AUT Journal of Mathematics and Computing

AUT J. Math. Comput., 6(3) (2025) 223-239
https://doi.org/10.22060/AJMC.2024.22576.1176

Original Article

Machine learning approach for bipolar disorder analysis and recognition based on handwriting digital images

Ahmadali Jamali*[a], Reza Kargar[a], Shahin Alipour[b], Mohsen Rostamy-Malkhalifeh[a]

[a] *Science and Research Branch, Islamic Azad University (IAU), Tehran, Iran*
[b] *Department of Biomedical Engineering, University of Houston, USA*

ABSTRACT: In some cases, handwriting is a manifestation of the human mind, and it can reveal various psychological characteristics and mental disorders. Among these disorders, bipolar disorder is a well-known and widely studied condition in cognitive science and psychotherapy, and it can be detected in handwriting. In this research, we applied image processing techniques to analyze the handwriting characteristics of people with bipolar disorder based on their responses to a survey. We also proposed a machine learning model that can classify whether a person has bipolar disorder or not by using their handwriting as an input.

Review History:

Received:24 July 2023
Revised:12 December 2023
Accepted:14 April 2024
Available Online:01 July 2025

Keywords:

Bipolar disorder
Handwriting
Image processing
Machine learning

MSC (2020):

62H30; 68T07; 68U10

1. Introduction

One of the most recent applications of artificial intelligence (AI) is to aim at other science to improve and get better performance, such as complex calculating in mathematics and physics or weather prediction in meteorology and disease analysis, forecasting and prevention in medical science [33, 34].

Arrival AI in humanities is a challenging approach because emotional human behavior features are not simple and exact to measure. Finding features of emotional attributes is not as facile as natural science features because instead of material concepts, we are fronted with mental and abstract subjects. Clinical modern psychology encompasses extended mental disorders. Recognizing, prevention, and treatment are the main concern of psychological studies. One of the most substantial topics in mental disorders is Bi-Polar disorder(BD), which is being studied and is well-known as one of the hot topics in psychology and cognitive science. The general symptoms of this disease have been studied in terms of cognitive science and clinical psychology. The recognition of Bipolar disorder is one of

*Corresponding author.
E-mail addresses: ahmadali.jamali@srbiau, kargarreza57@yahoo.com, salipour@uh.edu, Mohsen_Rostamy@yahoo.com

the important studies as help to prevent and treatment of this mental disorder. This type of disorder has been studied and searched by both cognitive science and clinical psychology, moreover, in some cases, there are some computational and statistical methods for mental disorder recognition. In this paper, we intend to introduce a newly proposed method as BD recognizing with AI. As we know, before any disease recognition, we definitely need the symptoms of the disease as attributes, in this research the visual symptoms of BD are the handwriting of the patient. To better understand the problem definition; in the next part of the introduction, we are going to take a look at a short history of BD recognition research and its symptoms then we will express a little overview of the handwriting of Bi-Polar patient's symptoms and the methods of BD predictions and at the end of the introduction, the problem plan is presented.

1.1. Related works

This review consists of clinical psychology and computational statistics research methods.

1.1.1. A review of BD psychology

One of the most substantial research methods in BD is the MEDLINE search in 1997 [13] which surveyed the BD diagnosis and treatment and introduced BD as a major public health problem. Studies in 2013 by B-Birmaher et al. [4] have asserted that more than 3% of youngsters suffer from this type of familial mental disorder. In 2015 MK-Singh et al. [36] displayed that BD is a common and debilitating condition, often beginning in childhood and adolescence. Besides, based on current research the age and gender of the persons have got a relationship. According to the research of F-DE-Crescenzo et al. in 2017 [8], young people and teenagers are more susceptible to BD. In recent psychology paper by Luis-Ayerbe in 2018 et al. [2] the main systematic studies addressing symptoms of BD have been declared. Based on the T-Messer research in 2017 et al. [26], the risk behavior of BD involves; alcohol and drug abuse, sexual indiscretion and interpersonal problems, etc. The mortality risk of BD has been announced as suicide according to P-Dome in 2019 et al. [10]. This mental disorder encompasses two main mental symptoms; unipolar hyper manic/manic is one part, and depressive symptoms are the other unipolar symptom part. Both of these symptoms together characterized this affective disorder. In recent studies by E-Vieta et al. in 2018 [43] and FJA Gordoves et al. review paper in 2020 [15], the hyper-manic/manic and depressive of BD have been analyzed in terms of behavioral and genetic, in this research, the challenges and future directions of BD are the recognition and treatment method.

1.1.2. A review of BD handwriting symptom

The recent search of neuroscience by P-Stullet et al. in 2018 [39] displays that, BD affects hormones and neurons behavior. In fact, because of the switching process between mania and depressive mood, the hormones and neurons of BD patients are affected. In such a case, hormonal changes reach the brain and cause to change in neuronal activity in a region receptor-dependent manner according EHL-Umeoka et al. in 2021 [42]. This hormonal switching effects on lifestyle and behavior of the patient. New research in 2022 by A-Nustret et al. [29] has shown that these switching moods can be traced in the handwriting of patients. In this particular study revealed significant alterations in the handwriting of individuals experiencing manic episodes due to bipolar disorder. These handwriting characteristics can serve as a means of screening patients for release and predicting when a patient may transition into a manic state.

1.1.3. A review of Bi-Polar analysis and prediction by statistics and computational methods

Based on the last recent studies, BD recognition has been analyzed by statistics and computational methods, such as machine learning on tabulated datasets. According to recent reviews, the data mining of this recent research has been aggregated from electrophysiological techniques, genetic data or clinical measures surveys. These methods work based on surveys and biological tests and classification dataset [3, 7, 27]. Handwriting analysis is another approach whose data is collected based on digital images, in the earliest result in 2018 Perez-Arribas et al. [32] showed that the signature pattern of BD and borderline disorder persons are specific attributes. Besides, according to these studies with modelling of the signature, it can possible to predict the mood of a patient BD person with more than 70% accuracy. Around the new articles, there is a lot of study about the psychology of handwriting, these research do not relate to the specific mental disorder but about the behavior of handwriting. The main feature extraction of these studies is about the size of the letter, the length of space between the words, and the size of the margin. Personality detection by machine learning is the main approach of this type of research [5, 6, 28].

1.2. Problem plan

1.2.1. Main proposed problem definition

In this paper, we attempt to introduce a proposed artificial intelligence method for Bipolar disorder prediction from the digital image of handwriting survey approach. In such cases, the handwriting in the survey of a person by image processing and machine learning techniques will be predicted whether the person is a BD patient or not.

1.2.2. Population

In this research the surveys fill out by 50 persons which consists of 17 handwriting of different Bipolar persons and 33 non-Bipolar and because the access to the BD person is so rare and hard, in this research the number of handwriting are augmented to 113, which involves 45 bipolar handwriting and 68 non-bipolar.

1.2.3. Research summary

In this research, the handwriting of BD patience and non-BD has been aggregated and labelled under the supervision of a psychiatrist. The images of the manuscript are noise reduced and the features of handwriting are extracted by image processing techniques as data gathering. Besides, the data are analyzed and pre-processed for modelling with machine learning to predict whether the person is BD or non-BD.

