Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2016) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

+20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2016, 5-7 September 2016, York, United Kingdom

Clock Drawing Test Digit Recognition Using Static and Dynamic Features

Zainab Harbi[a,b]*, Yulia Hicks[a], Rossitza Setchi[a]

[a]Cardiff School of Engineering, Queen's Buildings, CF24 3AA, Cardiff, UK.
[b]Computer Science Department, Faculty of Education for Women, Kufa University, Najaf, Iraq.

Abstract

The clock drawing test (CDT) is a standard neurological test for detection of cognitive impairment. A computerised version of the test promises to improve the accessibility of the test in addition to obtaining more detailed data about the subject's performance. Automatic handwriting recognition is one of the first stages in the analysis of the computerised test, which produces a set of recognized digits and symbols together with their positions on the clock face. Subsequently, these are used in the test scoring. This is a challenging problem because the average CDT taker has a high likelihood of cognitive impairment, and writing is one of the first functional activities to be affected. Current handwritten digit recognition system perform less well on this kind of data due to its unintelligibility. In this paper, a new system for numeral handwriting recognition in the CDT is proposed. The system is based on two complementary sources of data, namely static and dynamic features extracted from handwritten data. The main novelty of this paper is the new handwriting digit recognition system, which combines two classifiers—fuzzy *k*-nearest neighbour for dynamic stroke-based features and convolutional neural network for static image-based features, which can take advantage of both static and dynamic data. The proposed digit recognition system is tested on two sets of data: first, Pendigits online handwriting digits; and second, digits from the actual CDTs. The latter data set came from 65 drawings made by healthy people and 100 drawings reproduced from the drawings by dementia patients. The test on both data sets shows that the proposed combination system can outperform each classifier individually in terms of recognition accuracy, especially when assessing the handwriting of people with dementia.

Keywords: Clock Drawing Test (CDT); cognitive impairment diagnosis; dementia; static and dynamic representation

* Corresponding author. Tel.: +447941076932;
E-mail address: HarbiZa@cardiff.ac.uk

1. Introduction

The clock drawing test (CDT) is a neurological test widely used as part of a general assessment of cognitive function [1]. It is well known for its simplicity and effectiveness for measuring multidimensional cognitive function. It can help to differentiate normal ageing from cognitive dysfunction associated with psychiatric and neurological disorders such as dementia. The CDT is conducted by asking the test participant to draw on paper a clock set to a specific time, such as five minute to three.

When the participant finishes, the test administrator scores the drawn clock. There is a number of scoring syst-ems; however, most of them rely on the clinician's subjective judgement.

In recent research [2], machine learning and computational algorithms were used to score the tests automatically. The developed system was able to classify CDT drawings into several classes, including healthy and several kinds of dementia with accuracy of 89.5%, which is comparable to that of medical practitioners. In addition, the author identified new CDT drawing features important for such classification. In two other articles [3,4], computerised CDT tests were administered using digitiser tablets and pens. This research showed that such systems obtain additional valuable information related to the process of drawing itself. Since the data is time-stamped, the system capture both the result (the drawing) and the behaviour that produced it: every pause, hesitation, and time spent simply holding the pen and writing time. Such information helps in assessing the cognitive function of the patients even when the test result appears superficially normal.

The data processing in the computerised system starts with sketch interpretation, i.e., classifying each pen stroke as digits, clock hands or other drawing artefacts. Consequently, this information is used in the scoring of the drawings. Information about any missing or repeated numbers, or numbers in the wrong position, is also used in the scoring. Therefore, a reliable handwriting recognition system needs to be applied. However, recognising the handwriting of people with cognitive impairment is challenging, since their writing skills are often affected by cognitive impairment [5]. In the past, very many algorithms have been developed for handwriting recognition [6], but to the best of the authors' knowledge, there is no actual evaluation of such algorithms using data sets collected from the elderly or persons with mild cognitive impairment.

Handwriting recognition algorithms can be classified by whether they use static or dynamic representations of data. The first type of algorithm considers the features extracted from static images [7]. The second kind of algorithm takes into consideration features related to the process of writing, such as sequence and direction writing [8]. This kind of representation is applicable for online handwriting recognition, since the data is captured using a computerised system and stored as time-stamped data.

