SEMINAR REPORT

ON

CNNs FOR MULTI - SPECTRAL IMAGE CLASSIFICATION

Submitted by

SUBITH O U

(KSD18CS084)

To the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING **LBS COLLEGE OF ENGINEERIING KASARAGOD KASARAGOD – 671542, KERALA**

NOVEMBER 2021

DECLARATION

"I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of this Institute or other Institute of higher learning, except where due acknowledgement has been made in the text."

PLACE: KASARAGOD
DATE: 26/11/2022

NAME: SUBITH O U
REG NO: KSD18CS084

CERTIFICATE

This is to certify that the seminar report entitled **CNNs FOR MULTI - SPECTRAL IMAGE CLASSIFICATION** *submitted by* **SUBITH O U** *to the APJ Abdul Kalam TechnologicalUniversity in partial fulfillment of the requirements for the award of the degree of* **Bachelor of Technology in Computer Science and Engineering** *is a bonafide record of the work done by him under my supervision and guidance*.

Mr. Sarith Divakar M	**Mrs. Sudheesh K P**	**Mrs. Smitha Mol MB**
Assistant Professor	Assistant Professor	Department of CSE
Department of CSE	Department of CSE	(Head of Department)
(Guide)	(Seminar coordinator)	

Place: Kasaragod
Date: 26/11/2022

# ACKNOWLEDGEMENT

It is a really a momentous opportunity and privilege to express my deep sense of Gratitude to all those who have us to accomplish this task. I express humble God Almighty for his incessant blessing on us during this seminar.

I have taken efforts in this seminar. However, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend our sincere to all of them.

I sincerely thank our principal **Dr. MOHAMMED SHEKOOR T** for providing me facilities in order to go ahead with our seminar development. I express our sincere gratitude to head of the department **Dr. SMITHAMOL M B** and I also express heartiest gratitude to seminar coordinator **Mrs. SUDHEESH K P** in computer science and engineering for their valuable advice and guidance. I would always oblige for the helping hands of all other staff members of the department and all my friends and well-wishers who directly or indirectly contributed to this venture.

Last but not least, I am indebted to God almighty for being the guiding light throughout this work and for helping me to complete the same within the stipulated time.

SUBITH O U

ABSTRACT

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. However, several works demonstrated that low-quality or noisy data (even including perceptually not visible noises) may have a huge impact on the accuracy of CNN models. But feedback features in CNNs have improved over the existing feed-forward CNNs. These recent works on the integration of recurrence and/or feedback to CNNs mostly tested deep networks on natural scenes with relatively perceptually good resolution color images. In this work, we explore the effectiveness of baseline CNN, Feedback and Recurrent CNN using the classification of mid-resolution (1 pixel - 30×30 square meters per pixel) multi-spectral satellite images.

