# Handedness Detection Based on Drawing Patterns using Machine Learning Techniques

Jungpil Shin
School of Computer Science and Engineering
University of Aizu
Aizuwakamatsu, Fukushima, Japan
Email: jpshin@u-aizu.ac.jp

Md Abdur Rahim
School of Computer Science and Engineering
University of Aizu
Aizuwakamatsu, Fukushima, Japan
Email: rahim_bds@yahoo.com

***Abstract*—Handedness detection has an effective role in classifying different criminal suspects into specific categories according to soft biometric properties. It determines human motor skills that are performed with the dominant hand while doing everyday activities such as writing and throwing. In this context, this paper offers a system that extracts the characteristics of a person's drawing patterns and uses these features to perform handwriting classifications with regards to handedness. For this, we collect left and right hand data and derive various types of parameters such as elapsed time, x-coordinate, y-coordinate, pen pressure, pen orientation, and pen height. We define different features like mean, maximum, minimum, writing pressure, speed, and Dynamic Programming (DP) for handwriting data for analysis. P-value and t-test are calculated for handwriting evaluation. Furthermore, handedness detection is achieved by using a Support Vector Machine (SVM) classifier. The result shows quite encouraging performance that highlights the effectiveness of the proposed system.**

***Keywords–Handedness; Handwriting; Drawing Pattern; Support Vector Machine (SVM).***

## I. INTRODUCTION

Handedness is a complex characteristic in human behavior that reflects the domination of the brain that is related to quantitative change. The use of the dominant hand instead of the other hand makes it more efficient and comfortable for everyday tasks. This became increasingly evident during childhood and persists throughout life. Handedness can be used to identify a person's cognitive ability and personality concerning fine motor skills and manual functionality and help to find suspects in criminal investigations. Many scholars have explained that psychological traits i.e., developmental process, cognitive abilities [1] and personality [2] are related to hand preferences. In [3], Nicholls et al. proposed a cognitive ability scale and hand preference test, which are subtle and sensitive measures of hand performance. Gradient feature-based biometric traits, such as handedness, age, gender prediction suggested in [4]. This work is addressed how to automatically predict these soft biometric features from the handwritten text. Moreover, in [5], Morera et al. proposed an offline handwriting-based gender and handedness prediction system. From the handwriting, they introduced a deep network in order to solve demographic classification problems. However, handwriting recognition can cause some problems when writing with inter- and intra-individual variations. This can be assessed by analyzing the correlation between the writing velocities of each test for each participant. Pressure and stroke sequences, use of different pen types, or background noise are also major concerns. Wang and Chuang [6] proposed a pen-type input device to detect trajectories for handwriting digits and gestures. They extract time and frequency-domain features from acceleration signals and then identify important features by a hybrid system. We, therefore, focus on the detection of handedness based on drawing pattern by analyzing elapsed time, x-coordinate, y-coordinate, pen pressure, pen orientation, and pen height. We consider different statistical methods and DP distances as a feature for handedness analysis and we make a classification using the SVM classifier.

The rest of this paper is structured as follows. Section II describes the proposed methodology. In Section III, we describe the data acquisition process and analyze the results. Section IV concludes this paper.

## II. METHOD OF HANDWRITING ANALYSIS

In this section, we explain the overall process of the handedness detection system. Figure 1 shows the basic flow diagram of the proposed system. Data is acquired using a pen tablet and processed for feature extraction. Then, we calculate the t-test and p-value for each feature. Finally, we classify handedness using different kernels of SVM.

### *A. Handwriting Feature Analysis*

To extract the handwriting features, we create a reference model. In this work, $T$ represents the length of each input data and $N_T$ represents the length of the coordinate points of the reference model. t(n) is the elapsed time after the start of the test, p(n) is the writing pressure, x(n) is the positional coordinate in the horizontal direction, and y(n) is the positional coordinate in the vertical direction. The experimental data is acquired by following (1), where $n = 1, 2, 3, \ldots, N$.

$$S(n) = [t(n), p(n), x(n), y(n)] \tag{1}$$

The reference model is defined using (2) where $x_T(n_T)$ represents the positional coordinate in the horizontal direction and $y_T(n_T)$ is the positional coordinate in the vertical direction.

$$S_T(n_T) = [X_T(n_T), y_T(n_T)] \tag{2}$$

where $n_T = 1, 2, 3, ..., N_T$. To evaluate the differences between right handed and left handed persons, we obtained some statistical features that are shown in Table I. Moreover, we had measured the DP distance using the DP matching algorithm as a feature.

