DYSIGN: Towards Computational Screening of Dyslexia and Dysgraphia Based on Handwriting Quality

Hana Ali Rashid*
hr05940@st.habib.edu.pk
Habib University
Karachi, Pakistan

Tasmiya Malik*
tm06183@st.habib.edu.pk
Habib University
Karachi, Pakistan

Iqra Siddiqui*
is06176@st.habib.edu.pk
Habib University
Karachi, Pakistan

Neelma Bhatti
neelma.bhatti@sse.habib.edu.pk
Habib University
Karachi, Pakistan

Abdul Samad
abdul.samad@sse.habib.edu.pk
Habib University
Karachi, Pakistan

ABSTRACT

Specific Learning Difficulties such as Dyslexia and Dysgraphia are characterized by struggles in reading and writing. Their diagnosis and intervention are critical as if left unattended, they can cause hindrance in academic activity, self-esteem, and long-term quality of life. Owing to the complex traditional processes for diagnosis, social stigma, and the general lack of availability of remedial therapists and clinical psychologists in Pakistan, this study explores the potential of handwriting quality features to be used in computationally screening for SLDs to make screening more accessible. This project consists of exploratory data analysis of handwriting scans of 25 children thus far, in the age group of 5 to 15, generating various handwriting quality features and using classification models to assess their potential. Our preliminary results are promising, with approximately 80% accuracy, thus showing potential for increased accuracy when paired with larger data samples and further feature generation.

CCS CONCEPTS

• **Applied computing → Education**; • **Human-centered computing → Accessibility**.

KEYWORDS

Specific Learning Difficulties, Pakistan, Handwriting, Dyslexia, Dysgraphia

ACM Reference Format:
Hana Ali Rashid, Tasmiya Malik, Iqra Siddiqui, Neelma Bhatti, and Abdul Samad. 2023. DYSIGN: Towards Computational Screening of Dyslexia and Dysgraphia Based on Handwriting Quality. In *Interaction Design and Children (IDC '23), June 19–23, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3585088.3593890

*Authors contributed equally to this research.

IDC '23, June 19–23, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0131-3/23/06.
https://doi.org/10.1145/3585088.3593890

1 INTRODUCTION

In Pakistani societies, academic success is considered one of the most significant markers of future success due to the potential for social mobility. Consequently, there is much pressure on students to perform well (i.e. get good grades) from parents and teachers alike, but not much consideration for students' actual learning which is reflected in the rote-learning and mark-scheme-oriented nature of education. As such, different styles of learning and Learning Difficulties (LDs) are hardly acknowledged, and struggling children are labeled lazy and unmotivated. This can have dire implications for the children such as lower self-esteem, motivation to learn, mental health, and long-term quality of life. In Pakistan, there is no standard system in place in schools to screen and support such students, and the prevalence of LDs is officially undocumented. There is a general lack of awareness and there are few centers and professionals in the main cities of Pakistan that cater to the diagnosis and intervention of Learning Difficulties, limiting the access of Pakistanis to these services.

We approached this problem by designing and developing an online screening system for LDs, which may potentially be accessed by parents and teachers anywhere in the country, and open the floor for discussions on LDs within society. This paper explores the potential of using handwriting as a basis for such a screener, due to the rich information that it provides as well as the relatively straightforward acquisition of handwriting data. We aim to explore the handwriting quality of children with **Specific Learning Difficulties** (SLDs), primarily Dyslexia and Dysgraphia which are directly linked to writing difficulties and identify features that may be used to classify handwriting scans into classes based on the risk of SLDs. These may then be used by future researchers to train classifiers and develop screening systems for use at homes and schools, thereby addressing the aforementioned issues.

1.1 Study Aims and Research Question

This project aimed to explore the potential of handwriting features, such as stroke width, curvature, text alignment, etc., in identifying the risk of SLDs Dyslexia and Dysgraphia. The research question guiding our explorations is "*what is the potential of using handwriting quality as a metric to computationally identify signs of the Specific Learning Difficulties Dyslexia and Dysgraphia in children?*". To the best of our knowledge, no existing study explores the handwriting

Table 1: Existing work using handwriting to screen for SLDs. Most are from recent years and have been conducted in a variety of contexts and languages like Hindi, Slovak and French.