Although the handwriting recognition problem has been addressed in previous research on computerised CDTs [4,9], the developed algorithms rely on information, which, in the case of CDT drawings from people with moderate or severe dementia, would likely lead to misclassifications due to unusual positions or shape of the clock digits. In [9] a solution based on the geometrical shape representation and the sequence of writing was developed. In [4], a different approach was used based on visual features that describe what the strokes look like and thus the available dynamic data is not used. In addition, information related to the position of the digits in the clock face is employed for augmentation of the classifier result.

In this paper, a new system for handwritten digit recognition is proposed. The system exploits both static and dynamic information by combining two classifiers: one static and the other dynamic digit representations. In the evaluation section, the proposed system is shown to be more effective than using each representation alone. Without including additional information about the position of the digits within a clock, the proposed system has proven more robust and tolerant of incorrect CDT digit positions, which is often the case for people with dementia. In order to test the proposed system's ability to recognise digits in both normal and abnormal drawings, it is evaluated on two data sets. The first was a publically available online Pendigits data set and the second consisted of digits extracted from 65 CDT drawings made by healthy participants and 100 copies of actual CDT drawings from dementia patients.

The paper is structured as follows: Section 2 introduces the background and review related work. Section 3 describes the proposed system, along with other related classifiers. Section 4 outlines the data set used in the evaluation. Section 5 presents comparative results and analysis, while the conclusion and future works are given in section 6.

2. Related Work

Handwriting recognition systems can be divided into two categories: online and offline. In an online system, a sequence of time-stamped coordinates representing the movement of a pen tip is transformed into a meaningful text. By contrast, in an offline recognition system, only the scanned image of the text is available [6]. Over the last few decades, a large number of algorithms have been proposed in both categories; for a detailed review the reader is referred to the broader survey in [6,7] .

Recently, convolutional neural networks (CNNs) [10] showed state-of-the-art performance in various domains such as speech recognition and visual object detection and recognition [11]. They consist of multiple alternating convolutional and pooling layers. Their successive alternating layers structure is designed to learn progressively higher-level features, where the last layer produces the classification results. CNNs were applied to recognising offline-handwritten digits with an error of just 0.23%, as reported on MINST handwritten digit database [12]. This is comparable to human performance.

With the recent development of smartphones and digital tablets, interest in online handwriting recognition has increased. Several new algorithms have been proposed [13, 14], with good results reported. Most of these algorithms use a set of features extracted from (x,y) coordinates such as the normalised first and second derivatives of the (x,y) coordinates, the distance and angles between pairs of points, the curvature, the start and end point positions and the pen on and off (if there are multiple strokes). Online recognition methods have been shown to be more robust against variations in input shape than offline algorithms. A state-of-the-art result reported in [14] for online handwritten digit recognition used dynamic time warping and a *k*-nearest neighbour (KNN) classification algorithm; it considered only the histogram of directional features extracted from the online handwritten digits data set. The *k*-nearest neighbour [15] is one of the most famous statistical methods used for classification and pattern recognition. During the training stage, a copy of each sample is stored with its label. Given a sample to classify, KNN works by selecting the *k* most similar samples from the training set. To calculate the distance, various distances can be used such as Euclidean or Mahalanobis distances. After all the distances are calculated, they are sorted and the nearest *k* samples are determined. Through a voting scheme, i.e. using the majority of the label of nearest neighbours, a label is assigned to the new sample

Combining multiple classifier systems is an approach proposed in [16], where a system based on hidden Markov models was used with online and offline digit representations. The result was derived using a voting strategy. The result showed a 2% improvement over the best individual recogniser. Other methods [17] achieved an improvement of 1.2% by combining the results of two multi-layer perceptron (MLP) classifiers, one trained on dynamic (x,y) coordinates the other on the static images of digits.

In the area of computerised CDT research, two approaches to digit recognition have been proposed. The first is by Cho [9], in which the system recognises the unique shape of each number by comparing the curved and straight lines and the writing sequences of strokes. The system has been tested on twenty clocks drawn by healthy participants and achieved 97.67% recognition accuracy. The second approach[4] based on adaptation of a symbol recognition algorithm originally developed by Ouyang [18], by using K nearest neighbor as a classifier with five features images. Four of the features images describe the stroke from horizontal, vertical, and diagonal point of view while the fifth feature represents the stroke endpoint. The reported result was over 96% recognition accuracy when the algorithm trained and tested on clocks from healthy individuals.

The convolutional neural networks achieved promising results in offline handwriting, but their application is limited due to the huge amount of data required for training to avoid overfitting. Nonetheless, their performance can be improved by using it in a combination with other classifiers, which are not highly dependent on the size of the training data set. An example of such classifier, KNN, will be presented in the next section.