1.2.4. Research plan

In the background of the paper, the prerequisites of the research are introduced. In the methodology, the main idea and proposed method are presented, which consists, of the handwriting survey data gathering way, and how to extract the features by image processing and pre-processing and modelling. In the result part, the data are analyzed and modelled with the dataset, and after that, there is a discuss the result and comparisons with other methods. And at the end, the conclusion and the future of the research are described.

2. Background

In this section the prerequisites of the methodology are presented which consists of two parts, psychological and computational background. If the topic of background is known the reader can skip.

2.1. Psychologic background

2.1.1. Bi-Polar disorder (BD)

Bi-Polar disorder (BD) is one of the eminent mental disorders which sometimes is known as manic-depressive disorder or abnormal mood swings. According to the new psychologic definition [25, 43]; Bi-Polar disorder (BD) is a recurrent cycle mood between hypomania/mania and severe depressive episodes. This constant sinusoidal behavior changes between these polar which causes Bipolar disorder. Hypomania or mania occurs when a person feels increased high energy and reluctance to sleep, abnormal talkative, lack of concentration, overthinking, confusion, and high libido. For more specialized mania or hypo mania definitions, it can be referred to the new clinical psychologic articles [11, 30]. On the other side, depression is another unipolar of this mental disorder whose general symptom involves; despair, emotional emptiness, sadness, anhedonia, and guilt. For more specialized depression information, the new clinical psychologic articles [18, 25] suggested. In general, if a person constantly feels this duality which consists polar of hypomania/mania and polar depression, she/he suffers from BD.

Besides the review of handwriting symptoms research which we have declared in 1.2.1, one of the visual symptoms of Bi-Polar patients is their handwriting as this disorder relates to the neurons, it affects the hand's neurons, besides, the mood of Bipolar patients is not stable, the thought process of Bi-Polar patient is not constant. In summary, when the person's mood during her/his lifestyle is sinusoidal, this unbalanced mood behavior shows itself in her/his daily work. The handwriting of bipolar patients is one of the traceable behavior.

2.2. Computational Background

2.2.1. Thresholding

Thresholding is one of the most popular simple and classic ways to discriminate between some objects. This concept is used for clustering or noise and saturation reduction between data. Besides, when we are sure that one variable cannot be more or less than the value we can be sure that our variable is limited between a range of values. In this case, we can consider the values which are out of this range as noise data. This discrimination condition can exist with one statement to more:

```
If value> n {
value = constant amount of a;}
else If value<m {
value = constant amount of b;}
.
.
.
/*The amount of m, n, a, and b, etc. are defined by the condition of the problem.
```

In fact, thresholding crop and determinate data between some linear or non-linear borders. The recent methods used thresholding as supplementary algorithm as segmentation [12, 31].

2.2.2. LOG(Laplacian of Gaussian)

LOG(Laplacian of Gaussian) is one of the eminent edge detection methods in digital image processing which is famous Marr-Hildreth model in 1980 [37]. It encompasses the Laplacian of the Gaussian filter, which is based on the Second derivative of the Gaussian filter as image sharping. As the derivatives of each function are linear, the LOG is a linear operator. The second derivative in matrix function defines so:

$$d^2 f = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y)$$

Which f is a matrix function, and in order to LOG, the Gaussian function calculated so:

$$G(x,y) = e^{-(x^2+y^2)/2\epsilon^2}$$

According to the second derivatives the LOG calculates so:

$$d^2 G(x,y) = \big((x^2+y^2-2\epsilon^2)/\epsilon^4\big) e^{-(x^2+y^2)/2c^2} \quad ; \quad d^2 G(x,y) \quad \text{called LOG}$$

LOG can be used as a noise reduction and image edge segmentation in 2D and 3D image processing. Which can relax the image from salt-pepper and Gaussian noises. The recent articles utilized it as pre-processing deep learning or image segmentation [40, 44].

2.2.3. K-means

K-means is one of the most famous and well-known unsupervised clustering algorithms, which is used for the Partitioning clustering data with a pre-defined k number of class. The method aims to partition the data set (x) into k number classes with k candidates' elbow points(the points created randomly) so that the sum of the square distance between data and mean(μ) of each class tend to be minimum, in fact finding the best centre clusters when the variant of each cluster becomes to the minimum value is the main goal of k-means clustering.

$$\text{Class}(ki) = \min \sum \sum (xi - \mu i)^2$$

To find the optimal center class, there are some algorithms with the extended application which are used in the recent article [19, 22, 45] K-means can be used in the digital image color or intensity clustering; in such a case or data are pixels with RGB values.

2.2.4. Linear Regression

Linear regression is one of the most basic statistics modeling supervised algorithms which can be used for the prediction of continuous data or data analysis and simulating data to a linear function of $Y = aX + b$. In fact, it finds a relationship between independent data and the linear function. The best fit function with data is created when the distance between data and predicted function (yi) becomes the minimum value.

Linear regression function attends to find the minimum of $\big(1/n(\sum(Yi - yi))^2\big)$ rely on linear algebra solution [16].

2.2.5. Entropy

Entropy in applied mathematics is one of the most famous measurement tools of ambiguity or disorder. In short, it can be said that the information obtained from observing an event is defined as the negative of the logarithm of the probability of its occurrence. The symbol of the Entropy function is $H(x)$ which defines the rate of disorder of discrete variables of x with a probability function of even of x.

$$H(X) = \sum_{i=1}^{n} P(x_i) I(x_i) = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i)$$

Furthermore, we can define the Entropy function with the expected value.

$$H(X) = E[I(X)] = E[-\log_b(P(X))]$$

Here I is the information content of X.

The Entropy in image processing is calculating the average information of a digital image pixels randomness degree [16]. In fact, the entropy of image signals can be expressed from the histogram of pixels value; the histogram of the image displays the probabilities of the intensity level of each pixel. By this means, we can calculate the measurement of ambiguity and disorder of signals. Recent articles have shown how to calculate the entropy of images by some different methods [16, 24].

2.2.6. Image augmentation

Image augmentation is common techniques to make new images from the original images. Making synthesis data depends on the dataset. By changing the size and making the Gaussian noises randomly and changing the light intensity of pixels, and rotation as long as the content of the photo is not lost and the information retains its meaning. Making a supervised synthesis dataset not only damaged the accuracy but only it avoids under fitting and overfitting of machine learning. For more information and study about various image augmentation techniques, it can be referred to the new review [46].

2.2.7. PCA

PCA or Principal Component Analysis, is an algebraic technique for dimension reduction or feature scoring to find and understand feature importance, and feature selection. It is used in data visualization and analysis and boosting dataset as pre-processing of machine learning [1, 23].

2.2.8. Random Forest

Random forest is based on a classic decision tree, the decision tree is a hyper-heuristics method, but the ensemble model of that work as hyper metaheuristics. Random Forest is an ensemble of decision trees; this method gets a vote from n decision trees which have been made from n random samples of the data set. Random Forest is useful to fill miss data and select the best features inherently based on the entropy score. For more information on random forests, it can be referred the new articles [38, 41].

2.2.9. Cat-Boosting

Cat-Boosting is another famous ensemble new method, which makes a decision tree and each time boosts the tree by the learning rate. This method avoids overfitting and can deal with miss data. CatBoost works by combining multiple decision trees to create a strong predictive model. It uses a technique called gradient boosting, which involves iteratively adding new decision trees to the model and adjusting their weights based on how well they perform on the training data. This Machine learning was used for handwriting character recognition by S-Ghosh et conference paper in 2022 [14]. For more information about Cat-boosing, check the new articles [9, 35].

