iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** VI **Month of publication:** June 2018

DOI: http://doi.org/10.22214/ijraset.2018.6129

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue VI, June 2018- Available at www.ijraset.com

Predicting Human Behaviour through Handwriting

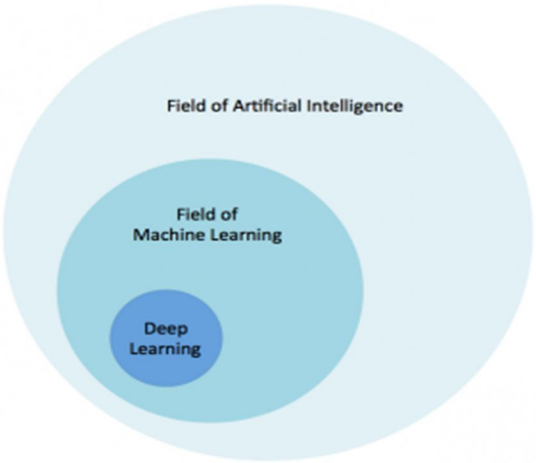
Prof. Akshita Chanchlani[1], Aakanksha Jaitly[2], Pratima Kharade[3], Rutuja Kapase[4], Sonal Janvalkar[5]
[1, 2, 3, 4, 5]*Computer Engineering, Sinhgad Academy Of Engineering, Kondhwa, Pune*

Abstract: ***Handwriting Analysis is a process that involves the analysis of person's handwriting. The scientific name for handwriting analysis is Graphology. It is a method of interpreting a person's behaviour and character traits from peculiarities in his handwriting. The mind is reflected by an individual person in many ways, writing is one of them. Professional handwriting examiners, called graphologists, can help understand and predict the personality of the writer with simply a sample of his handwriting. The accuracy of handwriting analysis depends on the skills of the graphologists. Nevertheless, the manual process of handwriting analysis is costly. Hence by using this system we can predict the personality traits automatically. The proposed system predicts the behaviour of a person from the characters. Using Convolution neural Network we train the dataset. Real time image is compared with validation dataset, which gives us output as behaviour prediction of person.***

Keywords: Graphology, Handwriting Analysis, Behaviour Analysis, Convolution Neural Network, Behaviour Prediction, Tensorflow.

I. INTRODUCTION

Graphology is the methodology of distinction, evaluating and understanding a person's temperament through the strokes and patterns unconcealed by handwriting. Handwriting reveals the true personality like emotional outlay, fears, honesty among many others. Graphologists provide a platform to the personality structure of an individual. By analyzing all characters of handwriting and interpreting them separately and together graphologists generate a view of the writer's character traits. throughout an individual's entire life period. behaviour is impacted by certain traits each individual has. There are different traits for different person and can produce different actions or behaviour from each person . Handwriting analysis one of the technique of predicting human behaviour. In our project we use the NIST dataset. The NIST stands for National Institute of Standard & Technology. Which provides the standard reference data for various purpose. We analyze this dataset and extracted those images which are going to required for our project. And for testing purpose we pass the real time handwriting image of person.

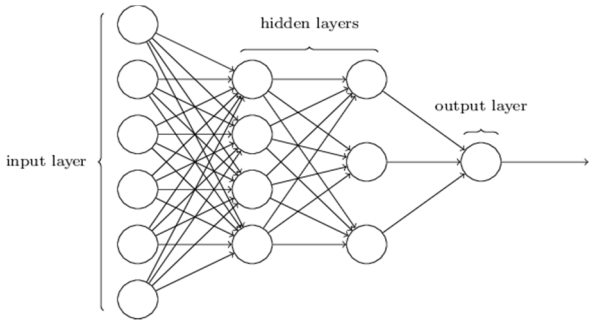


Deep learning is a specific subset of Machine Learning, which is a specific subset of Artificial Intelligence. For individual definitions:

A. Artificial Intelligence is for creating machines that can think intelligently.
B. Machine Learning is the subset of ANN, without explicitly programmed computer has ability to learn from data.
C. Deep Learning uses specific algorithm called a Neural Network.