Paper	Year	SLDs	Data Type	Data Collection	Sample Size	Ages	Models	Results	Language/Region
Development of CNN Transfer Learning for Dyslexia Handwriting Recognition [16]	2021	Dyslexia	Handwriting (images)	Self-simulated	138500 characters	Not Specified	Transfer learning, CNN LeNet-5	95.34%, 3 classes	Malaysia
Deep Learning Approach to Automated Detection of Dyslexia-Dysgraphia [20]	2020	Dyslexia, Dysgraphia	Handwriting (images)	Collected from schools existing	54 children, 267 characters	7-9	CNN	(86.14 ± 1.02)%	Hindi
Dyslexia and Dysgraphia prediction: A new machine learning approach - Dystech [15]	2020	Dysgraphia	Handwriting (images)	Collected using mobile app	1481 paragraphs	Not Specified	Majority class, Naıve Bayes, Logistic regression,Random Forest	Majority class 74.7%, Naive Bayes 90.8%, Logistic regression 95.6%, Random Forest 96.2%	Roman alphabet
Acquisition of handwriting in children with and without dysgraphia: A computational approach [8]	2020	Dysgraphia	Handwriting (digital tablet)	Collected from schools	280 children	9	Linear models, manually extracted features, K-means clustering	3/12 features significantly correlated	French
Dysgraphia detection through machine learning [6]	2020	Dysgraphia	Handwriting (digital tablet)	Collected from schools	120 schoolchildren	8-15	PCA tSNE, WkNN-FS, AdaBoost, Random Forest, SVMs	AdaBoost classifier 79.5%, SVM 78.8%, Random Forest 77.6%	Slovak
Pubudu: Deep Learning Based Screening And Intervention of Dyslexia Dysgraphia And Dyscalculia [10]	2019	Dysgraphia	Handwriting	Collected from schools	5000 characters	6-7	CNNs, SVMs	85%	Sri Lanka
Automated Detection of Dyslexia Symptom Based on Handwriting Image for Primary School Children [9]	2019	Dyslexia	Handwriting (images)	Collected from Association of Dyslexia Malaysia	30	Not Specified	ANNs	50 - 75%	English

quality features of children in Pakistan for this purpose, and there is currently no online screening service for SLDs in Pakistan.

2 LITERATURE REVIEW

2.1 Computationally Screening for Learning Difficulties

Existing work for screening for LDs (Learning Difficulties) consists of the use of various data types such as brain activity from EEG (electroencephalogram) scans [11, 21], eye-tracking [5, 14], web-based serious games [7, 10, 13], and handwriting [17]. Most of these works relied on primary data due to a lack of available datasets and achieved from 60% to 90% accuracy in their approaches. Most studies aimed to develop a screening system, instead of a definitive diagnosis, for use before seeking a professional diagnosis, due to the sensitive nature of the topic. Other than the variance in the type of data analyzed, existing work also varies in user age groups and the languages they are conducted in, including Spanish [14], Sri Lankan [10], German, English, Catalan [13], and more. This makes it challenging to compare work as SLDs, especially Dyslexia which includes trouble linking letters and words to their sounds, have slightly different manifestations in different languages given that they are closely linked to the characteristics of the language used to read and write. There appear to be no benchmark models, methods, or data sets so far. Instead, there is a high variance in experimental parameters across these studies which makes them challenging to compare.

We also came across a few commercially available screening tools for SLDs, including Lexplore [1] which is based on eye-tracking using specialized hardware, Dystech [3] which is a mobile app with a series of tasks that children attempt, and Dytective [2] which is a web-based game based in Spanish. For our work, we have referred to all of these and aim to develop a web-based prototype that is similar to these in nature.

2.2 Handwriting for Screening SLDs

We picked handwriting as a potential metric for detecting signs of SLDs given its convenient nature to acquire data (compared to EEG and MRI scans), as well as the possible signs of Dyslexia and Dysgraphia manifesting in handwriting. In particular, we used offline (paper-based) handwriting instead of digital which would have had a hardware requirement (digital tablet and stylus) and decreased access to the public that does not use tablets and/or are not used to writing on them.