### *B. Dynamic Programming (DP) Matching*

DP matching is a pattern recognition methodology used in many studies in the field of signature authentication and speech recognition [7][8]. It calculates the distance between sample data and input data which optimizes all route-to-route measurements through backtracking. In this study, the differences between the right handed person and left handed

 

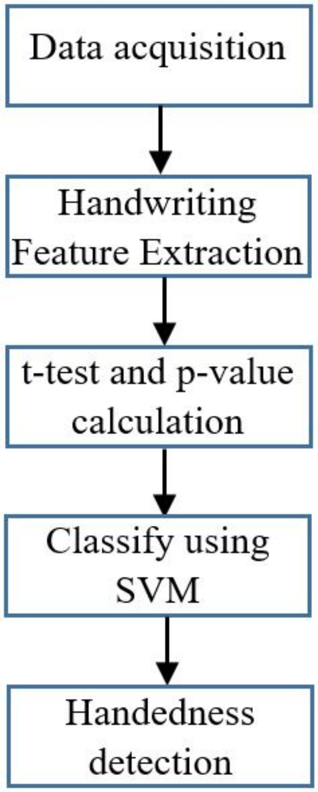


Figure 1. Basic block diagram of a handedness detection system.

TABLE I. STATISTICAL FEATURES FOR HANDEDNESS DETECTION

| Feature | Equation |
|---|---|
| Average pressure | $P_{mean} = \frac{1}{N} \times \sum_{n=1}^{N} p_{(n)}$ |
| Maximum pressure | $P_{\max} = \max_{1 \le n \le N} (p_{(n)})$ |
| Minimum pressure | $P_{\min} = \min_{1 \le n \le N} (p_{(n)})$ |
| Average velocity | $V_{mean} = \frac{1}{N} \times \sum_{n=1}^{N} \frac{\sqrt{\{x_{(n)} - x_{(n+1)}\}^2 + \{y_{(n)} - y_{(n+1)}\}^2}}{t_{(n+1)} - t_{(n)}}$ |

person are defined as the total cost $DP_x$ of the optimum route and are calculated as the difference between the cost and the sample data. $d_{(i,j)}$ was measured by the distance between the x-coordinates of input data and reference data, as shown in (3) and (4). The cost $C_x$ of DP matching was calculated from using (5).

$$d_{(i,j)} = x_{\gamma(i)} - y_j \quad (3)$$

$$g_{(i,j)} = d(i,j) + min \begin{bmatrix} C(i-1,j) \\ 2C(i-1,j-1) \\ C(i,j-1) \end{bmatrix} \quad (4)$$

$$DP_x = g(N_T, N) \quad (5)$$

$d_{(i,j)}$ defines the distance between the $i$th coordinate point of the sample data and the $j$th coordinate point of the input data. $DP_y$ is defined similarly. After that, these analysis methods are used to recognize the differences between a right handed person and a left handed person.

### C. Handedness Classification using SVM

Support Vector Machine (SVM) differentiates the advanced features into different domains based on the concept of finding a hyperplane [9]. In this work, we use different kernels, such as linear, polynomial, Radial Basis Function (RBF) to classify handedness. Table II shows the function of different SVM kernels.

TABLE II. DIFFERENT SVM KERNELS METHOD

| SVM kernels | Functions |
|---|---|
| Linear | $f(X) = B(0) + sum(a_i * (X, X_i))$ |
| Polynomial | $K(X_1, X_2) = (a + X_1^T X_2)^b$ |
| RBF | $K(X_1, X_2) = \text{exponent}(-\gamma \|X_1 - X_2\|^2)$ |

## III. EXPERIMENTAL RESULTS

The experiment aims to classify handedness, provide handwriting features for analysis, and evaluate performance through machine learning techniques.