849

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue VI, June 2018- Available at www.ijraset.com



Input values get passed through this "network" of hidden layers until they eventually converge to the output layer. There are three layers in ANN, that are input layer, output layer and hidden layer .The output layer is our prediction: it might be one node if the model just outputs a number, or a few nodes if it's a multiclass classification problem. The hidden layers of a Neural Network perform modifications on the data to eventually feel out what its relationship with the target variable is. Each node has a weight, and it multiplies its input value by that weight. The different types of Neural Network are available for use. Models like Convolution Neural Networks, Recurrent Neural Networks, Recursive Neural Networks. In our project, we are implementing Convolution Neural Network for behaviour prediction. A covnets is a sequence of layers, and every layer transforms one volume to another through differentiable function. Covnet has different layers such as input layer, convolution layer, activation function layer, pool layer and fully connected layer. The row input is pre-processed and passed to the CNN. In pre-processing various operations performed on image such as resize and gray-scale. And this is passes to the CNN layers at which different function is performed at different layer. The input layer holds the row input of the image with width 128 , height 128, depth 24. Convolution layer computes the output volume by computing dot product between all filters and image patch. At activation function layer, we apply the element wise activation function to the output of convolution layer. Some common activation functions are RELU: max(0, x), Sigmoid: 1/(1+e^-x), Tanh, Leaky RELU, etc. The pool layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents from overfitting. Two common types of pooling layers are max pooling and average pooling**.** Fully-Connected Layer is regular neural network layer which takes input from the previous layer and gives the class of output.

II. LITERATURE SURVEY

There are various techniques and algorithm for the predicting human behaviour through the handwriting. the algorithm such as Artificial neural network , Deep Learning , Machine Learning.

"Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM"Bayu Yudha Pratama Riyanart Department of Informatics Institut Teknologi Sepuluh Nopember Surabaya, 60111, Indonesia

Personality is a fundamental basis of human behaviour. Personality affects the interaction and preferences of an individual. People are required to take a personality test to find out their personality. Social media is a place where users express themselves to the world. Posts made by users of social media can be analyzed to obtain their personal information. This experiment

uses text classification to predict personality based on text written by Twitter users. The languages used are English and Indonesian. Classification methods implemented are Naive Bayes, K-Nearest

Neighbors and Support Vector Machine. Testing results showed Naive Bayes slightly outperformed the other methods.

The system will retrieve a collection of traits from users. Text from user then preprocessed into vector data. Classification process will classify user's text into a labeled dataset. The results are

predictions for each Big Five traits, primary personality characteristics and secondary personality characteristics which obtained from the combination between two traits. System developed is a web application.

A. "Vacuity Measure for Handwritten Character Analysis"

in this paper, we propose a study on the complexity measure of an object. It is based on the analysis of different details that may be limited by object contours. They may be holes or convexity evolution along the contour line. We focus in the same way on empty zones and filled zones. This study leads to a novel easure of the topology complexity – vacuity measure – that

quantifies the relation between emptiness or space and objects. Based on the vacuity measure, we propose to define a novel shape descriptor and the associated dissimilarity measure. They can be

850

applied in handwritten character analysis and in object recognition in general. The experiments are performed on a handwritten character dataset (ORIFLAMMS) and the object shape dataset.

However, human recognition process considers not only the object itself but also considers the elements not belonging to the object, for example, the holes and contour details of the object.

Such elements are defined the vacuity part of the objet. Human perception will capture both the vacuity and the object into the process. Therefore, we are working on a vacuity measure and

its application on pattern recognition. This measure quantifies the relation of object and its empty parts disposition.