Table 1 summarizes relevant studies involving handwriting that informed this study. We noticed a high variance in the languages used in these studies, such as [20] that worked with Hindi words, [8] who used French, and [10] that used regional Sri Lankan languages. Interestingly, [15] defined their samples based on the use of the Roman alphabet rather than language. This variance poses a limitation in comparing approaches and results as Dyslexia manifests differently in different languages, based on its “orthography” i.e. the relationship between letters and their sounds. For example, in languages like Italian which have a “shallow” orthography, children are less likely to make spelling errors compared to English, which has a "deep" orthography [4]. Moreover, the use of different scripts, such as the Hindi and Roman alphabet, also has the potential to impact handwriting quality, and spelling fluency [19]. Given the limited amount of studies that leverage handwriting to screen SLDs, there is potential for more work to be done in the area. It is also interesting to note that most studies use standard Machine Learning or Deep Learning approaches, calling attention to the need to explore optimized and specialized models to ensure better screening performance. The method of acquiring handwriting samples also varies across studies, as some use offline samples (on paper), and others use digital tablets for digital handwriting. While some insights are transferable across both methods, there are distinct characteristics (such as pressure in digital handwriting) of each that may make their approach and results differently. Therefore, reviewing existing literature reinforces the limitation of not having benchmark data sets and models to work with. However, given that most of the work is recent, it shows that this area has potential and is an open area of research.

3 METHODOLOGY

This section details our approach, from data collection, dataset formation, feature generation and analysis, and finally training binary classifiers.

Table 2: Features related to handwriting quality extracted from our datasets.

Feature	Description	Dataset
Stroke width	The average stroke width of all strokes in the input image, a measure of smoothness	words, characters, contours
Stroke curvature	Writing smoothness and confidence	words, characters
Character size	Size of each character written based on area, as a measure of consistency in handwriting, neatness, and confidence	words, characters
Axial alignment	Distance from midline of each word, for level of spatial awareness	words

3.1 Parameter Definition

The key focus of this study is on children in Pakistan aged 5-15 with Dyslexia, chosen as this age is when children are developing key reading and writing skills. Handwriting acquired in the English language will be used for analysis, for comparability of results with other work in English. Offline handwriting will be used as opposed to digital, due to its universal nature and lack of hardware requirements which is a convenience in collection. Also, a larger aim of this project is to help develop a screening tool that may be used by the public, and offline handwriting would allow a wider audience to access the tool.

3.2 Data Collection and Dataset Formation

Due to unavailability of English handwriting datasets for children with SLDs, we collected de-identified existing handwriting samples with the help of remedial therapists. For further data collection we aimed to conduct group writing activities with children diagnosed with SLDs, but due to issues of parental consent and stigma surrounding learning differences, we changed our approach and designed a remote data collection kit in the form of a booklet, which was distributed to parents/caregivers to complete with their children. The booklet consisted of three writing tasks: copying, writing days of the week and months in a year, and free writing based on a picture, selected as per recommendations of remedial therapists. The booklet also had written instructions for parents/caregivers as well as video instructions linked via a QR code. The consent form was attached with this booklet as well, and the child's age and diagnosis (if any) was mentioned by the parent.

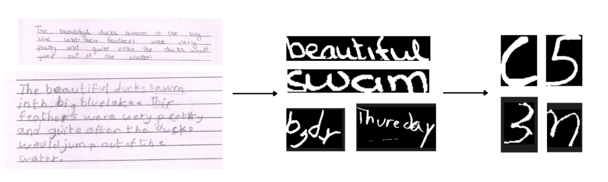

Figure 1: Raw handwriting data in the form of scanned pages are preprocessed and converted into two datasets: words and characters.

We have collected 121 pages of handwriting from 25 participants thus far, which were scanned and preprocessed as images (denoised and binarized) and converted into two types of datasets: characters (≈7000) and words (≈1500). Figure 1 shows samples from the datasets. Approximately 75% of the samples were from children diagnosed with SLDs, and the rest were from neurotypical children (children performing at the expected level as their age group).