2.2.10. Artificial Neural Network(ANN)

Artificial Neural Network(ANN) is a well-known supervised and unsupervised regression/classification machine learning approach which can model systems as forecasting and clustering the dataset. In fact, it is a function, which gets inputs of a set of X and returns the value of Y as an output. This machine encompasses the input layer, hidden layer/s and output layer. Hidden layer/s involves weights neuron which is updated with backpropagation, to be near to the output. This updating means neuron learning. ANN has got vast architectures depending on the dataset. The new studies are still working on the new architects and backpropagation methods [17, 35].

3. Methodology

As the main goal of this project is to detect the BD from their handwriting, we need absolutely the changing rate of each phrase in the survey in this particular we need image processing and feature extraction to measure the changing rate.

The prerequisites of this part have been described in the background part. This methodology involves; The proposed survey structure and data gathering, noise reduction and feature extraction by image processing and dataset statistical analysis, pre-processing and modelling.

3.1. The survey structure and data gathering method

The proposed model is based on supervised learning, in this case, our data should be labelled with expert and reliable supervisors, such as a psychologist, psychotherapist, psychiatrist, or neurologist who can recognize the BD patient from her or his behavior in an interview. The survey form should be filled out by the volunteers. The main goal of this survey is to show the traces of the Bipolarity of volunteers in the best way. In this research, we have designed a new survey method which can show better the feature of BP from their handwriting. The form with the environmental conditions should be written down. Besides, all the volunteers should write one predefined phrase. The sentences should be written with the below specific conditions:

3.1.1. Survey environmental conditions

- **Pen:** All of the volunteers should use one specific pen with the same color.
- **Paper:** All of the paper should be the same size as A4 and have the same color and texture. The paper should be clean and flat not crumpled.
- **Supervisor:** All of the volunteers should be interviewed by the same expert.
- **Environmental Room interview:** All volunteers should be in the normal room without any tension and stress. They should fill out the paper voluntarily, without any force and impose

The main reason for these conditions: All the *environmental conditions* must be *equal* for each person.

3.1.2. Survey structure conditions

- The phrase must be a predefined simple sentence without any complex words and grammar, the sense of the phrase should be clear. All of the volunteers should write this common phrase. The phrase sense should not induce the mind of volunteers. As an example "The apples are red and beautiful, when I eat them I feel happy".
- The predefined sentence should be repeated 6 times under each other so that between each iteration the volunteers should not see the last sentence.

The main reason for these conditions: The main proposed idea of this survey is to make a form which can show the Bipolarity pattern of the persons from their handwriting. As discussed in Psychologic background 2.1.1 and review of research 1.2.2, the handwriting of Bipolar persons during writing changes a lot. In such a case, the survey should be designed so that the changes could be measurable. This means all the persons should write the *same phrase* as a *fair comparison*, and they should repeat the sentences *6 times*. Because we want to show that each time the *sentences how change* and *compare* the 6 sentences' differences of the person to each other rather than word by word. The 6 number is an experienced choice and is not an exact number of repetitions.

3.1.3. Survey questions

According to the review of search 1.2.1 the **gender** and **age** of people could make a relationship with BD. In such a case the age and gender of a person should be defined in her/his form.

The general form of the survey is as follows.

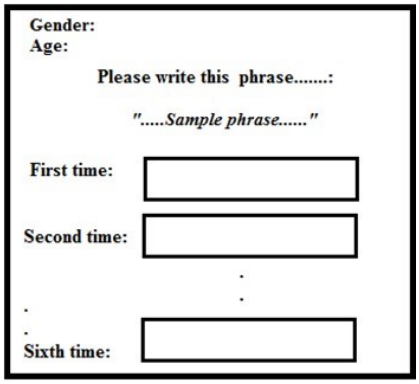

Figure 1: The sample of survey

3.1.4. Survey paper to digital image

As a data gathering, each survey should be scanned by the *same scanner machine*, not with a camera because all of the paper should have got the same light, intensity, contrast and same size.

3.2. Image Processing Part

In this part the images are cropped and pre-processed with noise-saturation reduction and the main component will be segmented. Then the images are prepared for feature extraction.

3.2.1. Image augmentation

As the number of Bipolar disorders is so limited and access to these people and getting information is not easy way, we can augment images from original images. As the data are unique we should be careful to do not to make useless, weak and outlier data. In this case, we augment just 25% of the image dataset for both non Bipolar and Bipolar with $\pm 4 - 6\%$ changes in the images, with close supervision image by image to the extent that the original contents are not damaged. The augmentation just implemented for train dataset as avoiding data leak prevention.

3.2.2. Survey phrase cropping

In this part, all the phrases of the survey should be cropped to the same size. For our survey plan, we have considered a size of 700*100. In the cropped image just the written phrase must exist not the margins board and typed element.

3.2.3. Dimension reduction

All the cropped images of the survey should be transferred from 3D to 2D images. In this research the colour component is not important, we need just the intensity of light.

3.2.4. Noise reduction

In this part image should be relaxed from noise and saturated pixels. The first noise reduction method is *thresholding*, the second part is *Laplace of Gaussian(LOG)*, and the third part consists of *K-means* and at the end *image masking* as semantic segmentation. The main goal of noise reduction is that clear noises and saturation pixels without damaging the handwriting.

Thresholding: In this part, the survey phrases image should be denoised. The main idea of this part is if any pixel is more than 220, becomes 225, so that:

```
(for any pixel >= 220)  {
pixel = 255;
}
```

This thresholding causes to remove the villi, lint and spots on the white paper and the paper background denoised. The border of 220 is an experimented number.

LOG: In this part, the image is denoised with the laplacian of the gaussian method which separates the edges of the image content. By this means the edge of the handwriting is separated from the paper.

K-means: in this section, the image intensity is clustered by k-means. In such a case the image is clustered in k class and the image pixels class number is reduced. This method causes a reduction of the intensity feature. The K number in this image processing part is 4.

Image masking: With image noise reduction methods we have separated the main information of handwriting and defined what is handwriting and what is background. After image segmentation methods by upper methods, the handwriting in the original image should be segmented from all the noises and saturated pixels. In the separated image, we have got just two parts: the *background* which has got just 255 pixels (white color) and the *handwriting* with non-white pixels($<= 240$). With masking, we transfer the main information of the handwriting pixels in the original image without any background and noise information.

In figure 2, the main noise reduction and handwriting separation is shown.

