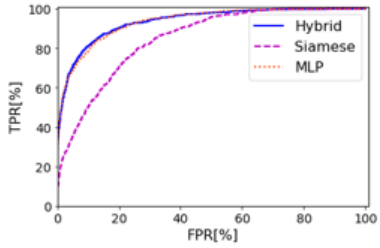

図 17　ハイブリッド認証モデルの ROC 曲線

表 7　筆記画像モデルと動的情報モデルの結合法

	距離ベクトル結合	類似度結合
AUC	**0.9355**	0.8848
EER[%]	**14.01**	17.90

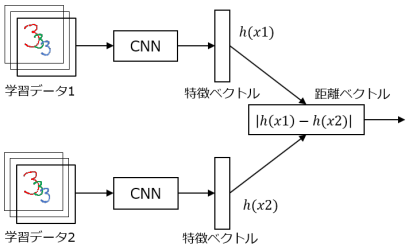

図 18　出力を距離ベクトルにした Siamese Network

類似度を用いた場合の比較結果を表 7 に示す．結果より，距離ベクトル結合の方が類似度結合よりも，AUC，EER ともに大きく精度が向上していることがわかる．

距離ベクトルを用いることで 128 次元のベクトル空間で分類できるため，特徴空間での非線形な変換を学習することが可能になる．ゆえに，高い表現能力を獲得できるといえる．一方，図 12 で示したように出力を類似度とすると，一次元の情報のみしか扱えなくなる．したがって，距離ベクトル結合にした方が高次元で分類できるため，精度が向上したと考えられる．

6. むすび

本研究では Web-testing において，受験者認証を先行研究 [2] [3] よりも高精度に行うことで，なりすましを防止することを目的とした．先行研究 [2] [3] は，動的情報のみを用いて認証をしていたため，認証精度に課題が残るものであった．そこで本研究では，Siamese Network による筆記画像モデルと MLP を用いた動的情報モデルを結合したハイブリッド認証モデルを提案した．

実験結果より，ハイブリッド認証モデルの AUC は 0.9355，EER は 14.01% となっている．これは，筆記画像モデルと動的情報モデルを単体で用いるよりも優れた精度になっている．したがって，ハイブリッド認証モデルの有効性を確認できた．

今後の課題としては，学習データ数を増やした場合に，認証精度がどのように変化するかを分析することなどがあげられる．

文　献

[1] 半谷精一郎，“バイオメトリクス教科書～原理からプログラミングまで～，”（社）映像メディア学会，（編）コロナ社，東京，2012.

[2] 林 大介，赤倉貴子，“e-Testing におけるタブレット PC とオンライン筆記情報を用いた筆記認証法の提案，”日本教育工学会論文誌，vol.42，no.Suppl，pp.101-104，2018.

[3] D. Hayashi, T. Akakura, “Proposal for Writing Authentication Method Using Tablet PC and Online Information in e-Testing,” Springer International Publishing AG, vol.10905, LNCS，pp.253-265，2018.

[4] 中村善一，木戸出正継，“筆跡鑑定の知見に基づく特性値を用いたオンライン筆者照合，”システム制御情報学会論文誌，vol.22，no.1，pp.37-47，2009.

[5] G. Koch, R. Zemel, R. Salakhutdinov, “Siamese Neural Networks for One-shot Image Recognition,” Proceedings of the 32nd international Conference on Machine Learning，2015.

[6] S. Maheshwary，S. Ganguly，V. Pudi，“Deep Secure: A Fast and Simple Neural Network based approach for User Authentication and Identification via Keystroke Dynamics，” International Joint Conference on Artificial Intelligence，Aug. 2017.

[7] A. Akobeng，“Understanding diagnostic tests 3: Receiver operating characteristic curves，” Acta Paediatr，vol.96，no.5，pp.644-647，2007.

[8] 川又泰介，赤倉貴子，“e-Testing における Web カメラとペンタブレットを用いた逐次受験者認証システムの開発，”電子情報通信学会論文誌 D，vol.J102-D，no.3，pp.163-172，2019.

[9] S. Raschka，V. Mirjalili，“Python 機械学習プログラミング達人データサイエンティストによる理論と実践，”高橋隆志（編），インプレス，東京，2018.

[10] A. L. Mass，A. Y. Hannun and A. Y. Ng，“Rectifier Nonlinearities Improve Neural Network Acoustic Models，” International Conference on Machine Learning，2013.

[11] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov，“Dropout: a simple way to prevent neural networks from overfitting，” Journal of Machine Learning Research，vol.15，no.1，pp.1929-1958，2014.

[12] D. P. Kingma，J. L. Ba，“Adam: A Method for Stochastic Optimization，” arXiv:1412.6980 [cs. LG], 2014..

[13] F. Chollet，“Keras Documentary，” https://keras.io，Nov 2019.

[14] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T. Y. Liu，“LightGBM: A Highly Efficient Gradient Boosting Decision Tree，” Advances in Neural Information Processing Systems，pp.3149-3157，2017.

[15] H. Zhang，“The Optimality of Naïve Bayes，” in Proceedings of the Seventeenth International Artificial Intelligence Research Society Conference，(V. Barr and Z. Markov, eds.)，Miami Beach，FL: AAAI Press，2004.

[16] T. Harris，“Credit scoring using the clustered support vector machine，” Expert Systems with Applications，vol.42，no.2，pp.741-750，2015.

[17] J. Ham，Y. Chen，M. M. Crawford and J. Ghosh，“Investigation of the random forest framework for classification of hyperspectral data，” IEEE Trans. Geoscience and Remote Sensing，vol.43，no.3，pp.492-501，2005.

[18] D. A. Adeniyi, Z. Wei, Y. Yongquan，“Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method，” Applied Computing and Informatics，vol.12，no.1，pp.90-108，2016.

[19] 伊藤康一，岡野健久，青木孝文，“畳み込みニューラルネットワークを用いた生体検知手法，”信学論 A，vol.J100-A，no.12，pp.455-464，2017.

[20] S. Ioffe，C. Szegedy，“Batch Normalization: Accelerating Deep Network Training by Reducing nternal Covariate Shift，” Proceedings of the 32nd International Conference on International Conference on Machine Learning，vol.37，pp.448-456，2015.