CONTENTS

List of Figures

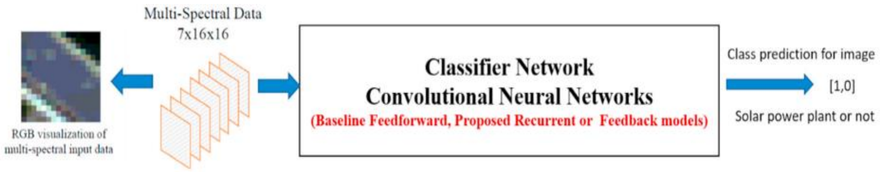
List of Abbreviations

CNN : Convolutional Neural Network

1. INTRODUCTION

To process large scale data, recent advancement on machine learning and especially on deep learning applications (e.g. convolutional neural networks: CNNs) have resulted in huge improvements on taking advantage of large-scale big data on remote sensing researches. Recently many deep learning applications have been done in wide variety of topics on remote sensing such as image classification, object detection, semantic labeling or pixel-level classification of images, super-resolution, etc. [2-12]. Despite having wide variety of approaches using deep learning algorithms for satellite image classification, Ishii et al. [2] did the one of the few large-scale surveys on solar-power plant classification and detection by analyzing multi-spectral data from Landsat8 satellite observations in Japan. They [2] made a survey of solar power plants in Japan by large-scale multi-spectral images (obtained from Landsat8) to be trained and tested by a relatively small (shallow) CNN that allows classification of small image patches. These small patches consist of a spatial resolution of 16x16 pixels where one pixel corresponds to approximately 30×30 square meters area in real measures, and 7 multi-spectral channels that cover different wavelengths or spectral bands [2, 3]. In the experimental results of [2], it is also demonstrated that relatively low resolution multi-spectral images can be classified with shallow CNNs with even better accuracy than the state-of-the-art deep CNN based work in [4]. In addition, most of the CNN models in remote sensing and computer vision applications have been using the feedforward approach as in the work by Ishii et al. [2] and similar tasks for satellite image classification approaches by using state-of-the-art approaches [2, 4-6, 8, 9]. However, several works [13, 14] demonstrated that low-quality or noisy data (even including perceptually not visible noises) may have a huge impact on the accuracy of CNN models. Therefore, the use of recurrent and feedback features in CNNs as an improvement to the existing feed-forward CNNs [15-20, 3].

effectiveness of shallow ConvRNNs by using the feed-forward CNN model INet [2] as a baseline model for shallow recurrent CNN model (R-Net) inspired by Recurrent-CNN [15] as a reference. Moreover, we extend INet [2] and Recurrent-CNN [15] implementation to feedback-CNN (F-Net) for the solar-power plant classification task on mid-resolution (1 pixel - 30×30 square meters) multi-spectral satellite images (see Figure 1 for an overview of the framework in this study). Then, we compare the accuracy of feed-forward, recurrent and feedback CNNs on solar power plant classification by using the Intersection over Union (IoU) evaluation metric.



2. Multi-spectral Image

A multi-spectral image is a collection of several monochrome images of the same scene, each of them taken with a different sensor. Each image is referred to as a *band*. A well known multi-spectral (or multi-band image) is a RGB color image, consisting of a red, a green and a blue image, each of them taken with a sensor sensitive to a different wavelength. In image processing, multi-spectral images are most commonly used for Remote Sensing applications. Satellites usually take several images from frequency bands in the visual and non-visual range. *Landsat 8*, for example, produces 11 bands of different wavelengths and spatial resolutions obtained from two sensors: Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) [22, 23]. OLI data corresponds to visible and near-infrared light, while TIRS corresponds to thermal infrared light [22, 23]. An overview of imaging sensors and bands of Landsat 8 is given in Table 1 [22, 23, 2].

Table 1. Observation wavelength and spatial resolution of Landsat 8 imaging sensors.

Sensor	Band	Wavelength (μm)	Resolution(m/p)
OLI	1	0.43 – 0.45	30
	2 (B)	0.45 – 0.51	30
	3 (G)	0.53 – 0.59	30
	4 (R)	0.64 – 0.67	30
	5	0.85 – 0.88	30
	6	1.57 – 1.65	30
	7	2.11 – 2.29	30
	8	0.53 – 0.68	15
	9	1.36 – 1.38	30
TIRS	10	10.60 – 11.19	100
	11	11.50 – 12.51	100

3. Multi-spectral Image Classification

3.1. FeedForward CNN (INet) for multi-spectral image classification

In Ishii et. al [2] model (I-Net), they propose a detection by classification approach on 16×16 spatial resolution patches with 7 channels. In their work [2], detection by classification refers to labelling all pixels in the whole patch as the class prediction obtained from classification of the input data. Therefore, I-Net as baseline model for comparison with variations of it for classification using MUSIC4P3 dataset [21].

I-Net has 3 convolutional layers followed by a fully connected layer. INet [2] model is given in Fig. 2 showing layers and feedforward processes (see Eq. 1) with the given multi-spectral image. All the convolutional layers have 3×3 kernels with padding. ReLU (Rectified Linear Unit) is used as transfer function after all convolutional layers. ReLU is followed by a normalization procedure learned during training with batch-normalization (BN) function; this normalization operation is denoted by the ζ( ) function in Equation 1 where U is the input multi-spectral image, wforward are weights for the convolutional filters for forward processes at each layer, and CNN feature representations F1, F2 and F3 (Eq. 1) all have 16×16 spatial resolution with 32 channels. Finally a fully connected layer using the F3 feature representations enables classification with a softmax function as the class probabilities. Unlike most of the standard models, no pooling layer is used, and all convolutional layers have a stride of 1 due to the small size of input images.