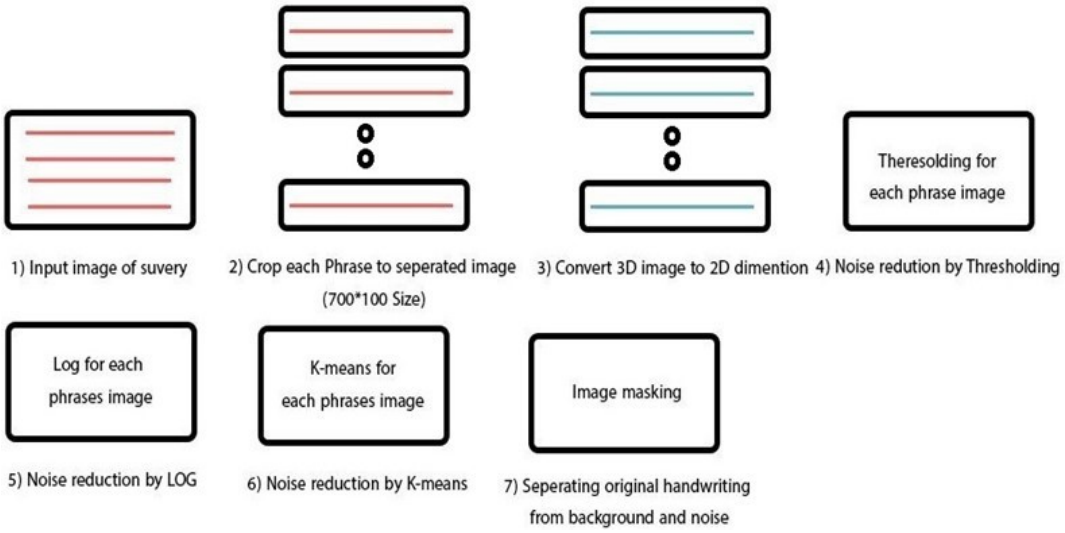

Figure 2: The main image preprocessing flowchart

3.3. Feature Extraction Part

In the image processing part, the handwriting is segmented from the background, as the segmentation is semantic we can access the main information in handwriting pixels and extract information. The main goal is to extract features of each 6 phrase which belongs to the survey.

Each feature extraction is based on the main idea and goal of the project. As we are finding a way to recognize the BD from handwriting we are going to extract features which can be related to this problem, as we have talked about in the introduction the BD patient's handwriting pattern changed in this order we should find a way to measure this changing rate.

From each handwriting, we can consider a lot of features such as color average, the variance of color, handwriting density, the entropy of handwriting, length of handwriting and the angle of handwriting. As we want to find the changes rate of each phrase in the survey we use to calculate the *variance* of each feature as a benchmark of handwriting *changing rate*.

In the following section, the extracted features are explained step by step:

- **Variance of Handwriting Density(VHD):** The handwriting density is calculated by the ratio of the pixel numbers of the handwriting of each phrase on the area., which is constant for each phrase surface (700*100). As we want to calculate the rate of changing of each phrase we consider the variance of Handwriting Density of every single phrase in one survey.

  In summary:
  VHD = Variance of 6 density phrases of the survey = V[Pixels number of the phrase(i)/area of phrase]
  We consider this feature as changing the handwriting density in the survey can make a correlation to the output.

- **Variance of Light Intensity Variance(VLV):** The variance of light intensity of each phrase is calculated and the total variance of 6 phrases is considered as the ratio rating of light intensity variance of each segmentation.

  In summary:
  VLV = Variance of total 6 phrases light intensity variance = V[light Intensity Variance of phrase(i)]
  We consider this type of feature as changing the variance of handwriting can make a pattern to the output.

- **Variance of Average of light intensity(VAI):** For this feature, the Variance of the average of light intensity of each phrase is calculated.

  In summary:
  VAL = Variance of total 6 phrases average light intensity variance = V[Average of phrase(i)]
  Considering Variance of color or light intensity can make sense to BP detection.

- **Variance of Entropy(VE):** The variance of Entropy of each phrase is calculated separately.
  As entropy of the system is most commonly associated with a state of disorder, this attribute can display the measure of handwriting disorder.

One of the most important features of handwriting of BP disorder is the changing rate of length and orientation. For calculation of these attributes, each phrase can convert to the linear mode by Linear Regression techniques.

- **Variance of Length(VL):** The length of each phrase should be calculated. One way to measure line of handwriting is using Linear regression. In this particular, the data of each phrase convert to the linear model. As each pixel of the phrase consists of X and Y coordinates, in this case, we can convert the phrase as a linear model of $Y = \alpha(x)$. By calculating the variance of the length of each phrase, we can get the amount of changes in the length of the handwriting, which parameter can relate to the Bi-Polar disorder symptoms.
- **Variance of slope(VS):** Based on Linear regression, we can find the orientation of each phrase. When we simulate a phrase to the line we can consider the slope of the line as the criterion for measuring the direction of handwriting. The slope is the gradient of $Y(x)$, which is the coefficient of X in $Y = \alpha(x)$ equation. The variance of the Slope can display the changing rate of orientation of each phrase.

These are the main features extracted from the image of the survey. Besides, we add more information from the voluntaries as a feature which can be related to the dataset results which involves, the *age* and the *gender* of the volunteers.

All features of each survey should be extracted and imported to the tabular dataset, for analysis and preprocessing and modelling.

3.4. Tabular dataset preprocessing

Powerful modelling needs definitely powerful preprocessing. The dataset should be managed by outliers. values and each feature should be verified by feature selection methods, here we utilize PCA to select the best features.

3.5. Modelling

The dataset is modelled by ensemble machine learning: Neural networks and Catboosting and Random Forest. For avoiding overfitting, the data. The best model selected is based on the average and variance of accuracy learning.

In the following figure, the road map of methodology is shown.

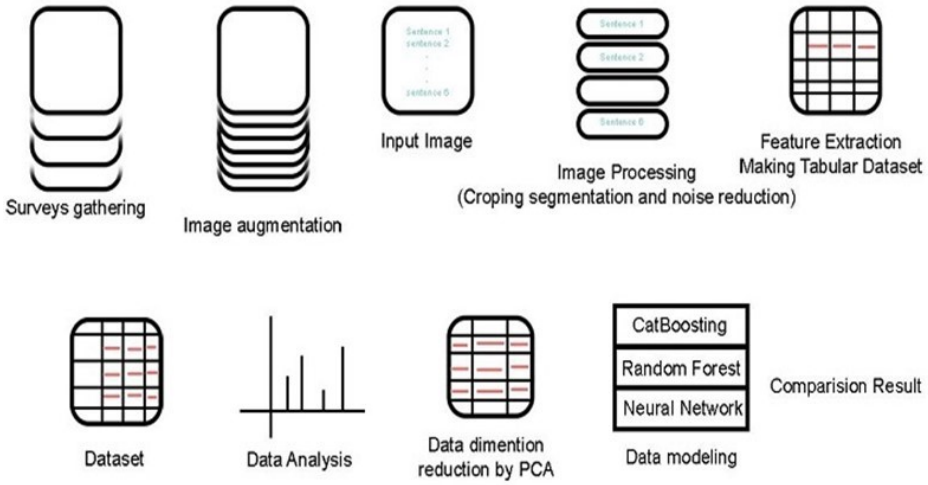

Figure 3: Main Idea of Methodology

4. Result

In this part, we are going to show the results of the outputs. The result encompasses image processing and dataset analysis and visualization and modelling. Before it, let's show the phrase which we selected for our survey. All phrases should be written in one language. In this paper, we used the Persian language with a specific meaningful neutral phrase:

دیروز هوا ابری بود باران بارید و زمین خیس شد.

It means: *Yesterday the weather was cloudy, it rained and the ground was wet.*

4.1. Image processing part

As the first computational step in the methodology is image processing 3.2 here we display the noise reduction example of one phrase of one survey.

Here is the original image of the first phrase of one survey example:

Figure 4: Original image of one phrase of the survey

In the image processing part the phrases should be denoised by forenamed techniquesin the 3.2.3:

Figure 5: Output result of image noise reduction by thresholding and LOG+k-means masked.

