

A Fast 3D CNN for Hyperspectral Image Classification

Muhammad Ahmad,

Abstract—Hyperspectral imaging (HSI) has been extensively utilized for a number of real-world applications. HSI classification (HSIC) is a challenging task due to high inter-class similarity, high intra-class variability, overlapping, and nested regions. A 2D Convolutional Neural Network (CNN) is a viable approach whereby HSIC highly depends on both Spectral-Spatial information, therefore, 3D CNN can be an alternative but highly computational complex due to the volume and spectral dimensions. Furthermore, these models do not extract quality feature maps and may underperform over the regions having similar textures. Therefore, this work proposed a 3D CNN model that utilizes both spatial-spectral feature maps to attain good performance. In order to achieve the said performance, the HSI cube is first divided into small overlapping 3D patches. Later these patches are processed to generate 3D feature maps using a 3D kernel function over multiple contiguous bands which preserve the spectral information as well. Benchmark HSI datasets (Pavia University, Salinas and Indian Pines) are considered to validate the performance of our proposed method. The results are further compared with several state-of-the-art methods.

Index Terms—3D Convolutional Neural Network (CNN); Kernel Function; Classification; Hyperspectral Images (HSI);

I. INTRODUCTION

HYPERSPECTRAL SENSOR collects the information (reflectance) in several hundreds of contiguous bands with a very high spectral resolution which enables us to classify the objects based on their spectral signatures. However, these images are in relatively low spatial resolution due to the sensor limitations, SNR, and complexity constraints which significantly affect the performance for several real-world applications [1]. The traditional classifiers, for instance, KNN [2], SVM [3], Maximum Likelihood [4], Logistic Regression [2] and Extreme Learning Machine (ELM) [5] are only works based on spectral information. These classifiers do not perform well due to spectral redundancy and high correlation among the spectral bands. Furthermore, these classifiers fail to preserve the important spatial variability of Hyperspectral data which also results in low performance.

The simplest way to improve the classification performance is to design a classifier that should incorporate both spectral and spatial information. Spatial information is considered as additional discriminative information associated with the size, shape, and structure of the object which if provided correctly brings more competitive results. Spatial spectral classifiers can generally be classified into two groups. First category explores the spatial and spectral information separately. The

spatial information is extracted in advance using entropy [6], morphological operations [7], [8], low rank representation [9] and attribute profiles [10]. Later this information is combined with spectral information to perform pixel-level classification.

The second category fuses the spatial-spectral information to get the joint features [11], for instance, 3D wavelet, scattering wavelet and Gabor filter [12], [13] are generated at different frequencies and scales to extract the joint spatial-spectral features for classification. Hyperspectral images are in 3D cubes thus the former category results in several 3D features i.e., spatial-spectra feature cubes comprising key information, thus preserving joint spatial-spectral correlations while extracting features can produce better results. However, the classical feature extraction methods are based on shallow learning and handcrafted features which largely depend on domain knowledge [14]. Therefore, the Deep models have been proposed to address the aforementioned issues i.e., automatically learn low to high-level features from raw HSI data which have attained incredible success for Hyperspectral Image Classification (HSIC).

The last few years witnessed an intensive improvement in Convolutional Neural Network (CNN) for HSIC where the spatial features are tailored by a 2D CNN model [15], [16], [17]. However, these spatial features are usually extracted separately which to some extent void the reason to jointly exploit the spatial-spectral information for HSIC. Therefore, in this paper, a novel 3D CNN for HSIC method is proposed. This work first divides the HSI cube into small overlapping 3D patches. These patches are processed to generate 3D feature maps using 3D kernel function over multiple contiguous bands to preserve the joint spatial and spectral information for the feature learning process which exploits important discrimination information for HSIC. As a preprocessing, incremental Principle Component Analysis (iPCA) is deployed to reduce the redundancy among the bands to process the few important wavelengths out of the entire HSI cube. Later the 3D CNN classifier is trained in an end-to-end fashion which involves fewer parameters than other 2D/3D CNN models. The comparative study is also carried out with several state-of-the-art 2D/3D CNN based HSIC methods proposed in the literature. Experimental/comparative results revealed that the proposed method outperforms the compared ones.