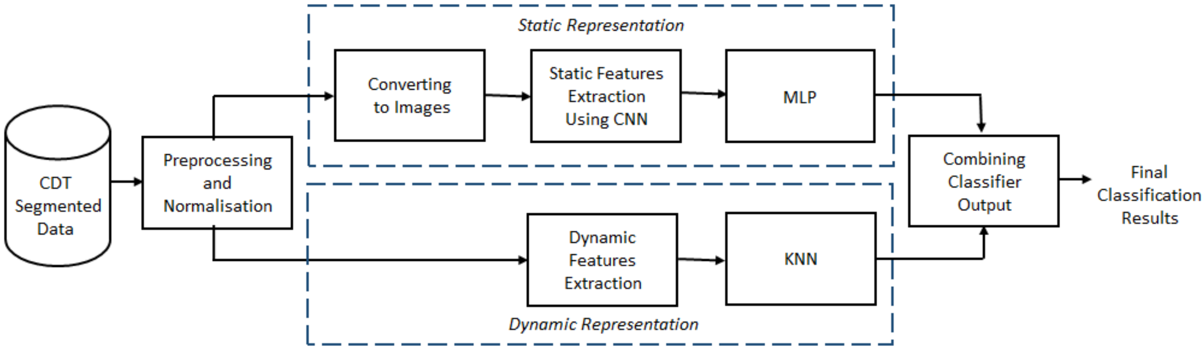

Fig. 1.Proposed handwriting digit recognition system.

3. The Proposed Handwriting Digit Recognition System

This section presents the proposed digit handwriting recognition system (Fig.1) and its components in detail. The proposed system consists of a preprocessing and normalization component, static and dynamic feature classification components and the unit for combining output of these components. In the following subsections, the components of the system are explained in detail.

3.1. Preprocessing and Normalisation

The data were captured using a Wacom Intuos Pro digitising tablet, which has a wireless electronic inking pen and a recording area of approximately 32.5 cm by 20.3 cm. The pen has a pressure-sensitive tip, and its shape and size are similar to regular pens, offering an experience that is no different from the normal paper and pen test. The data were captured when the clock are drawn as a set of (x,y) coordinates. These points were segmented into a set of objects using a segmentation algorithm previously proposed in [19]. The segmented object was further separated into hands and digits based on their position from the clock's centre.

The coordinate sequence received from the tablet was normalised to eliminate differences due to sampling, scale, and translation. This improve the robustness of the recognition system, ensuring that all the digits centered and scaled appropriately. All digits are transformed so that they have the same bounding box dimensions while preserving their aspect ratio. The dimension used in this paper is 100x100 following the same approach in Pendigits data set.

Different writer write with different speed and the online stroke are typically sampled at a constant temporal frequency, thus the distance between neighboring points in the pen trajectory varies based on the speed of the pen. In addition, more samples could be found in the corner or regions of high curvature, where the pen is typically slower. In order to eliminate these variations Bresenham's line algorithm is used to find the sequence of points that form a line between each coordinates.

3.2. Static Feature Classification

The static feature classification starts with converting the preprocessed and normalized data into gray scale images. In order to obtain the images from sequence of x-y coordinates, these coordinates further down sampled and mapped to 20x20 pixels box to store each image. Furthermore, the images are smoothed using Gaussian-smoothing function to increase tolerance to local shift and distortion. The Gaussian filter used is 3x3 pixel with 0.75 uniform standard deviation. Finally, the images were centered in a 28x28 image by computing the center of mass of the pixels, and translating the image to position this point at the center of the 28x28 field.

After the digits data are converted into images, the next step is feature extraction and classification. CNNs are well known for their invariance to distortion and simple geometrical transformations such as translation, scaling and rotation [12]. This feature makes a CNN an appropriate candidate for tackling the problem of CDT handwritten digit.

However CNN success stories notwithstanding, there are some limitations to their application, such as the large quantity of training data, which is required in order to avoid overfitting the CNN to the training data set. Since the available CDT data set is relatively small, one can considered using CNN as features extractor by employing the advantage that a CNN network converges from high-resolution information to reduced but highly informative space recognition. First CNN was trained on a data set that was close to the CDT data set, which is Pendigits online digits data set and this network more augmented by adding a data from our data set. Then the pre-trained network used as a feature extractor to the CDT data sets. Investigating the physical meaning of these features is outside of the scope of this work. These features are feed to MLP classifier. The MLP with backpropagation training is the standard algorithm for any supervised learning pattern recognition process, and it can fit well with the CNN architecture. Moreover the MLP's output is considered as posteriori class probabilities, with useful properties (e.g. positivity, summing to one), providing an efficient framework for a classifier combination [20].