As figure displays, the output is clear from noise and saturation pixels without damaging the main information of handwriting. For a better understanding the noise reduction if we plot the entropy of the input image and output we can realize the differences.

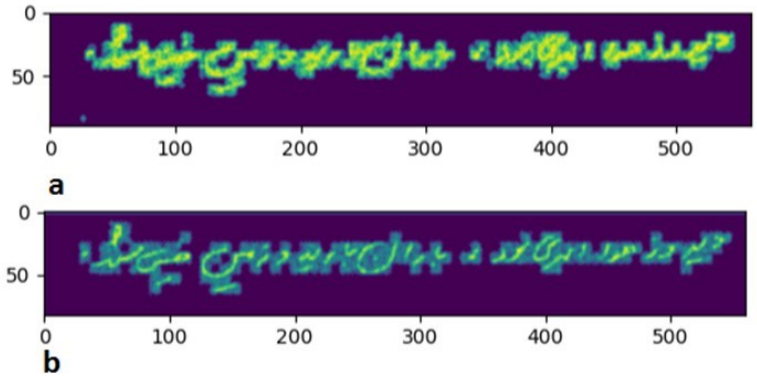

Figure 6: a: Entropy scheme of the original image example. b: Entropy of noise reduction output image.

Based on the figure the entropy of original image a is higher than noise-reduced output. The amount of entropy of the original image is 1.29 and the output image is 0.95. It means the entropy is 33% reduced without damaging the main information of handwriting.

The next table shows the mean entropy of the whole input original image and noise-reduced output image:

Table 1: Mean entropy of input/output images

Image	**Mean-Entropy**
Original images	1.34
Noised reduced images	0.97

Based on the table, the entropy of the original images is 38% reduced after noise reduction. After image pre-processing, we are going to feature extraction. All forenamed features in the 3.3 part are extracted and inserted in the dataset.

4.2. Data analysis and visualization

The tabular dataset is created by all survey images. As we discussed in the feature extraction in 3.3 we simulated each phase of the survey to the line by linear regression. In this case, we can reach the length of the lines and the scope of them. By this means, as data visualization, we can visualize each survey in terms of lines and scopes of each phrase. In this case, we can understand some differences between bipolar disorder and non-bipolar persons' handwriting. In figures 7 and 8, the handwriting linear simulation of the four bipolar disorders and four non-bipolar disorders has been displayed.

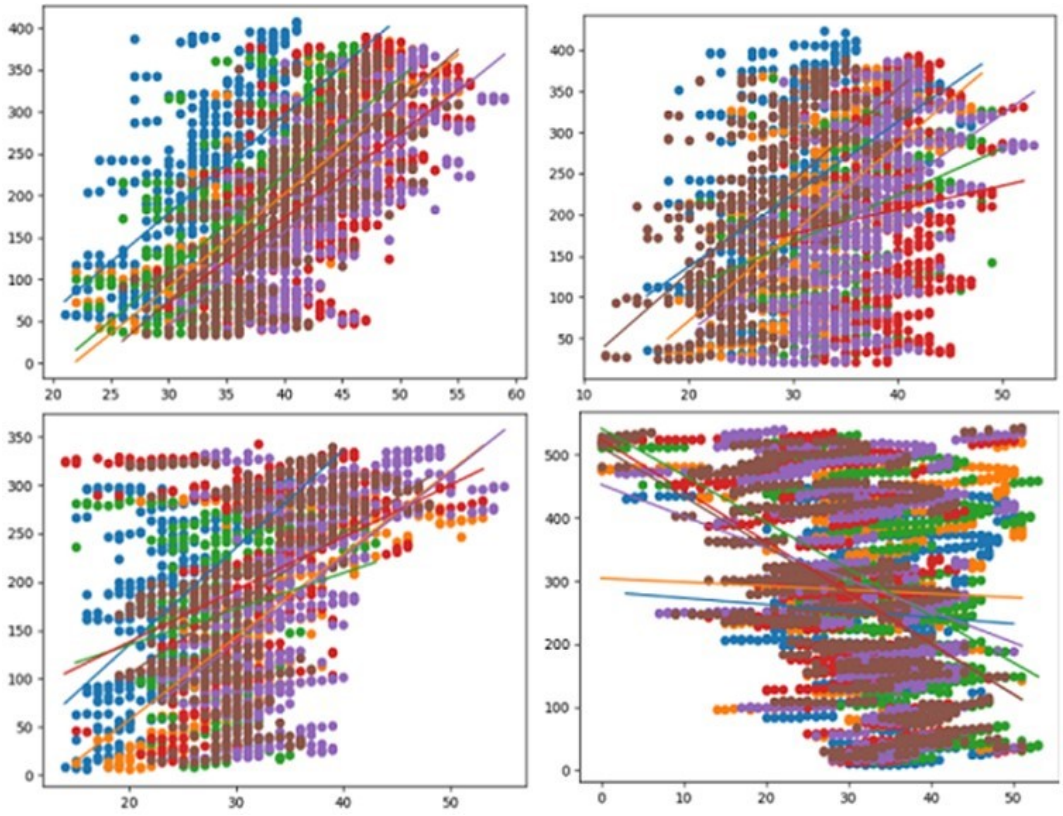

Figure 7: Linear regression visualization of four non-bipolar persons.

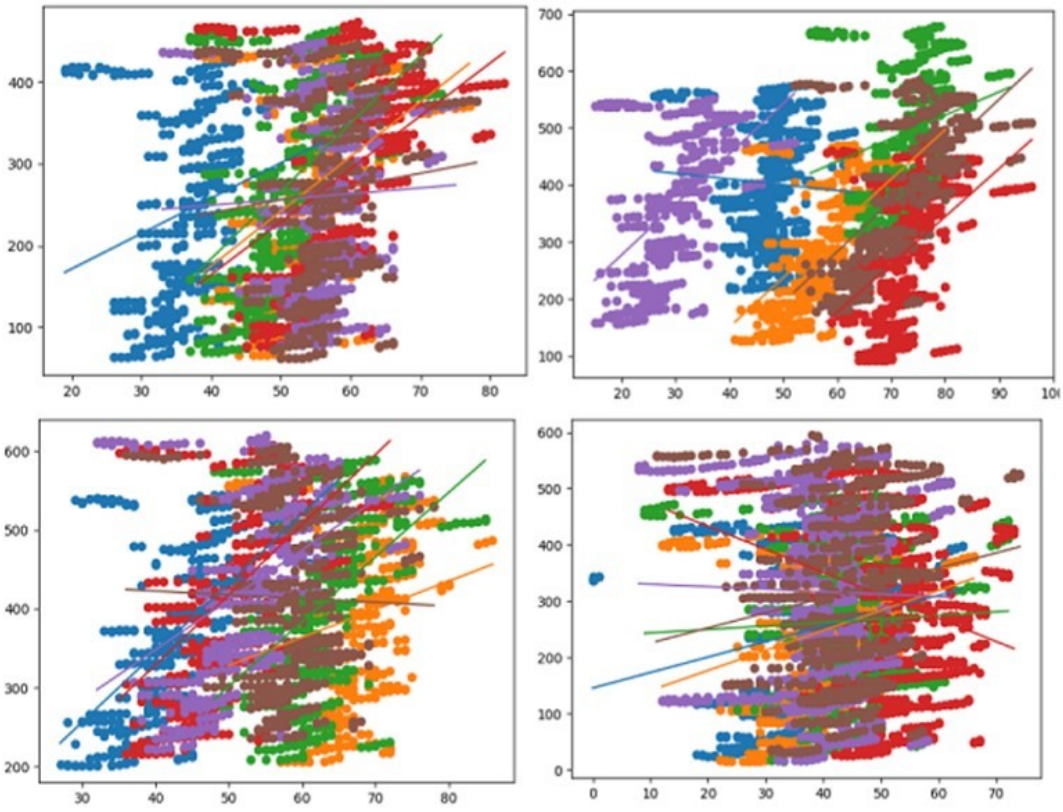

Figure 8: Linear regression visualization of four bipolar disorders.

