

“© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Age Estimation using Disconnectedness Features in Handwriting

¹V. Basavaraja, ²Palaiahnakote Shivakumara, ¹D. S. Guru, ³Umapada Pal, ⁴Tong Lu and ⁵Michael Blumenstein
¹Department of Studies in Computer Science, University of Mysore, Mysore, Karnataka, India.

Email: basavarajavyadav@gmail.com, dsguruji@yahoo.com.

²Faculty of Computer System and Information Technology, University of Malaya, Kuala Lumpur, Malaysia.

Email: shiva@um.edu.my.

³Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India. Email: umapada@isical.ac.in.

⁴National Key Lab for Novel Software Technology, Nanjing University, Nanjing, China. Email: lutong@nju.edu.cn

⁵University of Technology Sydney, Australia. Email: michael.blumenstein@uts.edu.au

Abstract—Real-time applications of handwriting analysis have increased drastically in the fields of forensic and information security because of accurate cues. One of such applications is human age estimation based on handwriting for the purpose of immigrant checking. In this paper, we have proposed a new method for age estimation using handwriting analysis using Hu invariant moments and disconnectedness features. To make the proposed method robust to both ruled and un-ruled documents, we propose to explore intersection point detection in Canny edge images of each input document, which results in text components. For each text component pair, we propose Hu invariant moments for extracting disconnectedness features, which in fact measure multi-shape components based on distance, shape and mutual position analysis of components. Furthermore, iterative k-means clustering is proposed for the classification of different age groups. Experimental results on our dataset and some standard datasets, namely, IAM and KHATT, show that the proposed method is effective and outperforms the state-of-the-art methods.

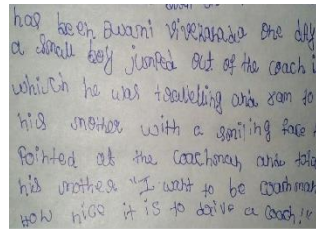
Keywords—Hu moments, Invariant moments, Disconnectedness, Age estimation, Handwriting analysis.

I. INTRODUCTION

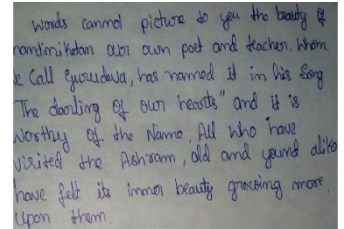
Age estimation is not a new problem for face recognition and other biometric fields. It has been used for improving face recognition performance in video surveillance to overcome the issues caused by uncontrolled environments [1-4]. However, due to innovations in health care solutions or fitness centers, it is hard to predict a person's age irrespective of gender with conventional methods. As a result, appearance does not match with age. In addition, when a human poses to an open environment, such as sun light and background variations, traditional methods may not perform well. In addition, the methods are said to be computationally expensive. These factors motivated us to introduce an alternate way for age estimation through handwriting analysis, which is simple and will not be affected by external influences compared to the face or skin of a human [5].

It is noted from the sample images shown in Fig.1 that as age changes, writing style changes in terms of font size, speed, force and shape of characters. This is justifiable because as a person grows in terms of maturity, responsibilities increase and therefore it reflects in the form of psychological stress during writing. As a result, we can see irregular or "choppy" writing in terms of disconnectedness between characters, words and lines as age increases. On the other hand, when age is low, writing is

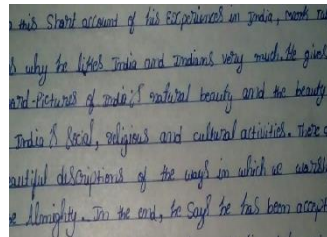
not affected much by psychological stress because individuals are less influenced by heavy responsibilities and/or stress. Therefore, one can expect legible writing with force, a uniform pattern and connectedness (regular pattern) between characters, words and lines when a writer is young. Based on these observations, a few methods have been developed in the literature [6, 7], which work well for a limited number of classes and the specific nature of datasets.