The CNN architecture used in this paper is LeNet [12], which is a deep convolutional neural network known to work well on MNIST handwritten digit classification tasks. LeNet architecture used in this paper is the default one using MatConvNet toolbox [21] which consist of eight layers. The first one is a convolutional layer with a filter bank of 20 single-channel filters of 5 × 5 size. The second one is a max pooling layer. Third is another convolutional layer with 50 different filters of 5 × 5 followed by another max pooling layer. The fifth layer contains a filter bank of 500 filters of 4 × 4 pixels. The sixth layer contains a rectifier linear unit followed by another convolutional layer of 10 filters with one single pixel and at the end; the eighth layer applies the softmaxloss operation. The CNN network is trained first and after training, it is used as features extractor by replacing the last two layers with a MLP. The MLP used in this paper is a simple two-layer perceptron with a logistic sigmoid activation function.

3.3. Dynamic Feature Classification

The first step in dynamic feature classification is dynamic features extraction, the dynamic features used in this paper is the normalised x-y coordinates. After preprocessing and normalization step, the data set consists of variable numbers of sequence points. In order to have constant- length feature vectors, the data was spatially resampled into sequence of points regularly spaced in arc length. Following the same approach of Pendigts data set, we used 8 points per digit.

Using *k*-nearest neighbours in online handwritten digit recognition demonstrated impressive results even when considering only simple directional features and a small data set [14]. In this work, a fuzzy KNN [22] algorithm is used where the algorithm assigns label probabilities to a sample rather than assigning the sample to a particular class. The inclusion of fuzzy set theory into these classifiers deals with imprecision when defining the classes, which is caused by the large variability of the samples belonging to the same class. Consequently, it improves the results.

The following relationship assigns class labels to the sample as a function of the sample's distance from its KNN training samples:

$$\boldsymbol{u_i(x)} = \frac{\sum_{j=1}^{k} \boldsymbol{u_{ij}} \left(\frac{\mathbf{1}}{\|\boldsymbol{x}-\boldsymbol{x_j}\|}\right)^{\frac{2}{m-1}}}{\sum_{j=1}^{k} \left(\frac{\mathbf{1}}{\|\boldsymbol{x}-\boldsymbol{x_j}\|}\right)^{\frac{2}{m-1}}} \qquad \textbf{(1)}$$

where $u_i(x)$ is the membership probability of the test sample x to class i, and m is the 'fuzzifier' which is a fuzzy strength parameter that determines how the distance is weighted when calculating each neighbour's contribution to the membership value. The variable k is the number of nearest neighbours; u_{ij} is the membership value of the j-th neighbour to the i-th class, which can be defined by giving them complete membership in their own class and no membership in all other classes. This is because the prototypes should naturally be assigned complete membership in the class that they represent. As seen from (1), the assigned memberships of x are influenced by the inverse of the distances from the nearest neighbours and their class memberships, this inverse distance serves to give more weigh to vector's membership if it is closer to, and less if it is further from, the vector under consideration. To calculate the distance there are many distance algorithms. Dynamic time wrapping is widely used with time serious data, but since each digits represent by eight points only with a fair distance between them, Euclidean distance is most appropriate and less computational complexity and time consuming.

3.4. Combining Classifier Output

One effective approach to improve the performance of handwriting recognition is to combine multiple classifiers [23]. Following this approach, a combination of two classifiers is used in order to obtain better digit recognition accuracy. CNN and KNN are built as individual classifiers for recognising offline and online patterns respectively. Data processed by each classifier are very different: dynamic representation (online) contains spatial and temporal information (stroke coordinates and order), while static (offline) representation consists of the image of a digit.

The advantage of the CNN classifier is that it automatically extracts the salient features of the input image. The features are largely invariant to the shift and shape distortions of the input characters. This invariance occurs because CNN adopts a weight-sharing technique on one feature map. CNNs are efficient at learning invariant features from the offline patterns, but do not always produce optimal classification results, particularly when there are small data sets or unbalanced training data. Conversely, KNNs, with their distance measuring, cannot learn complicated invariance. However, they do produce good decisions when considering the sequencing of points in an online pattern, which can be achieved with a small number of patterns.

