



HHS Public Access

Author manuscript

Med Image Comput Comput Assist Interv. Author manuscript; available in PMC 2019 December 04.

Published in final edited form as:

Med Image Comput Comput Assist Interv. 2014 ; 17(Pt 3): 401–408. doi:10.1007/978-3-319-10443-0_51.

A Novel Multi-relation Regularization Method for Regression and Classification in AD Diagnosis

Xiaofeng Zhu, Heung-Il Suk, Dinggang Shen*

Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, USA

Abstract

In this paper, we consider the joint regression and classification in Alzheimer's disease diagnosis and propose a novel multi-relation regularization method that exploits the relational information inherent in the observations and then combines it with an $\ell_{2,1}$ -norm within a least square regression framework for feature selection. Specifically, we use three kinds of relationships: feature-feature relation, response-response relation, and sample-sample relation. By imposing these three relational characteristics along with the $\ell_{2,1}$ -norm on the weight coefficients, we formulate a new objective function. After feature selection based on the optimal weight coefficients, we train two support vector regression models to predict the clinical scores of Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) and Mini-Mental State Examination (MMSE), respectively, and a support vector classification model to identify the clinical label. We conducted clinical score prediction and disease status identification jointly on the Alzheimer's Disease Neuroimaging Initiative dataset. The experimental results showed that the proposed regularization method outperforms the state-of-the-art methods, in the metrics of correlation coefficient and root mean squared error in regression and classification accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve in classification.

Keywords

Alzheimer's disease; feature selection; sparse coding; manifold learning; MCI conversion

1 Introduction

For the computer-aided Alzheimer's Disease (AD) or Mild Cognitive Impairment (MCI) diagnosis, the available sample size is usually small, but the feature dimension is high. For example, the sample size used in [7,21] was less than one hundred, while the feature dimension (including both Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) features) was hundreds or even thousands. The small sample size makes it difficult to build an effective model, and the high dimensionality of data leads to an overfitting problem. To this end, researchers mostly predefined the disease-related features and used such low-dimensional features in clinical label identification or clinical score prediction.

*Corresponding author. dgshen@med.unc.edu.

In the meantime, recent studies have shown that the feature selection helps overcome both problems of high dimensionality and small sample size, by removing uninformative features [14,16,13,19,20,18]. Moreover, among various feature selection techniques, manifold learning has been successfully used in either regression or classification [9,13,12,17]. For example, Cho *et al.* adopted a manifold harmonic transformation method on the cortical thickness data and achieved a sensitivity of 63% and a specificity of 76% on the dataset with 72 MCI Converters (MCI-C) and 131 MCI Non-Converters (MCI-NC) [3]. While most of the previous studies focused on identifying brain disease and estimating clinical scores separately [4], there have been also efforts to select joint features that could be used for both tasks simultaneously. For example, Zhang and Shen proposed a multi-task sparse feature selection method for joint disease status identification and clinical scores prediction, and showed that such combination can achieve better performance than performing them separately [15,21].

In line with Zhang and Shen's work, in this paper, we consider the prediction of both clinical scores and disease status jointly in a unified framework, as in [7,9]. However, unlike the previous manifold-based feature selection methods that considered only the manifold of the samples, but not manifold of either the features or the response variables, we propose a novel multi-relation regularization method. Specifically, we use the relational information inherent in the observations and combine it with an $\ell_{2,1}$ -norm within a least square regression framework. The rationale for the proposed multi-relation regularization method is as follows: (1) If some features are related to each other, then the same or similar relation is expected to be preserved between the respective weight coefficients in a least square regression model. (2) Due to the algebraic operation in least square regression, *i.e.*, matrix multiplication, the weight coefficients are linked to the response variables via regressors, *i.e.*, feature vectors in our work. Therefore, it is natural to impose the relation between a pair of weight coefficients to be similar to the relation between the corresponding pair of target response variables. (3) As considered in many manifold learning methods [1,6,17], if a pair of samples are similar to each other, then their respective response values should be also similar to each other. By imposing these three relational characteristics along with the $\ell_{2,1}$ -norm on the weight coefficients, we formulate a new objective function. We then select features to build classification and regression models for clinical label identification and clinical scores (Alzheimer's Disease Assessment Scale-Cognitive subscale: ADAS-Cog, Mini-Mental State Examination: MMSE) prediction, respectively.

