

Curvelet-Based Classification of Brain MRI Images

Rafat Damseh¹(✉) and M. Omair Ahmad²

¹ Institut de génie biomédical, École Polytechnique de Montréal,
Montreal, QC H3T 1J4, Canada
rafat.damseh@polymtl.ca

² Department of Electrical and Computer Engineering, Concordia University,
Montreal, QC H3G 1M8, Canada
omair@encs.concordia.ca

Abstract. Classification of brain MRI images is crucial in medical diagnosis. Automatic classification of these images helps in developing effective non-invasive procedures. In this paper, based on curvelet transform, a novel classification scheme of brain MRI images is proposed and a technique for extracting and selecting curvelet features is provided. To study the effectiveness of their use, the proposed features are employed into three different prediction algorithms, namely, K-nearest neighbours, support vector machine and decision tree. The method of K-fold stratified cross validation is used to assess the efficacy of the proposed classification solutions and the results are compared with those of various state-of-the-art classification schemes available in the literature. The experimental results demonstrate the superiority of the proposed decision tree classification scheme in terms of accuracy, generalization capability, and real-time reliability.

Keywords: MRI imaging · Curvelet transform · Feature extraction and classification

1 Introduction

Magnetic resonance imaging (MRI) is an imaging modality that captures high quality anatomical images of the human body, especially the brain. It provides a huge information that facilitate clinical diagnosis and surgical operations. Brain MR images help in the evaluation of neoplasms and plays a critical role in making decisions regarding initial and evolving treatment strategies. Automatic classification of normal and abnormal brain MR images can make one avoid invasive procedures and anticipate the diagnosis without time-consuming histological examinations.

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and in part by the Regroupement Strategique en Micro-electronique du Quebec (ReSMiQ).

Sparse representation as an approach for feature extraction has been successfully applied to numerous applications, especially in image processing, computer vision and pattern recognition. One well-known simple sparse representation using pre-specified transform functions, namely, discrete wavelet transform (DWT), has been widely used [1–6]. Although the DWTs has an impressive reputation as a good tool for signal processing, it has the drawback of poor directionality, which make its usage limited in many applications. The development of directional wavelets has been investigated in recent years. Ridgelet transform, which is an anisotropic geometric wavelet transform was proposed by Candés and Donoho [7]. To analyze local line or curve singularities, the ridgelet transform is applied to partitions (subsignals) of the original signal. This block ridgelet-based transform is named as curvelet transform (CT) [8]. First-generation curvelet transform is limited because the geometry of ridgelets is unclear. A simpler second-generation curvelet transform based on a frequency partition technique has been proposed later [9]. The second-generation curvelet transform has been shown to be an efficient tool for different applications in signal processing and its discrete version has been proven to be efficient in representing curve-like edges in digital images. Several other developments of directional sparse representations have been proposed with the goal of optimal representation of directional features of signals. However, none of these approaches has reached the publicity of the curvelet transform. In the literature, many classification and detection schemes have exploited the curvelet transform as a tool for extracting meaningful features from signals with applications in medical studies, remote sensing and vehicle verification [10–12].

In this paper, a novel curvelet-based feature extraction scheme for Brain MR images is proposed. Up to the author knowledge, the curvelet coefficients have never been tested as features of the brain MR images. The effectiveness of the proposed feature extraction scheme is examined by combining it with different prediction algorithms and testing on a benchmark dataset consisting of 160 brain MR images to classify normal and abnormal brains. The performance of the proposed classification schemes is compared with that of various state-of-the-art classification schemes proposed in the literature. This paper is organized as follows. In Sect. 2, a brief overview of the related work is presented. In Sect. 3, the proposed mechanism of feature extraction is illustrated. In Sect. 4, experiments are performed. Finally, some concluding remarks highlighting the contribution of the paper and scope of further research are provided in Sect. 5.

2 Related Work

Many recent schemes for classification of normal and abnormal brain MR images have been proposed in the literature. These schemes vary in the pre-processing procedures used for constructing the feature descriptors of the data and the learning algorithm applied to build the classifier. Chaplot, et al. [3] used the approximation coefficients obtained by DWT, and employed the self-organizing map (SOM) neural network and support vector machine (SVM). More promising results have been obtained by El-Dahshan, et al. [4]. In [4], the coefficients