Overwriting is one of the problems that can cause misclassifications to the dynamic classifier, as the stroke point sequencing will be changed dramatically while the final shape of writing will be the same. In this case, the advantage will be for the static classifier. In other cases, the shape of the written digits may be overly distorted but the same sequencing information is preserved, which will give credit to the dynamic classifier.

In the proposed combination system, the CNN is trained with normalised images and used as features extractor, while the KNN classifier is trained with the normalised (x,y) coordinates that represent the dynamic data. The two classification results are then processed by a combination scheme, and this scheme generates a ranked list of predictions for the input image. Classifier combination techniques operate on the outputs of individual classifiers, while a function or a rule combines the classifier scores in a predetermined manner. The formula is defined as follows:

$$P\ (c_i|S) = f\{P(c_i\ |C, P(c_i|K)\} \qquad i = 1..m \qquad \textbf{(2)}$$

where $P(c_i \mid C)$ is a posterior probability for one class (i), computed from the CNN model; $P(c_i \mid K)$ represents a probability for the same class (i) given by the KNN model; $P(c_i \mid S)$ is the combination probability for the class (i); and f represents the function applied to the classifier probability results. The average is used in this paper’s experiment as it generated better results than other combination methods such as maximum, product and weighted sum. Finally, a ranked list of candidates is obtained with a decreasing order of probabilities after the combination process. The top candidate is then chosen as the predicted class for the input pattern.

To assess the accuracy of the combination CNN and KNN model, as well as of separate classifiers, they are applied to the Pendigits data set and the clock drawing data set. Details of the data sets, experiments and results are described in the next section.

4. Data Sets and Experimental Set up

Two different sets of online isolated handwritten digits were used in the set of implemented experiments: Pendigits and CDT digits. Pendigits [17] is an online data set available from the UCI Machine Learning Repository [24] . This resource contains handwritten instances of 10 digits from several writers. 7,494 glyphs from 30 writers are used as a training set and 3,498 glyphs from 14 different writers are used as test data. Each digit is represented by eight successive pen points in a two-dimensional coordinates system. The second data set is the CDT digits. These digits were extracted from the clock drawings collected by the author according to an ethically-approved experiment. The CDT data set can be subdivided into two further categories:

1. Digits extracted from the clocks drawn by 65 healthy volunteers aged between 25 and 87 years. The group included 15 individuals who were older than 60. The participants included 32 females and 33 males, and their educational attainments ranged between basic and college graduate. The participants were asked to draw a clock on a paper sheet laid on the surface of the digitiser. Each person drew one clock, so 65 clock images were collected. There were 975 digits extracted from these drawings, which will be referenced as ‘normal digits’ in the next sections.
2. The second subset came from dementia patients’ drawings. These patients were diagnosed with dementia during their examination at Llandough Hospital in Cardiff UK. Each patient’s diagnosis was one of the following: mild cognitive impairment, Alzheimer’s dementia and vascular dementia. A volunteer reproduced 100 drawings using a digitiser copying the original drawings from the patients. Fig. 2 shows several examples of the reproduced clocks. There were 1435 digits extracted from these drawings. These digits are referred to in the next sections as ‘abnormal digits’.

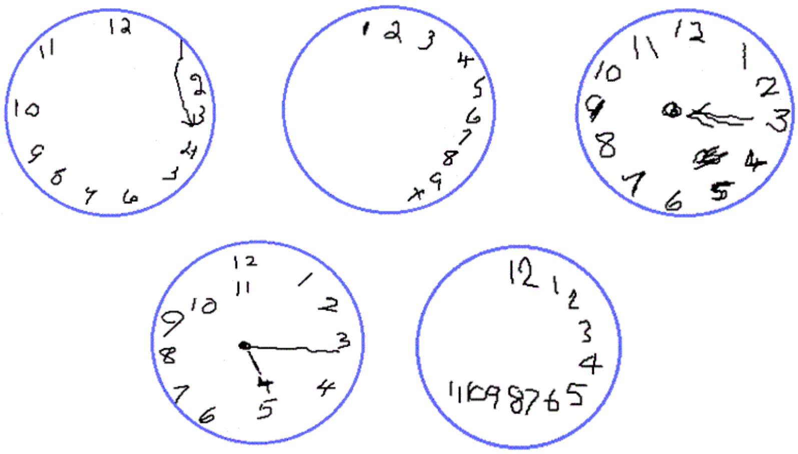

Fig. 2. Examples of clock drawings produced by people diagnosed with dementia.