The rest of the paper is structured as follows; section II presents the proposed methodology. Section III describes the experimental Datasets, Results and discussion. Finally Section IV concludes the paper with possible future research directions.

M. Ahmad is with the Department of Computer Engineering, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, 64200, Pakistan, E-mail: mahmad00@gmail.com

which are not mutually exclusive. The water absorption bands have been removed before the experiments thus the remaining 200 bands are used in this experiment.

Pavia University Dataset (PUD) gathered over Pavia in northern Italy using a Reflective Optics System Imaging Spectrometer (ROSIS) optical sensor. PUD consists of 610×610 spatial and 103 spectral bands with a spatial resolution of 1.3 meters. PUD ground truth classes are 9. Further details about the experimental datasets can be found at [23]. The ground images of all the experimental datasets are shown in Figure 3.

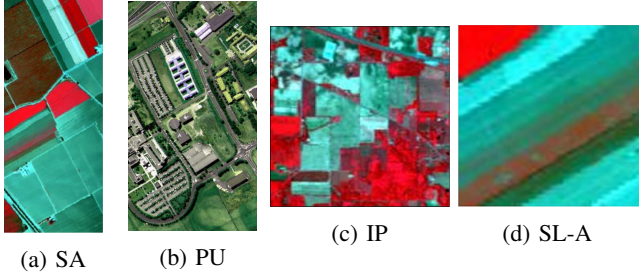


Fig. 3: Ground images of experimental datasets used in this work.

All the experiments were performed on an online platform known as Google Colab [24]. Google Colab is an online platform that requires a good speed of internet to run any environment. Google Colab provides an option to execute the codes on python 3 notebook with Graphical Processing Unit (GPU), 25 GB of Random Access Memory (RAM) and 358.27 GB of could storage for data computation. In all the experiments, the initial Test/Train set is divided into a 30/70% ratio on which Training samples (70% of the entire population) are further divided into 50/50% for the Training and Validation set.

To make the fair comparisons, the learning rate for all the experiments is set to 0.001, *relu* as an activation function is used for all layers except last on which *softmax* is used, the patch sizes are set as $11 \times 11 \times 20$, $13 \times 13 \times 20$, $15 \times 15 \times 20$, $17 \times 17 \times 20$, $19 \times 19 \times 20$, $21 \times 21 \times 20$ and $25 \times 25 \times 20$, respectively with 20 most informative bands selected by iPCA method. For evaluation purposes, Average Accuracy (AA), Overall Accuracy (OA) and Kappa (κ) coefficient have been computed from the confusion matrices. AA represents the average class-wise classification performance, OA is computed as the number of correctly classified examples out of the total test examples and finally, κ is known as a statistical metric that considered the mutual information regarding a strong agreement among classification and ground-truth maps. Along with OA, AA and κ metrics, several statistical tests are also being considered such as F1-Score, Precision and Recall.

The convergence loss and accuracy of our proposed 3D CNN model for a 50 number of epochs are shown in Figure 4. From these figures, one can conclude that the proposed model is converged almost around 32 echos. Whereas, the computational time of our proposed model is shown in Table II which also reveals a fast convergence and computational efficiency of our proposed model. The computational time

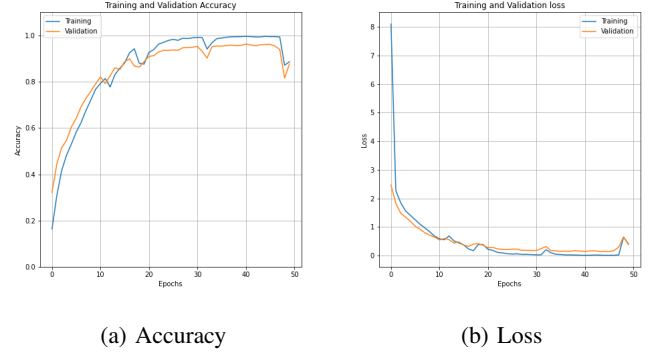


Fig. 4: Accuracy and Loss for Training and Validation sets on Indian Pine Dataset with 11×11 window patch corresponds to the 50 number of Epochs.

highly depends on the speed of the internet and available RAM.