Based on the figures each plot belongs to the one survey and each line is the simulated phrase and each point is the pixels of handwriting. Each color represents one of the six phrases. In fact, each plot (with 6 lines and dots) is abstract of the image survey (with 6 phrases). By converting image surveys to the liner regression visualization we can focus more on analysis.

One of the most obvious parameters in non-bipolar is that parallel lines, by contrast the length of the lines and slopes in bipolar charts are sparse with high variance rather than in non-bipolar persons. In figure 9 covariance matrix image shows the information on all features and acknowledges this context.

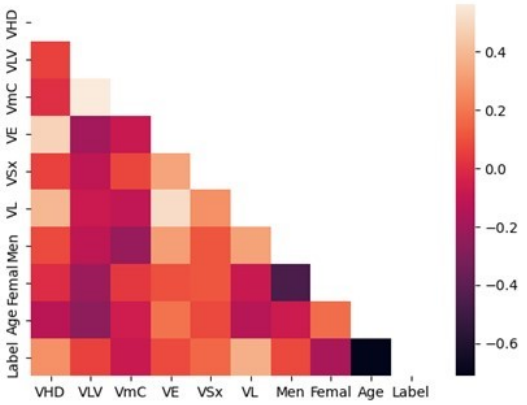

Figure 9: Matrix-covariance of the dataset

The covariance matrix displays the variance of the phrase length(VL) is highly correlated to the label and the feature of the age with negatively correlated to the label.

In order to score each column, PCA is a great tool for scoring and feature selection. The next table shows the score of each column.

Table 2: PCA score

Feature name	PC score
0 VLV	9.999748e-01
1 VHD	1.020577e-09
2 VmC	1.204409e-04
3 VE	-2.910059e-07
4 VsX	-8.050358e-05
5 VL	-7.093596e-03
6 Men	-3.041674e-06
7 Femal	-4.912702e-06
8 Age	-6.197852e-06

Based on the PCA table the VLV is the best column with high score. It shows that the Variance of Light Intensity Variance plays an effective role in recognition by modelling, furthermore age and gender have got low scores.

In figure 10 the T-SNE of the data set is demonstrated.

4.3. Modelling

The noises and outliers are managed and the dataset is prepared for modelling. In this step, we are using ensemble modelling and compare to each other. The 30% of dataset (31 random sample) separated for the test. The model candidates are *CatBoosting, Random Forest* and *neural network*. The parameter of each model are chosen by Keras open source library. Here is the parameter of each model:

- CatBoostClassifier `{ learning_rate=0.1, n_estimators=100, subsample=0.075, max_depth=2, Verbos = 8, l2_leaf_reg = 3, bootstrap_type= ‘‘Bernoulli’’, }`
- RandomForestClassifier `{ max_depth=2, n_estimators=100, min_samples_split=3, max_leaf_nodes=5, random_state=22 }`

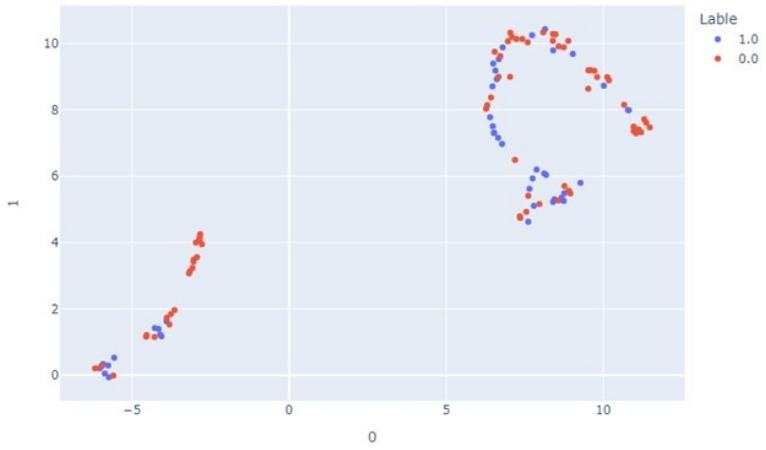

Figure 10: T-SNE of the data set

- Neural_network_MLPClassifier `{ solver='sgd', alpha=1e-5, hidden_layer_sizes=(5, 2, 2), random_state=0, drop_out = True, activator = sigmoid }`

The next Tables show the accuracy of the three models.

Table 3: The result of Train/Train

Model	test 1	test 2	test 3	test 4	test 5
Cat-Boost	0.91	0.91	0.96	0.92	0.94
Random Forest	0.89	0.91	0.91	0.88	0.93
Neural network	0.87	0.87	0.88	0.9	0.89

Table 4: The results of Train/Test

Model	test 1	test 2	test 3	test 4	test 5
Cat-Boost	0.9	0.92	0.9	0.91	0.92
Random Forest	0.94	0.94	0.9	0.96	0.94
Neural network	0.91	0.92	0.85	0.93	0.92

To select the best model, we should consider the overfitting and accuracy of each model. Each distance of the train/train would be higher than the test/train the model leads to overfitting, in this case we are going to select a model which except for high accuracy, it involves less variance accuracy and less distance with train/train.

Table 5: The mean of train/train and train/test of three candidate models

Model	Mean accuracy train/train	Mean accuracy train/test	Mean F1 score	Overfitting error
Cat-Boost	92.6	91	90.02	1.6
Random-Forest	92	89.8	89.3	2.2
Neural-network	91	88.8	87.7	2.2

As the results and the above table, the Cat-Boost model with low overfitting error and high accuracy is the best model. The parameters of max-dep and l2_leaf_reg are very effective in determining the accuracy in Cat-Boost.

In the following image the confusing matrix of the Cat-Boost model is displayed.

5. Discuss

By considering data visualization, we can understand that usually people who are not bipolar, write phrases approximately parallel with invariance shape, by contrast, BD write sentences sparsely and with high variance

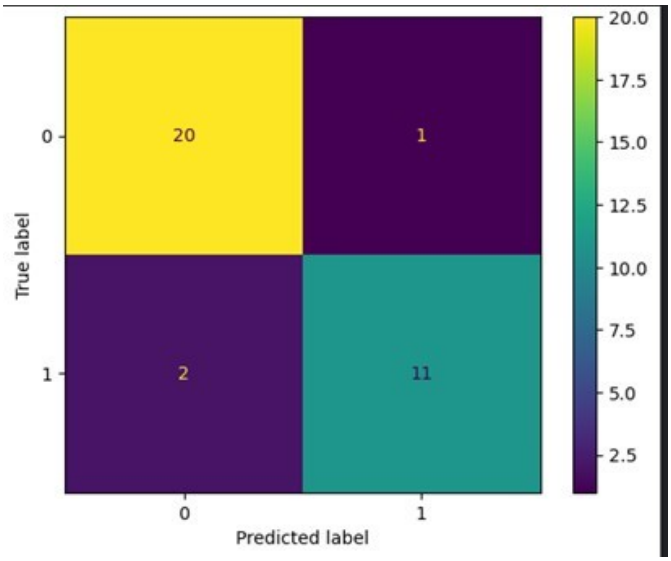

Figure 11: The Cohesion matrix of Cat-boost model

length and slops. It acknowledges the instability minds of bipolar disorder. According to the matrix- covariance and PCA, the VL and VLV are so effective to target recognition, on the other side the Age and gender features do not have as many points as others, but they are considered in the modelling. Based on the ML result Cat-Boost by forenamed parameters get better mean accuracy with less tendency to overfitting error.