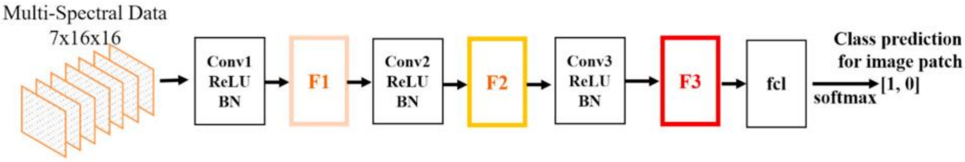
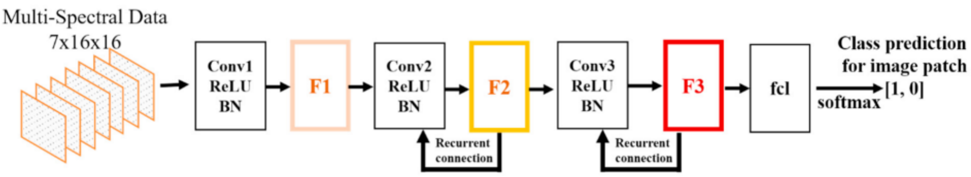


Figure 2 INet Model: Feed-forward baseline CNN for solar power plan classification on multi-spectral satellite imagery

$$\begin{aligned} \mathbf{F1} &= \zeta\left(\left(\mathbf{w}_{conv1}^{forward}\right) * \mathbf{U}\right) \\ \mathbf{F2} &= \zeta\left(\left(\mathbf{w}_{conv2}^{forward}\right) * \mathbf{F1}\right) \\ \mathbf{F3} &= \zeta\left(\left(\mathbf{w}_{conv3}^{forward}\right) * \mathbf{F2}\right) \end{aligned} \qquad (1)$$

3.2. Recurrent CNN (R-Net) for multi-spectral image classification

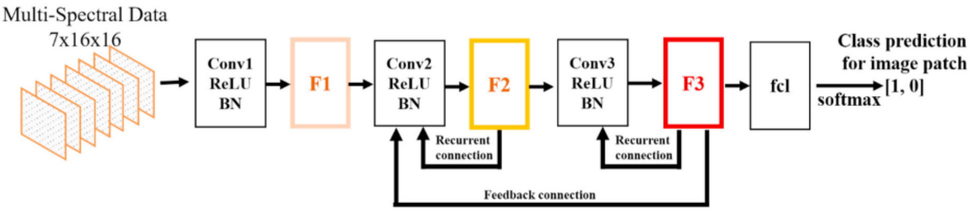
Recurrent-CNN (R-Net) is implemented on the INet [2] base model by having 3 convolutional layers followed by a fully connected layer (see Fig.3). As in INet [2], all the convolutional layers in R-Net also have convolutional filters with 3×3 kernels. ReLU (Rectified Linear Unit) is used as transfer function after all convolutional layers. Again, ReLU is followed by a normalization procedure learned during training with the batch-normalization (BN) function. However, in the R-Net (Fig.3) case, we have recurrent connections added at conv2 and conv3 layers (i.e. F2t-1 features input to conv2 layer and F3t-1 features to conv3 layer). This is achieved by adding extra convolutional filters for these recurrent input signals as given in the computational process of F2t and F3t (see Equation 2).



$$
\begin{aligned}
\mathbf{F1}_t &= \zeta\left(\left(\mathbf{w}_{conv1}^{forward}\right) * \mathbf{U}_t\right) \\
\mathbf{F2}_t &= \zeta\left(\left(\mathbf{w}_{conv2}^{forward}\right) * \mathbf{F1}_t + \left(\mathbf{w}_{conv2}^{recurrent}\right) * \mathbf{F2}_{t-1}\right) \\
\mathbf{F3}_t &= \zeta\left(\left(\mathbf{w}_{conv3}^{forward}\right) * \mathbf{F2}_t + \left(\mathbf{w}_{conv3}^{recurrent}\right) * \mathbf{F3}_{t-1}\right)
\end{aligned}
\qquad (2)
$$