Two representations of handwritten digit samples are used in the proposed system: the static representation, where each digit is represented by a fixed size 28 × 28 pixels grayscale image with each pixel value ∈ {0,... , 255}., and the dynamic representation where each digit is represented by a sequence of (x_t,y_t) coordinates in time. Pendigits is an already normalised data set in dynamic representation; the only pre-processing applied on them is converting them into grayscale images for static representation. The CDT data set are followed the all preprocessing and normalization as explained in Section 3.

The CDT data were collected using software developed by the authors to interface the digitiser with the computer system. Data pre-processing and classification models were all implemented using MATLAB. The Pendigits data set was divided into training and testing using the original settings. Following the same approach, the CDT data set was divided into fixed 70% training and 30% testing sets for fair comparison between the proposed classification systems.

5. Experimental Results

This section presents the performance evaluation and comparison of the proposed classification system. To evaluate which method would be more accurate—the individual classifiers or the combination system—a number of experiments were conducted on a public Pendigits data set as well as CDT digits.

5.1. CNN and Static Representation

In order to train a CNN, a large amount of data is required. To overcome this problem the CNN used as a feature extractor as explained in Section 3. LeNet was first trained on Pendigits, which is an online handwritten digits data set. Next, the pretrained network was used as a feature extractor for digits that were extracted from The CDT data set. These features were fed to the MLP classifier with a simple structure consisting of one hidden layer with 200 neuron and one output layer. An implementation from the MatConvNet MATLAB library [21] was used. The model was trained using a stochastic gradient descent with a batch size of 100 samples and 0.001 learning rate. The network was trained for 100 epochs on an NVIDIA GeForce 970 GX 4-GB GPU.

In addition, data augmentation was applied to the data sets. In particular, the data set were increased by a factor of ten by rotating the image through ten different angles [-25,-20,-15,-10,-5, 5, 10, 15, 20, 25]. Rotation is the most suitable augmentation, which can be applied to the clock drawing digits, since most people try to write with some rotation according to the circle of the clocks. Table 1 shows the recognition accuracy of CNN as a classifier and using it as a feature extractor by replacing the last fully-connected layer with MLP.The first row of the table represents the result of training and testing the network from scratch on the data set. This experiment is applicable only to the Pendigits data; the CDT data set is not large enough for such experiments. The result for MLP shows the case when the MLP was trained directly on the row images rather than using the features map extracted by CNN.

Table 1. Recognition accuracy of Pendigits, normal and abnormal digits with MLP and CNN.

	Pendigits	Normal digits	Abnormal digits
CNN	97.2%	-	-
MLP	94%	89%	84%
CNN + MLP	98%	97.3%	93%

It is clear from the recognition accuracy that using CNN as a feature extractor outperforms the MLP when trained on the row images. CNN is a considerable feature learner, even in the case of a small data set such as CDT data set digits. Moreover, the difference in accuracy between normal and abnormal digits indicates that there is a remarkable effect in recognition algorithm performance when considering digits drawn by healthy people and others with cognitive impairment.

5.2. KNN and Dynamic Representation

In this set of experiments, the accuracy of KNN using the dynamic representation of data (i.e. a set of temporal x,y-coordinates) were compared with other machine learning classifiers: LIB SVM, Naïve Bayes and RBF Network. Weka data mining software [25] was used for the implementation of the classifier with the same setting parameters. The training and testing data size is the same as used for all previous experiments. There is no data augmentation here. As shown in Table 2, the experiments indicate that KNN outperforms other classifiers for different k values.

Table 2. Recognition accuracy of Pendigits, normal and abnormal digits with dynamic representation and KNN, MLP, Lib SVM, Naïve Bayes and RBF Network.

	Pendigits	Normal digits	Abnormal digits
KNN (k=1)	97.68%	95.2%	92.5%
KNN (k=3)	97.8%	96.7%	92.6%
KNN (k=5)	97.7%	95.5%	92.7%
MLP	94.5%	90.1%	89.1%
Lib SVM	96.9%	93.8%	91.4.%
Naïve Bayes	89%	87.5%	85.2%
RBF Network	95.6%	92.3%	88.3%

The performance of SVM is very is close to the RBF network; however the SVM is affected by the small data set. As in the case of the CDT data set with normal digits, the size of the data is smaller than in other cases. All the classifier results show differences in recognition performance between normal and abnormal digits. KNN with three nearest neighbors has achieved a slightly better result than other values, so this setting will be used in the combination classifier.