6. Summary and Conclusion

As the main idea of this research is Bipolar disorder (BD) detection from handwriting, the definition of BD and its relationship with handwriting has been declared. The data set is gathered by filling out the proposed survey and labeled by psychiatrist. The images of survey were segmented from noise by criteria of entropy, based on the proposed combined noise reduction technic the entropy of original image 38% reduced after segmentation on average. After image processing, the related features were extracted and the tabulate dataset was created and scored by covariance matrix and PCA. Dataset modeled by Cat-boosting, Random-forest, Neural-network and compared. All model tested by train/train and train/test. On average, Cat-boosting model with high accuracy and less overfitting error is the best model with 92.6% accuracy in Train by Train and 91% in Train by Test and 90.02% F1 score. The main approach of this research, is that by filling out the provided survey form and introduced modeling, it is possible to determine with significant accuracy whether the person suffers from BD or not.

Application of this research helps to recognize bipolar disorder by handwriting, and it speeds up the diagnosis process of a psychiatrist or psychologist. On the other hand, by filling out the survey in schools it can show how many per cent of pupils suffer from bipolar, this fast statistics result can be beneficial from a sociological point of view to how to control, prevent and therapy.

Future work recommendation: Analysis and recognizing of bipolar disorder from the point of view of Natural Language Processing(NLP). By writing or typing some random meaningful words by the bipolar patient and normal persons and making a supervised dataset, it can be possible to find the relationship between the words by Markov chains of each person. By this particular test, it can be understood that random words by bipolar persons are sparser than normal persons or not, if it is, it can make a NLP model for recognizing bipolar disorder by random words. Furthermore, inspired by this research, other neurological disorders such as schizophrenia and obsessive-compulsive disorder recognition, can be investigated.

Acknowledgements

Special thanks to *Dr. Kazem Golchin* (Psychiatrist) in Clinic Gharb of Tehran for helping and supporting filling surveys from his patients. And advance thanks for its valuable information about bipolar disorder. And thanks to all patience and persons who spends her/his time on the survey filling.

Data and code availability

The original image of surveys and the dataset is available in Kaggle [20]. The main code of preprocessing and modelling is accessible as open scours Kaggle note-book [21].

References

[1] F. Anowar, S. Sadaoui, and B. Selim, *Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)*, Computer Science Review, 40 (2021), p. 100378.

[2] L. Ayerbe, I. Forgnone, J. Addo, A. Siguero, S. Gelati, and S. Ayis, *Hypertension risk and clinical care in patients with bipolar disorder or schizophrenia; a systematic review and meta-analysis*, Journal of affective disorders, 225 (2018), pp. 665–670.

[3] C. C. Bennett, M. K. Ross, E. Baek, D. Kim, and A. D. Leow, *Predicting clinically relevant changes in bipolar disorder outside the clinic walls based on pervasive technology interactions via smartphone typing dynamics*, Pervasive and Mobile Computing, 83 (2022), p. 101598.

[4] B. Birmaher, *Bipolar disorder in children and adolescents*, Child and adolescent mental health, 18 (2013), pp. 140–148.

[5] K. Chaudhari and A. Thakkar, *Survey on handwriting-based personality trait identification*, Expert systems with applications, 124 (2019), pp. 282–308.

[6] Y. Chernov and C. Caspers, *Formalized computer-aided handwriting psychology: Validation and integration into psychological assessment*, Behavioral Sciences, 10 (2020), p. 27.

[7] F. Colombo, F. Calesella, M. G. Mazza, E. M. T. Melloni, M. J. Morelli, G. M. Scotti, F. Benedetti, I. Bollettini, and B. Vai, *Machine learning approaches for prediction of bipolar disorder based on biological, clinical and neuropsychological markers: A systematic review and meta-analysis*, Neuroscience & Biobehavioral Reviews, 135 (2022), p. 104552.

[8] F. De Crescenzo, G. Serra, F. Maisto, M. Uchida, H. Woodworth, M. P. Casini, R. J. Baldessarini, and S. Vicari, *Suicide attempts in juvenile bipolar versus major depressive disorders: systematic review and meta-analysis*, Journal of the American Academy of Child & Adolescent Psychiatry, 56 (2017), pp. 825–831.

[9] J. de Prado-Gil, C. Palencia, P. Jagadesh, and R. Martínez-García, *A comparison of machine learning tools that model the splitting tensile strength of self-compacting recycled aggregate concrete*, Materials, 15 (2022), p. 4164.

[10] P. Dome, Z. Rihmer, and X. Gonda, *Suicide risk in bipolar disorder: a brief review*, Medicina, 55 (2019), p. 403.

[11] M. Faurholt-Jepsen, E. M. Christensen, M. Frost, J. E. Bardram, M. Vinberg, and L. V. Kessing, *Hypomania/mania by dsm-5 definition based on daily smartphone-based patient-reported assessments*, Journal of affective disorders, 264 (2020), pp. 272–278.

[12] J. Gao, B. Wang, Z. Wang, Y. Wang, and F. Kong, *A wavelet transform-based image segmentation method*, Optik, 208 (2020), p. 164123.

[13] B. Geller and J. Luby, *Child and adolescent bipolar disorder: a review of the past 10 years*, Journal of the American Academy of Child & Adolescent Psychiatry, 36 (1997), pp. 1168–76.

[14] S. Ghosh, *Comparative analysis of boosting algorithms over mnist handwritten digit dataset*, in Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2021, Springer, 2022, pp. 985–995.

[15] F. J. A. Gordovez and F. J. McMahon, *The genetics of bipolar disorder*, Molecular psychiatry, 25 (2020), pp. 544–559.

[16] B. S. Graham and C. C. de Xavier Pinto, *Semiparametrically efficient estimation of the average linear regression function*, Journal of Econometrics, 226 (2022), pp. 115–138.

[17] T. K. Gupta and K. Raza, *Optimization of ANN architecture: a review on nature-inspired techniques*, Machine learning in bio-signal analysis and diagnostic imaging, (2019), pp. 159–182.

[18] Y. Huang, L. Li, Y. Gan, C. Wang, H. Jiang, S. Cao, and Z. Lu, *Sedentary behaviors and risk of depression: a meta-analysis of prospective studies*, Translational psychiatry, 10 (2020), p. 26.

[19] A. F. Jahwar and A. M. Abdulazeez, *Meta-heuristic algorithms for K-means clustering: A review*, PalArch's Journal of Archaeology of Egypt/Egyptology, 17 (2020), pp. 12002–12020.