3.3. FeedBack CNN (F-Net) for multi-spectral image classification

In addition to recurrent connections in R-Net, the addition of a feedback connection at conv2 layer. Call this network a Feedback-CNN (F-Net) and we display its architecture in Fig. 4. The output from conv3 layer, $F3_{t-1}$ features, is used as an additional feedback signal to conv2 layer along with the feedforward and recurrent inputs. By this way, feedback from higher level representation to lower representations at conv2 by adding extra convolutional filters for these feedback input signals as given in the computational process of $F2_t$ (see Equation 3).



$$
\begin{aligned}
&\mathbf{F1}_t = \zeta\left(\left(\mathbf{w}_{conv1}^{forward}\right) * \mathbf{U}_t\right) \\
&\mathbf{F2}_t = \zeta\left(\left(\mathbf{w}_{conv2}^{forward}\right) * \mathbf{F1}_t + \left(\mathbf{w}_{conv2}^{recurrent}\right) * \mathbf{F2}_{t-1} + \left(\mathbf{w}_{conv2}^{feedback}\right) * \mathbf{F3}_{t-1}\right) \\
&\mathbf{F3}_t = \zeta\left(\left(\mathbf{w}_{conv3}^{forward}\right) * \mathbf{F2}_t + \left(\mathbf{w}_{conv3}^{recurrent}\right) * \mathbf{F3}_{t-1}\right)
\end{aligned}
\tag{3}
$$

4. Performance Analysis

For the multi-spectral image classification performance evaluation, we use the Intersection over Union (IoU) metric as in [2, 3], which is defined as

IoU = True Positive / (True Positive + False Negative + False Positive)

We compare I-Net, R-Net, and the proposed F-Net to explore how top-down signals improve the baseline feed-forward model.

The lowest IoU results being the least performing approach.

5. CONCLUSION

The recurrent-CNN (R-Net) and Feedback-CNN (F-Net) based on a state-of-the-art feed-forward model for multi-spectral image classification. Our experiments demonstrated that using top-down signals (especially recurrent and feedback features together) on CNNs can provide good representation of multi-spectral images which can in turn improve classification accuracy drastically.

6. REFERENCES

[1] Exploring Recurrent and Feedback CNNs for Multi-Spectral Satellite Image Classification.

[2] Ishii, Tomohiro, Edgar Simo-Serra, Satoshi Iizuka, Yoshihiko Mochizuki, Akihiro Sugimoto, Hiroshi Ishikawa, and Ryosuke Nakamura (2016) “Detection by Classification of Buildings in Multispectral Satellite Imagery”, International Conference on Pattern Recognition (ICPR).

[3] İmamoğlu, Nevrez, Motoki Kimura, Hiroki Miyamoto, Aito Fujita, Ryosuke Nakamura (2017) “Solar Power Plant Detection on Multi-Spectral Imagery using Weakly-Supervised CNN with Feedback Features and m-PCNN Fusion”, in Proceedings of the British Machine Vision Conference (BMVC), Imperial College London, UK.

[4] Otavio A. B. Penatti, Keiller Nogueira, and Jefersson. A. dos Santos (2015) "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?" International Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops).

[5] Hu, Fan, Gu-Song Xia, Jingwei Hu, and Liangpei Zhang (2015) “Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery”, Remote Sensing 7: 14680-14707.