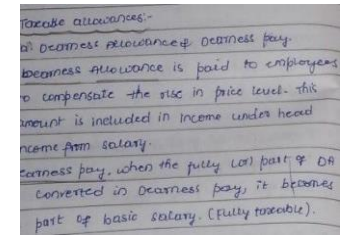
Class-1 [11-12]



Class-2 [13-16]



Class-3 [17-20]



Class-4 [21-24]

Fig. 1. Example of handwritten documents for each class of different age group.

For example, Maadeed and Hassaine [6] proposed automatic prediction of age, gender, and nationality in offline handwriting. The method extracts geometric features for handwriting components, such as direction, curvature, tortuosity, chain code and edge-based features. The features are then passed to a random forest classifier for classification. It is noted from the results that the method reports poor results for age estimation due to large intra-class variations. In addition, the focus of the method is not only focused on age estimation but also gender and nationality. Bouadjenek et al. [7] proposed age, gender and handedness prediction from handwriting using gradient features. The method explores histogram-oriented gradients and local binary pattern features, which are fed to an Support Vector Machine (SVM) classifier for age estimation, gender identification and handedness classification. However, the method considers only two classes, namely, under 35 years and above 35 years. Even though it considers the two-class

problem, the method reports poor results. In another work, Bouadjenek et al. [8] proposed a fuzzy integral for combining SVM-based handwritten soft biometrics prediction. The method proposes topological features and a gradient local binary pattern, and then it combines a fuzzy approach with an SVM classifier for classification. However, the method only works well for two classes with a large interval. Bouadjenek et al. [9] proposed robust soft-biometric prediction from offline handwriting analysis. The method is an extension of the previous method [8], hence it considers only two classes with large intervals for age estimation. Zouaoui et al. [10] proposed a co-training approach for improving age range prediction from handwritten texts. The method extracts the above-mentioned features [8], and then includes the concept of co-training, which is a semi-supervised learning model, to improve classification results. The method includes more samples and increases the age range for experimentation.

In light of the above discussions, it is noted that a few methods are proposed for age estimation, which are limited to two classes of age ranges from less than 35 years and greater than 35 years. Besides, these methods score poorly for two-class age classification. This shows that the problem is complex and it is hard to achieve better results. Most of the existing age estimation methods rely on traditional descriptors such as Histogram Oriented Gradients (HOG), Gradient Local Binary Pattern (GLBP), etc and classifiers [7, 9]. It is true that these features are good when there are significant differences in images. Therefore, we can conclude that there is scope for estimation with good results on more classes. In this work, we consider class-1 of age 11-12, class-2 of age 13-16, class-3 of age 17-20 and class-4 of age 21-24 year old students as shown in sample images of Fig.1. The reason to consider these classes is that estimating ages for these classes is challenging compared to the classes of older people because variations in writing for classes of young ages are not much.

Hence, in this paper, we extract a new feature called disconnectedness at Canny edge components of the input images. As mentioned in the Introduction Section, it is expected legible and clean writing with force, as well as a uniform speed when the writer is younger. It is the opposite as age increases. This is understandable because when writers are young, they are generally less-constrained, anxious, stressed, etc. But when a writer is older, due to psychological influences, one cannot expect good/legible writing when compared to a younger individual. In order to extract such observations, inspired by the method [11] where disconnectedness is studied for multiple component shapes, we explore the same disconnectedness for studying variations in handwritten documents that are written by different people at different ages. It is expected that variations are extracted by studying shapes, mutual positions, mutual values and spatial relationships between edge components of the same image. For the purpose of extracting disconnectedness, we propose to explore a new invariant moment [11].

II. PROPOSED METHODOLOGY

As age changes, one can see the effect in writing as shown in Fig.2, where we can see changes in font size of characters from class-1 to class-4. To validate the above observation, we

calculate the aspect ratio for edge components as defined by the height of an edge component divided by its width. The edge component is defined as pixels in the component that are connected each other without disconnections. The proposed method performs histogram operations on aspect ratio values as shown in Fig.3, which considers the aspect ratio on the x-axis and its frequencies are on the y-axis. It is observed from Fig.3 that the number of edge components that have uniform sizes between 0 and 1 decreases as the class number increases. This shows that font size variation increases as age increases. In addition, the curve, which represents class-4, has more variations compared to the other three classes as shown in Fig 3. In Fig. 3, there are components which have zero aspect ratio because we consider too small value of aspect ratio as zero. This is the basis for proposing a new disconnectedness measure to study the change in multiple shape edge components for age estimation.