III. PROPOSED METHOD

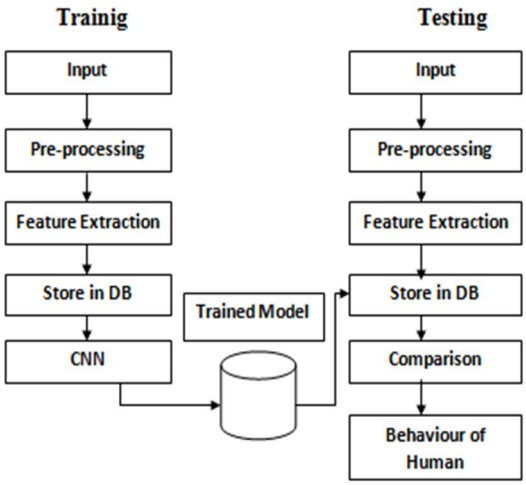


Figure 1: System Architecture

A. The system Architecture is Divided Into Two Parts

1) Training part

2) Testing part.

For giving input to the CNN dataset is prepared. The images in the dataset are pre-processed. We created five labels for dataset which is Criminal Intent, Excitable, Honest, Narcissistic and self centered, Persistent.

From the whole dataset 20% dataset is used for validation and remaining is used for train the system. For that various operations are performed on that images. The images are resized in particular size i.e 128*128. Then images are converted to the numpy array. After that gray-scale is performed by multiplying by 1.0/255.0. We created batch size of images as 64 that means at a time 64 images are analyzed. As we increase batch size generating number of epoch are decrease.

In training module , we used tensorflow for implementing the prediction system. In this module we imported the tensor flow libraries. Random seed is set and after that we configure the device i.e GPU:1 , CPU:4.

We start the session. In that set the batch size to 16. Five labels are created which are ‘criminal_intent’,’excitable’,’honest’, ’narcissist’, ’persistent’.

20% dataset is automatically used for validation. We load the all training and validation images and labels into memory using openCV and used that during training.

openCV is machine learning software library. Now placeholder is created for holding the input training image and all the images are read in the dataset and resized. Then we define the filter. After that we create weights and biases which are the parameter of the network. The objective of the training is to get the best possible values of the all these parameters which solve the problem reliably. The set of parameters used for process called Backward propagation, i.e. we start with a random set of parameters and keep changing these weights such that for every training image we get the correct output. In our project we created 2D tensor flow whose specific shape is predefined as part of network design.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue VI, June 2018- Available at www.ijraset.com

In Convolution Layer, all neurons apply convolution operation on the inputs, hence they are called convolutions neurons. Filter size is most important parameter of convolution neuron. We define the convolution layers which have the parameters i.e input, num_input_channels, con,_filter_size, num_filters. We slide convolutional filter over whole input image to calculate the output of convolution layer. In this case, we slide our window by 1 pixel at a time. This number is called *STRIDE*. 'padding=SAME' means we shall 0 pad the input such a way that output x, y dimensions are same as that of input. After convolution, we add the biases of that neuron, which are also learnable/trainable. Again we start with random normal distribution and learn these values during training.

Now, we apply max-pooling using tf.nn.max_pool function that has a very similar signature as that of conv2d function.

Pooling layer is mostly used immediately after the convolutional layer to reduce the spatial size(only width and height, not depth). This reduces the number of parameters, hence computation is reduced. The most common form of pooling is Max pooling where we take a filter of size F*F and apply the maximum operation over the F*F sized part of the image. After that, we use a RELU as our activation function which simply takes the output of max_pool and applies RELU using tf.nn.relu. There are various activation functions such as 'relu', 'sigmoid', 'tanh', 'elu'. From that we are using RELU because it is best for deep learning applications and also control the positive value and eliminate the negativity.

The Output of a convolutional layer is a multi-dimensional Tensor. We want to convert this into a one-dimensional tensor. This is done in the flattening layer. We use the reshape operation to convert to one dimensional tensor.

The next layer is the Fully Connected Layer. If each neuron in a layer receives input from all the neurons in the previous layer, then this layer is called fully connected layer. The output of this layer is computed by matrix multiplication followed by bias offset. In that, we declare weights and biases as random normal distributions. In fully connected layer, we take all the inputs, do the standard z=wx+b operation on it. Also RELU is used to activate the neurons. Here we finished defining building blocks of the network.Now, using these functions defined above we create the layers.

Then we get the probability of each class by applying softmax to the output of the fully connected layer. y_pred contains the predicted probability of each class for each input image.