[6] Gao, Qishuo, Samsung Lim and Xiuping Jia (2018) “Hyperspectral Image Classification Using Convolutional Neural Networks and Multiple Feature Learning”, Remote Sensing 10 (299): 18 pages

[7] Cheng, Gong, Zhenpeng Li, Junwei Han, Xiwen Yao and Lei Guo (2018) “Exploring Hierarchical Convolutional Features for Hyperspectral Image Classification”, IEEE Transactions on Geoscience and Remote Sensing: 11 pages. (accepted to appear)

[8] Hao, Siyuan , Wei Wang, Yuanxin Ye, Enyu Li and Lorenzo Bruzzone (2018), “A Deep Network Architecture for Super-Resolution-Aided Hyperspectral Image Classification with Classwise Loss”, IEEE Transactions on Geoscience and Remote Sensing: 14 pages. (accepted to appear)

[9] Yuan, Yuan, Xiangtao Zheng and Xiaoqiang Lu (2017) “Hyperspectral Image Superresolution by Transfer Learning”, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (5): 1963-1974.

[10] Luo, Yimin , Liguo Zhou, Shu Wang and Zhongyuan Wang (2017) “Video Satellite Imagery Super Resolution via Convolutional Neural Networks” IEEE Geoscience and Remote Sensing Letters 14 (12): 2398 – 2402.

[11] Li, Xiaodong, Feng Ling, Yun Du and Yihang Zhang (2014) “Spatially Adaptive Superresolution Land Cover Mapping With Multispectral and Panchromatic Images” IEEE Transactions on Geoscience and Remote Sensing 52 (5): 2810 – 2823.

[12] Zhao, Wei, Licheng Jiao, Wenping Ma, Jiaqi Zhao, Jin Zhao, Hongying Liu, Xianghai Cao and Shuyuan Yang (2017) “Superpixel-Based Multiple Local CNN for Panchromatic and Multispectral Image Classification” IEEE Transactions on Geoscience and Remote Sensing 55 (7): 4141 – 4156.

[13] Nguyen, Anh, Jason Yosinski and Jeff Clune (2015) “Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images”, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR’15).

[14] Moosavi-Dezfooli, Seyed-Mohsen, and Alhussein Fawzi (2017) “Universal adversarial perturbations”, IEEE International Conference on Computer Vision and Pattern Recognition (CVPR’17).

[15] Liang, Ming and Xiaolin Hu (2015) “Recurrent Convolutional Neural Network for Object Recognition” IEEE International Conference on Computer Vision and Pattern Recognition (CVPR’15).

[16] Tang, Youbao, Xiangqian Wu and Wei Bu (2016) “Deeply-Supervised Recurrent Convolutional Neural Network for Saliency Detection” in Proceedings of the 2016 ACM on Multimedia Conference (ACMMM 2016).

[17] Zamir, Amir R., Te-Lin Wu, Lin Sun, William B. Shen, Bertram E. Shi, Jitendra Malik. Silvio Savarese (2017) “Feedback Networks” IEEE International Conference on Computer Vision and Pattern Recognition (CVPR’17).

[18] Stollenga, Marjin F., Jonathan Masci, Faustino Gomez and Juergen Schmidhuber (2014) “Deep Networks with Internal SelectiveAttention through Feedback Connections” Conference on Neural Information Processing Systems (NIPS).

[19] Liu, Xianming, Amy Zhang, Tobias Tiecke, Andreas Gros and Thomas S. Huang (2016) “Feedback Neural Network for Weakly Supervised Geo-Semantic Segmentation” arXiv preprint arXiv:1612.02766.

[20] Aaron van den Oord, Nal Kalchbrenner and Koray Kavukcuoglu (2016) “Pixel Recurrent Neural Networks” International Conference on Machine Learning (ICML).

[21] MUltiband Satellite Imagery for object Classification (MUSIC) to detect Photovoltaic Power Plants (MUSIC4P3) dataset online available at: https://github.com/gistairc/MUSIC4P3, [Online accessed June 2018]

[22] Roy, D.P., M. A. Wulder, T.R.Loveland, C. E. Woodcock, R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons, D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote, A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C. O. Justice, A. Kilic, V. Kovalskyy ,Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y.Shuai, R. Trezza, J. Vogelmann, R. H. Wynne and Z. Zhu “Landsat-8: Science and Product Vision for Terrestrial Global Change Research”, Remote Sensing of Environment 145: 154-172, 2014.

[23] Landsat-8: https://landsat.usgs.gov/landsat-8, [Online accessed June 2018]