2 Method

By taking the features as regressors and the concatenation of clinical scores (*e.g.*, ADAS-Cog, MMSE) and a class label as responses, we apply the proposed method to select features that are jointly used to represent clinical scores and class labels. Based on the selected features, we finally build clinical scores regression models and a clinical label identification model with Support Vector Regression (SVR) and Support Vector Classification (SVC), respectively.

Let $\mathbf{X} \in \mathbb{R}^{n \times d}$ and $\mathbf{Y} \in \mathbb{R}^{n \times c}$ denote d neuroimaging features and c clinical response values of n subjects or samples, respectively. In this work, we assume that the response values of

clinical scores and a clinical label can be represented by a linear combination of the features. Then, the problems of regressing clinical scores and identifying a class label can be formulated by a least square regression model as follows:

$$\mathcal{L}(\mathbf{W}) = \|\mathbf{Y} - \mathbf{X}\mathbf{W}\|_F^2 = \|\mathbf{Y} - \widehat{\mathbf{Y}}\|_F^2 = \sum_{i=1}^n \sum_{j=1}^c (y_{ij} - \hat{y}_{ij})^2 \quad (1)$$

where $\|\cdot\|_F$ denotes a Frobenius norm, $\mathbf{W} \in \mathbb{R}^{d \times c}$ is a weight coefficient matrix, and $\widehat{\mathbf{Y}} = \mathbf{X}\mathbf{W}$. While the least square regression model has been successfully used in many fields, it is well known that the solution is generally overfitted to the training samples in its original form. To overcome the overfitting problem and to find a more generalized solution, a variety of its variants using different types of regularizations have been proposed [5], which can be mathematically simplified as follows:

$$\min_{\mathbf{W}} \mathcal{L}(\mathbf{W}) + \mathcal{R}(\mathbf{W}) \quad (2)$$

where $\mathcal{R}(\mathbf{W})$ denotes a set of regularization terms.

From a machine learning point of view, a well-defined regularization term helps find a generalized solution to the objective function, and thus result in a better performance for the final goal. In this paper, we devise novel regularization terms that effectively utilize various pieces of information inherent in the observations. Note that since, in this work, we extract features from the parcellated brain areas or Regions-Of-Interest (ROIs), which are structurally or functionally related to each other, it is natural to assume that there exist relations among features. Meanwhile, if two features are highly related to each other, then it is reasonable to have the respective weight coefficients also related. However, to the best of our knowledge, none of the previous regression methods in the literature considered and guaranteed this characteristic in their solutions. To this end, we devise a new regularization term with the claim that, if some features are related to each other, the same or the similar relation is expected to be preserved between the respective weight coefficients. To utilize this 'feature-feature' relation, we impose the relation between columns in \mathbf{X} to be reflected in the relation between the corresponding rows in \mathbf{W} , by defining the following regularization term:

$$\mathcal{R}_1(\mathbf{W}) = \frac{1}{2} \sum_{i,j} m_{ij} \|\mathbf{w}^i - \mathbf{w}^j\|_2^2 \quad (3)$$

where m_{ij} denotes an element in the feature similarity matrix $\mathbf{M} = [m_{ij}] \in \mathbb{R}^{d \times d}$ that encodes the relation between features in the samples. Throughout this paper, we use a radial basis function kernel to measure the similarity between vectors.

Meanwhile, given a feature vector \mathbf{x}^i , in our joint regression and classification framework, we use a different set of weight coefficients to regress the elements in the response vector \mathbf{y}^i . In other words, the elements of each column in \mathbf{W} are linked to the elements of each column in \mathbf{Y} via feature vectors. By taking this mathematical property into account, we further

impose the relation between column vectors in \mathbf{W} to be similar to the relation between the respective target response variables (i.e., respective column vectors) in \mathbf{Y} , which we call as ‘*response-response*’ relation:

$$\mathcal{R}_2(\mathbf{W}) = \frac{1}{2} \sum_{i,j}^c g_{ij} \|\mathbf{w}_i - \mathbf{w}_j\|_2^2 \quad (4)$$

where g_{ij} denotes an element in the matrix $\mathbf{G} = [g_{ij}] \in \mathbb{R}^{c \times c}$ that represents the similarity between every pair of target response variables (i.e., every pair of column vectors in \mathbf{Y}). Due to the algebraic operation in the least square regression, i.e., matrix multiplication, the weight coefficients are linked to the response variables via regressors, i.e., feature vectors in our work. Therefore, it is meaningful to impose the relation between a pair of weight coefficients to be similar to the relation between the respective pair of target response variables.