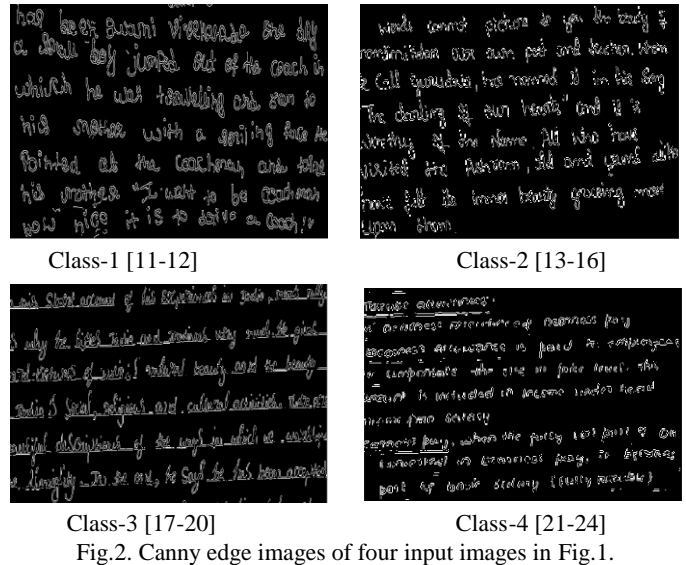


Fig.2. Canny edge images of four input images in Fig.1.

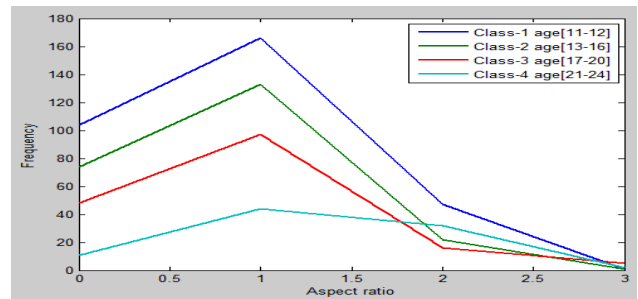


Fig.3. Aspect ratio vs. frequency of edge components of the images shown in Fig.2.

A. Text Components Detection

It is common practice that for handwriting, one can use ruled or un-ruled papers as shown in the sample images in Fig.1. This difference makes the problem more interesting because ruled lines in the image affect the shapes of edge components as in class-3 shown in Fig.2, where the ruled line connects all the edge components of the line as one component and hence it affects the shapes of the edge components. This is not true for un-ruled

images. Therefore, we propose to explore intersection point detection in Canny and Sobel edge images of the input image to remove ruled lines as defined in Equation (1). The proposed method checks the condition that if any pixel is connected to more than three neighboring pixels in edge images, the pixel is considered as an intersection point. It is true that these lines intersect at boundary points of character components. At the same time, Canny is good for both low and high quality images, while Sobel is good for only high quality images [12]. As a result, we can expect more disconnections in the case of Sobel edge images compared to Canny edge images as shown in Fig.4(a), where it is evident that Canny does not produce many disconnections compared to Sobel edge images. The effects can be seen in the output of intersection point detection for the edge components in Canny and Sobel edge images as shown in Fig.4(b), where Canny gives more intersection points while Sobel loses intersection points. Interestingly, intersection point detection eliminates the pixels that represent ruled lines as shown in Fig.4(b). Since one cannot predict the quality of a paper, the result of intersection point detection alone may not be sufficiently accurate for eliminating the ruled lines in an image. To overcome this issue, the proposed method performs an intersection operation as defined in Equation (2) for the results of intersection point detection of Canny and Sobel edge images as shown in Fig.4(c); there we can see pixels that are unlikely to represent ruled lines in the image, which results in seed pixels. The main advantage of this step is that it works well irrespective of ruled and un-ruled images because the seed pixels given by this step represent character components but not ruled lines in the image.

$$I_{ij} = \begin{cases} 1 & \text{If } X_{ij} \text{ is connected to three} \\ & \text{neighboring white pixels} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$SE = A \cap B \quad (2)$$

where, X_{ij} is the Canny edge image of the input image, SE denotes seed pixels, which are common to Canny (A) and Sobel (B) intersection points.