The class having higher probability is the prediction of the network. We define the cost that will be minimized to reach the optimum value of weights. We use a simple cost that will be calculated using a

Tensorflow function softmax_cross_entropy with logits which takes the output of last fully connected layer and actual labels to calculate cross_entropy whose average will give us the cost.

Then we use the optimization fuction. We use AdamOptimizer for gradient calculation and weight optimization. We specify that we are trying to minimise cost with a learning rate of 0.0001. we run optimizer operation inside session.run(), in order to calculate the value of cost, the whole network will have to be run and we will pass the trainingimages in a feed_dict() in each iteration.where next_batch is a simple python function in dataset.py file that returns the next 16 images to be passed for training. Similarly, we pass the validation batch of images independently to in another session.run() call.

We are passing cost in the session.run() with a batch of validation images as opposed to training images. In order to calculate the cost, the whole network(3 convolution+1 flattening+2 fc layers) will have to be executed to produce layer_fc2(which is required to calculate cross_entropy, hence cost).

We can calculate the accuracy on validation set using true labels(y_true) and predicted labels(y_pred). We calculate the validation accuracy by passing accuracy in session_run() and providing validation images in a feed_dict. Similarly, we also report the accuracy for the training images. As, training images along with labels are used for training, so in general training accuracy is higher than validation.

We report training accuracy to know that we are at least moving in the right direction and are at least improving accuracy in the training dataset. After each Epoch, we report the accuracy numbers and save the model using saver object in Tensorflow.

In the Testing , we predict the real time handwritten image of user.

For that take the image pass it to the system. Same as pre-processing is performed on the training is performed in testing phase. After that network is define. Restore the save model which is created in training phase.

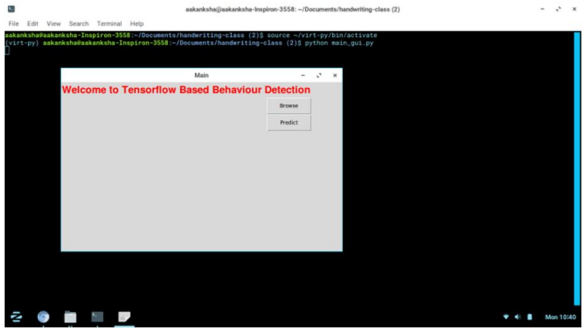
In the original network y_pred is the tensor that is the prediction of the network. Feed the image to the input placeholder . Then we create the feed_dict that is required to be fed to calculate y_pred.

Then we take the Maximum probability Class as output and it is show to the user.

852

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue VI, June 2018- Available at www.ijraset.com

IV. RESULT SET



There are option for user i.e. browse and predict. In the browse we take the input of handwritten image.

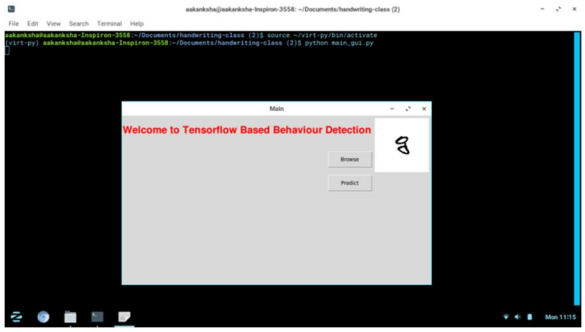
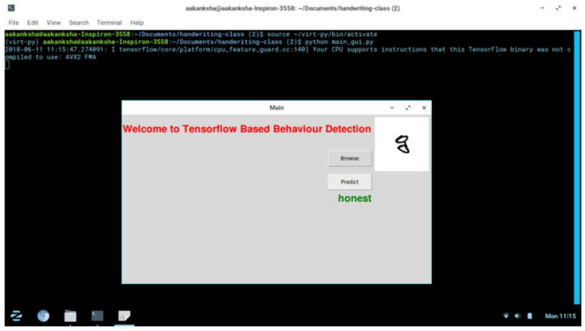


Image is loaded to the system.



Output is given as honest which predicted by system with maximum probability.