TABLE II: Computational time in minutes for all the experimental datasets with several window sizes.

Dataset	Window Size						
	11 × 11	13 × 13	15 × 15	17 × 17	19 × 19	21 × 21	25 × 25
SL-A	0.22	0.23	0.56	0.28	0.98	0.37	0.45
SL	1.34	1.41	1.60	2.00	3.17	2.63	3.52
IP	0.33	0.33	0.61	0.78	0.62	0.58	0.76
PU	2.16	5.26	1.35	2.00	2.46	2.14	2.83

The accuracy analysis i.e., OA, AA, and κ based on the impact of spatial dimensions ¹ processed by the proposed model is presented in Table III. While looking into the Table III, one can conclude that the window size of 11×11 is enough for Pavia University, Salinas and Salinas-A dataset whereas the window size of 13×13 and 25×25 both works almost the same.

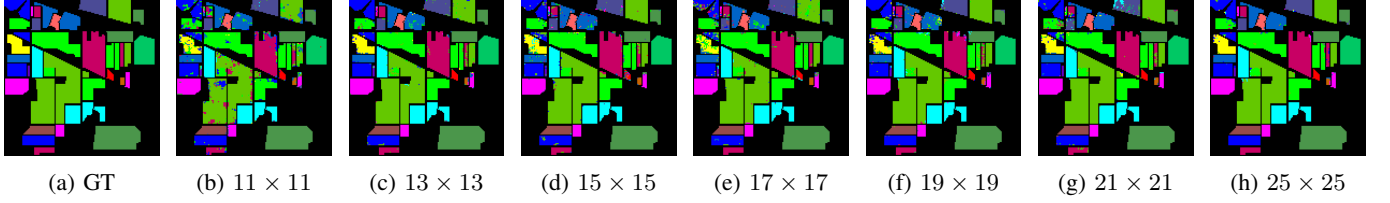
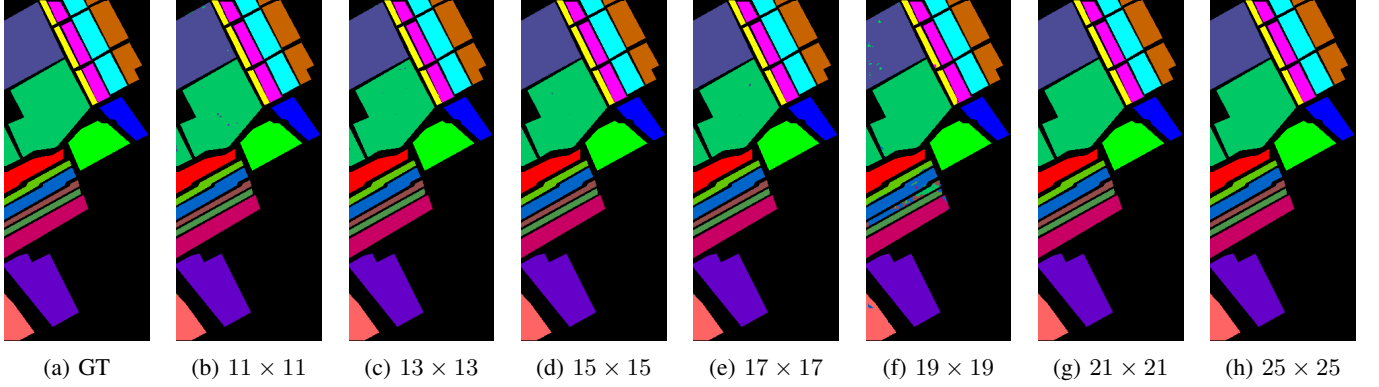
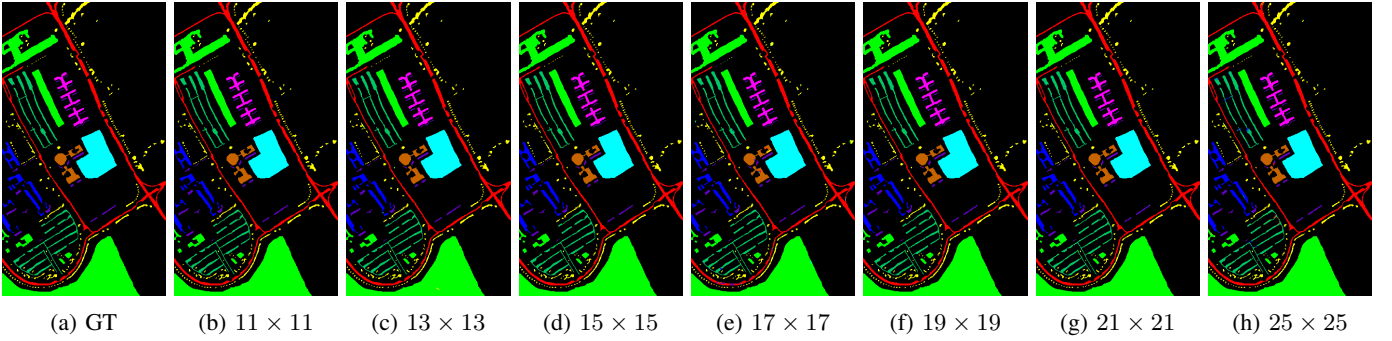
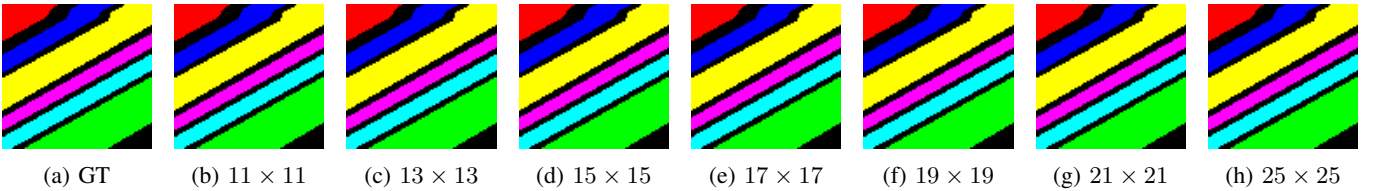
Furthermore, the classification maps (geographical locations for each class) according to the different number of window sizes (spatial dimensions) are shown in Figures 6-8. In regards to comparison, the proposed model is compared with several state-of-the-art