of 3-level DWT are reduced via principal component analysis (PCA), and a hybrid combination of feed-forward back-propagation artificial neural network (FP-ANN) and K-nearest neighbour (KNN) classifiers has been used. Zhang, et al. [5] proposed another scheme for feature extraction and selection, using DWT and PCA, respectively, and implemented forward neural network (FNN) with scaled chaotic artificial bee colony (SCABC) as a classifier. However, superior results have been obtained in [6] by Zhang and Wu using kernel SVMs; the kernels of homogeneous polynomial, inhomogeneous polynomial, and Gaussian radial basis have been suggested. To reduce the storage memory required to obtain the DWT coefficients, Yang et al. [13] proposed to use the wavelet energy (WE) calculated from the wavelet coefficients as a technique for features extraction. Despite the success in reducing the required memory storage, their solution has limited generalization capability without a big improvement in the classification accuracy. To overcome the limited directionality and non-supportiveness to anisotropy of the DWT, which make it not able to capture the subtle and intrinsic details of the brain MRI images, Das et al. [14] developed a scheme for feature extraction using the Ripplet transform Type-I (RT). On the other hand, to have a translation-invariant feature extraction, Wang et al. [15] employed the stationary wavelet transform (SWT) instead of DWT. However, the two lastly mentioned works have had some improvement on the classification accuracy but at the expense of increasing the prediction computational time due to their complex feature extraction scheme.

From the aforementioned classification schemes recently proposed in the literature, it is noticed that there is still a room for further investigations to develop a more powerful classification scheme that can achieve a better trade-off between obtaining a higher accuracy rate, reducing the prediction time and providing better generalization capability. The aim of this work is to develop a novel classification solution that can meet this concern.

3 Proposed Scheme for Feature Extraction

As known, in the process of developing a solution for a classification problem, one should design a mechanism for feature extraction. In this section, The proposed feature extraction scheme using CT is illustrated. It is to be noted that CT is originally defined in the continuous domain but implemented for the discrete domain. Derived from the spatial mother curvelet function, curvelets spatially differ in scale parameter j , rotation parameter $l \in \mathbb{N}_0$, and translation parameter $k \in \mathbb{Z}^2$. CT coefficients of a function $f \in \mathbb{R}^2$ is given as the inner product of the function with the curvelets $\varphi_{j,l,k}$ as in

$$c_{j,l,k} = \langle f, \varphi_{j,l,k} \rangle = \int_{\mathbb{R}^2} f(x) \overline{\varphi_{j,l,k}(x)} dx \quad (1)$$

Discrete curvelet tiling in the frequency domain differs from its continuous counterpart as its ring windowing function is defined as square rings and its angular windowing function is defined as shear functions.

In general, the number of CT coefficients of an image is huge, and using all these coefficients as features of the corresponding image can drastically decrease the classification performance. Hence, in the proposed feature extraction scheme, the use of PCA is proposed for reducing the dimension of the feature vector. The full schematic diagram that describes the proposed feature extraction scheme is depicted in Fig. 1. The process of feature extraction is performed by two steps. First, the normalized five-scale curvelet coefficients of an input MRI image is obtained. Second, PCA is applied on these coefficients to produce the final feature vector. In Fig. 1, the input image is in gray-scale format with the size 256×256 and its pixel values are normalized to $[0, 1]$. As suggested in [16], the five-scale CT is selected based on the following formula

$$\text{number of scales} = \log_2(N) - 3 \tag{2}$$

where N defines the size $N \times N$ of an image. It is to be noted that the number of orientations in the curvelet transform depends on the image size. In our case, there are 4×8 orientations at scales 4 and 5, whereas there are 4×16 orientations at scale 3. In the proposed feature extraction scheme, the CT coefficients are partitioned into subsets and then a first-stage PCA is separately applied on the CT coefficients of each subset. Partitioning is done based on the scale and the orientation of the CT coefficients. Specifically, the CT coefficients at scale S and orientation in $O_i, i = 1, 2, 3, 4$, where $O_1 \in [-\pi/2, \pi/2)$, $O_2 \in [\pi/2, 3\pi/2)$, $O_3 \in [3\pi/2, -3\pi/2)$ and, $O_4 \in [-3\pi/2, -\pi/2)$, are clustered as one partition denoted as $P_{(S,i)}$. The approximation coefficients of CT are clustered in a separate partition denoted as P_0 .

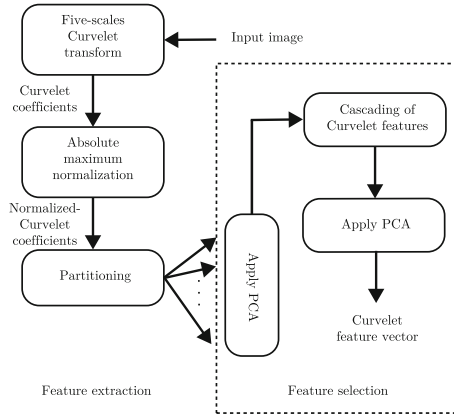


Fig. 1. Proposed curvelet-based feature extraction scheme.