We can also utilize the relational information between samples, called as ‘*sample-sample*’ relation. That is, if samples are similar to each other, then their respective response values should be also similar to each other. In this regard, we define a regularization term as follows:

$$\mathcal{R}_3(\mathbf{W}) = \frac{1}{2} \sum_{i,j}^n s_{ij} \|\hat{y}^i - \hat{y}^j\|_2^2 = \frac{1}{2} \sum_{i,j}^n s_{ij} \|\mathbf{x}^i \mathbf{W} - \mathbf{x}^j \mathbf{W}\|_2^2 \quad (5)$$

where s_{ij} is an element in the matrix $\mathbf{S} = [s_{ij}] \in \mathbb{R}^{n \times n}$ that measures the similarity between every pair of samples. We should note that this kind of sample-sample relation has been successfully used in many manifold learning methods [1,6]. We argue that the simultaneous consideration of these newly devised regularization terms, i.e., feature-feature relation, sample-sample relation, and response-response relation, can effectively reflect the relational information inherent in observations in finding an optimal solution.

Regarding feature selection, we believe that due to the underlying brain mechanisms that determine clinical scores or a clinical label, i.e., response variables, if one feature plays a role in predicting one response variable, then it also devotes to the prediction of the other response variables. So, we further impose to use the same features across the tasks of clinical scores and clinical label prediction. Mathematically, this can be formulated by a $\mathcal{L}_{2,1}$ -norm on \mathbf{W} , i.e., $\|\mathbf{W}\|_{2,1}$.

Therefore, our final objective function is formulated as follows:

$$\min_{\mathbf{W}} \mathcal{L}(\mathbf{W}) + \alpha_1 \mathcal{R}_1(\mathbf{W}) + \alpha_2 \mathcal{R}_2(\mathbf{W}) + \alpha_3 \mathcal{R}_3(\mathbf{W}) + \lambda \|\mathbf{W}\|_{2,1} \quad (6)$$

where α_1 , α_2 , α_3 , and λ denote control parameters of the respective regularization terms, respectively. This objective function can be efficiently optimized using the framework in [22].

It is noteworthy that unlike the previous regularization methods such as local linear embedding [10], locality preserving projection [6], predictive space aggregated regression [2], and high-order graph matching [9] that focused on the sample similarities by imposing nearby samples to be still nearby in the transformed space, the proposed method utilizes richer information inherent in the observations. Thus, it is expected that the proposed method can find a generalized solution, which can be robust to noise or outliers.

3 Experimental Analysis

We compared the performance of the proposed method and the state-of-the-art methods on a subset of the ADNI dataset. Our dataset has 202 subjects including 51 AD, 52 NC, and 99 MCI. Moreover, 99 MCI contains 43 MCI-C and 56 MCI-NC.

3.1 Image Processing and Feature Extraction

We conducted image pre-processing for MRI and PET images by the sequential application of spatial distortion, skull-stripping, and removal of cerebellum. Then, for structural MRI images, we segmented them into three different tissues: gray matter (GM), white matter (WM), and CSF. By warping Kabani *et al.*'s atlas [8] into a subject's MRI image, we further dissected the GM tissue into 93 ROIs by HAMMER[11]. We then regarded the volume of the GM tissue of each ROI as a feature. We aligned each PET image to its corresponding MRI image, and then took the average intensity of each ROI as a feature. Thus, we extracted 93 features from MR and PET images, respectively.