3.3 Feature Generation

The datasets were then used for exploratory data analysis. The aim is to identify significant identifying features within handwriting that may help identify SLDs in children. We used handwriting and sketch features as a starting point for this stage in the work, inspired by work of [12], [18] and [8]. Table 2 lists the features we have explored so far. These were chosen as they give insights in the development of reading, writing and memory skills such as confidence in writing, motor control, neatness and spatial awareness. Some more features for ongoing work include consistency of spacing between words and letters, density of writing, spelling errors, signs of crossed out or erased writing and more. Each feature was generated for both datasets (characters and words), as well as a third dataset formed to augment the characters dataset by splitting each written character into its contours, with ≈26000 datapoints, to explore how that impacted results, if at all. The generated features were saved as CSV files for further analysis.

3.4 Models and Preliminary Results

To evaluate the potential of our intended approach, we trained several Deep Learning Models on characters and word datasets as images to detect patterns. They were labeled as "LD" and "NT" for Learning Difficulty (Dyslexia and Dysgraphia) and Neurotypical respectively. The models included 5-layered CNN and 3 pre-trained models for transfer learning: InceptionV3, EfficientNetB0, and VGG16. The train-test-split was 80-20. Table 4 summarizes the results.

The results show that handwriting quality does manifest enough patterns to screen for SLDs with 80% accuracy using InceptionV3. Other models also indicated rapidly decreasing loss function, and we believe the results can become better with increased epochs, but for our project, it gave us a green flag to carry on with our manual feature based analysis. To start with Machine Learning, we extracted various characteristics (Cartesian coordinates, frequency of angles, character size, etc.) and stored them in CSVs. Each of these characteristics was extracted from words and characters, which were then evaluated by SVMs, Naive Bayes (NB), and Random Forests (RF); the same datasets were also evaluated

Table 3: Example of experiments conducted using generated handwriting features to train various classifiers. The values are percentage accuracies for Support Vector Machines (SVMs), Random Forest (RF), Naive Bayes (NB), Multi-Layer Perceptron (MLP), and long short-term memory networks (LSTMs).

	Curvature (%)					StrokeWidth (%)					Spatial (%)				
	SVM	*RF*	*NB*	*MLP*	*LSTM*	*SVM*	*RF*	*NB*	*MLP*	*LSTM*	*SVM*	*SNN*	*NB*	*MLP*	*LSTM*
Characters Dataset	76.1	62.4	70.7	66	70.9	75.3	63.4	71.5	79	71.9	-	-	-	76	75
Words Dataset	77.3	62.7	77.3	74	61.3	76.2	73	52.3	63.4	80	80.0	80	60.3	79	78.9
Contours Dataset	80.3	73.5	79.9	79.88	76.9	79.8	76.9	76.2	-	-	-	-	-	-	-

Table 4: The table summarizes the accuracy of deep learning models trained on English words and characters from handwriting data.

Model	Words dataset	Characters dataset
CNN	76.29%	76.6%
InceptionV3	80.4%	73.1%
EfficientNetB0	77.3%	75.8%
VGG16	75.2%	74.4%

through neural networks (MLP and LSTM). Table 3 shows some of these experiments and their results.

4 DISCUSSION

This exploratory research was performed on handwriting data in the form of extracted characters, words, and contours. We used these to train binary classifiers using Machine Learning and neural networks. We also extracted various features corresponding to curvature, stroke width, and spatial awareness, and ran all the mentioned models displayed in Table 3. The best results we got for the curvature feature were through the contours dataset. However, SVMs performed the best with 80.3%. This implies that the contours' features (angle, edges, etc.), differ for children with SLDs and not. Children with SLDs often struggle with writing smooth curves and lines, resulting in a higher number of edges, strokes, and angle frequency, and also extreme angles between two strokes.

For the stroke width feature, models ran on the words dataset performed the best with 80% accuracy using LSTMs. The stroke width of characters taken per word gives a good measure of the variance of width in each sample. Children with LDs usually apply more pressure when writing which results in wider strokes. Spatial Awareness was another set of faetures tested on words and characters datasets. We achieved 80% accuracy with SVMs on characters and 80% accuracy with CNN. We recorded the text alignment (x, y coordinates, width, height, the distance of each character from center to top and bottom, etc.) and the size of characters (area of the bounding box). Children with SLDs struggle with the alignment of text and character consistency, which was indicated by these results. Overall, the results were promising that SLDs can be screened through handwriting. To form a single screening algorithm, we can either generate a collective list of features or use an ensemble approach and feed our models' results to a single neural network to give a single combined result.

