

# MCS HOG Features and SVM Based Handwritten Digit Recognition System

Hamayun A. Khan

Faculty of Computer Studies, Arab Open University, Industrial Ardiya, Kuwait

Email: H.Khan@arabou.edu.kw

**How to cite this paper:** Khan, H.A. (2017) MCS HOG Features and SVM Based Handwritten Digit Recognition System. *Journal of Intelligent Learning Systems and Applications*, 9, 21-33.  
<https://doi.org/10.4236/jilsa.2017.92003>

**Received:** March 7, 2017

**Accepted:** May 9, 2017

**Published:** May 12, 2017

Copyright © 2017 by author and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Digit Recognition is an essential element of the process of scanning and converting documents into electronic format. In this work, a new Multiple-Cell Size (MCS) approach is being proposed for utilizing Histogram of Oriented Gradient (HOG) features and a Support Vector Machine (SVM) based classifier for efficient classification of Handwritten Digits. The HOG based technique is sensitive to the cell size selection used in the relevant feature extraction computations. Hence a new MCS approach has been used to perform HOG analysis and compute the HOG features. The system has been tested on the Benchmark MNIST Digit Database of handwritten digits and a classification accuracy of **99.36%** has been achieved using an Independent Test set strategy. A Cross-Validation analysis of the classification system has also been performed using the 10-Fold Cross-Validation strategy and a 10-Fold classification accuracy of **99.26%** has been obtained. The classification performance of the proposed system is superior to existing techniques using complex procedures since it has achieved at par or better results using simple operations in both the Feature Space and in the Classifier Space. The plots of the system's Confusion Matrix and the Receiver Operating Characteristics (ROC) show evidence of the superior performance of the proposed new MCS HOG and SVM based digit classification system.

## Keywords

Handwritten Digit Recognition, MNIST Benchmark Database, HOG Analysis, Multiple-Cell Size HOG Analysis, SVM Classifier, 10-Fold Cross-Validation, Confusion Matrix, Receiver Operating Characteristics

## 1. Introduction

Digit Recognition and Character Recognition are important activities associated with document scanning and converting them into an electronic format. Recog-

dition of digits and characters, whether machine typed or handwritten, belongs to the important field of Optical Character Recognition (OCR) which is one of the earliest applications of the general Pattern Recognition (PR) and Artificial Intelligence (AI) techniques [1]. Despite the fact that the Handwritten Digit Recognition problem has been studied extensively for a number of years, one of its benchmark databases is still actively used by researchers till date [2] [3] [4]. The conventional applications of OCR are numerous and include diverse areas such as automatic bank cheque processing, immigration data processing, health data record conversion into electronic format, tax forms data conversion and many other applications. Despite being one of the earliest research and application areas of AI, the digit and character recognition remains an active research topic. The present day popularity and mass availability of smart phones with sophisticated camera technology enable users to capture images of handwritten notes containing digits and characters. Subsequently there is a need to extract the handwritten notes information from these images to convert them into text files. Analysis of documents and images with texts continues to be active research topics [5] [6] [7] [8] [9]. Hence the need for developing efficient handwritten digit and character recognition algorithms and techniques continues to exist even today.

Handwritten documents present more challenges for digit and character recognition compared to typed documents due to the variations in the handwriting of individual persons. The adoption of ITC systems and the continuous migration from manual systems to electronic formats, as part of automation efforts, especially in the developing and emerging markets, present opportunity to develop efficient document scanning and conversion systems. This in turn gives rise to the need for more accurate digit and character recognition techniques. Our present work is aimed at finding an improved classification system for handwritten digit recognition purposes.

The breakdown of this paper is as follows: A review of existing techniques for handwritten digit recognition problem is given in Section 2. The proposed handwritten digit recognition system is presented in Section 3. Section 4 presents the experimental setup for the experiments conducted in this work in order to assess the performance of the system and this section also presents the selection of relevant parameters associated with the experiments. Results and discussions have been presented in Section 5. Finally, conclusions and future work ideas are discussed in Section 6.

## **2. Review of Existing Techniques**

The Character Recognition problem has been studied extensively by the Pattern Recognition community and it is commonly known as the Optical Character Recognition (OCR) problem [1]. The handwritten digit recognition and character recognition are part of the wider and more general OCR research. As a Pattern Recognition problem, the handwritten digit images are usually processed by a feature extraction block and features are extracted from the digit images in the

Feature Space. The recognition or classification is based on the set of feature vectors that have been extracted from the digit images. The features are then fed to a classification block in which classification decision is taken. The Mixed National Institute of Standards and Technology (MNIST) database is a benchmark database that is commonly used to assess the performance of algorithms for recognition of hand written digits [10]. This database is also used as a benchmark database for evaluating and testing various Artificial Intelligence (AI) and Machine Learning (ML) algorithms [11]. The MNIST Handwritten digit database is maintained by Lecun *et al.* [10] and its website contains an excellent record of the various classifier techniques that have been applied by researchers to this problem and the associated test error rates that have been achieved. A large number of classifier techniques have been applied to this problem including Linear Classifiers, K-Nearest Neighbor, Boosted Stumps, Non-Linear Classifiers, SVMs, Neural Nets and Convolutional Nets [10].

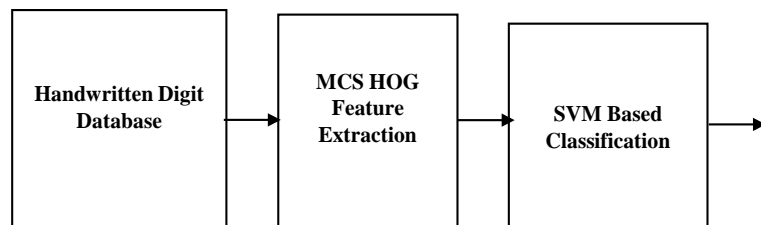
Lecun *et al.* achieved excellent classification results in their seminal work on document recognition using gradient-based learning in [12]. Their research can easily be considered amongst the pioneering work in the applications of Convolutional Neural Networks (CNNs) for image analysis and recognition. In their ground breaking work they achieved error rates of 0.8% using SVM employing degree 9 polynomial functions and 0.7% using a Convolutional Neural Network architecture. The CNNs are the main stay of the currently popular technique of Deep Learning for image recognition [3] [4]. Decoste *et al.* have reported error rates of 0.56% by applying SVMs to the digit recognition problem using degree 9 polynomials [13]. Simard *et al.* have also considered the best practices for the application of CNNs to visual document analysis [14]. Keyser *et al.* [15] have applied non-linear deformation models for image recognition tasks and have achieved an error rate of 0.54% in their work. Increasingly low error rates have been achieved by researchers by employing ensemble techniques to the digit recognition problem. Meier *et al.* have reported an error rate of 0.39% using a committee of Neural Networks [16]. A low error rate of 0.23% has been achieved by Ciresan *et al.* in [17] [18] using a committee of CNNs.

Our current research contributions include the utilization of the proposed new Multiple-Cell Size (MCS) concept for HOG feature extraction and utilization of the multiclass classification capabilities of SVM classifiers through the Error Correcting Output Codes (ECOC) methodology to achieve efficient handwritten digit recognition. Hence the proposed optimizations in both the Feature Space as well as in the Classifier Space have enabled us to achieve at par or better performance compared to existing handwritten digit recognition techniques while employing comparatively simple computational methods in both domains.

### 3. The Proposed Handwritten Digit Recognition System

The proposed handwritten digit recognition system follows the standard model of feature based classification systems consisting of the digit image database, an

essential feature extraction sub-block and a main classification sub-block. The MNIST Benchmark database of handwritten digits has been considered in this work. The Histogram of Oriented Gradient (HOG) technique [19] has been extended in this work by using a new Multiple-Cell Size (MCS) HOG approach to extract features from the database images and the classification sub-block is based on the Support Vector Machine (SVM) classification methodology [21]. The performance of the classification system depends on a number of factors including the type of feature extraction techniques used and the kind of classifiers used for the classification task. Other important factors that affect the classification performance include pre and post processing procedures. In this work the pre or post processing steps have not been employed in the feature extraction process since the HOG descriptors based classification systems are not sensitive to pre-processing operations [19]. The number of images available in the dataset for training, validation and test purposes also plays dominant role in the performance of the classification system [3] [4]. Both the Independent Test Set as well as the 10-Fold Cross-Validation strategies have been used to evaluate the performance of the system. The Block Diagram of the proposed Handwritten Digit Recognition system is shown in **Figure 1**. The detailed description of the various sub-blocks is given below.



**Figure 1.** Proposed MCS HOG and SVM based digit recognition system.

### 3.1. MNIST Handwritten Benchmark Digit Database

For the handwritten digit database, the Benchmark MNIST Digit Database has been considered in this work to test and validate the digit recognition system. The MNIST Digit Database consists of 60,000 images of 10 digit classes in the training set and another 10,000 digit images in the test set for a total of 70,000 images in the database. This database is available at the website maintained by Lecun *et al.* at [10]. The spatial resolution of the images is  $28 \times 28$  pixels and all images are grayscale images. Details about the MNIST digit database are available at [10]. The Independent Test Set and the Cross-Validation strategies have been employed to assess the performance of the system on this database. For the Independent Test set, the given 10,000 images of the test set have been used. For the Cross-Validation Strategy purposes, the 10-Fold Cross-Validation has been used on the set of 70,000 images obtained by combining the given two sets. The 10-fold Cross-Validation Strategy has been used to obtain a realistic performance determination of the proposed digit recognition system. A random selection of images from the MNIST digit database is shown in **Figure 2** below.



**Figure 2.** Random images from the MNIST digit data set.

### 3.2. New Multiple-Cell Size HOG Feature Extraction

The feature extraction sub-block of the proposed system is based on the Histogram of Oriented Gradient (HOG) feature extraction methodology [19]. The Histogram of Oriented Gradient (HOG) technique was suggested by Navneet *et al.* in [19] for the task of human detection in images. As reported by Navneet *et al.*, the HOG technique is similar to the edge orientation histograms [20] but the computations involved in HOG analysis are based on utilizing a dense grid of uniformly spaced cells along with overlapping local contrast normalizations to achieve better performance. The HOG technique exploits the property of the edge directions *i.e.* the local orientation measurements that carry useful information about object shape and form [19]. Hence the HOG descriptors can be used for object and shape recognition and classification. Apart from their initial application to the problem of human detection in images, the HOG descriptors have also been extended to shape and object recognition problems. The HOG descriptors have been applied to digit recognition in [21] using synthetic data set. In this work the HOG application has been extended to real world handwritten digit recognition problem using the Benchmark MNIST Handwritten digit database and we have also introduced a new concept of utilizing a Multiple-Cell Size (MCS) HOG approach to achieve high classification accuracy. Classification and recognition tasks based on the HOG descriptors are sensitive to the Cell Size that is required in the relevant computations [19]. Finding an appropriate Cell Size for the computation of HOG descriptors plays crucial role in the classification performance. In this work, two different Cell Sizes have been used in the computations and the feature vector is a concatenation of the HOG descriptors computed using the 2 cell sizes. It is proposed to use the new Multiple-Cell Size (MCS) concept for computation of HOG descriptors to form the feature vector for classification.

### 3.3. SVM Based Classification

The Support Vector Machine (SVM) algorithm is a powerful classification tool

that is used extensively in Artificial Intelligence (AI) and Machine Learning (ML) tasks. The SVM algorithm was developed by Vapnik *et al.* [22] for ML tasks. In its original form, the SVM algorithm results in a binary classification solution. The extension of SVMs to multiclass problems is possible through the Error Correcting Output Codes (ECOC) technique [23]. Multiclass classification has been achieved in this work using SVMs by utilizing the ECOC technique available in MATLAB. The ECOC technique breaks down the multiclass classification problem into a series of binary classification tasks. A One-vs-One (OVO) coding design strategy has been used for the multiclass SVM implementation using the ECOC model.

## 4. Experimental Setup and Parameter Selection

The performance of the Digit Recognition system depends on both the Feature Extraction as well as the SVM Classification steps. These steps in turn depend on the relevant parameters that must be properly selected in order to achieve high classification accuracy. Hence the design of an efficient recognition system depends on the proper selection of parameters both in the Feature Space as well as in the Classifier Space.

### 4.1. Parameter Selection for the New MCS HOG Descriptors

The HOG based feature extraction process is sensitive to the cell size that is used in HOG analysis of images [19]. Choosing a small cell size compared to the size of the image will result in a larger feature vector which may confuse the classifier system leading to poor recognition performance. On the other hand, choosing a larger cell size may result in insufficient features which will also result in poor classifier performance. Hence choosing an appropriate cell size is crucial for classification accuracy. It has been observed in this work that concatenating feature vectors of two different cell sizes results in superior classification performance compared to individual feature vectors. Hence a new Multiple-Cell Size (MCS) approach has been adopted in this work and the following two cell sizes have been used to compute the HOG descriptors to achieve superior classification accuracy:

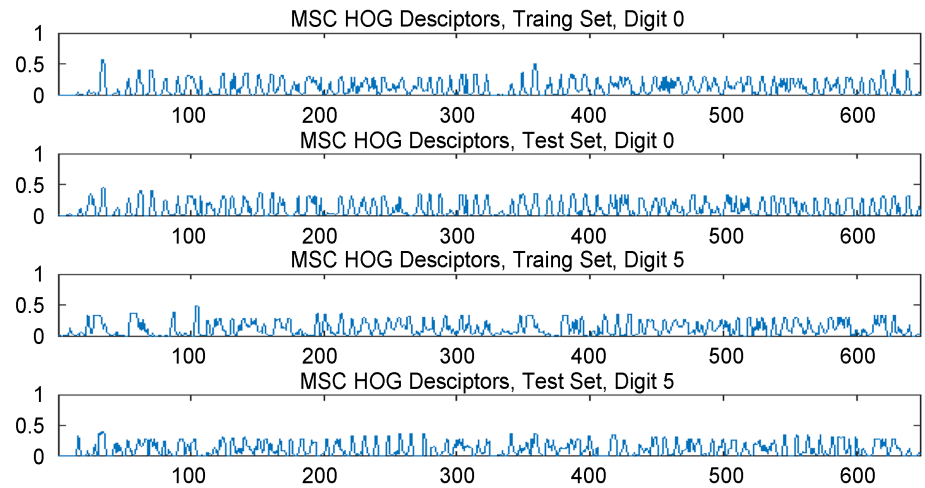
- Cell\_Size1 ( $6 \times 6$  pixels) for computing the feature vector HOG\_Cell\_Size1
- Cell\_Size2 ( $7 \times 7$  pixels) for computing the feature vector HOG\_Cell\_Size2

The HOG features computed using the two cell sizes have been concatenated together to form the overall feature vector for the classifier system:

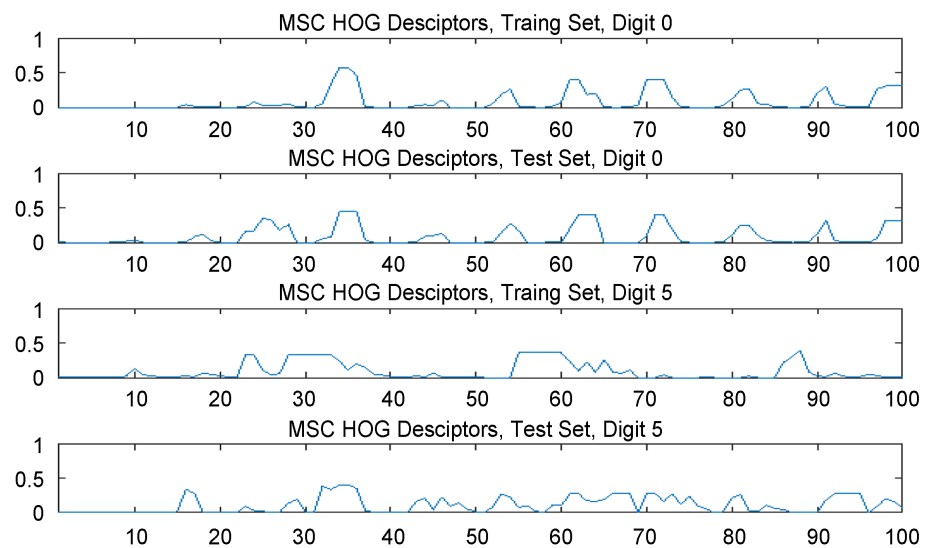
$$\text{New\_MCS\_HOG} = [\text{HOG\_Cell\_Size1 HOG\_Cell\_Size2}] \quad (1)$$

The concatenation of the feature vectors combines the individual discriminating powers of the two feature vectors into one overall powerful feature vector.

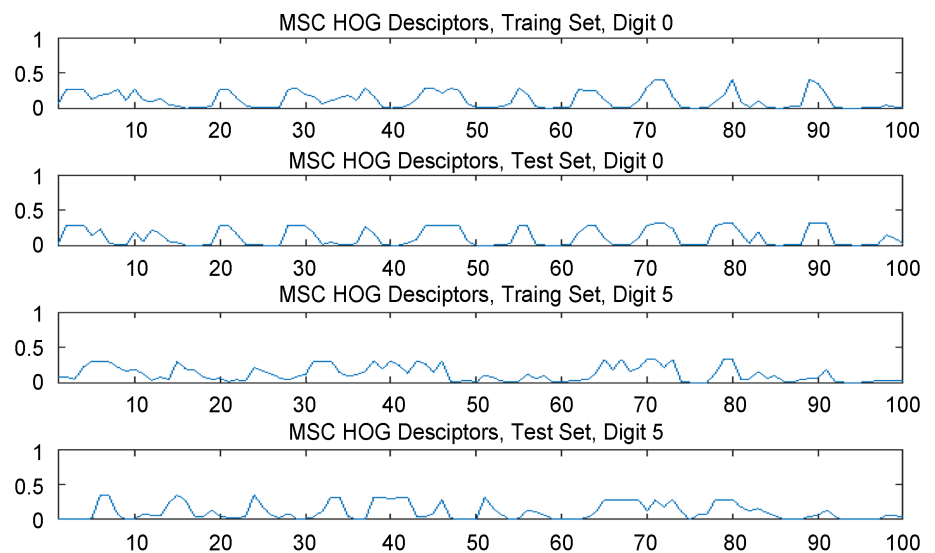
The HOG descriptors associated with the MCS methodology have been plotted in **Figures 3-5** below as sequence of values to highlight their discriminating power. It is clear from the figures that MCS HOG features computed for different digits carry different information thus highlighting the discriminating power



**Figure 3.** New MCS HOG descriptors for digit “0” and digit “5”.



**Figure 4.** First 100 values of new MCS HOG descriptors for digit “0” and digit “5”.



**Figure 5.** Last 100 values of new MCS HOG descriptors for digit “0” and digit “5”.



of this technique. **Figure 3** shows the full or complete feature vector whereas the first and last hundred values of the feature vector have been plotted in **Figure 4** and **Figure 5** to show the values at expanded scale. It is mentioned here that the selection of the two cell sizes  $6 \times 6$  pixels and  $7 \times 7$  pixels is related to the spatial resolution of the digit images of the MNIST database. Selecting smaller or larger cell sizes relative to the digit image size will result in larger or smaller feature vectors which will affect the classification performance of the recognition system. Arriving at the optimum cell sizes has been based on achieving the best classification performance through experimentation.

From **Figure 4** and **Figure 5**, it is evident that for different digit images *i.e.* Digit “0” and “5” the HOG descriptors carry relevant discriminating information highlighting the discriminating powers of these descriptors. At the expanded scale, the same-class pattern similarity is clear from the plots of **Figure 4** for the first 100 values. Similarly, the different-class pattern differences are also clear from the subject plots.

The same-class and different-class trends, as mentioned above, are obvious from **Figure 5** as well.

#### 4.2. Parameter Selection for the SVM Classifiers

As mentioned before, the ECOC model for multiclass SVM implementation has been used in this work. The ECOC model uses a number of binary SVMs to achieve multiclass classification [23]. Hence the parameters associated with the SVM Based Classification sub-block of the system are related to the multiclass extension technique used *i.e.* the ECOC model and the individual binary SVMs that are used within the ECOC model to achieve the multiclass extension. For the individual binary SVMs in the model, a polynomial Kernel of order 4 and the Sequential Minimal Optimization (SMO) algorithm have been used. For the ECOC model itself, a One-vs-One (OVO) coding design methodology has been used on the binary learners in the model. The number of binary learners needed for the  $k = 10$  class handwritten digit recognition problem are 45 as given below:

$$\text{Binary Learners} = \frac{k(k-1)}{2} = 45 \quad (2)$$

#### 4.3. Summary of Parameter Selection

**Table 1** below provides a summary of the parameter selection and system setup for the proposed classification system. The table shows the selected values for both the Feature Space as well as the Classifier Space parameters.

**Table 1.** Summary of parameter selection.

| Item/Entity                                    | Parameter Type              | Parameter Value                   |
|--|-----------------------------|-----------------------------------|
| New MCS HOG Descriptors<br>(Feature Space)     | Cell Size                   | Cell_Size1 ( $6 \times 6$ pixels) |
|  |                             | Cell_Size2 ( $7 \times 7$ pixels) |
| SVM Multiclass Extension<br>(Classifier Space) | ECOC Coding Design          | One-vs-One (OVO)                  |
|  | SVM Polynomial Kernel Order | 4                                 |
|  | SVM Optimization Algorithm  | SMO                               |



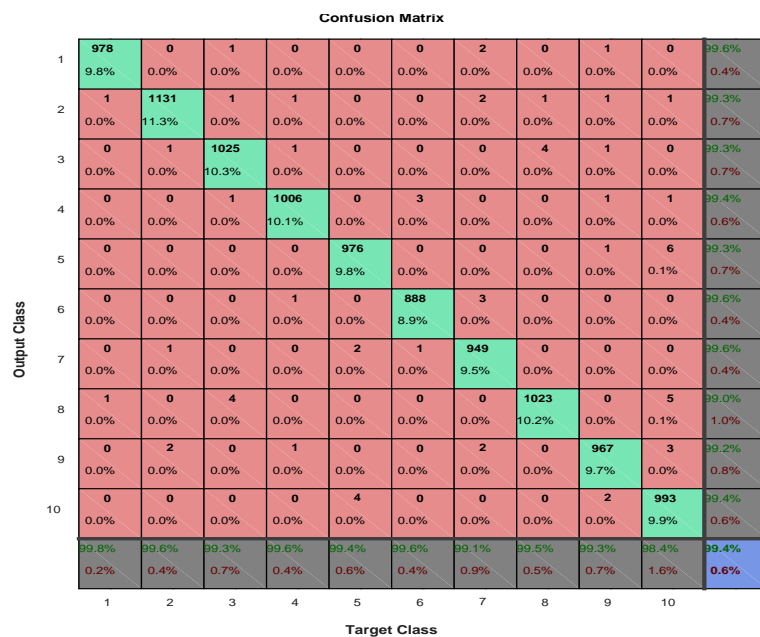
## 5. Results and Discussions

A large number of experiments have been conducted by us in order to determine the performance of the Handwritten Digit Recognition system. Both the Independent Test Set as well as the 10-Fold cross-validation methodology have been used in order to determine the classification performance of the system. The classifier performance has been measured in terms of Classification Accuracy. Classification accuracies of **99.36%** and **99.26%** have been achieved for the Independent Test Set and the 10-Fold cross-validation strategies respectively.

### 5.1. Performance of the System on the Independent Test Set

The Independent Test Set structure of the MNIST Handwritten Benchmark Digit database consists of a Training Set of 60,000 images and a Test Set of 10,000 images. Using the Independent Test Set, our proposed system has been able to achieve a classification accuracy of **99.36%** based on a concatenated set of HOG features computed using two cell sizes *i.e.* a cell size of  $6 \times 6$  pixels and a second cell size of  $7 \times 7$  pixels. For the individual HOG descriptors the system has obtained classification accuracies of 99.17% and 99.14% for the 2 cell sizes respectively. But using the Multiple-Cell Size concept, the system has been able to obtain improved classification performance of 99.36%.

The Confusion Matrix and the Receiver Operating Curve (ROC) have also been computed to analyze the performance of the proposed system on the Independent Test Set. The Confusion Matrix plot is shown in **Figure 6**. The Confusion Matrix plot shows that the system has achieved an overall classification accuracy of 99.4% *i.e.* 99.36% rounded off by the system. The Confusion Matrix Plot also shows the performance of the classification system for individual classes. The main diagonal shows classification performance of the system on the



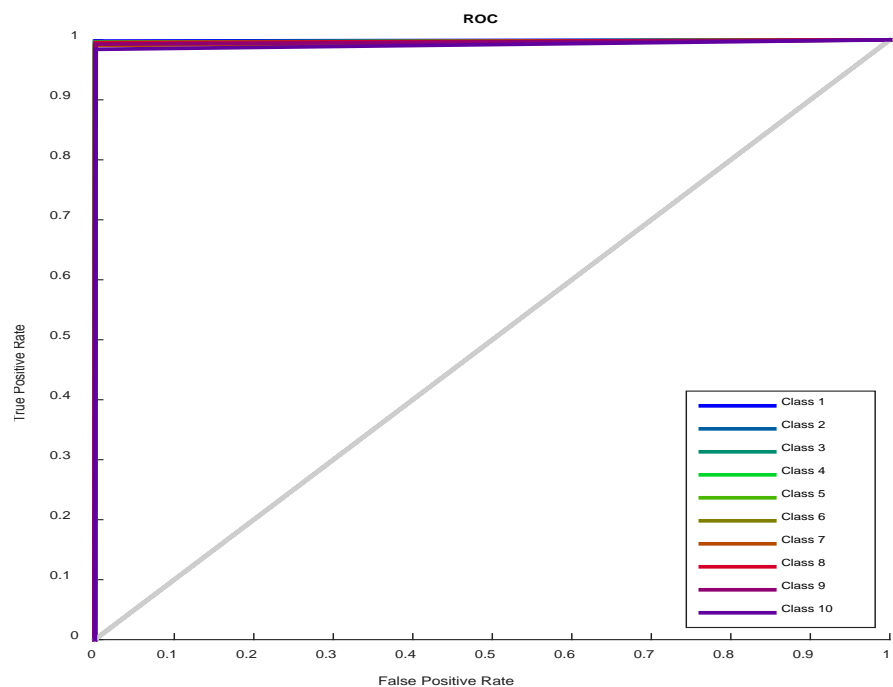
**Figure 6.** Confusion matrix for independent test set.

individual classes. The off diagonal values indicate the challenges encountered by the system in classifying individual classes. It is clear from the Confusion Matrix plot that the classification system has performed excellently on Classes 1, 6 and 7. The plot also shows that Class 8 has been the toughest class for classification with classification accuracy of 90%.

The Receiver Operating Characteristics (ROC) curve has been plotted in **Figure 7** and it shows the concentration of values towards the top left hand corner of the plot indicating excellent performance of the proposed system. It is clear from the plots of both the Confusion Matrix and the ROC that the performance of the proposed system on the Independent Test Set has been excellent.

### 5.2. Performance of the System on the 10-Fold Cross-Validation

In order to obtain the 10-Fold Cross-Validation performance, the available Training Set and Test Set are first concatenated to obtain an overall dataset of 70,000 digit images. We then performed a 10-fold cross-validation of the system and obtained classification performance averaged over the 10 folds. In the 10-fold cross-validation strategy, the available data set is divided into 10-folds with 9 folds used for training purposes and the remaining 10th fold used for testing purposes. This process is repeated for all 10 folds in turn and the performance is averaged over the 10 folds. Similar to the independent test set case, the performance of the system has been evaluated on the 10-fold case for both the individual HOG Descriptors as well as the new MCS HOG descriptors. For the Cell\_Size1 HOG descriptors of size  $6 \times 6$  pixels, a 10-fold classification accuracy of 99.10% has been obtained and for the second Cell\_Size2 HOG descriptors of size  $7 \times 7$  pixels, the 10-fold classification accuracy achieved has



**Figure 7.** ROC Plot showing performance on independent test set.

been 99.17%. For the MCS HOG descriptor feature vector, a 10-fold classification accuracy of **99.26%** has been achieved by the system. The 99.26% classification accuracy for 10-fold cross-validation indicates excellent performance of the proposed system.

### 5.3. Summary of the Results

The summary of the results of the experiments for the Independent Test Set and the 10-fold cross-validation are presented below in **Table 2**. It is clear from the table that the proposed new MCS HOG technique gives improved results compared to single cell size based HOG descriptors.

### 5.4. Performance Comparison with State-of-the-Art

As discussed in Section 2, the error rates reported by prominent researchers for the MNIST benchmark database range from 0.8% to 0.54% corresponding to classification accuracy of 99.2% to 99.46% respectively for non-ensemble based systems. We will not consider ensemble methods in our comparison and will only compare our non-ensemble technique's performance to peers of similar nature. The classification accuracy of **99.36%** of the proposed system on the Independent Test Set is comparable to the state-of-the-art in handwritten digit recognition systems. Also, extra performance analysis through the 10-fold cross-validation methodology has been provided and the system has achieved **99.26%** classification accuracy thus highlighting the strength of the proposed technique. It is emphasized here that the proposed technique has achieved this improved performance by exploiting the Feature Space as well as the Classifier Space through use of optimum parameter selection and by specifically using the new MCS HOG concept of utilizing different cell sizes in HOG descriptor calculations to harness the discriminating powers of the MCS HOG descriptors.

We mention here, that the other existing methods that have achieved recognition accuracy like our proposed method (*i.e.* greater than 99% recognition accuracy on the MNIST Benchmark database) include Deep Learning techniques employing individual CNNs and CNN ensembles. But, our proposed technique is preferable to Deep Learning systems since Deep Learning Handwritten Digit Recognition systems based on Convolutional Neural Networks (CNNs) have complex network architectures and are difficult to train compared to our proposed SVM Based scheme. CNNs usually require GPU processing instead of the

**Table 2.** Classification performance of the proposed system.

| Testing/Validation Strategy | HOG Cell Size             | Classifier Accuracy |
|-----------------------------|---------------------------|---------------------|
| Independent Test Set        | Cell_Size1 (6 × 6 pixels) | 99.17%              |
|                             | Cell_Size2 (7 × 7 pixels) | 99.14%              |
|                             | <b>New MCS HOG</b>        | <b>99.36%</b>       |
| 10-Fold Cross-Validation    | Cell_Size1 (6 × 6 pixels) | 99.10%              |
|                             | Cell_Size2 (7 × 7 pixels) | 99.17%              |
|                             | <b>New MCS HOG</b>        | <b>99.26%</b>       |

regular CPU and this requirement is a burden on the computing resources. We have achieved State-of-the-Art results without requiring GPU processing. Our proposed system is superior in the sense that it achieves a State-of-the-Art recognition accuracy without needing GPUs and complicated CNNs. The MCS HOG based Feature Extraction process of our system for Handwritten Digit Recognition is preferable to any complicated CNN based Feature Extraction scheme.

## 6. Conclusion and Future Work

It has been demonstrated in this work that using a new Multiple-Cell Size approach to HOG descriptor computation and the SVM based classification scheme, a classification accuracy of **99.36%** has been achieved on the Independent Test Set for the Handwritten Digit Recognition problem. It has been demonstrated that using a 10-fold cross-validation strategy, a classification accuracy of **99.26%** has been obtained thus highlighting the generalization property of the proposed scheme. It has been shown that using the new MCS HOG approach results in superior classification performance compared to individual cell size HOG descriptors. Comparing the performance of the proposed system with the eminent results of existing techniques, it is concluded that the new MCS HOG based proposed system has achieved comparable performance to most of the existing techniques as well as it has exceeded the performance of others. Hence our results are comparable to state-of-the-art and this research will contribute positively to the research effort in the field of handwritten digit recognition. In future, other feature extraction methods and classification schemes will be considered for the digit recognition system.

## Acknowledgements

This research effort has been greatly supported by the Arab Open University (AOU). This support is highly appreciated and acknowledged.

## References

- [1] Schantz, H.F. (1982) The History of OCR. Recognition Technologies Users Association, VT.
- [2] Zhou, Z.H. and Feng, J. (2017) Deep Forest: Towards an Alternative to Deep Neural Networks. arXiv preprint arXiv:1702.08835. <https://arxiv.org/pdf/1702.08835.pdf>
- [3] Nielsen, M.A. (2015) Neural Networks and Deep Learning. <http://neuralnetworksanddeeplearning.com/>
- [4] Bengio, Y.I., Goodfellow, J. and Courville, A. (2016) Deep Learning. MIT Press, Cambridge, MA. <http://www.deeplearningbook.org>
- [5] Farulla, G.A., Murru, N. and Rossini, R. (2016) A Fuzzy Approach for Segmentation of Touching Characters. arXiv:1612.04862v1. <https://arxiv.org/abs/1612.04862>
- [6] Chen, G., Li, Y. and Srihari, S.N. (2016) Word Recognition with Deep Conditional Random Fields. *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, 25-28 September 2016, 1928-1932. <https://doi.org/10.1109/icip.2016.7532694>

- [7] Liu, Z., Li, Y., Ren, F. and Yu, H. (2016) A Binary Convolutional Encoder-Decoder Network for Real-Time Natural Scene Text Processing. arXiv:1612.03630v1. <https://arxiv.org/abs/1612.03630v1>
- [8] Jaderberg, M., Simonyan, K., Vedaldi, A. and Zisserman, A. (2014) Reading Text in the Wild with Convolutional Neural Networks. arXiv preprint arXiv:1412.1842.
- [9] Yin, X.C., Yin, X., Huang, K. and Hao, H.W. (2014) Robust Text Detection in Natural Scene Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **36**, 970-983. <https://doi.org/10.1109/TPAMI.2013.182>
- [10] LeCun, Y., Cortes, C. and Burges, C.J. (1998) The MNIST Database of Handwritten Digits. <http://yann.lecun.com/exdb/mnist/>
- [11] Tabacof, P., Tavares, J. and Valle, E. (2016) Adversarial Images for Variational Autoencoders. <https://arxiv.org/pdf/1612.00155v1.pdf>
- [12] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [13] Decoste, D. and Schölkopf, B. (2002) Training Invariant Support Vector Machines. *Machine Learning*, **46**, 161-190. <https://doi.org/10.1023/A:1012454411458>
- [14] Simard, P.Y., Steinkraus, D. and Platt, J.C. (2003) Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. *Proceedings of the 7th International Conference on Document Analysis and Recognition*, **3**, 958-963. <https://doi.org/10.1109/icdar.2003.1227801>
- [15] Keysers, D., Deselaers, T., Gollan, C. and Ney, H. (2007) Deformation Models for Image Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**, 1422-1435. <https://doi.org/10.1109/TPAMI.2007.1153>
- [16] Meier, U., Ciresan, D.C., Gambardella, L.M. and Schmidhuber, J. (2011) Better Digit Recognition with a Committee of Simple Neural Nets. *IEEE International Conference on Document Analysis and Recognition*, Beijing, 18-21 September 2011, 1250-1254.
- [17] Ciresan, D., Meier, U. and Schmidhuber, J. (2012) Multi-Column Deep Neural Networks for Image Classification. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, RI, 16-21 June 2012, 3642-3649. <https://doi.org/10.1109/cvpr.2012.6248110>
- [18] Ciresan, D.C., Meier, U., Gambardella, L.M. and Schmidhuber, J. (2011) Convolutional Neural Network Committees for Handwritten Character Classification. *IEEE International Conference on Document Analysis and Recognition*, Beijing, 18-21 September 2011, 1135-1139.
- [19] Dalal, N. and Triggs, B. (2005) Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, **1**, 886-893.
- [20] Freeman, W.T. and Roth, M. (1995) Orientation Histograms for Hand Gesture Recognition. *International Workshop on Automatic Face and Gesture Recognition*, **12**, 296-301.
- [21] Machine Learning with MATLAB, MathsWorks. [www.mathworks.com/solutions/machine-learning.html](http://www.mathworks.com/solutions/machine-learning.html)
- [22] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [23] Dietterich, T.G. and Bakiri, G. (1995) Solving Multiclass Learning Problems via Error-Correcting Output Codes. *Journal of Artificial Intelligence Research*, **2**, 263-286.

**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact [jilsa@scirp.org](mailto:jilsa@scirp.org)