3.2. Experimental Setting

We considered three binary classification problems: AD vs. NC, MCI vs. NC, and MCI-C vs. MCI-NC. For MCI vs. NC, both MCI-C and MCI-NC were labeled as MCI. For each set of experiments, we used 93 MRI features or 93 PET features as regressors, and 2 clinical scores along with 1 class label as responses in the least square regression model. We employed the metrics of Correlation Coefficient (CC) and Root Mean Squared Error (RMSE) between the target clinical scores and the predicted ones in regression, and also the metrics of classification ACCuracy (ACC), SENSitivity (SEN), SPEcificity (SPE), and Area Under the receiver operating characteristic Curve (AUC) in classification.

To validate the effectiveness of the proposed method, we considered rigorous experimental conditions: (1) In order to show the validity of the feature selection strategy, we performed the tasks of regression and classification without precedent feature selection, and considered them as a baseline method. Hereafter, we use the suffix "N" to indicate that no feature selection was involved in. For example, by MRI-N, we mean that either the classification or regression was performed using the full MRI features. (2) One of the main arguments in our work is to select features that can be jointly used for both regression and classification. To this end, we compare the multi-task based method with a single-task based method, in which the feature selection was carried out for regression and classification independently. In the following, the suffix "S" manifests a single-task based method. For example, MRI-S represents single-task based feature selection on MRI features. (3) We compare with two state-of-the-art methods: High-Order Graph Matching (HOGM) [9] and Multi-Modal Multi-

Task (M3T) [15]. The former used a sample-sample relation along with an ℓ_1 -norm in an optimization of single-task learning. The latter used multi-task learning with an $\ell_{2,1}$ -norm only to select a common set of features for the tasks of regression and classification.

3.3 Classification Results

Table 1 shows the classification performances of all the competing methods. From these results, we can draw three conclusions. First, it is important to conduct feature selection on the high-dimensional features before training a classifier since the baseline methods with no feature selection, *i.e.*, MRI-N, and PET-N, reported the worst performances. Second, it is beneficial to use joint regression and classification framework, *i.e.*, multi-task learning, for feature selection. As shown in Table 1, M3T and our method, which utilized the multi-task learning, achieved better classification performances than the single-task based method. Specifically, the proposed method showed the superiority to the single-task based method, *i.e.*, MRI-S and PET-S, improving the accuracies by 2.5% (AD vs. NC), 3.0% (MCI vs. NC), and 7.3% (MCI-C vs. MCI-NC) with MRI, and by 3.9% (AD vs. NC), 10.2% (MCI vs. NC), and 9.0% (MCI-C vs. MCI-NC) with PET, respectively. Lastly, based on the fact that the best performances over the three binary classifications were all obtained by our method, we can say that the proposed regularization terms were effective to find class-discriminative features. It is worth noting that compared to the state-of-the-art methods, the accuracy enhancements by our method were 5% (vs. HOGM) and 4.7% (vs. M3T) with MRI, and 4.6% (vs. HOGM) and 4.2% (vs. M3T) with PET for MCI-C vs. MCI-NC classification, which is the most important for early diagnosis and treatment.

3.4 Regression Results

Regarding the prediction of two clinical scores of MMSE and ADAS-Cog, we summarized the results in Table 2, we can see that, similar to the classification results, the regression performance of the methods without feature selection (MRI-N and PET-N) was worse than any of the other methods with feature selection. Moreover, our method consistently outperformed the competing methods for the cases of different pairs of clinical labels.

In the regression with MRI for AD vs. NC, our method showed the best CCs of 0.669 for ADAS-Cog and 0.679 for MMSE, and the best RMSEs of 4.43 for ADAS-Cog and 1.79 for MMSE. The next best performances in terms of CCs were obtained by M3T, *i.e.*, 0.649 for ADAS-Cog and 0.638 for MMSE, and those in terms of RMSEs were obtained by HOGM, *i.e.*, 4.53 for ADAS-Cog and 1.91 for MMSE. In the regression with MRI for MCI vs. NC, our method also achieved the best CCs of 0.472 for ADAS-Cog and 0.50 for MMSE, and the best RMSEs of 4.23 for ADAS-Cog and 1.62 for MMSE. For the case of MCI-C vs. MCI-NC with MRI, the proposed method improved the CCs by 0.092 for ADAS-Cog and 0.053 for MMSE compared to the next best CCs of 0.497 for ADAS-Cog and 0.550 for MMSE by M3T. Note that the proposed method with PET also reported the best CCs and RMSEs for both ADAS-Cog and MMSE over the three regression problems, *i.e.*, AD vs. NC, MCI vs. NC, and MCI-C vs. MCI-NC.