It is true that intersection point detection for Canny edge images provides more details about character components compared to Sobel edge images as shown in Fig.4(b) on the left side. To restore character components without ruled lines, for every pixel in the seed image, the proposed method connects corresponding neighboring pixels in the intersection point detection of the Canny edge image, which fall in the range of a three pixel radius through morphological operations. Next, the proposed method fixes the bounding box for the grouped pixels as shown in Fig.4(d), where it can be seen that the bounding boxes cover almost all the character components without ruled lines. Furthermore, the proposed method extracts components in the input image corresponding to bounding boxes as shown in Fig.4(e), which results in text components.

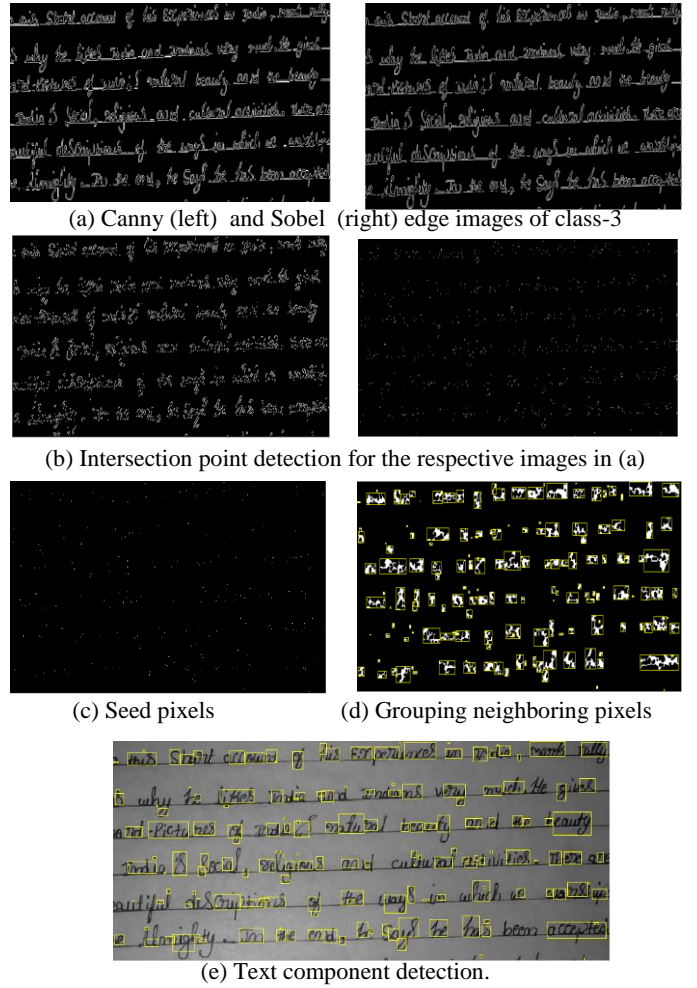


Fig.4. Text component detection by removing ruled lines.

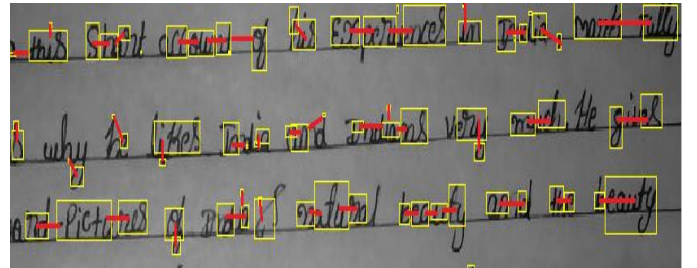


Fig.5. Grouping the first nearest neighbor of each text component in Fig.4 (e).