A limitation of this study is the age group of participants which is from 5 to 15 due to lack of data. This is a broad group as children are developing rapidly at this time. As such, there is the possibility of noise in the data and models, and future work would do well to collect more data from a narrower age group so as to more effectively compare the information within them. The impact of children being multilingual on handwriting is also unexplored in this study, and may be worth pursuing in future work.

Our work shows the potential of handwriting as a simple yet effective medium to use to screen for signs of SLDs in children, which opens possibilities of developing online screeners so as to address the lack of professionals and resources in Pakistan. Such screeners may be used not only by caregivers at home, but may also be scaled to be used by schools due to their online nature, exposing more people to the concept of LDs and SLDs, and possibly leading to more work done for those struggling with them.

5 ONGOING AND FUTURE WORK

This study aimed to explore the potential of handwriting quality as a measure to identify signs of SLDs in children, particularly Dyslexia and Dysgraphia which are directly linked to handwriting. The project thus far set up the pipeline for data collection, dataset formation, feature generation, and analysis. We explored several features related to handwriting quality, including smoothness, curvature, axial alignment, and consistency in writing size. Our preliminary results are promising, with up to 80% accuracy of classifiers trained on the data from 25 participants. To draw our final conclusions for the project, we are in the process of collecting more handwriting data, with a goal of ≈ 30 more participants, so as to draw generalizable and robust conclusions on the performance of these features to identify signs of Dyslexia and Dysgraphia in children's handwriting.

ACKNOWLEDGMENTS

We would like to thank remedial therapist Maria Ali and clinical psychologist Raazieh Fatima for their guidance and expertise throughout this project.

REFERENCES

[1] 2015. The Research. https://www.lexplore.com/how-does-it-work/the-research/
[2] 2016. Dytective por Change Dyslexia. https://www.changedyslexia.org/
[3] 2020. Dystech - Assess and Monitor Reading Performance Online. https://dystech.com.au/
[4] Barbara Arfé, Francesca Corato, Eleonora Pizzocaro, and Anne Merella. 2020. The Effects of Script and Orthographic Complexity on the Handwriting and Spelling Performance of Children With Dyslexia. *Journal of Learning Disabilities* 53, 2 (April 2020), 96–108. https://doi.org/10.1177/0022219419892845

[5] Thomais Asvestopoulou, Victoria Manousaki, Antonis Psistakis, Erjona Nikolli, Vassilios Andreadakis, Ioannis M. Aslanides, Yannis Pantazis, Ioannis Smyrnakis, and Maria Papadopouli. 2019. Towards a Robust and Accurate Screening Tool for Dyslexia with Data Augmentation using GANs. In *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*. 775–782. https://doi.org/10.1109/BIBE.2019.00145 ISSN: 2471-7819.

[6] Peter Drotár and Marek Dobeš. 2020. Dysgraphia detection through machine learning. *Scientific Reports* 10, 1 (Dec. 2020), 21541. https://doi.org/10.1038/s41598-020-78611-9 Number: 1 Publisher: Nature Publishing Group.

[7] Ombretta Gaggi, Claudio Enrico Palazzi, Matteo Ciman, Giorgia Galiazzo, Sandro Franceschini, Milena Ruffino, Simone Gori, and Andrea Facoetti. 2017. Serious Games for Early Identification of Developmental Dyslexia. *Computers in Entertainment* 15, 2 (April 2017), 4:1–4:24. https://doi.org/10.1145/2629558

[8] Thomas Gargot, Thibault Asselborn, Hugues Pellerin, Ingrid Zammouri, Salvatore M. Anzalone, Laurence Casteran, Wafa Johal, Pierre Dillenbourg, David Cohen, and Caroline Jolly. 2020. Acquisition of handwriting in children with and without dysgraphia: A computational approach. *PLOS ONE* 15, 9 (Sept. 2020), e0237575. https://doi.org/10.1371/journal.pone.0237575 Publisher: Public Library of Science.