It is to be noted that partitioning of CT coefficients is done in order to reduce the complexity of the PCA computation since it is not feasible to perform PCA on the whole CT coefficients at once. The CT coefficients selected from the

various partitions using the first-stage PCA are then aggregated and a second PCA is applied on them. It is to be mentioned that in each stage of PCA, 80% of the data variance is chosen as a threshold to determine the desired number of features. This threshold is chosen based on experimental observations, through which it is demonstrated that a higher threshold, leading into more features, does not improve the classification performance. Based on the dataset on-hand, the result obtained when applying the first-stage PCA on partition $P_{(3,1)}$ (this partition is arbitrary chosen for explanation purpose only) of the CT coefficients is depicted in Fig. 2. It is seen from this figure that 9 principle components (features) are sufficient to represent 80% of the data variance, however, this result could vary when applying PCA on a different partition. Figure 2 shows the result obtained after applying the second-stage PCA. It is seen from this figure that, for an MRI image of size 256×256 , the final feature vector obtained using the proposed feature extraction scheme contains only 6 features. Having such a small number of features can substantially reduce the computational burden of the classification solution and make it more feasible. In the next section, the effectiveness of utilizing these features to classify different brain MRI images is studied with implementations of various prediction algorithms.

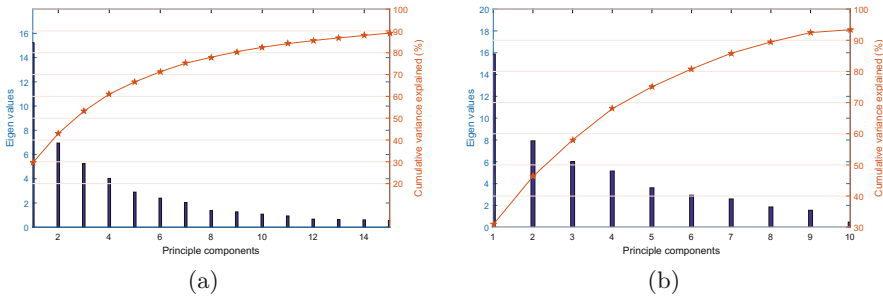


Fig. 2. Results of applying PCA on the 5-scale CT coefficients as described in the proposed feature extraction scheme: (a) first-stage PCA; (b) second-stage PCA.

4 Experiments and Discussion

In this section, extensive experimentations are conducted to study the effectiveness of the proposed feature extraction scheme. Using the proposed features, three prediction algorithms, namely, KNN, SVM with linear and gaussian kernels, and decision tree, are implemented to identify abnormal MRI brains images. These proposed classification schemes will be denoted as CT-KNN, CT-LSVM, CT-GSVM and CT-DT, respectively. The performance of the proposed classification schemes is compared with that of other classification schemes available in the literature, namely, the schemes of DWT-GSVM [3], DWT-PCA-KNN [4], DWT-PCA-GSVM [6], Ripplet-PCA-LSSVM [14] and SWT-PCA-FNN [15]. Experiments are carried out on the basis of averaging the results of 10 executions of 2-, 4-, 5-, 10- and 20-fold cross validation.

4.1 Data Collection and Analysis

The dataset contains 160 T2-weighted MR brain images in axial plain with 256×256 in-plane resolution, which were down-loaded from the website of Harvard Medical School (URL: <http://med.harvard.edu/AANLIB/>), OASIS dataset (URL: <http://www.oasis-brains.org/>), and ADNI dataset (URL: <http://adni.loni.uc-la.edu/>). The abnormal brain MR images in the dataset are of the following diseases: Alzheimer’s disease, meningioma, glioma, Pick’s disease, sarcoma, Huntington, and Alzheimer’s disease with visual agnosia. Samples of the various disease are shown in Fig. 3. In the dataset, 20 images capture normal brains and the rest capture various brain abnormalities (20 images/each disease). The statistical characteristics and cross validation settings of the dataset are listed in Table 1.