B. New Disconnectedness Measure for Multi-Shape Text Components

We believe as age changes, it reflects in handwriting in terms of character shape, stroke, font size, contrast, writing speed, force, etc. As a result, one can expect disconnectedness increases as age increases. Inspired by the method in [11] where Hu invariant moments are explored for measuring disconnectedness for multiple object shapes in different situations, we explore Hu invariant moments for extracting disconnectedness for multiple shape text components in handwriting images of different classes. Since Hu invariant moments for measuring disconnectedness requires more than

one component, the proposed method finds the first nearest neighbor component for each component in the text component image as shown in Fig.5, where we can see text component pairs.

For this, the proposed method finds the centroid for each text component and estimates the Euclidean distance between a text component and its neighbors. The component that gives the minimum distance to each text component is considered as the first nearest neighbor component. In this way, for every text component pair, the proposed method estimates disconnectedness $D(S)$ using Hu invariant moments as defined in Equation (3), where S denotes a text component.

$$D(S) = H(S_1 \cup S_2) - \frac{1}{M_{0,0}(S_1 \cup S_2)^3} \sum_{i=1}^2 M_{0,0}(S_i)^3 H(S_i) \quad (3)$$

where, S_1 and S_2 are two text components of the pair, $M_{0,0}$ denotes geometric moments, and $H(S)$ is the first Hu invariant moment. The first Hu invariant moment is calculated as defined in Equation (4).

$$H(S) = \frac{1}{M_{0,0}(S)^2} (\bar{M}_{2,0}(S) + \bar{M}_{0,2}(S)) \quad (4)$$

Geometric moments and central moments are calculated as defined in Equation (5)-Equation (7) and Equation (8)-Equation (10), respectively. Note that here $f(x, y)$ is a text component image, where x, y denote coordinates of a pixel, and $f(x, y)$ is the value of the pixel, and M, N are the dimensions of it.

$$M_{0,0}(S) = \sum_{x=1}^M \sum_{y=1}^N f(x, y) \quad (5)$$

$$M_{1,0}(S) = \sum_{x=1}^M \sum_{y=1}^N x f(x, y) \quad (6)$$

$$M_{0,1}(S) = \sum_{x=1}^M \sum_{y=1}^N y f(x, y) \quad (7)$$

Central Moments:

$$\bar{M}_{2,0}(S) = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^2 f(x, y) \quad (8)$$

$$\bar{M}_{0,2}(S) = \sum_{x=1}^M \sum_{y=1}^N (y - \bar{y})^2 f(x, y) \quad (9)$$

Where

$$\bar{x} = \frac{M_{1,0}(S)}{M_{0,0}(S)} \quad \text{and} \quad \bar{y} = \frac{M_{0,1}(S)}{M_{0,0}(S)} \quad (10)$$

For every text component pair in the image, the proposed method estimates disconnectedness measures as discussed

above, which gives a Feature Vector (FV) for the whole image. FV includes the Hu invariant moments corresponding to text components. The size of FV is the same as the number of text components. In order to extract degree of disconnectedness measures in handwritten images written by different aged groups, we propose to explore iterative k-means clustering with $k=2$, which gives two clusters. The reason for choosing k-means clustering is that the number of components in images affects iterative clustering process. Therefore, the number of iterations increases as disconnectedness increases. It is expected that as the number of classes increases, the number of iterations of k-means clustering process increases. This is justifiable because when a feature vector contains uniform values, most of the values are classified into one cluster and a few values are classified into another cluster, which is iteration-1. For the next iteration, when we apply k-means clustering again with $k=2$ for the cluster that has a few values, it is expected a cluster containing either a single value or null out of two clusters. When a cluster gets a single value or null, the iterative process stops. The process is illustrated in Fig. 6 for the image shown in Fig. 5, which is of class-3. Fig. 6 shows that iterative k-means clustering gives empty clusters marked by a red color at iteration-3. However, at iteration-4, the process gives empty clusters for both parent clusters. Therefore, the process stops at iteration-4. Note that if any one of the child clusters (output of k-means clustering) out of two clusters obtained by the parent cluster is empty (single value or null), the process stops and the cluster does not grow further. Iterative k-means clustering with $k=2$ is performed when the parent has two clusters without an empty one as shown at iteration-3 in Fig. 6.