methods published in the recent few years. From experimental results listed in Table IV one can conclude that the proposed model has competitive results and to some extent better in regards to the other methods. The comparative methods includes Multi-scale-3D-CNN [25], 3D/2D-CNN [22], [26], [27], [28]. From experiments listed in Table IV

TABLE III: Impact of window size on our proposed model

Window	PU			IP			SA			SL-A		
	OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
11×11	99.94	99.89	99.92	88.65	83.52	87.11	99.80	99.91	99.78	100	100	100
13×13	99.81	99.65	99.75	95.38	94.14	94.72	99.93	99.94	99.93	100	100	100
15×15	99.85	99.62	99.80	93.69	93.09	92.79	99.99	99.99	99.99	100	100	100
17×17	99.05	98.49	98.75	91.80	91.74	90.62	99.95	99.97	99.95	99.93	99.93	99.92
19×19	99.93	99.78	99.91	93.13	93.42	92.15	98.04	94.02	97.81	100	100	100
21×21	99.78	99.43	99.72	94.34	91.31	93.52	99.99	99.99	99.99	100	100	100
25×25	98.79	97.67	98.39	97.75	96.17	97.44	99.96	99.93	99.95	100	100	100

shows the proposed method improves the results significantly then the state-of-the-art methods with even less number of training samples.