[9] Iza Sazanita Isa, Wan Nurazwin Syazwani Rahimi, Siti Azura Ramlan, and Siti Noraini Sulaiman. 2019. Automated Detection of Dyslexia Symptom Based on Handwriting Image for Primary School Children. *Procedia Computer Science* 163 (Jan. 2019), 440–449. https://doi.org/10.1016/j.procs.2019.12.127

[10] Ruchira Kariyawasam, Madhuka Nadeeshani, Tuan Hamid, Inisha Subasinghe, Pradeepa Samarasinghe, and Pasangi Ratnayake. 2019. Pubudu: Deep Learning Based Screening And Intervention of Dyslexia, Dysgraphia And Dyscalculia. In *2019 14th Conference on Industrial and Information Systems (ICIIS)*. 476–481. https://doi.org/10.1109/ICIIS47346.2019.9063301 ISSN: 2164-7011.

[11] Shankar K. Parmar, Oias A. Ramwala, and Chirag N. Paunwala. 2021. Performance Evaluation of SVM with Non-Linear Kernels for EEG-based Dyslexia Detection. In *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)* (2021-09). 1–6. https://doi.org/10.1109/R10-HTC53172.2021.9641696 ISSN: 2572-7621.

[12] Seth Polsley, Larry Powell, Hong-Hoe Kim, Xien Thomas, Jeffrey Liew, and Tracy Hammond. 2021. Detecting Children's Fine Motor Skill Development using Machine Learning. *International Journal of Artificial Intelligence in Education* (Oct. 2021). https://doi.org/10.1007/s40593-021-00279-7

[13] Maria Rauschenberger, Luz Rello, Ricardo Baeza-Yates, and Jeffrey P. Bigham. 2018. Towards Language Independent Detection of Dyslexia with a Web-based Game. In *Proceedings of the 15th International Web for All Conference* (New York, NY, USA, 2018-04-23) *(W4A '18)*. Association for Computing Machinery, 1–10. https://doi.org/10.1145/3192714.3192816

[14] Luz Rello and Miguel Ballesteros. 2015. Detecting readers with dyslexia using machine learning with eye tracking measures. In *Proceedings of the 12th International Web for All Conference* (New York, NY, USA, 2015-05-18) *(W4A '15)*. Association for Computing Machinery, 1–8. https://doi.org/10.1145/2745555.2746644

[15] Gilles Richard and Mathieu Serrurier. 2020. Dyslexia and Dysgraphia prediction: A new machine learning approach. https://doi.org/10.48550/arXiv.2005.06401 arXiv:2005.06401 [cs, stat].

[16] Mohamed Syazwan Asyraf Bin Rosli, Iza Sazanita Isa, Siti Azura Ramlan, Siti Noraini Sulaiman, and Mohd Ikmal Fitri Maruzuki. 2021. Development of CNN Transfer Learning for Dyslexia Handwriting Recognition. In *2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 194–199. https://doi.org/10.1109/ICCSCE52189.2021.9530971

[17] Katherine Spoon, David J. Crandall, and K Siek. 2019. Towards Detecting Dyslexia in Children's Handwriting Using Neural Networks.

[18] Xien Thomas, Larry Powell, Seth Polsley, Samantha Ray, and Tracy Hammond. 2022. Identifying Features that Characterize Children's Free-Hand Sketches using Machine Learning. In *Interaction Design and Children (IDC '22)*. Association for Computing Machinery, New York, NY, USA, 529–535. https://doi.org/10.1145/3501712.3535281

[19] Rana Yassin, David L. Share, and Yasmin Shalhoub-Awwad. 2020. Learning to Spell in Arabic: The Impact of Script-Specific Visual-Orthographic Features. *Frontiers in Psychology* 11 (2020). https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02059

[20] Pratheepan Yogarajah and Braj Bhushan. 2020. Deep Learning Approach to Automated Detection of Dyslexia-Dysgraphia: The 25th IEEE International Conference on Pattern Recognition. https://pure.ulster.ac.uk/ws/portalfiles/portal/88801667/GOOD_paperid_2_new.pdf

[21] A. Z. A Zainuddin, Khuan Y. Lee, W. Mansor, and Z. Mahmoodin. 2018. Extreme Learning Machine for Distinction of EEG Signal Pattern of Dyslexic Children in Writing. In *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*. 652–656. https://doi.org/10.1109/IECBES.2018.8626700