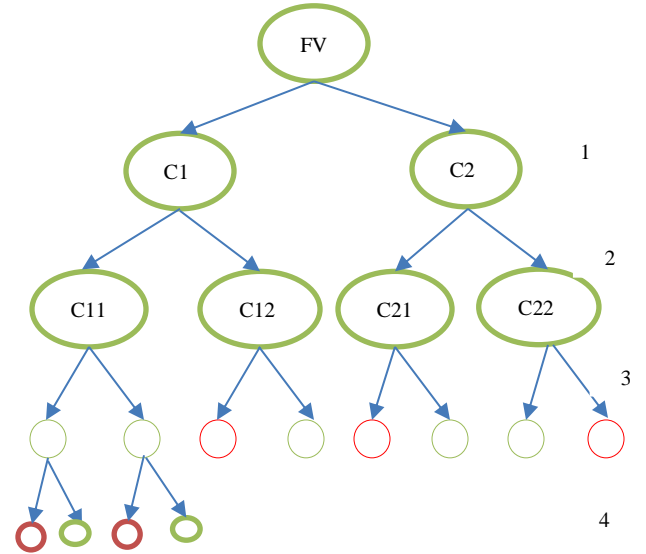


Fig. 6. Iterative k-means clustering for classification of four classes.

Based on the experimental results, we determine the number of iterations for classifying the four classes as defined in Equation (11), where it is noted that class-1 terminates at 1 or 2

iterations, class-2 terminates at 3 iterations, class-3 terminates at 4 iterations, and class-4 terminates at 5 iterations.

$$\text{Class} = \begin{cases} \text{Class - 1 age}[11 - 12], & \text{if iteration} \leq 2 \\ \text{Class - 2 age}[13 - 16], & \text{if iteration} = 3 \\ \text{Class - 3 age}[17 - 20], & \text{if iteration} = 4 \\ \text{Class - 4 age}[21 - 24], & \text{otherwise} \end{cases} \quad (11)$$

III. EXPERIMENTAL RESULTS

For experimentation, we consider our dataset by collecting handwriting images of students from 5th standard to post graduate, which includes age ranges from 11 to 24 years. We request each student in the class to write at least a half page on any topic in English without giving any instructions. As a result, the dataset consists of images with ruled lines and un-ruled lines, and different ink, pens, etc. Therefore, we divide the images into four classes, namely, class-1 with ages between 11-12 years, class-2 with ages between 13-16 years, class-3 with ages between 17-20 years and class-4 with ages between 21-24 years. Each class contains 100 images, which gives a total of 400 images. To test the objectiveness of the proposed method, we also use standard datasets, namely, IAM [7, 9] and KHATT datasets [7, 9], which consist of text lines of English and Arabic scripts written by different aged persons ranging from 16-56 years of ages. The IAM dataset includes 84 samples for each class, and the dataset is divided into two classes, namely, class-1 with ages between 25-34 years and class-2 with ages between 35-56 years [7]. Similarly, the KHATT dataset includes 135 samples for each class and the dataset is divided into two classes, namely, class-1 with ages between 16-25 years and class-2 with ages between 26-56 years [9]. In total, there are 400+168+270 = 838 images for evaluating the proposed method. We divide the classes as per the instructions and ground-truth given in [7, 9] for both IAM and KHATT datasets.

As per the instructions given in [7, 9], we use Classification Rate (CR) and a confusion matrix for measuring the performance of the proposed method. Therefore, we calculate CR and the confusion matrix for all the experiments. Sample successful classifications of the proposed method on four classes of our dataset, two classes of IAM and two classes of the KHATT datasets are shown in Fig.7. It is clear from Fig.7 that the proposed method works well for half page/full-page documents and text lines of different scripts. The reason for obtaining good results is that the proposed features for text component detection and Hu invariant moments for disconnectedness features are robust to both ruled and un-ruled documents/scripts. In addition, the proposed method does not depend much on the number of components in images as it involves iterative k-means clustering for classification.