5.3. Classifier Combination

In this section, two classifiers model is combined: CNN as features extractor with MLP and KNN. Each classifier was trained individually on the same training data; however, the training data were in two different representations. Dynamic representation used the set of sequential (x,y) points for the KNN, while static representation took the form of images in the case of CNN. In testing, the patterns were presented to both classifiers simultaneously. The output probability of both classifiers was combined using the average, which was given the best result in this paper. Finally, a ranked list of candidates was obtained with a decreasing order of probabilities. The top candidate was then chosen as the predicted class for the input pattern. By combining two pieces of knowledge, the accuracy was increased considerably: 98.8% for the Pendigits data, 98% for normal digits and 96.5% for abnormal cases. A significant improvement can be reported, especially with abnormal digits: about 3%. This was the most challenging task (see Table 3). The advantage comes from the diversity of the classifiers' strengths on different input patterns. Moreover most of the abnormal digits are overwritten, that confused the KNN classifier, while the CNN has a higher certainty of it is classification result, this can improve the final classification results. In another hands some ab-normal digits are badly distorted but still the sequencing information preserved, In such cases the KNN has some confidence of the classification results and that's improve the final classification results subsequently. However, the small size of CDT normal digit data set has an impact on recognition accuracy, in comparison between normal digits and the Pendigits data set. In addition, the context is different: writing in a round clock is different from writing in a linear document. The examples in Fig.3 show that many of missed digits could be considered difficult even for the human.

Table 3. Recognition accuracy of Pendigits, normal and abnormal digits for the combination system.

	Pendigits	Normal digits	Abnormal digits
CNN	98%	97.3%	93%
KNN	97.8%	96.7%	92.6%
CNN + KNN	98.9%	98%	96.5%

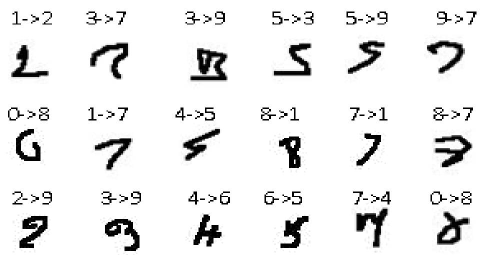

Fig. 3. Examples of incorrectly classified digits by the proposed handwriting digit recognition system. First line examples from pendigits data set, second line example from normal digits while third one from abnormal digits. The label is corresponding truth->predicted.

6. Conclusion and Future Work

This paper investigated the unconstrained handwritten digits recognition problem in computerised CDTs. In order to build a reliable computer-based test for diagnosing dementia, a robust digit recognition system was developed. Different data representations and classification techniques were employed. In addition, a new combination of classifiers is proposed by combining KNN and CNN. The combination system has the advantages of both classifiers and static and dynamic data representations. The experimental results showed the effectiveness of the proposed system in terms of improved recognition accuracy.

From experimental observations, it was found that the task of recognising the handwritten digits from dementia-afflicted people requires special consideration compared to normal handwriting. Therefore, future studies should address the question of how to improve the recognition performance in such sensitive cases. Moreover, there are some limitations to the proposed system, such as the small size of the CDT data set and the simplicity of the features used in the recognition algorithms. Investigating more sophisticated features such as the direction and angles of writing and using larger data sets can further improve recognition accuracy.