¹The Confusion matrices (with per class accuracy) for each window size and every dataset is provided in the supplementary material.

Fig. 5: **Indian Pines Dataset** Ground Truths for each spatial dimensions processed through our proposed model.Fig. 6: **Salinas Dataset** Ground Truths for each spatial dimensions processed through our proposed model.Fig. 7: **Pavia University Dataset** Ground Truths for each spatial dimensions processed through our proposed model.Fig. 8: **Salinas-A Dataset** Ground Truths for each spatial dimensions processed through our proposed model.TABLE IV: Comparative evaluations with State-of-the-art methods while considering 11×11 Spatial dimensions and 10% of training samples.

dataset	Multi-scale-3D-CNN			3D-CNN			2D-CNN			Proposed		
	OA	AA	Kappa	OA	AA	Kappa	OA	AA	Kappa	OA	AA	Kappa
PU	95.95	97.52	93.40	96.34	97.03	94.90	96.63	94.84	95.53	98.40	97.89	97.89
IP	81.39	75.22	81.20	82.62	76.51	79.25	80.27	68.32	75.26	97.75	94.54	97.44
SA	94.20	96.66	93.61	85.00	89.63	83.20	96.34	94.36	95.93	98.06	98.80	97.85

IV. CONCLUSION

Hyperspectral Image Classification (HSIC) is a challenging task due to high inter-class similarity and high intra-class variability. Therefore, this paper proposed a lightweight fast

3D CNN model which not only overcome the abovesaid challenges but also provide state of the art experimental results in a computationally efficient fashion on three benchmark Hyperspectral datasets. The inter-class similarity and high intra-class variability issues are being resolved using a spatial-spectral information using 3D convolutions. The experimental results reveal that the proposed method outperformed the state of the art methods, furthermore, the proposed model is less complex than the conventional 3D CNN models.

RUNNING CODE

The Running Demo can be found at Github.