[20] A. Jamali, *Bipolar vs non bipolar handwriting.* `https://www.kaggle.com/datasets/ahmadalijamali/bipolar-vs-non-bipolar-handwriting`. Accessed: 2023.

[21] ——, *Data analysis and modelling bipolar vs non bipolar.* `https://www.kaggle.com/code/ahmadalijamali/data-analysis-and-modelling-bipolar-vs-non-bipolar`. Accessed: 2023.

[22] A. Jamali, M. Rostamy-Malkhalifeh, and R. Kargar, *The novel self-organizing map combined with fuzzy C-means and K-means convolution for a soft and hard natural digital image segmentation*, AUT Journal of Mathematics and Computing, 5 (2024), pp. 151–165.

[23] C. A. Kieslich, F. Alimirzaei, H. Song, M. Do, and P. Hall, *Data-driven prediction of antiviral peptides based on periodicities of amino acid properties*, in Computer Aided Chemical Engineering, vol. 50, Elsevier, 2021, pp. 2019–2024.

[24] M. Kopec, G. Garbacz, A. Brodecki, and Z. L. Kowalewski, *Metric entropy and digital image correlation in deformation dynamics analysis of fibre glass reinforced composite under uniaxial tension*, Measurement, 205 (2022), p. 112196.

[25] A. McGuinness, J. A. Davis, S. Dawson, A. Loughman, F. Collier, M. O'hely, C. Simpson, J. Green, W. Marx, C. Hair, et al., *A systematic review of gut microbiota composition in observational studies of major depressive disorder, bipolar disorder and schizophrenia*, Molecular psychiatry, 27 (2022), pp. 1920–1935.

[26] T. Messer, G. Lammers, F. Müller-Siecheneder, R.-F. Schmidt, and S. Latifi, *Substance abuse in patients with bipolar disorder: A systematic review and meta-analysis*, Psychiatry research, 253 (2017), pp. 338–350.

[27] B. Mwangi, M.-J. Wu, B. Cao, I. C. Passos, L. Lavagnino, Z. Keser, G. B. Zunta-Soares, K. M. Hasan, F. Kapczinski, and J. C. Soares, *Individualized prediction and clinical staging of bipolar disorders using neuroanatomical biomarkers*, Biological psychiatry: cognitive neuroscience and neuroimaging, 1 (2016), pp. 186–194.

[28] J. A. Nolazco-Flores, M. Faundez-Zanuy, O. A. Velázquez-Flores, C. Del-Valle-Soto, G. Cordasco, and A. Esposito, *Mood state detection in handwritten tasks using PCA–mFCBF and automated machine learning*, Sensors, 22 (2022), p. 1686.

[29] A. Nusret, O. Celbis, E. P. Zayman, R. Karlidağ, and B. S. Önar, *The use of handwriting changes for the follow-up of patients with bipolar disorder*, Archives of Neuropsychiatry, 59 (2022), p. 3.

[30] G. Olivito, M. Lupo, A. Gragnani, M. Saettoni, L. Siciliano, C. Pancheri, M. Panfili, M. Cercignani, M. Bozzali, R. D. Chiaie, et al., *Aberrant cerebello-cerebral connectivity in remitted bipolar patients 1 and 2: new insight into understanding the cerebellar role in mania and hypomania*, The Cerebellum, 21 (2022), pp. 647–656.

[31] S. Pare, A. Kumar, G. K. Singh, and V. Bajaj, *Image segmentation using multilevel thresholding: a research review*, Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 44 (2020), pp. 1–29.

[32] I. Perez Arribas, G. M. Goodwin, J. R. Geddes, T. Lyons, and K. E. Saunders, *A signature-based machine learning model for distinguishing bipolar disorder and borderline personality disorder*, Translational psychiatry, 8 (2018), p. 274.

[33] M. Rajabi, H. Golshan, and R. P. Hasanzadeh, *Non-local adaptive hysteresis despeckling approach for medical ultrasound images*, Biomedical Signal Processing and Control, 85 (2023), p. 105042.

[34] M. Rajabi and R. P. Hasanzadeh, *A modified adaptive hysteresis smoothing approach for image denoising based on spatial domain redundancy*, Sensing and Imaging, 22 (2021), p. 42.

[35] G. Ryan, P. Katarina, and D. Suhartono, *Mbti personality prediction using machine learning and smote for balancing data based on statement sentences*, Information, 14 (2023), p. 217.

[36] M. K. Singh, R. G. Kelley, K. D. Chang, and I. H. Gotlib, *Intrinsic amygdala functional connectivity in youth with bipolar I disorder*, Journal of the American Academy of Child & Adolescent Psychiatry, 54 (2015), pp. 763–770.

[37] T. Smith Jr, W. Marks, G. Lange, W. Sheriff Jr, and E. Neale, *Edge detection in images using Marr-Hildreth filtering techniques*, Journal of neuroscience methods, 26 (1988), pp. 75–81.

[38] J. L. Speiser, M. E. Miller, J. Tooze, and E. Ip, *A comparison of random forest variable selection methods for classification prediction modeling*, Expert systems with applications, 134 (2019), pp. 93–101.

[39] P. Steullet, J.-H. Cabungcal, S. A. Bukhari, M. I. Ardelt, H. Pantazopoulos, F. Hamati, T. E. Salt, M. Cuenod, K. Q. Do, and S. Berretta, *The thalamic reticular nucleus in schizophrenia and bipolar disorder: role of parvalbumin-expressing neuron networks and oxidative stress*, Molecular psychiatry, 23 (2018), pp. 2057–2065.

[40] R. Thillaikkarasi and S. Saravanan, *An enhancement of deep learning algorithm for brain tumor segmentation using kernel based CNN with M-SVM*, Journal of medical systems, 43 (2019), p. 84.

[41] H. Tyralis, G. Papacharalampous, and A. Langousis, *A brief review of random forests for water scientists and practitioners and their recent history in water resources*, Water, 11 (2019), p. 910.

[42] E. H. Umeoka, J. M. van Leeuwen, C. H. Vinkers, and M. Joëls, *The role of stress in bipolar disorder*, Bipolar Disorder: From Neuroscience to Treatment, (2021), pp. 21–39.

[43] E. Vieta, M. Berk, T. G. Schulze, A. F. Carvalho, T. Suppes, J. R. Calabrese, K. Gao, K. W. Miskowiak, and I. Grande, *Bipolar disorders*, Nature reviews Disease primers, 4 (2018), pp. 1–16.

[44] G. Wang, C. Lopez-Molina, and B. De Baets, *Automated blob detection using iterative laplacian of Gaussian filtering and unilateral second-order Gaussian kernels*, Digital Signal Processing, 96 (2020), p. 102592.

[45] C. Yuan and H. Yang, *Research on K-value selection method of K-means clustering algorithm*, J, 2 (2019), pp. 226–235.

[46] T. Zhou, S. Ruan, and S. Canu, *A review: Deep learning for medical image segmentation using multi-modality fusion*, Array, 3 (2019), p. 100004.

Please cite this article using:

Ahmadali Jamali, Reza Kargar, Shahin Alipour, Mohsen Rostamy-Malkhalifeh, Machine learning approach for bipolar disorder analysis and recognition based on handwriting digital images, AUT J. Math. Comput., 6(3) (2025) 223-239
https://doi.org/10.22060/AJMC.2024.22576.1176