Local Fourier features for handwriting digit images classification

Djimeli-Tsajio Alain Bernard[1], Noulamo Thierry[2], Lienou T. Jean-Pierre[2], Tchiotsop Daniel[3], Panduranga Nagabhushan[4]

[1]Department of Telecommunication and Network Engineering, URAIA, IUT-FV, University of Dschang, Bandjoun, Cameroon
[2]Department of Computer Engineering, URAIA, IUT-FV, University of Dschang, Bandjoun, Cameroon
[3]Department of Electrical Engineering, URAIA, IUT-FV, University of Dschang, Bandjoun, Cameroon
[4]Department of Information Technology, Indian Institute of Information Technology, Allahabad, India

Article Info

ABSTRACT

Multiple choice questions (MCQ) are effective in normative assessment and offline testing is still relevant due to the lack of efficient mass infrastructures and maintenance. For the automatic correction of MCQ paper form and reporting of the grade, it is generally necessary to read and recognize a handwriting digit in a box. This paper focuses on local feature extraction in the frequency domain using Fourier transform. The pre-process begins with the extraction of the fields from the entity map, followed by the application of 2D fast Fourier transform (2DFFT) and the reduction of computed coefficients to obtain the corresponding final local characteristic in the representation. The experimental results of the Modified National Institute of Standards and Technology (MNIST) handwriting digits dataset show that the local characteristics extracted in the frequency domain used as input to a support vector machine (SVM) classifier are efficient in terms of 99.51% accuracy. The proposed system successfully helped in the reporting of all the marks for seven subjects in a class of 98 students during the automatic correction of the MCQ exam papers.

Corresponding Author:

Djimeli-Tsajio Alain Bernard
Department of Telecommunication and Network Engineering, URAIA, IUT-FV, University of Dschang
P.O. Box 134 Bandjoun, Cameroon
Email: adjimeli@yahoo.fr

1. INTRODUCTION

Multiple choice questions (MCQ) or unique choice questions (UCQ) has proven to be an objective, fair, and efficient assessment process in an environment where we want to select a few candidates out of thousands of applicants, such as entrance examinations. Thanks to information and communication technology tools that reduce the number of participants in the marking process, this form of assessment is now automated and accessible online. However, due to the lack of infrastructure and maintenance, online assessment under such conditions cannot yet be applied in all the countries of the world and traditional paper-based assessment still has a bright future ahead. To reduce human intervention in the correction processes, how to automate the correction of a traditional paper evaluation and automatically publish the results as in the case of an online evaluation? Multiple forms for automatic correction of MCQ are available [1], [2]. In the exam database, each candidate's name is associated with a special code called an anonymity number which is pending the corresponding grade. This anonymity number is associated with the candidate's copy. During automatic correction, this anonymity number field is extracted and the corresponding handwriting digits are digitized to match the candidate's name and grade in the information system. Although handwriting

digit recognition is an old problem [3], automatic and accurate correspondence between the candidate mark and name is a challenge as an anonymity number permutation would result in a permutation of the corresponding candidate's grade and a duplicate of the anonymity number would lead to a verification process that could be long and tedious, with manual correction.

The classification problem is a popular application of machine learning where the reduction of the noise of the dataset in preprocessing has an effective consequence on the expected result. At the input of a classifier, we introduce the description of an object and, hopefully, at the output, a label for the object. A classifier inducer is a supervised learning algorithm that uses a set of observations to learn a hypothesis that can map inputs to correct class labels. For a classifier to be able to generate a good discriminative model, the representation of the input space is of key importance. Many classification algorithms cannot achieve the expected performance when faced with difficult real-world problems, because most of these groups of classifiers are inherently incapable of transforming their input space to gain class separability. Constructed features act as functions that transform data points in the original input space to points in a new space which provides better reparability between classes. Feature selection, also known as variable selection is commonly used as a preprocessing step to machine learning. This is a technique commonly used when selecting a subset of relevant features for building robust learning models, as it is found to be very effective in removing irrelevant and redundant features, increasing the efficiency of learning tasks, improving learning performance such as predictive accuracy and improve the comprehensibility of learned results [4]. However, it is difficult to link the features from the original feature space to new features. Therefore, further analysis of new features is problematic since there is no physical meaning for the transformed features obtained from feature extraction techniques [5]. A popular approach to classification today is to use a pre-trained deep neural network that is complex and training expensive, to extract useful and informative features of natural images and use it as a starting point for training, assuming that the source and target domains are related to each other [6]–[8]. Feature obtained in a transform domain is a promising way to improve network performances [9]. Kou *et al.* [10] have used local features for the age estimation of facial images. The global feature was a cascade of the real part and imaginary part of the Fourier transform of all images. The image was divided into *n* blocks of the same size the local features were obtained by applying the principal component analysis (PCA) to each block. Their experiment gives reason to a local feature compared to a global feature on the age information on which they concentrated. Li *et al.* [11] were convinced that, because the image of the palm print in the spatial domain has a strong line, in the frequency domain, there will be more information in other directions. They proposed a process of searching for palm prints in the database by comparing the global characteristics extracted in the frequency domain used as an index. Su *et al.* [12] has proposed a face recognition method that combines both global and local discriminative features. Global features are real and imaginary components of a low frequency of the image 2D fast Fourier transform (2DFFT). They describe the characteristics of the whole face. Gabor wavelet transform is firstly applied to the whole face image, and then the resulting Gabor features are spatially partitioned into N subsets. These local features capture more details in local face areas. The hierarchical feature of global features and local features gives acceptable results. The authors conclude that local features based on Gabor filters can achieve excellent performance better than the best-known results. This idea has been extended with good results to palmprint recognition [11].