### A. Data Acquisition Process

We used a liquid crystal tablet (Wacom Cintiq Pro 16) to collect handwriting data. Random participants are requested to draw and write characters that can be obtained as 6-dimensional data sets such elapsed time, x-coordinate, y-coordinate, pen pressure, pen orientation, pen height. The elapsed time since application startup (ms) and the x-coordinates and y-coordinates are represented by the pixel value. The writing pressure is represented by a $2^{15}$ scale; the value decreases as the writing pressure becomes weaker and the value becomes larger as the writing pressure becomes stronger. The pen orientation is represented by 90 degrees while the pen is positioned vertically on the surface of the board and the pen was at close to 0 degrees while it was horizontally aligned to the right of the board surface. As the tip of the pen points to the top of the display, the horizontal component of the height of the pen increases and the value decreases when pointing down. The value range is expressed between 0 and 1800, and the height angle can be calculated by dividing the value obtained by 10. However, these data were acquired as time-series data at an average of 40 ms. We used these parameters to detect and analyze the differences between handedness. Ten people (7 right handed and 3 left handed) participated in the handwriting experiment as we collecting data. Figure 2 shows some examples of handwriting samples in handedness detection. Samples 1 to 3 are continuous spiral writing, 4 to 6 writing a line in different directions, 7 is writing a square continuously, and 8 and 9 are dotted lines. The participants we asked to write the same character in the blank space in 10 and 11.

### B. Results Analysis

We calculated handwriting characteristics from each sample data of each individual. Also, the t-test was performed and the p-value was calculated for each feature. Table III shows the results of the t-test analysis from each feature. In all results, the difference between right handed and left handed