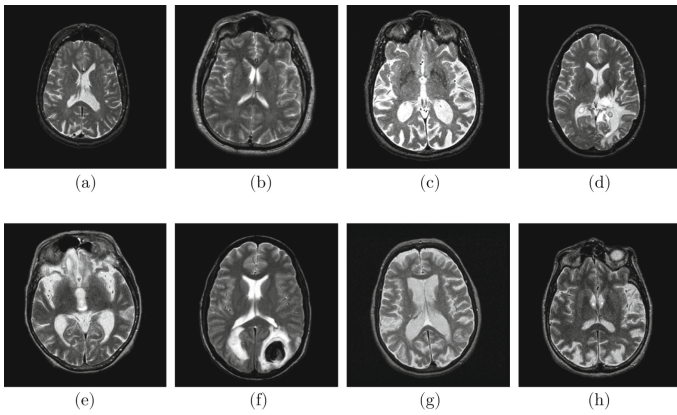


Fig. 3. MRI brain dataset consisting of images representing normal and 7 types of different ubnormal brains: (a) Normal; (b) Alzheimer; (c) Meningioma; (d) Glioma; (e) Pick’s disease; (f) Sarcoma; (g) Huntington’s disease; (h) Alzheimer with visual agnosia.

Table 1. Cross-validation with various K-fold settings.

K-fold	Total		Training		Validation		Run
	Normal	Abnormal	Normal	Abnormal	Normal	Abnormal	
2	20	140	10	70	10	70	10
4	20	140	15	105	5	35	10
5	20	140	16	112	4	28	10
10	20	140	18	126	2	14	10
20	20	140	19	133	1	7	10

4.2 Classification Performance

Now, after designing the experimental procedure, a comparison between the various classification schemes is carried out in terms of their accuracy rates. The comparison results are shown in Table 2. Some entries of the table are extracted from the corresponding original papers. It is clear from the table that, when the proposed CT-DT scheme is used, the highest average accuracy rate is achieved. The second-best accuracy rate belongs also to the proposed CT-GSVM scheme. Since the experiments have been carried out using different K-fold settings, the average accuracy rates obtained in Table 2 provide a strong indication about the generalization capabilities of the various classification schemes. Thus, the performance of the proposed CT-DT scheme is expected to be superior to those of the other various schemes when predicting new brain MRI datasets.

Table 2. Accuracy rates of the various classification schemes.

Classification schemes	Number of Features	Data partitioning for cross validation					Average accuracy (%)
		2-fold	4-fold	5-fold	10-fold	20-fold	
DWT-GSVM [3]	4761	97.23	97.48	97.33	97.13	96.9	97.214
DWT-PCA-KNN [4]	7	97.77	97.47	97.54	97.73	97.71	97.644
DWT-PCA-GSVM [6]	19	98.52	98.96	99.38	99.07	98.13	98.812
Ripplet-PCA-LSSVM [14]	9	99.08	99.89	100	99.21	99.24	99.484
SWT-PCA-FNN [15]	7	99.46	100	100	99.68	99.18	99.664
CT-KNN	6	98.93	98.42	98.37	98.36	98.13	98.442
CT-LSVM	6	98.35	98.9	99.25	99.1	99.07	98.934
CT-GSVM	6	99.91	99.33	100	99.76	99.77	99.754
CT-DT	6	100	100	100	99.83	100	99.966

Since the aim of this paper is to develop a computer-aided diagnosis tool for identifying abnormal brains, the development of such tool implies that it be reliable in real-time environments. Thus, studying the computational-time complexity of the proposed classification scheme is very important. It is known that the classification solution for an input is obtained by firstly training the classifier on a dataset on-hand. Since the training process can be performed off-line, the computational time associated with this process is not a significant factor in assessing the performance of the proposed classification scheme. Hence, only the prediction time-complexity of the proposed classification scheme is studied. By performing the prediction process on all the images in the dataset using MATLAB implementation on a 2.9-GHz 8-GB machine, the average time of prediction for one image is recorded as 0.053 s. This time includes the time of feature extraction (0.049 s), feature selection (0.003 s) and classification decision (0.001 s). It is obvious that the proposed classification scheme can successfully meet the time requirement for real-time diagnosis. Despite that the time-complexity for feature extraction in the proposed CT-DT scheme is slightly greater than that of the

other various schemes, the use of CT-DT which employs only 6 dimensional feature vector results in lower storage cost. Also, from Table 2, it is proven that the use of this short features vector provides a positive impact on the generalization capability of the classifier.

5 Conclusion

Manual classification of brain MRI images is time consuming and not reliable in some cases. Therefore, the development of automatic diagnostic tools to identify normal and abnormal brains in such images is necessary. Many schemes for automatic classification of normal and abnormal brain MRI images have been proposed in the literature. However, These schemes have limitation in their accuracy or generalization capability. In this paper, a novel scheme for feature extraction using the curvelet transform (CT) has been proposed, and the discrimination ability of the extracted features has been explored by employing them into three prediction algorithms, namely, K-nearest neighbours (KNN), support vector machine (SVM) and decision tree. In order to extract the curvelet features from an MRI image, five-scale curvelet transform has been applied and followed by two stages of principle component analysis (PCA). By applying the proposed feature extraction scheme on MRI images of size 256×256 , the size of the classification problem has been successfully reduced to include only 6 dimensions. The method of K-fold stratified cross validation has been used to assess the performance of the proposed classification schemes. By comparing with other state-of-the-art classification schemes available in the literature, the experimental results have demonstrated the superiority of the proposed DT classification scheme in terms of accuracy and generalization capability, and have proven the reliability of the proposed scheme in real-time environments. One can further extend the research work undertaken in this paper by studying the classification performance of the proposed scheme on other types of datasets that include different MRI modalities, such as T1-weighted, Proton-density weighted, and diffusion weighted images. Also, the behaviour of the proposed scheme could be examined on noisy datasets with different noise characteristics; this examination can provide more insight about the robustness of the proposed classification scheme.

Acknowledgments. The authors would like to thank Yudong Zhang for providing a portion of the MRI dataset.

References

1. Yao, J., Chen, J., Chow, C.: Breast tumor analysis in dynamic contrast enhanced MRI using texture features and wavelet transform. *IEEE J. Sel. Top. Sig. Process.* **3**(1), 94–100 (2009)
2. Nanthagopal, A.P., Sukanesh, R.: Wavelet statistical texture features-based segmentation and classification of brain computed tomography images. *IET Image Process.* **7**(1), 25–32 (2013)

3. Chaplot, S., Patnaik, L., Jagannathan, N.: Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomed. Sig. Process. Control* **1**(1), 86–92 (2006)
4. El-Dahshan, E.S.A., Hosny, T., Salem, A.B.M.: Hybrid intelligent techniques for MRI brain images classification. *Digit. Sig. Proc.* **20**(2), 433–441 (2010)
5. Zhang, Y., Wang, S., Wu, L.: A novel method for magnetic resonance brain image classification based on adaptive chaotic PSO. *Prog. Electromagnet. Res.* **109**, 325–343 (2010)
6. Zhang, Y., Wu, L.: An MR brain images classifier via principal component analysis and kernel support vector machine. *Prog. Electromagnet. Res.* **130**, 369–388 (2012)
7. Candès, E.J., Donoho, D.L.: Ridgelets: a key to higher-dimensional intermittency? *Philos. Trans. R. Soc. Lond. A: Math. Phys. Eng. Sci.* **357**(1760), 2495–2509 (1999)
8. Candès, E.J., Donoho, D.L.: Curvelets: a surprisingly effective nonadaptive representation for objects with edges. Technical report, DTIC Document (2000)
9. Candès, E.J., Donoho, D.L.: Continuous curvelet transform: II. discretization and frames. *Appl. Comput. Harmonic Anal.* **19**(2), 198–222 (2005)
10. Bekker, A.J., Shalhon, M., Greenspan, H., Goldberger, J.: Multi-view probabilistic classification of breast microcalcifications. *IEEE Trans. Med. Imaging* **35**(2), 645–653 (2016)
11. Uslu, E., Albayrak, S.: Curvelet-based synthetic aperture radar image classification. *IEEE Geosci. Remote Sens. Lett.* **11**(6), 1071–1075 (2014)
12. Guo, J.M., Prasetyo, H., Farfoura, M.E., Lee, H.: Vehicle verification using features from curvelet transform and generalized Gaussian distribution modeling. *IEEE Trans. Intell. Transp. Syst.* **16**(4), 1989–1998 (2015)
13. Yang, G., Zhang, Y., Yang, J., Ji, G., Dong, Z., Wang, S., Feng, C., Wang, Q.: Automated classification of brain images using wavelet-energy and biogeography-based optimization. *Multimedia Tools Appl.* **75**(23), 15601–15617 (2015)
14. Das, S., Chowdhury, M., Kundu, M.K.: Brain MR image classification using multiscale geometric analysis of ripplelet. *Prog. In Electromagn. Res.* **137**, 1–17 (2013)
15. Wang, S., Zhang, Y., Dong, Z., Du, S., Ji, G., Yan, J., Yang, J., Wang, Q., Feng, C., Phillips, P.: Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. *Int. J. Imaging Syst. Technol.* **25**(2), 153–164 (2015)
16. Candès, E., Demanet, L., Donoho, D., Ying, L.: Curvelab toolbox, version 2.0. CIT (2005)