To show the usefulness of the proposed method, we implement the method in [7], which extracts features based on Histogram Oriented Gradients, Local Binary Pattern and gradient information, and uses an SVM classifier for age estimation. However, the method is developed for two-class classification by considering text lines as the input but not for four classes as in the proposed work. As a result, the classification step of the existing method is tuned such that it can

be used for four classes. For standard datasets, namely, IAM and KHATT, we use the existing method [7] as the classification step of the proposed method, which is modified to run on two classes. The reason to choose this method for a comparative study is that the method is the state-of-the-art for predicting ages using handwriting analysis at the text line level.

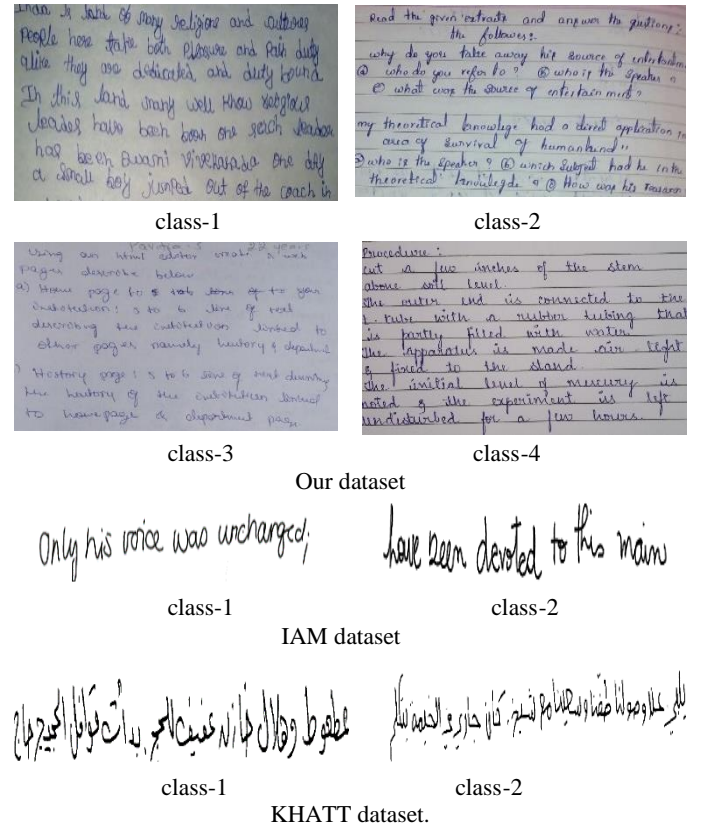


Fig. 7. Qualitative results of the proposed method for successful classification on our dataset and benchmark datasets.

Quantitative results of the proposed and existing methods for our dataset, IAM and KHATT datasets are reported in Table I, Table II and Table III, respectively. Table I, Table II and Table III show that the proposed method scores the best classification rate compared to the existing method. When we compare the results of our dataset and standard datasets, both the proposed and existing methods obtain better results for our dataset than on the standard datasets. This is because our dataset provides at least a half-page and maximum full-page documents for feature extraction, while the standard datasets provide text lines with a few components. In the same way, the reason that poor results are obtained on the existing method is that the features are not as robust as the Hu invariant moments proposed in this work, which extract disconnectedness. Therefore, overall, the proposed method performs the best since it is robust to script, ruled documents and different numbers of components in documents. Thus, it can be extended to more classes and other applications in the future.

Table I. Age classification rates of the proposed and existing methods on our dataset.