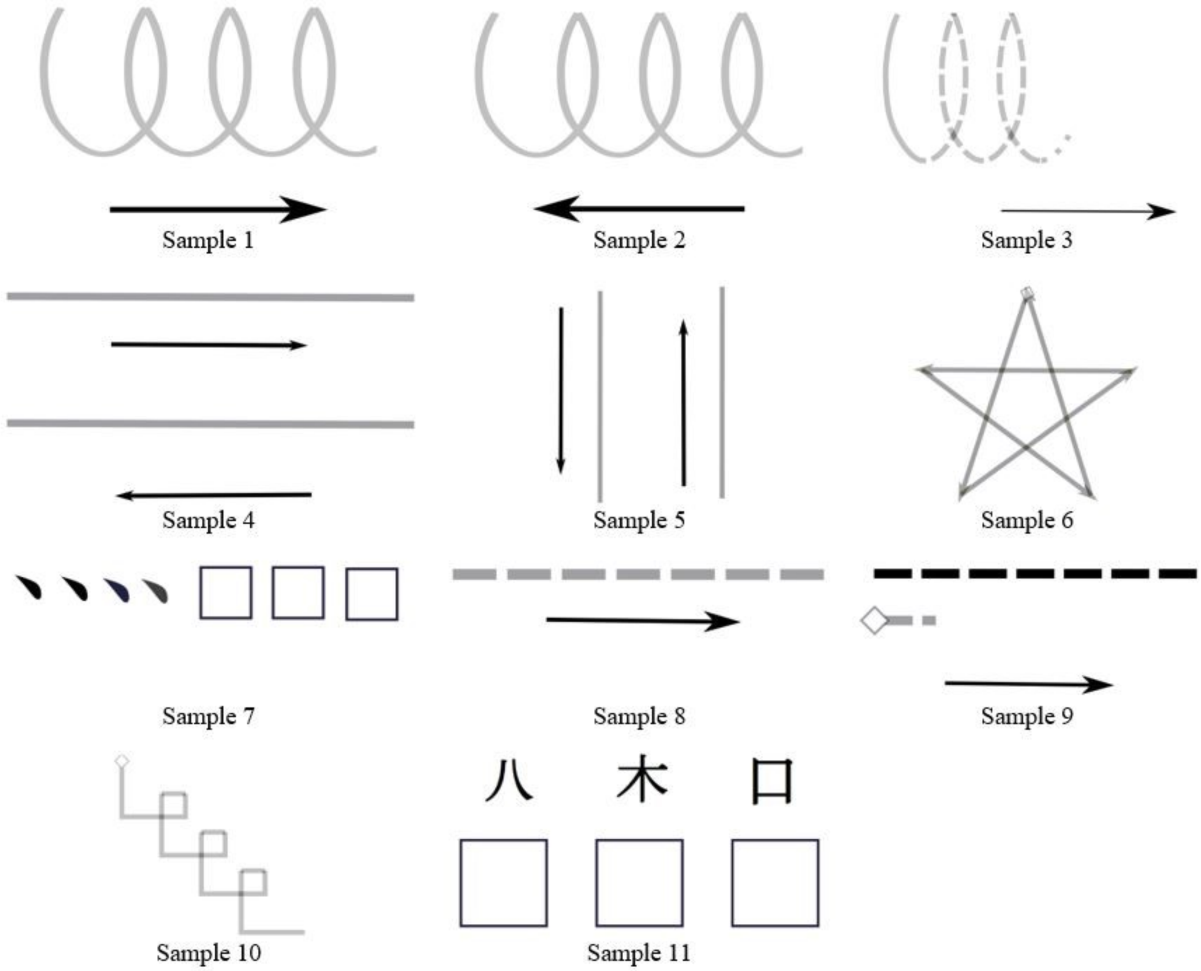


Figure 2. Example of handwriting samples in handedness detection.

individuals was significant at the 5% level. Finally, the results of significant differences were found to be the maximum pressure, the average pressure, and average velocity. For the classification process, we used and compared the different SVM classifications techniques shown in Table IV. The highest recognition accuracy at the highest pressure and the average pressure is about 95.20%. In the Polynomial kernel, we have achieved the highest accuracy in analyzing $Pmax$ and $Pmean$ features. Thus, a major difference between the right hand and the left hand has been observed from the pressure analysis of the pen. Table V presents the comparison of recognition accuracy with a state-of-the-art system.

## IV. Conclusion

This work addresses the use of a statistical and DP distance feature to detect handedness. The SVM classifier is used to identify left hand and right hand individuals. We used built-in datasets in the experiments. There are 10 people who participated to create the dataset. We collected left handed and right handed data for each individual. We analyzed six parameter values, such as elapsed time, x-coordinate, y-coordinate, pen pressure, pen orientation, and pen height. From the results, we can say that the pressure of the pen is an important feature of distinguishing the handedness. The average classification accuracy of handedness is 95.20%. The obtained results reveal that the proposed system provides a significant improvement compared to the state of the art. In future work, we will explore more robust optical features and review various integrated feature identities and create a feature ranking to improve the detection of handedness.

## References

[1] M. Papadatou-Pastou, "Handedness and cognitive ability: Using meta-analysis to make sense of the data," in Progress in brain research, vol. 238, pp. 179-206, 2018.

[2] F. M. Bryson, G. M. Grimshaw, and M. S. Wilson, "The role of intellectual openness in the relationship between hand preference and positive schizotypy," Laterality, vol. 14, no. 5, pp. 441-456, 2009.

[3] M. E. Nicholls, H. L. Chapman, T. Loetscher, and G. M. Grimshaw, "The relationship between hand preference, hand performance, and general cognitive ability," Journal of the International Neuropsychological Society, vol. 16, no. 4, pp. 585-592, 2010.

[4] N. Bouadjenek, H. Nemmour, and Y. Chibani, "Age, gender and handedness prediction from handwriting using gradient features," in 2015 13th International Conference on Document Analysis and Recognition (ICDAR), IEEE, pp. 1116-1120, 2015.

[5] A. Morera, A. Sánchez, J. F. Vélez, and A. B. Moreno, "Gender and handedness prediction from offline handwriting using convolutional neural networks," Hindawi Complexity, 2018.

[6] J. S. Wang and F. C. Chuang, "An Accelerometer-Based Digital Pen With a Trajectory Recognition Algorithm for Handwritten Digit and