References

[1] E. Pinto and R. Peters, "Literature Review of the Clock Drawing Test as a Tool for Cognitive Screening," Dementia Geriatric Cognitive Disorder, vol. 27, pp. 201–213, 2009.
[2] M. Bennasar, "Clinical Decision Support System for Early Detection and Diagnosis of Dementia," PhD Dissertation, Cardiff University, UK, 2015.
[3] R. Davis, D. J. Libon, R. Au, D. Pitman, and D. L. Penney, "THink : Inferring Cognitive Status from Subtle Behaviors," Proceeding of Innovative Applications of Artificial Intelligence Conference, pp. 2898–2905, 2014.
[4] H. Kim , YS. Cho, and E. Y.-L. Do. "Computational Clock Drawing Analysis for Cognitive Impairment Screening," Proceedings of the International Conference on Tangible, Embedded, and Embodied Interaction, pp.297-300, 2011.
[5] S. Rosenblum, M. Samuel, S. Zlotnik, I. Erikh, and I. Schlesinger, "Handwriting as an Objective Tool for Parkinson's Disease Diagnosis," Journal of Neurology, vol. 260, pp. 2357–2361, 2013.
[6] R. Plamondon and S. N. Srihari, "On-Line and Off-Line Handwriting Recognition : A Comprehensive Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22 , pp. 63–84, 2000.
[7] N. Arica and F. T. Yarmanvural, "An Overview of Character Recognition Focused on Off-Line Handwriting," IEEE Transactions Systems, Man, Cybernetics, Application Reveiw, vol. 31, pp. 216–233, 2001.
[8] C. Bahlmann, "Directional Features in Online Handwriting Recognition," Pattern Recognition, vol. 39, pp. 115–225, 2006.
[9] Y. S. Cho, H. Kim, and E. Y.-L. Do, "Curved-Straight Lines-Analysis (CSLA) Algorithm for Handwritten Digit Recognition Enhancement," Proceeding of Intelligent Interactive Technologies and Multimedia- IITM '10, pp. 150–154, 2010.
[10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceeding of the IEEE, vol. 86, pp.2278–2324, 1998.
[11] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," Neural Networks, vol. 61, pp. 85–117, 2015.
[12] D. Cires, U. Meier, and J. Schmidhuber, "Multi-column Deep Neural Networks for Image Classification," Arxiv Prepr. arXiv1202.2745, 2012.
[13] O. D. Fenwa, E. O. Omidiora, and O. A. Fakolujo, "Development of a Feature Extraction Technique for Online Character Recognition System," Journal of Innovative Systems Design and Engineering, vol. 3, pp. 10–23, 2012.
[14] M. J. Castro-Bleda, S. España, J. Gorbe, F. Zamora, D. Llorens, A. Marzal, F. Prat, and J. M. Vilar, "Improving a DTW-Based Recognition Engine for On-line Handwritten Characters by Using MLPs," Proceeding of International Conference on Document Analysis and Recognition ICDAR, pp. 1260–1264, 2009.
[15] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," The American Statistician, vol 46 , pp. 175–185
[16] M. Liwicki and H. Bunke, "Combining On-line and Off-line Systems for Handwriting Recognition," Proceeding of International Conference on Document Analysis and Recognition ICDAR, pp. 372–376, 2007.
[17] F. Alimoglu and E. Alpaydin, "Combining Multiple Representations and Classifiers for Pen-Based Handwritten Digit Recognition," Proceeding of International Conference on Document Analysis and Recognition ICDAR, pp. 637–640, 1997.
[18] T. Y. Ouyang and R. Davis, "A Visual Approach to Sketched Symbol Recognition," International Joint Conference on Artificial Intelligence IJCAI, pp. 1463–1468, 2009.
[19] Z. Harbi, Y. Hicks, R. Setchi, and A. Bayer, "Segmentation of Clock Drawings Based on Spatial and Temporal Features," Proceeding of International Conference on Knowledge Based and Intelligent Information and Engineering Systems KES, vol. 60, pp.1640-1648, 2015.
[20] J. Pinto, S. Sivaram, M. Magimai-Doss, H. Hermansky, and H. Bourlard, "Analysis of MLP-Based Hierarchical Phoneme Posterior Probability Estimator," IEEE Transactions on Audio, Speech and Language Processing, vol. 19, pp. 225–241, 2011.
[21] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the Devil in the Details: Delving Deep into Convolutional Nets," Proceeding of British Machine Vision Conference, pp. 1–11, 2014
[22] J. M. Keller and M. R. Gray, "A Fuzzy K-Nearest Neighbor Algorithm," IEEE Transactions Systems, Man, Cybernetics, vol. SMC-15, pp. 580–585, 1985.
[23] L. Xu, A. Krzyżak, and C. Y. Suen, "Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition," IEEE Transactions, Systems, Man, Cybernetics, vol. 22, pp. 418–435, 1992..
[24] M. Lichman, 'UCI Machine Learning Repository', University of California, Irvine, School of Information and Computer Sciences, 2013. Available: http://archive.ics.uci.edu/ml.
[25] M. Hall, Pentaho, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA Data Mining Software: An Update," SIGKDD Exploration, vol. 11, 2009.