REFERENCES

- [1] S. Mei, X. Yuan, J. Ji, Y. Zhang, S. Wan, and Q. Du, "Hyperspectral image spatial super-resolution via 3d full convolutional neural network," *Remote Sensing*, vol. 9, p. 1139, 11 2017.
- [2] M. Ahmad, A. Khan, A. M. Khan, M. Mazzara, S. Distefano, A. Sohaib, and O. Nibouche, "Spatial prior fuzziness pool-based interactive classification of hyperspectral images," *Remote Sensing*, vol. 11, no. 9, May. 2019. [Online]. Available: <https://www.mdpi.com/2072-4292/11/9/1136>
- [3] Y. Wang, W. Yu, and Z. Fang, "Multiple kernel-based svm classification of hyperspectral images by combining spectral, spatial, and semantic information," *Remote Sensing*, vol. 12, no. 1, p. 120, 2020.
- [4] A. Alcolea, M. E. Paoletti, J. M. Haut, J. Resano, and A. Plaza, "Inference in supervised spectral classifiers for on-board hyperspectral imaging: An overview," *Remote Sensing*, vol. 12, no. 3, p. 534, 2020.
- [5] M. Ahmad, S. Shabbir, D. Oliva, M. Mazzara, and S. Distefano, "Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification," *Optik-International Journal for Light and Electron Optics*, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0030402619316109>
- [6] D. Tuia, M. Volpi, M. D. Mura, A. Rakotomamonjy, and R. Flamary, "Automatic feature learning for spatio-spectral image classification with sparse svm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 6062–6074, 2014.
- [7] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectralspatial classification techniques based on attribute profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2335–2353, 2015.
- [8] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 480–491, 2005.
- [9] S. Jia, X. Zhang, and Q. Li, "Spectralspatial hyperspectral image classification using $\ell_{1/2}$ regularized low-rank representation and sparse representation-based graph cuts," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2473–2484, 2015.
- [10] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 542–546, 2011.
- [11] Y. Zhong, A. Ma, and L. Zhang, "An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1235–1248, 2014.
- [12] L. Shen and S. Jia, "Three-dimensional gabor wavelets for pixel-based hyperspectral imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 5039–5046, 2011.
- [13] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 2276–2291, 2013.
- [14] S. Roy, G. Krishna, S. R. Dubey, and B. Chaudhuri, "Hybridsn: Exploring 3-d-2-d cnn feature hierarchy for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 277–281, 06 2019.
- [15] Y. Li and L. He, "An improved hybrid CNN for hyperspectral image classification," in *Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*, Z. Pan and X. Wang, Eds., vol. 11373, International Society for Optics and Photonics. SPIE, 2020, pp. 485 – 490. [Online]. Available: <https://doi.org/10.1117/12.2557384>
- [16] B. Fang, Y. Bai, and Y. Li, "Combining spectral unmixing and 3d/2d dense networks with early-exiting strategy for hyperspectral image classification," *Remote Sensing*, vol. 12, p. 779, 02 2020.
- [17] L. Huang and Y. Chen, "Dual-path siamese cnn for hyperspectral image classification with limited training samples," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2020.
- [18] M. Ahmad, D. Ulhaq, Q. Mushtaq, and M. Sohaib, "A new statistical approach for band clustering and band selection using k-means clustering," *International Journal of Engineering and Technology*, vol. 3, pp. 606–614, 12 2011.
- [19] M. Ahmad, D. Ulhaq, and Q. Mushtaq, "Aik method for band clustering using statistics of correlation and dispersion matrix," 01 2011, pp. 114–118.
- [20] M. Ahmad, D. Ihsan, and D. Ulhaq, "Linear unmixing and target detection of hyperspectral imagery using osp," 01 2011, pp. 179–183.
- [21] M. Ahmad, S. Lee, D. Ulhaq, and Q. Mushtaq, "Hyperspectral remote sensing: Dimensional reduction and end member extraction," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, pp. 2231–2307, 05 2012.
- [22] Y. Li, H. Zhang, and Q. Shen, "Spectralspatial classification of hyperspectral imagery with 3d convolutional neural network," *Remote Sensing*, vol. 9, p. 67, 01 2017.
- [23] *Hyperspectral Datasets Description*, 2020 (accessed 2020-01-12), http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes.
- [24] T. Carneiro, R. V. M. Da Nóbrega, T. Nepomuceno, G.-B. Bian, V. H. C. De Albuquerque, and P. P. Reboucas Filho, "Performance analysis of google colab as a tool for accelerating deep learning applications," *IEEE Access*, vol. 6, pp. 61 677–61 685, 2018.
- [25] M. He, B. Li, and H. Chen, "Multi-scale 3d deep convolutional neural network for hyperspectral image classification," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 3904–3908.
- [26] Y. Luo, J. Zou, C. Yao, T. Li, and G. Bai, "HSI-CNN: A novel convolution neural network for hyperspectral image," *CoRR*, vol. abs/1802.10478, 2018. [Online]. Available: <http://arxiv.org/abs/1802.10478>
- [27] A. Ben Hamida, A. Benoit, P. Lambert, and C. Ben Amar, "3-d deep learning approach for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4420–4434, 2018.
- [28] B. Liu, X. Yu, P. Zhang, X. Tan, A. Yu, and Z. Xue, "A semi-supervised convolutional neural network for hyperspectral image classification," *Remote Sensing Letters*, vol. 8, no. 9, pp. 839–848, 2017. [Online]. Available: <https://doi.org/10.1080/2150704X.2017.1331053>