TABLE III. REPRESENTATION OF T-TEST ANALYSIS RESULTS

| | | Sample1 | Sample2 | Sample3 | Sample4 | Sample5 | Sample6 | Sample7 | Sample8 | Sample9 | Sample10 | Sample11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | right | 21294.08 | 21335.62 | 22480 | 21056.85 | 21876.54 | 22048.46 | 21122 | 17229.08 | 17757.62 | 20119.23 | 20034.54 |
| Pmax (R) | left | 21614.17 | 21351.33 | 22880.33 | 20359.33 | 21832.17 | 22479 | 22384.83 | 16486.33 | 16715 | 18172.33 | 18296.33 |
| | p-value | 88% | 92% | 98% | 47% | 76% | 68% | 40% | 88% | 41% | 26% | 43% |
| | right | 2206.62 | 2137.46 | 1938.46 | 787.23 | 872.46 | 2121.77 | 1948.923 | 96.38 | 18.77 | 1177.69 | 237.54 |
| Pmin (R) | left | 1448.83 | 667.67 | 940.33 | 767.33 | 791.67 | 1683.67 | 142.67 | 1 | 11.5 | 239.67 | 1 |
| | p-value | 26% | 20% | 46% | 89% | 100% | 39% | 27% | 31% | 68% | 19% | 26% |
| | right | 18468.68 | 19408.64 | 20002.76 | 18338.79 | 19178.18 | 19660.54 | 19362.3 | 13618.27 | 14032.43 | 16254.72 | 15406.84 |
| Pmean (R) | left | 18971.08 | 19137.52 | 19861.17 | 17366.57 | 18547.55 | 19375.12 | 20359.69 | 13013.25 | 13166.97 | 14626.41 | 14065.96 |
| | p-value | 69% | 80% | 92% | 44% | 47% | 99% | 48% | 76% | 44% | 20% | 41% |
| | right | 19089.92 | 18680.38 | 19557 | 17839.62 | 18198.23 | 19367.23 | 19253.31 | 15862.23 | 15769.54 | 15480.69 | 16707.46 |
| Pmax (L) | left | 22749.67 | 23335.67 | 24430 | 23152.5 | 23458.5 | 23355.17 | 23612.5 | 20590 | 20732.17 | 19727.67 | 22189 |
| | p-value | 2.39% | 0.71% | 0.54% | 0.57% | 1.48% | 2.95% | 1.97% | 1.98% | 2.04% | 0.43% | 1% |
| | right | 595.85 | 1079 | 1029.77 | 1376.92 | 87.85 | 1043.46 | 1593.69 | 55.62 | 18 | 342.5 | 6.15 |
| Pmin (L) | left | 1812.83 | 2077.67 | 2131.33 | 443.33 | 98.83 | 2214.33 | 1444.83 | 232.67 | 1 | 1628.5 | 74.67 |
| | p-value | 10.42% | 46.89% | 15.82% | 12.73% | 7.58% | 46.13% | 78.34% | 37.04% | 40.62% | 13.69% | 19.34% |
| | right | 16672.45 | 16337.96 | 17220.88 | 15330.92 | 15522.78 | 17341.71 | 17283.53 | 11605.87 | 11051.81 | 11796.97 | 12497.45 |
| Pmean (L) | left | 20535.54 | 20929.17 | 21960.79 | 19713.1 | 20558.53 | 21357.55 | 21525.28 | 16089.76 | 16205.18 | 15187.84 | 17069.56 |
| | p-value | 1.56% | 0.55% | 0.78% | 0.91% | 1.98% | 3.57% | 2.13% | 0.80% | 0.17% | 0.83% | 0.78% |
| | right | 7425.67 | 2616 | 17273.58 | 8884.92 | 178365.1 | 2918 | 3521.5 | 84777.7 | 2378.83 | 2169.08 | 7652.83 |
| DPx (R) | left | 259435 | 252878.75 | 20743.25 | 377701.1 | 377818.38 | 17613.25 | 132931.12 | 251843.25 | 252063.62 | 126930.38 | 141956.62 |
| | p-value | 6.90% | 4.71% | 41.19% | 1.13% | 28.40% | 15.51% | 16.55% | 27.99% | 4.78% | 18.33% | 14.75% |
| | right | 4157.77 | 5274.38 | 18875.69 | 3147.31 | 45864.08 | 10191.31 | 5444.38 | 2470.92 | 3707.54 | 769544.69 | 469589.46 |
| DPy (R) | left | 3215.43 | 3761.14 | 11564.14 | 3657.43 | 2219.71 | 3530.29 | 4472.71 | 2357.71 | 2423.86 | 428877.57 | 4369.71 |
| | p-value | 4.97% | 25.98% | 5.20% | 52.74% | 29.18% | 44.19% | 35.85% | 86.13% | 3.83% | 14.13% | 2.88% |
| | right | 9058 | 2476.5 | 8170.92 | 2128.92 | 175808.25 | 2969.92 | 2763.92 | 84216.83 | 1975.08 | 1159.5 | 6197.67 |
| DPx (L) | left | 264135 | 254873.25 | 16478.75 | 377115.9 | 378215.75 | 10508.38 | 130495 | 250945.63 | 252336.13 | 126222.75 | 134194.63 |
| | p-value | 6.52% | 6.98% | 22.63% | 1.99% | 32.76% | 13.35% | 21.78% | 33.15% | 7.28% | 22.93% | 21.51% |
| | right | 5673.77 | 4810.7 | 12835.15 | 3950.92 | 28878.23 | 3899.85 | 5749.23 | 1293.31 | 1870.23 | 769445.92 | 467958.46 |
| DPy (L) | left | 4243.43 | 4024.57 | 13409.57 | 4164.14 | 2014.29 | 3657.86 | 4168.86 | 1356.57 | 1859.86 | 428799.57 | 5555 |
| | p-value | 37.17% | 41.98% | 77.19% | 87.69% | 30.40% | 73.74% | 14.55% | 88.00% | 98.42% | 14.15% | 3.00% |
| | right | 1.8 | 1.99 | 2.14 | 2.25 | 1.74 | 2.24 | 1.33 | 0.997 | 0.86 | 1.18 | 1.81 |
| Vmean (R) | left | 1.51 | 1.54 | 1.71 | 1.97 | 1.48 | 2.18 | 1.13 | 0.86 | 0.7 | 1.08 | 1.62 |
| | p-value | 29.15% | 16.18% | 20.70% | 46.83% | 43.65% | 90.51% | 24.40% | 39.41% | 25.54% | 68.62% | 48.46% |
| | right | 1.51 | 1.62 | 1.83 | 1.82 | 1.43 | 1.66 | 1.05 | 0.71 | 0.67 | 0.93 | 1.35 |
| Vmean (L) | left | 1.59 | 1.56 | 1.7 | 2.04 | 1.63 | 1.9 | 1.22 | 0.9 | 0.91 | 1.83 | 2.29 |
| | p-value | 76.02% | 84.95% | 67.97% | 61.26% | 60.25% | 50.98% | 24.70% | 20.28% | 14.09% | 0.04% | 0.36% |

TABLE IV. RECOGNITION ACCURACY USING DIFFERENT SVM KERNELS

| | *Linear* | *Polynomial* | *RBF* |
|---|---|---|---|
| $P_{max}$(R) | 66.70% | 61.90% | 61.90% |
| $P_{min}$(R) | 42.90% | 42.90% | 61.90% |
| $P_{mean}$(R) | 61.90% | 42.90% | 61.90% |
| $P_{max}$(L) | 81.00% | 76.20% | 76.20% |
| $P_{min}$(L) | 81.00% | 81.00% | 76.20% |
| $P_{mean}$(L) | 71.40% | 76.20% | 71.40% |
| $P_{max}$ | 90.50% | 95.20% | 90.05% |
| $P_{min}$ | 61.90% | 66.70% | 61.90% |
| $P_{mean}$ | 90.50% | 95.20% | 90.50% |
| $DP_x$(R) | 60.00% | 65.00% | 70.00% |
| $DP_y$(R) | 65.00% | 60.00% | 75.00% |
| $DP_x$(L) | 55.00% | 65.00% | 65.00% |
| $DP_y$(L) | 55.50% | 65.00% | 65.00% |
| $V_{mean}$(R) | 76.20% | 71.40% | 71.40% |
| $V_{mean}$(L) | 71.40% | 71.40% | 71.40% |

TABLE V. COMPARISON OF RECOGNITION ACCURACY.

| *Reference* | *Handedness Recognition Accuracy* |
|---|---|
| Ref. [4] | 90% |
| Ref. [5] | 90.70% |
| Ref. [10] | 95.97% |
| Proposed system | 95.20% |


Gesture Recognition," IEEE Transactions on Industrial Electronics, vol. 7, no. 59, pp. 2998-3007, 2012.

[7] J. Shin, "On-line cursive hangul recognition that uses DP matching to detect key segmentation points." Pattern Recognition, vol. 37, no. 11, pp. 2101-2112, 2004.

[8] W. D. Chang and J. Shin, "Dynamic positional warping: dynamic time warping for online handwriting," International Journal of Pattern Recognition and Artificial Intelligence, vol. 23, no. 05, pp. 967-986, 2009.

[9] I. Indyk and M. Zabarankin, "Adversarial and counter-adversarial support vector machines," Neurocomputing, vol. 356, pp. 1-8, 2019.

[10] E. Griechisch and E. Bencsik, "Handedness detection of online handwriting based on horizontal strokes," in 2015 13th International Conference on Document Analysis and Recognition (ICDAR), IEEE, pp. 1272-1277, 2015.