Classes	Proposed Method				Existing Method [7]			
	Class-1	Class-2	Class-3	Class-4	Class-1	Class-2	Class-3	Class-4
Class-1	68.0	26.0	6.0	0.0	64.3	17.2	5.1	13.4
Class-2	22.0	65.0	12.0	1.0	20.5	67.2	4.7	7.6
Class-3	3.0	8.0	70.0	19.0	14.6	10.1	60.7	14.6
Class-4	2.0	10.0	26.0	62.0	8.6	15.7	10.4	65.3
CR	66.25%				64.37%			

Table II. Age classification rates of the proposed and existing methods on the IAM dataset

Classes	Proposed Method		Existing Method [7]	
Classes	Class-1	Class-2	Class-1	Class-2
Class-1	62.7	37.3	60.2	39.8
Class-2	35.5	64.5	36.3	63.7
CR in (%)	63.6		61.9	

Table III. Age classification rates of the proposed and existing methods on the KHATT dataset.

Classes	Proposed Method		Existing Method [7]	
Classes	Class-1	Class-2	Class-1	Class-2
Class-1	67.40	32.59	65.2	34.8
Class-2	38.51	61.48	44.6	55.4
CR in (%)	64.44		60.3	

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a new method for age estimation using handwriting analysis based on Hu invariant moments and disconnectedness features. The proposed method works based on the fact that as age increases, variations in handwriting also increase. The proposed method explores intersection point detection, Canny and Sobel edge images for detecting text components regardless of ruled or un-ruled document images. For each text component pair in a handwritten document image, the proposed method introduces new Hu invariant moment features for extracting disconnectedness in terms of shape, font size, stroke, etc. The extracted features are passed to an iterative k-means clustering method for the classification of different age groups. Since the proposed features are robust to script and content, we have plans to extend the proposed idea for estimating different ages with a one-year gap. In addition, the extension of the work considers qualification factor for improving results.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of China under Grant 61672273 and Grant 61832008, the Science Foundation for Distinguished Young Scholars of Jiangsu under Grant BK20160021.

REFERENCES

- [1] J. Tang, Z. Li and X. Zhu, "Supervised deep hashing for scalable face image retrieval", *Pattern Recognition*, 75, 2018, pp 25-32.
- [2] C. C. Ng, M. H. Yap, Y. T. Cheng and G. S. Hsu, "Hybrid ageing patterns for face age estimation", *Image and Vision Computing*, 69, 2018, pp 92-102.
- [3] L. Liu, C. Xiong, H. Zhang, Z. Niu, M. Wang and S. Yan, "Deep aging face verification with large gaps", *IEEE. Trans. MM*, 18, 2016, pp 64-75.

- [4] G. Antipov, M. Baccouche, S. A. Berrani and J. L. Dugelay, "Effective training of convolutional neural networks for face based gender and age prediction", *Pattern Recognition*, 72, 2017, pp 15-26.
- [5] K. Chaudhari and A. Thakkar, "Survey on handwriting-based personality trait identification", *Expert Systems with Applications*, 2019.
- [6] S. A. Maadeed and A. Hassine, "Automatic prediction of age, gender and nationality in offline handwriting", *EURASIP Journal on Image Processing*, 2014, pp 1-10.
- [7] N. Bouadjeneq, H. Nemmour and Y. Chibani, "Age, gender and handedness prediction from handwriting using gradient features", *In Proc. ICDAR*, pp 1116-1120, 2015.
- [8] N. Bouadjeneq, H. Nemmour and Y. Chibani, "Fuzzy integral for combining SVM-based handwritten soft-biometrics prediction", *In Proc. DAS*, pp 311-316, 2016.
- [9] N. Bouadjeneq, H. Nemmour and Y. Chibani, "Robust soft biometrics prediction from offline handwriting analysis", *Applied Soft Computing*, 46, 2016, pp 980-990.
- [10] F. Zouaoui, N. Bouadjeneq, H. Nemmour and Y. Chibani, "Co-training approach for improving age range prediction from handwriting text", *In Proc. ICFHR*, 2017.
- [11] J. Zunic, P. L. Rosin and V. Ilic, "Disconnectedness: A new moment invariant for multi-component shapes", *Pattern Recognition*, 78, 2018, pp 91-102.
- [12] K. S. Raghunandan, P. Shivakumara, B. J. Navya, G. Pooja, Navya, N. Prakash, G. H. Kumar, U. Pal and T. Lu, "Fourier coefficients for fraud handwritten document classification through age analysis", *In Proc. ICFHR*, pp 25-30, 2016.