V. CONCLUSION

A method has being developed to predict the behaviour of a person from the features extracted from his handwriting. The personality traits determined by individual's handwriting are explored in this paper. This are input to the CNN which predicts the personality traits of the writer and displays those as the output.

VI. FUTURE WORK

The future work can include more handwriting features with the proposed method with cursive handwriting. A system can be designed and made user specific.

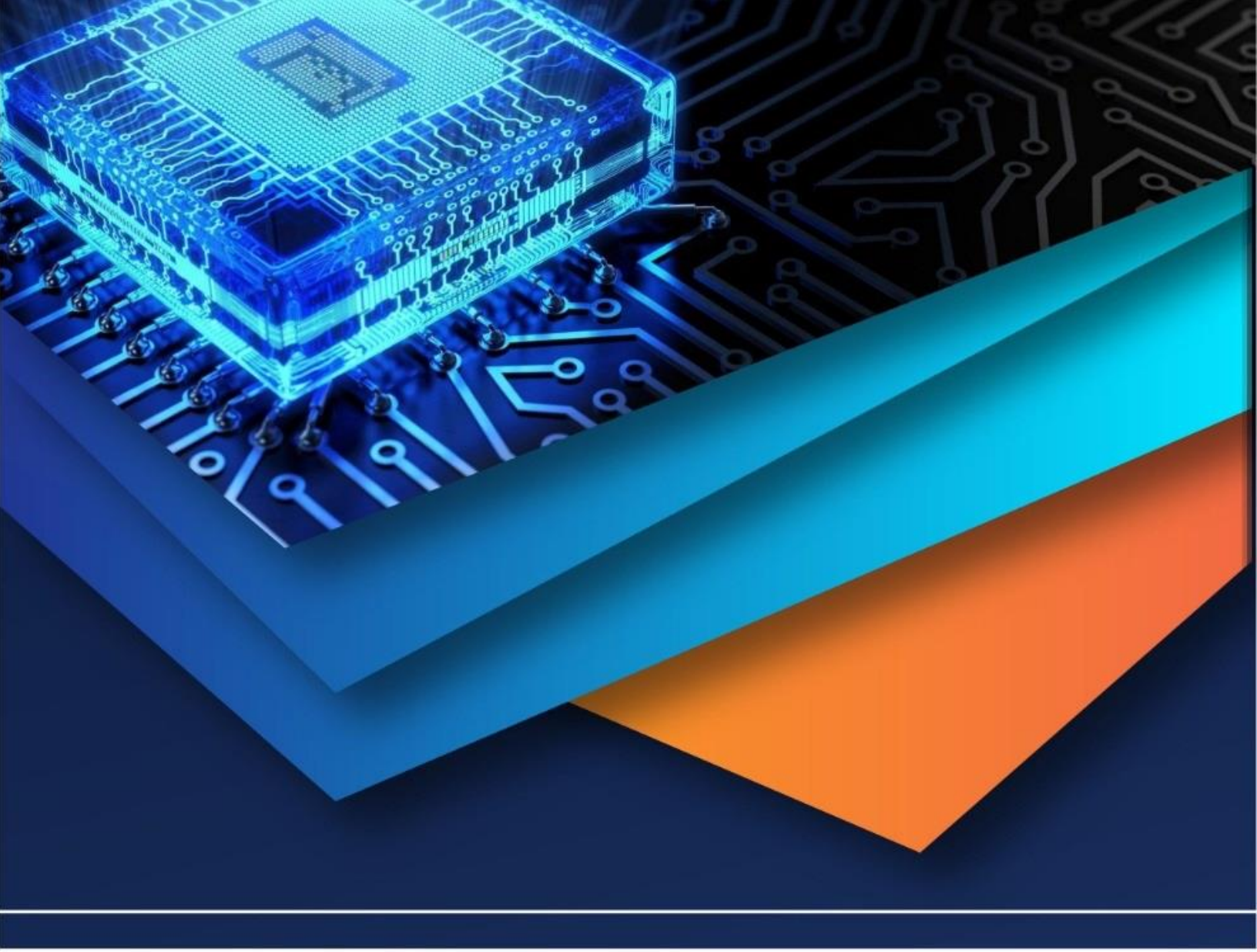
REFERENCES

[1] Behnam Fallah, Hassan Khotanlou "Identify human personality parameters based on handwriting using neural network"IEEE,2016
[2] Sh. Prasad, V. Kumar, A. Sapre, "Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine", International Journal of Computer Applications (0975 - 8887) , October 2010,Vol 8- No. 12.
[3] J. Fisher, A. Maredia, A. Nixon, N. Williams, J. Leet, "IdentifYing Personality Traits, and Especially Traits Resulting in Violent behaviour through Automatic Handwriting Analysis", Proceedings of StudentFaculty Research Day, CSIS, Pace University, May 4th, 2012.

853

[4] A. Rahiman, D. Varghese , M. Kumar," Handwritten Analysis Based Individualistic Traits Prediction", International Journal of Image Processing (TJIP),20 13, Volume , 2013.
[5] H.N. Champa , K.R.AnandaKumar , "Artiticial Neural Network for Human behaviour Prediction through Handwriting Analysis", International Journal of Computer Applications (0975 - 8887), May
[6] Butcher, 1. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A, & Kaemmer, B. (2010).The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring. Minneapolis, MN: University of Minnesota Press.
[7] Sh. Prasad, V. Kumar, A. Sapre, "Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine", International Journal of Computer Applications (0975 - 8887) , October 2010,Vol 8- No. 12
[8] R.Azmy, "recognition of printed texts Persian" doctoral thesis in electrical engineering, Tarbiat Modarres University, Summer (1387).
[9] S. M. E. Hossain and G. Chetty, “Human Identity Verification by Using Physiological and Behavioural Biometric Traits”, International Journal of Bioscience, Biochemistry and Bioinformatics, Vol. 1, No. 3, September 2011