4 Conclusions

In this work, we proposed a novel feature selection method by devising new regularization terms that consider relational information inherent in the observations for joint regression and classification in the computer-aided AD diagnosis. From our extensive experiments on the ADNI dataset, we found that the utilization of the devised three regularization terms, *i.e.*, sample-sample relation, feature-feature relation, and response-response relation, were helpful to improve the performances in the problem of joint regression and classification, outperforming the state-of-the-art methods.

Acknowledgements.

This study was supported by National Institutes of Health (EB006733, EB008374, EB009634, AG041721, AG042599, and MH100217). Xiaofeng Zhu was partly supported by the National Natural Science Foundation of China under grant 61263035.

References

1. Belkin M, Niyogi P, Sindhvani V: Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of Machine Learning Research* 7, 2399–2434 (2006)
2. Chen T, Kumar R, Troianowski GA, Syeda-Mahmood TF, Beymer D, Brannon K: Psar: Predictive space aggregated regression and its application in valvular heart disease classification. In: *ISBI*, pp. 1122–1125 (2013)
3. Cho Y, Seong JK, Jeong Y, Shin SY: Individual subject classification for Alzheimer’s disease based on incremental learning using a spatial frequency representation of cortical thickness data. *NeuroImage* 59(3), 2217–2230 (2012) [PubMed: 22008371]
4. Duchesne S, Caroli A, Geroldi C, Collins DL, Frisoni GB: Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *NeuroImage* 47(4), 1363–1370 (2009) [PubMed: 19371783]
5. Hastie T, Tibshirani R, Friedman J, Franklin J: The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer* 27(2), 83–85 (2005)
6. He X, Cai D, Niyogi P: Laplacian score for feature selection. In: *NIPS*, pp. 1–8 (2005)
7. Jie B, Zhang D, Cheng B, Shen D: Manifold regularized multi-task feature selection for multi-modality classification in Alzheimers disease In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N (eds.) *MICCAI 2013, Part I. LNCS*, vol. 8149, pp. 275–283. Springer, Heidelberg (2013)
8. Kabani NJ: 3D anatomical atlas of the human brain. *NeuroImage* 7, 0700–0717 (1998)
9. Liu F, Suk H-I, Wee C-Y, Chen H, Shen D: High-order graph matching based feature selection for Alzheimer’s disease identification In: Mori K, Sakuma I, Sato Y, Baril-lot C, Navab N (eds.) *MICCAI 2013, Part II. LNCS*, vol. 8150, pp. 311–318. Springer, Heidelberg (2013)
10. Roweis ST, Saul LK: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326 (2000) [PubMed: 11125150]
11. Shen D, Davatzikos C: HAMMER: hierarchical attribute matching mechanism for elastic registration. *IEEE Transactions on Medical Imaging* 21(11), 1421–1439 (2002) [PubMed: 12575879]
12. Suk HI, Lee SW: A novel Bayesian framework for discriminative feature extraction in brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(2), 286–299 (2013) [PubMed: 22431526]
13. Suk H-I, Shen D: Deep learning-based feature representation for AD/MCI classification In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N (eds.) *MICCAI 2013, Part II. LNCS*, vol. 8150, pp. 583–590. Springer, Heidelberg (2013)
14. Wee CY, Yap PT, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D: Identification of MCI individuals using structural and functional connectivity networks. *Neuroimage* 59(3), 2045–2056 (2012) [PubMed: 22019883]

15. Zhang D, Shen D: Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage* 59(2), 895–907 (2012) [PubMed: 21992749]
16. Zhang D, Wang Y, Zhou L, Yuan H, Shen D: Multimodal classification of Alzheimer's disease and mild cognitive impairment. *NeuroImage* 55(3), 856–867 (2011) [PubMed: 21236349]
17. Zhu X, Huang Z, Cheng H, Cui J, Shen HT: Sparse hashing for fast multimedia search. *ACM Transactions on Information Systems* 31(2), 9 (2013)
18. Zhu X, Huang Z, Cui J, Shen T: Video-to-shot tag propagation by graph sparse group lasso. *IEEE Transactions on Multimedia* 13(3), 633–646 (2013)
19. Zhu X, Huang Z, Shen HT, Cheng J, Xu C: Dimensionality reduction by mixed kernel canonical correlation analysis. *Pattern Recognition* 45(8), 3003–3016 (2012)
20. Zhu X, Huang Z, Yang Y, Shen HT, Xu C, Luo J: Self-taught dimensionality reduction on the high-dimensional small-sized data. *Pattern Recognition* 46(1), 215–229 (2013)
21. Zhu X, Suk HI, Shen D: Matrix-similarity based loss function and feature selection for Alzheimer's disease diagnosis. In: *CVPR* (2014)
22. Zhu X, Wu X, Ding W, Zhang S: Feature selection by joint graph sparse coding. In: *SDM*, pp. 803–811 (2013)

Table 1.

Comparison of classification performances (%) of the competing methods

Feature	Method	<u>AD vs. NC</u>				<u>MCI vs. NC</u>				<u>MCI-C vs. MCI-NC</u>			
		ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
MRI	MRI-N	89.5	82.7	86.3	95.3	68.3	92.6	39.2	82.5	60.3	15.5	92.3	68.7
	MRI-S	91.2	85.9	92.5	96.7	76.7	93.3	37.6	83.7	64.5	24.9	95.8	70.6
	HOGM	93.4	89.5	92.5	97.1	77.7	95.6	51.4	84.4	66.8	36.7	95.0	72.2
	M3T	92.6	87.2	95.9	97.5	78.1	94.5	54.0	83.1	67.1	37.7	92.0	72.5
	Proposed	93.7	88.6	97.8	97.6	79.7	94.8	56.9	84.7	71.8	48.0	92.8	81.4
PET	PET-N	86.2	83.5	84.8	94.8	69.0	95.0	30.8	77.9	62.2	21.6	93.1	71.3
	PET-S	87.9	85.7	90.9	94.7	73.8	96.5	36.2	78.7	65.1	31.0	95.5	73.5
	HOGM	91.7	91.1	92.8	95.6	74.7	96.5	43.2	79.3	66.6	35.5	95.5	72.4
	M3T	90.9	90.5	93.1	96.4	77.2	94.5	44.3	80.5	67.0	39.1	93.2	73.1
	Proposed	91.8	91.5	93.8	96.9	79.2	97.1	45.3	80.8	71.2	47.4	93.0	77.6

Table 2.

Comparison of regression performances of the competing methods in terms of Correlation Coefficient (CC) and Root Mean Square Error (RMSE)

Feature	Method	AD vs. NC				MCI vs. NC				MCI-C vs. MCI-NC			
		ADAS-Cog		MMSE		ADAS-Cog		MMSE		ADAS-Cog		MMSE	
		CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE	CC	RMSE
MRI	MRI-N	0.587	4.96	0.520	2.02	0.329	4.48	0.309	1.90	0.420	4.10	0.441	1.51
	MRI-S	0.591	4.85	0.566	1.95	0.347	4.27	0.367	1.64	0.426	4.01	0.482	1.44
	HOGM	0.625	4.53	0.598	1.91	0.352	4.26	0.371	1.63	0.435	3.94	0.521	1.41
	M3T	0.649	4.60	0.638	1.91	0.445	4.27	0.420	1.66	0.497	4.01	0.550	1.41
	Proposed	0.669	4.43	0.679	1.79	0.472	4.23	0.500	1.62	0.589	3.83	0.603	1.40
PET	PET-N	0.597	4.86	0.514	2.04	0.333	4.34	0.331	1.70	0.382	4.08	0.452	1.50
	PET-S	0.620	4.83	0.593	2.00	0.356	4.26	0.359	1.69	0.437	4.00	0.478	1.48
	HOGM	0.600	4.69	0.515	1.99	0.360	4.21	0.368	1.67	0.430	4.03	0.523	1.41
	M3T	0.647	4.67	0.593	1.92	0.447	4.24	0.432	1.68	0.520	3.91	0.569	1.45
	Proposed	0.671	4.41	0.620	1.90	0.513	4.13	0.485	1.66	0.526	3.87	0.570	1.37