The Fourier transform helps to recognize position-shifted characters in the magnitude spectrum. Fourier transform and its variants such as Fourier moments, Fourier descriptors, Local binary pattern Fourier histogram, polar Fourier transform, Fourier-Mellin transform, or wavelet-Fourier descriptors are widely used for feature extraction and shape classification [13]–[15]. This rich knowledge domain can also be used to extract local shapes in the image. Local Fourier decomposition has been successfully used for the analysis of microscopic images [16] or to improve deep learning performances [17]. This paper aims to provide the reader with a simplified approach that improves the performance of classical handwriting digits classifiers by providing an overview of its behavior with an entry presented as a local structure of the target entry in the frequency domain. Applications are made on the support vector machine (SVM) for the extraction of anonymity number during the process of automatic correction of MCQ. The rest of this paper is organized as follows: section 2 describes the material and the method, emphasizing the proposed approach. Section 3 includes the discussion of the experimental results obtained and section 4 concludes our work.

2. METHOD
2.1. Fourier transform
2.1.1. General basis of the frequency domain representation
The Fourier transform is a commonly used signal processing and control theory technique when examining information in the frequency domain. It provides a one-to-one transform of signals from a time-domain representation $f(t)$ to a frequency-domain representation $F(\xi)$, the reverse transformation is possible.

Frequency domain analysis uses mathematical functions or signals to extract information related to frequency rather than time. Usually, a pair of mathematical operators called transformations is needed. The 1D discrete Fourier transform (1DFT) is suitable for periodic signals. If the signal is not periodic then the Windowed FT or the linear integral transformation with time-localized basis function can be used. To calculate the discrete Fourier, transform and its inverse, we usually use the fast Fourier transform (FFT) algorithm. The 1DFT is used in the case of a finite-duration or discrete-time signal. $X(0) \dots X(N-1)$ the 1DFT of the signal $x(0) \dots x(N-1)$ is stated as (1):

$$X(v) = \sum_{n=0}^{N-1} x[n] \, exp(-2\pi j \frac{vn}{N}) \tag{1}$$

where $v = 0,1,\dots,N-1$

The frequency domain representation of a signal conveys information about the amplitude and phase of the signal at each frequency, accordingly, the output of the FFT is complex. The 2DFT is given by (2).

$$F(u,v) = \frac{1}{MN}\sum_{m=0}^{M-1\Sigma}\sum_{n=0}^{N-1\Sigma} f(m,n) \, exp\left[-2\pi j (\frac{mu}{M}\frac{nv}{N}[])\right] \tag{2}$$

where $u = 0,1,\dots,M-1$, $v = 0,1,\dots,N-1$. Power spectrum is given as (3).

$$P(u,v) = |F(u,v)^2| = F_{Real}^{2_{Imag}^2} \tag{3}$$

where F_{Real} and F_{Imag} are respectively the real part and imaginary of $F(u,v)$. One of the important properties of the Fourier transform is the conservation of energy between the time and frequency domain.

2.1.2. Local Fourier features

Feature representation in the frequency domain is a preprocessing approach that transfers local spatial information into its correspondent in the frequency domain. Because we are interested in the local structure of the image a window size (F) is needed. For compliance, we will only use square windows to preserve the relationship between the pixels by learning the image features using said small squares of input data. For the representation, the window is then shifted until it covers the entire image while computing the 2DFFT. If we start with the window in the upper left corner of the input table and then drag its global locations down and to the right, we can move our window more than one pixel at a time ignoring the intermediary locations. We refer to the number of rows and columns traversed per slide as the stride (S). The stride is a parameter that tells how many pixels the window is shifted. For example, if use the 6×6 image with a stride of 2, and a window of size 3, we end up with the local Fourier coefficients shown in Figure 1.
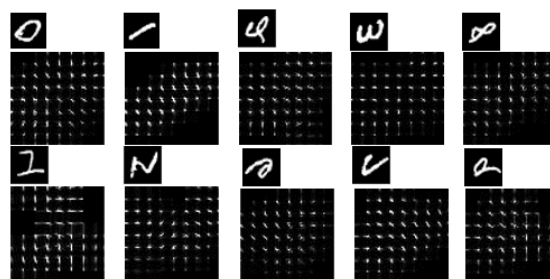

Figure 1. Local Fourier features for an element in each Modified National Institute of Standards and Technology (MNIST) dataset class ranging from 0 to 9 for stride 3 and window 10

In this example, the square window of size 3 is superimposed top left of the image so that the maximum number of the window boxes are used before the computation of the 2DFFT. Because we are interested in the preservation of the image energy, only the magnitude of the 2DFFT is reported. The window is then shifted successively by two pixels to the right and successively by two pixels (stride) downwards to obtain the required matrix which will undergo normalization before being used. The number of Fourier coefficients for the example given in Figure 1 is (10×10). In this figure, for stride 3, local Fourier features were computed for one element of each class of the MNIST dataset. We can see the decimal character at the

top and the representation of its characteristic at the bottom. Table 1 shows a fraction of the corresponding number of Fourier coefficients computed with respect to the filter size and the stride for a 28×28 image.

Table 1. Number of expected load frequency control (LFC) with respect to some window size and stride

Window size stride	3	5	7	9	11	13	15	17	19	21	23	25	27
1	6724	17956	33856	53824	77284	103684	132496	163216	195364	228484	262144	295936	329476
2	1681	4489	8464	13456	19321	25921	33124	40804	48841	57121	65536	73984	82369
3	784	2116	4096	6400	9216	12544	15876	19600	23716	27556	31684	36100	40000
5		784	1444	2304	3364	4489	5776	7056	8464	10000	11449	12996	14400
7			625	1024	1521	2116	2704	3249	3844	4489	5184	5929	6724
9				784	1156	1600	2116	2704	3364	3844	4356	4900	5476
11					784	1024	1296	1600	1936	2304	2704	3025	3364
13						784	1024	1296	1600	1936	2304	2704	3136
15							529	676	841	1024	1225	1444	1600

2.1.3. Coefficient reduction

The number of calculated coefficients explodes when the stride tends towards 1. We want to know how to reduce these coefficients and their impact on the performance of our representation. The Fourier transform is a window-specific task where one would like to keep higher frequencies which represent the edges of the image and lower frequencies which represent the details of the image while ignoring frequencies that correspond to the homogeneous area of the image. These frequencies, which give rise to non-crucial information, are located in the middle of the widths and lengths of the representation of the coefficients in the window. Figure 2 displays an example of an 11×11 window for the representation of the characteristics of the image where the higher brightness represents the importance of the pixel, and the higher darkness represents the non-importance of the pixel in the representation.
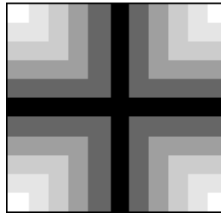

Figure 2. Window for feature representation where higher brightness represents the importance of the pixel and the higher darkness represents the non-importance of the pixel in the representation

2.2. Support vector machine classifier

Support vector machine (SVM) for classification uses hyperplanes for decision boundaries in the input space or in the high-dimensional feature space from a labeled training dataset [18]. The separator is usually constructed with a maximum distance between the hyperplane and the nearest negative and positive samples. Throughout the training phase, SVM takes each element in a labeled data matrix and treats it as a row in an input space or a high-dimensional feature space, where the number of attributes identifies the dimensionality of space. The goal of the SVM training algorithm is then to determine the best hyperplane that separates each positive and negative training sample. For the hyperplane defined as $w^T x + b = 0$, where w is the vector of hyperplane coefficients and b is a bias term, the margin between the hyperplane and the nearest point is to be maximized [18]. A suitable kernel function such as sigmoid function (SF), polynomial function (PF), and radial basis function (RBF) for a certain problem, depends on the specific data. Multi-class SVM includes several two-class subproblems that can be easily combined using one-over-all and one-over-one coding algorithm [19], [20]. In our case, we applied the $fitcecoc()$ function on the LFC features for the classification task.

2.3. Hardware and dataset

Our experiments have been carried out on the MNIST database for grayscale images. The MNIST dataset [21] has 70,000 28×28 binary handwritten digits. It gathers 60,000 learning images and 10,000 test images. Our experiment is carried out on an Intel (R) Core (TM) I5-4310M PC with 16 Go of RAM with Windows 7 operating system and MATLAB R2016 [22].

3.   RESULTS AND DISCUSSION

The objective of this section is to study the behavior of our presentation. The parameter used to test the effectiveness of our approach is accuracy, known as the percentage of correctly classified instances. It is calculated as (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

where *TP, TN, FP*, and *FN* represent the number of true positives, true negatives, false positives, and false negatives respectively. The error is also computed as (5).

$$Error = 1 - Accuracy \tag{5}$$

3.1. Performance of the local Fourier features for image classification on MNIST

This section provides best practices with intuition on the size of the window and the pitch for obtaining better accuracy with better computing performance. Our experiments were performed by filling out an accuracy table for some window size values and strides for the MNIST handwriting dataset. This result is presented in Table 2. Except for window size equal to 7 and smaller stride values, the behavior of the accuracy with respect to the window size is also observed. The accuracy increases with the size of the window and stabilizes for window values 15 before decreasing to a small value. It can then be noted that the best accuracies are obtained for window 7 with stride equal to 2 and for window 15 with stride 5. For a given window, it is mainly observed that the performances are better for smaller stride values. However, the accuracy tends to decrease when the stride tends toward one because of the presence of many noisy useful LFC. This observation is highlighted for window 3 and 21. We explored smaller strides for higher window size where the number of computed LFCs is very high and computed the error with respect to the percentage of LFC removed and found that there was no significant effect on the classification error when we reduced the coefficients. For example, for Window 7 and a coefficient reduction of 26.5%, we obtained no loss in accuracy and a loss in accuracy of 0.001 for a coefficient reduction of 81.6%. After experiments, Table 3 summarizes the best accuracies and shows the best accuracy is obtained when the stride is 3 and the coefficient reduction is 39.7%

Table 2. Classification error obtained with respect to some window size and stride

Window size Stride	3	5	7	9	11	13	15	17	19	21	23	25	27
1	1.29												
2	1.22	1	0.84										
3	1.5	1.11	0.84	0.93	0.96	0.49							
5		1.45	1.16	0.93	1.02	0.93	0.88	0.94	1.04				
7			1.64	1.28	1.14	1	1.15	1.17	1.09	1.2	1.15	1.19	1.27
9				2.45	1.3	1.28	1.18	1.19	1.12	1.18	1.23	1.22	1.40
11					3.32	2.32	1.90	1.39	1.17	1.29	1.34	1.29	1.36
13						5.02	3.49	2.61	1.97	1.45	1.27	1.40	1.35
15							7.75	5.74	4.41	333	2.59	2.05	1.93

Table 3. Best classification accuracy reported for the proposed approach on the MNIST handwriting digit dataset

Window size	5	7	9	9	11	13	15	17
Stride	2	2	5	2	2	3	4	5
Coefficients reduction (%)	36.0	36.0	30.6	46.3	69.3	39.7	33.9	0
Best accuracy	99.14	99.18	99.14	99.14	99.10	99.51	99.32	99.06

The local Fourier features representation is a preprocessing technique which aim to reduce inter-class correlation in the dataset. The confusion matrix of our implementation is shown in Figure 3. The maximum accuracy of 99.51 obtained is state-of-the-art. The obtained result could be improved by using the transformation and the combination of transfer learning features [23], [24]. Although works using convolutional neural networks (CNNs) are becoming the state-of-the-art techniques for the classification of hand writing digits with interesting accuracy [3], CNNs are slower than classical classifiers, the training process is CPU intensive and time consuming. In use for example and for the extraction of 10,000 features, the use of transfer learning with the ResNet-101 [25] architecture took us 8613.141627 seconds against only

4102.725112 seconds for the features extraction with the proposed approach for a window of size 9 and a reduction of the coefficient of 46.3%. Figure 4 displays an excerpt of half of the characters not recognized by the system. In this figure, there is the image of the characters and below the label associated with the images of said characters. To assess the degree of satisfaction of the recognition system with respect to unrecognized characters, the labels of the latter were masked, and the figure was submitted to a group of 50 teachers for their labelling. It was asked to assign a label to the image of the characters presented in the figure knowing that they are numbers ranging from 0 to 9. After counting, the results were very variable from a volunteer teacher to another, and everyone had their own understanding of said characters. A good labelling rate ranging from 17/24 to 6/24 with an average of 11.6/24 was obtained. These rates sufficiently show the polysemic character of the numbers not recognized by the system. One of the merits of the system is not to recognize the characters whose recognition and interpretation remain ambiguous for the human being, since machine learning is only the transfer of human knowledge to the machine.

Figure 3. Confusion matrix of LFF and SVM on MNIST handwriting digits dataset
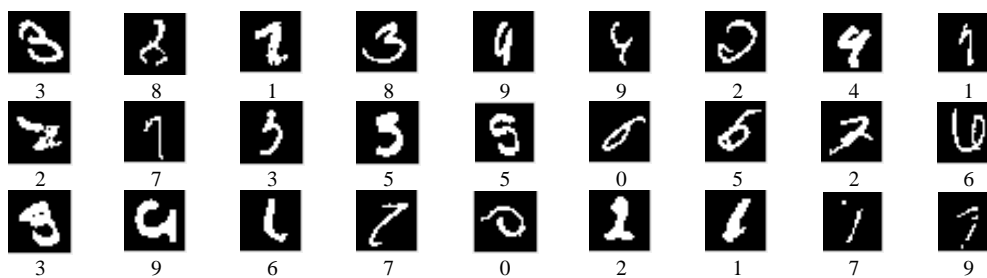
Figure 4. Half of the character's images not recognized by our system and their label below

Data augmentation techniques are generally used to increase the amount of data by combining it with slightly modified copies of previously existing data. Then, augmented data can include biases such as present in the testing dataset. As a result, among others, it resolve class imbalance issues in classification and improve the accuracy of model prediction. Therefore, comparing the results obtained with the data augmentation techniques with the results of those who do not use it is not fair and we have chosen not to compare the result of the proposed approach with those of the literature using data augmentation techniques for MNIST handwriting digits dataset. Table 4 presents the most competitive results (error rate < 1%) found in the state of the art including the proposed approach for the MNIST dataset without data augmentation and CNN [3].

Table 4. Comparison between the results of our studies with those found in the literature for the MNIST database without data augmentation and CNN [3]

Technique	Test Error Rate	Technique	Test Error Rate
NN 6-layer 5,700 hidden units	0.35%	Pooling + SVM	0.64%
MSRV C-SVDDNet	0.35%	Virtual SVM, deg-9 poly, 1-pixel jit	0.68%
Committee of 25 NN 2-layer 800 hidden units	0.39%	NN 2-layer 800 hidden units, XE loss	0.70%
HOPE+DNN with unsupervised learning features	0.40%	SOAE-σ with sparse connectivity and activity	0.75%
Proposed approach (Local Fourier features and SVM)	0.49%	Deep convex net	0.83%
K-NN (P2DHMDM)	0.52%	CDBN	0.82%
COSFIRE	0.52%	S-SC + linear SVM	0.84%
K-NN (IDM	0.54%	2-layer MP-DBM	0.88%
Virtual SVM, deg-9 poly, 2-pixel jit	0.56%	DNet-kNN	0.94%
RF-C-ELM, 15,000 hidden units	0.57%	2-layer Boltzmann machine	0.95%
PCANet (LDANet-2)	0.62%	NN 2-layer 800 hidden units, MSE loss	0.90%
PCANet (LDANet-2)	0.62%	DNet-kNN	0.94%
K-NN (shape context)	0.63%	2-layer Boltzmann machine	0.95%

3.2. Uses of local Fourier features for anonymity number digitization

The trained network is intended for practical use. The MCQ form used is shown in Figure 5. It has marks in the corners of the paper, which help to resize and automatically flip the image of the paper to a standard horizontal position. From this position, the correction system can extract the answer field for the calculation of the score while our module is interested in the field which contains the anonymity number as shown in Figure 6. The horizontal and vertical histogram profile then makes it possible to extract the characters that are introduced one after the other into the system. The character images are first preprocessed, resized and their local Fourier features are used as the input for classification. When all characters are classified, the anonymity number is known and is associated with the grade of the correction system before being inserted into the database. The entity relational model associated is represented in Figure 7. The designed system has 2 modes. In mode 1, the tables Exam and Examinee can be filled before the exam or OCRed and filled from the script.
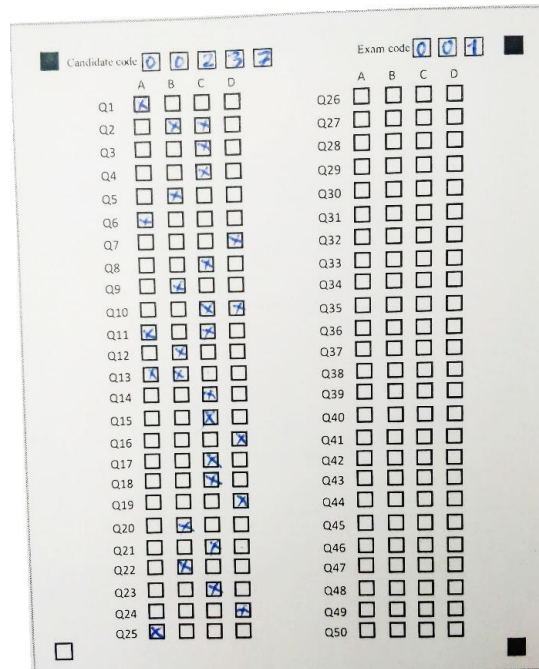

Figure 5. Example of MCQ

The relation *Anonymtable* has 2 foreign keys *IdExaminee* and *IdExam*. These 2 fields permit to create a form that permits to key in *AnonymNumber* which updates the relation and another form based on this field updates the mark. This is essential to keep the process away from malpractice cases. This mode is

recommended to be the manual mode in case there are errors during the marking process of a singular script. In mode 2, the *iidExam, anonymNumber* and the mark are retrieved automatically from the script and updates (inserts) the relation *Anonymtable*. This is recommended to be the batch mode.

Figure 6. Anonymity field extracted

The system was used to report the grades of seven subjects in a class of 98 students during the automatic correction of the exam papers by MCQ. Note that in the experimental phase, we asked anonymous people to encode anonymity and a person has not anonymized more than ten copies. All grades were reported successfully for a window size of 13, stride of 3 and 33.9% coefficient reduction.
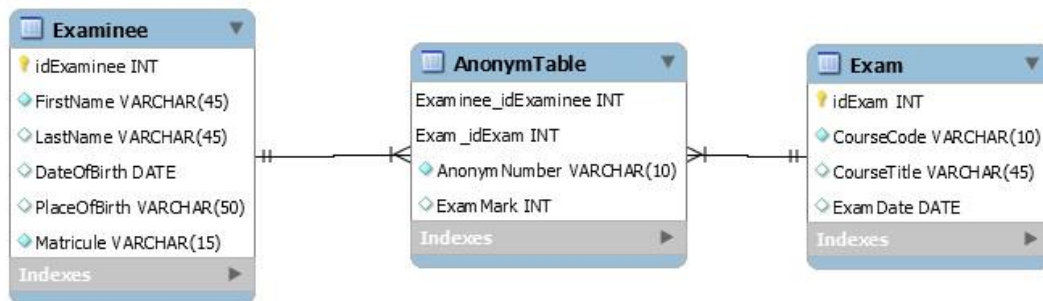
Figure 7. Relational model associated

4. CONCLUSION

This work proposed the use of a network of depth 1 of 2D local Fourier transform instead of convolving a 2D kernel with a network of several depths. This article presents the framework of the local Fourier features for handwriting digit image classification, a preprocessing approach which aims to reduce the interclass correlation in the data set. The parameters of the local Fourier features have been studied and their representation has been described as the reduction in the size of the matrix. Compared to the traditional approach where the grayscale intensities of the color channels are introduced into the classifier, the inputs taken as local characteristics in the frequency domain are more robust and work well. Experiments carried out on the MNIST handwriting dataset have shown very promising results for the proposed technique over the classical classifier in terms of character recognition accuracy and processing time. The obtained result could be improved by using the transformation and the combination of transfer learning features. This approach has been used successfully for anonymity number digitization for automatic grade reporting during offline auto-correction of the MCQ exam form and can be extended to other transformations for the exploration of hidden information for machine learning.

REFERENCES

[1] M. Habeek, C. E. Dridi, and M. Badeche, "Automatic correction of free format MCQ tests," *International Journal of Software Innovation*, vol. 8, no. 1, pp. 50–64, Jan. 2020, doi: 10.4018/IJSI.2020010103.

[2] J. Al-Sadi, D. Al-Halabi, and H. Al-Halabi, "MCQ exams correction in an offline network using XML," *Gstf International Journal on Computing*, vol. 1, no. 1, 2010, doi: 10.5176/2010-2283_1.1.29.

[3] A. Baldominos, Y. Saez, and P. Isasi, "A survey of handwritten character recognition with MNIST and EMNIST," *Applied Sciences (Switzerland)*, vol. 9, no. 15, p. 3169, Aug. 2019, doi: 10.3390/app9153169.

[4] Z. M. Hira and D. F. Gillies, "A review of feature selection and feature extraction methods applied on microarray data," *Advances in Bioinformatics*, vol. 2015, pp. 1–13, Jun. 2015, doi: 10.1155/2015/198363.

[5] C. C. Aggarwal, "Educational and software resources for data classification," *Data Classification: Algorithms and Applications*, pp. 657–665, Jan. 2014, doi: 10.1201/b17320.

[6] A. M. Hassan, M. B. El-Mashade, and A. Aboshosha, "Deep learning for cancer tumor classification using transfer learning and feature concatenation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 6, pp. 6736–6743, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6736-6743.

[7] M. Ghafoorian *et al.*, "Transfer learning for domain adaptation in MRI: Application in brain lesion segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10435 LNCS, Springer International Publishing, 2017, pp. 516–524.

[8] M. Huh, P. Agrawal, and A. A. Efros, "What makes ImageNet good for transfer learning?," *Arxiv.org/abs/1608.08614*, Aug. 2016.

[9] N. Orlov, L. Shamir, T. Macura, J. Johnston, D. M. Eckley, and I. G. Goldberg, "WND-CHARM: Multi-purpose image classification using compound image transforms," *Pattern Recognition Letters*, vol. 29, no. 11, pp. 1684–1693, Aug. 2008, doi: 10.1016/j.patrec.2008.04.013.

[10] J. Kou, J. X. Du, and C. M. Zhai, "Integration of global and local feature for age estimation of facial images," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7390 LNAI, Springer Berlin Heidelberg, 2012, pp. 455–462.

[11] W. Li, D. Zhang, and Z. Xu, "Palmprint identification by Fourier transform," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 16, no. 4, pp. 417–432, Jun. 2002, doi: 10.1142/S0218001402001757.

[12] Y. Su, S. Shan, X. Chen, and W. Gao, "Hierarchical ensemble of global and local classifiers for face recognition," *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1885–1896, Aug. 2009, doi: 10.1109/TIP.2009.2021737.

[13] K. K. Thyagharajan and I. Kiruba Raji, "A review of visual descriptors and classification techniques used in leaf species identification," *Archives of Computational Methods in Engineering*, vol. 26, no. 4, pp. 933–960, May 2019, doi: 10.1007/s11831-018-9266-3.

[14] K. Hameed, D. Chai, and A. Rassau, "A comprehensive review of fruit and vegetable classification techniques," *Image and Vision Computing*, vol. 80, pp. 24–44, Dec. 2018, doi: 10.1016/j.imavis.2018.09.016.

[15] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognition*, vol. 37, no. 1, pp. 1–19, Jan. 2004, doi: 10.1016/j.patcog.2003.07.008.

[16] B. R. Jany, A. Janas, and F. Krok, "Automatic microscopic image analysis by moving window local Fourier transform and machine Learning," *Micron*, vol. 130, Art. no. 102800, Mar. 2020, doi: 10.1016/j.micron.2019.102800.

[17] J. A. Stuchi, L. Boccato, and R. Attux, "Frequency learning for image classification," *Arxiv.org/abs/2006.15476*, Jun. 2020.

[18] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/bf00994018.

[19] W. Chmielnicki and K. Stąpor, "Combining one-versus-one and one-versus-all strategies to improve multiclass SVM classifier," in *Advances in Intelligent Systems and Computing*, vol. 403, Springer International Publishing, 2016, pp. 37–45.

[20] J. Luts, F. Ojeda, R. Van de Plas Raf, B. De Moor, S. Van Huffel, and J. A. K. Suykens, "A tutorial on support vector machine-based methods for classification problems in chemometrics," *Analytica Chimica Acta*, vol. 665, no. 2, pp. 129–145, Apr. 2010, doi: 10.1016/j.aca.2010.03.030.

[21] L. Yann, C. Corinna, and B. Christopher, "The MNIST database of handwritten digits," *The Courant Institute of Mathematical Sciences*. pp. 1–10, 1998.

[22] MATLAB Release 2013b, "The MathWorks Inc.," *Search*. 2013.

[23] A. B. Djimeli-Tsajio, N. Thierry, L. T. Jean-Pierre, T. F. Kapche, and P. Nagabhushan, "Improved detection and identification approach in tomato leaf disease using transformation and combination of transfer learning features," *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 665–674, Apr. 2022, doi: 10.1007/s41348-022-00608-5.

[24] R. A. Rahma, R. A. Nugroho, D. Kartini, M. R. Faisal, and F. Abadi, "Combination of texture feature extraction and forward selection for one-class support vector machine improvement in self-portrait classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 1, pp. 425–434, Feb. 2023, doi: 10.11591/ijece.v13i1.pp425-434.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

BIOGRAPHIES OF AUTHORS

Djimeli-Tsajio Alain Bernard received B.S. (2001) and M.S. with thesis (2004) from the Faculty of Science of the University of Yaoundé 1 and the Ph.D. (2016) from the Faculty of Science of the University of Dschang, all in Cameroon in the field of physics option electronics. Since 2006, he has joined Fotso Victor University Institute of Technology of the University of Dschang as a lecturer in the Department of Telecommunication and Network Engineering. He is a member of URAIA of the present university where he is carrying out research in the field of artificial intelligence for biomedical processes. He can be contacted at email: adjimeli@yahoo.fr.

Noulamo Thierry earned his master's degree, Diploma of Advanced Studies (DEA), and Ph.D. in software engineering science from the University of Yaounde I in 2000, 2001, and 2010, respectively. He is currently working as lecturer in Department of Computer Engineering from Fotso Victor University Institute of Technology (IUT-FV) of Bandjoun, since 2004. He is a member of EANG since 2009, Life member of URAIA since 2008. He has published more than 11 research papers in reputed international journals and conferences. His main research work focuses on process automation using MDE and muti-agent approach. He has 18 years of teaching experience and 11 years of research experience. He can be contacted at email: thierry.noulamo@gmail.com.
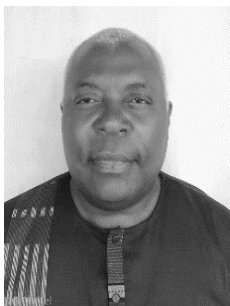
Lienou T. Jean-Pierre Lecturer/Researcher, he obtained his MSc in system engineering in Kiev Polytechnic Institute (National Technical University of Ukraine) and his PhD in University of Yaounde I (Cameroon). Former maintenance engineer at Labotech Medical, he was in charge of medical imaging equipment. He joining the University of Dschang since 2002. He is actually the head of the Department of Computer Engineering, College of Technology, The University of Bamenda. He is a member of EANG since 2018. His research fields are method engineering applied in control systems, multi agent systems applied in power electric systems, cyber resilience and the use of various artificial intelligent techniques in diagnostic of complex systems. He manages a grant related to cyber resilience funded by US ARL. He can be contacted at email: lienou@googlemail.com.
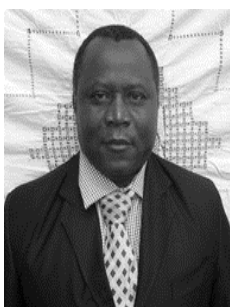
Tchiotsop Daniel was born in 1965 in Tombel, Cameroon. He graduated in electromechanical engineering from "(ENSP)" of Yaoundé-Cameroon in 1990, he obtained a MS degree in Solid Physics in 1992 from the Faculty of Science of the University of Yaoundé I, a MS degree in electrical engineering and Telecommunication in 2003 from ENSP-Yaoundé and a PhD at INPL Nancy-France, in 2007. Dr TCHIOTSOP teaches in the Department of Electrical Engineering of "IUT-FV of Bandjoun" – University of Dschang since 1999 where he is actually the head of Department. His main items of research include biomedical engineering, biomedical signal and image processing, telemedicine and intelligent systems. Dr TCHIOTSOP is partner with "(CRAN) Université de Lorraine", France, "(LETS)-ENSP, University of Yaoundé 1", and "(UR- AIA)" of the IUT-FV of Bandjoun of the University of Dschang, Cameroon. He can be contacted at email: daniel.tchiotsop@gmail.com.

Panduranga Nagabhushan received the B.Eng. degree in electrical. Panduranga Nagabhushan received a Ph.D. degree from the University of Mysore, with a focus on detecting text in compressed document images. He is currently the Vice-Chancellor of Vignan University, Guntur. His research interests include image processing and pattern recognition in compressed images, compressed data analytics, dimensionality reduction integrated classification, and character recognition by dimensionality reduction. He can be contacted at email: pnagabhushan@hotmail.com.