[10] S. M. Lajevardi, Zahir M. Hussain, "Facial Expression Recognition: Gabor Filters versus Higher-Order Correlators,"International Conference on Communication, Computer and Power (ICCCP'09), Muscat, Oman, 15- 18 Feb. 2011. 126
[11] Swarna Bajaj, SumeetKaur, “ Typing Speed Analysis of Human for Password Protection (Based On Keystrokes Dynamics”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-3, Issue-2, July 2013. Er. Ankita Sharma,Sadhana” Human Behaviour Modelling & Analysis Using Artificial Neural Network”IEEE,2015
[12] Esmeralda C Djamal, SheldyNurRamdlan, Jeri Saputra, “Recognition of Handwriting Based on Signature and Digit of Character Using Multiple of Artificial Neural Networks in Personality Identification”, Information Systems International Conference (ISICO), 2 – 4 December 2013.
[13] Hugo Gamboaa and Ana Fredb, “A Behavioural Biometric System Based on Human Computer Interaction”.
[14] Albert Ali Salah1, Theo Gevers1, Nicu Sebe2, and Alessandro Vinciarelli3, “Challenges of Human behaviour Understanding”.
[15] Prof. Seema Kedar, Ms. Vaishnavi Nair, Ms. Shweta Kulkarni, Dept. of Information Technology,JSPM‟s RSCOE, “ Personality Identification through Handwriting Analysis”, Volume 5, Issue 1, January 2015 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering.
[16] Janet Fisher, Anish Maredia, Anita Nixon, Nerissa Williams, and Jonathan Leet,” Identifying Personality Traits, and Especially Traits Resulting in Violent behaviour through Automatic Handwriting Analysis”, May 4th, 2012 .
[17] EC. Djamal and R. Darmawati, “Recognition of Human Personality Based on Handwriting Using Multi Structures Algorithm and Artificial Neural Networks”, 2nd IEEE Conference on Control, Systems & Industrial Informatics Bandung, Indonesia June 23-26, 2013
[18] P.D. Sunar, “Read the full review Graphology Personality Through Writing His handwriting”, Yogyakarta: Diva Press, 2010.
[19] Taufik, M., Rosette, “Step By Step to analyzed personality trait through handwriting“, Jakarta, 2010.

854



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY