POLSAR IMAGE CLASSIFICATION USING ATTENTION BASED SHALLOW TO DEEP CONVOLUTIONAL NEURAL NETWORK

Mohammed Q. Alkhatib, Mina Al-Saad, Nour Aburaed, M. Sami Zitouni, Hussain Al Ahmad
College of Engineering and IT, University of Dubai, UAE

Paper ID : 2823

INTRODUCTION

This paper introduces a novel network for PolSAR image classification and interpretation. The network incorporates multi-branch feature fusion using Complex-valued Convolutional Neural Networks (CV-CNNs). The proposed approach extracts polarimetric features from each branch to achieve accurate classification. Additionally, the model incorporates Squeeze and Excitation (SE) to enhance channel interdependencies without significant computational overhead. The effectiveness of the attention-based shallow to deep CV-CNN model is evaluated using the Flevoland benchmark dataset. Experimental results demonstrate the model's efficacy in PolSAR image classification, as measured by metrics such as Kappa Coefficient (k), Overall Accuracy (OA), and Average Accuracy (AA).

DATASET

- The scene for Flevoland dataset [1] is acquired by NASA/JPL AirSAR system over the agricultural area in Netherlands.
- The image size is 750 × 1024 pixels.
- The classification map contains fifteen classes and the unassigned pixels in the image are colored in black as shown in Fig. 1.
- The Hermitian coherency matrix (**T**) was used in this study.
- Three real-valued (diagonal) and Three complex-valued elements of the upper triangle of the coherency matrix (T_{11}, T_{22}, T_{33}, T_{12}, T_{13}, T_{23}) are used as the input feature of the models.

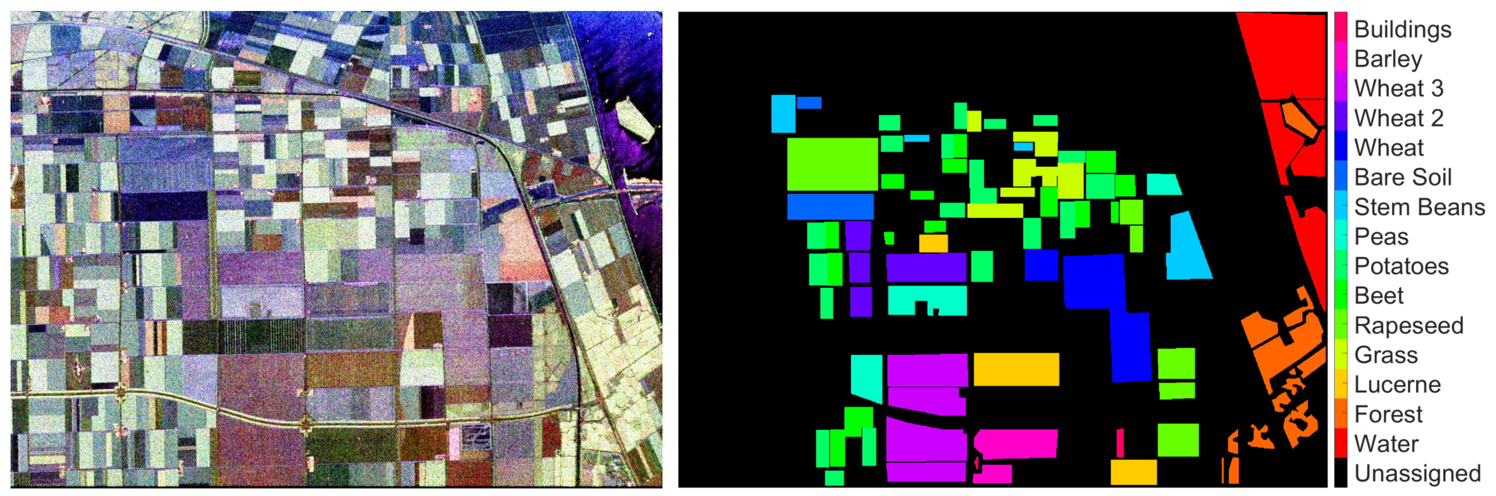

Fig. 1 Pauli RGB (Left) and Classification Map (Right) of Flevoland dataset.

NETWORK ARCHITECTURE

The PolSAR image classification model (Fig. 2) includes CV-CNNs [2], multi-branch feature fusion, and channel attention with SE [3]. The data is processed through a three-branch network to extract features, which are then concatenated. The SE block (Fig.3) enhances channel dependencies. Finally, classification is performed using fully connected layers with dropout for regularization and a softmax layer for prediction.

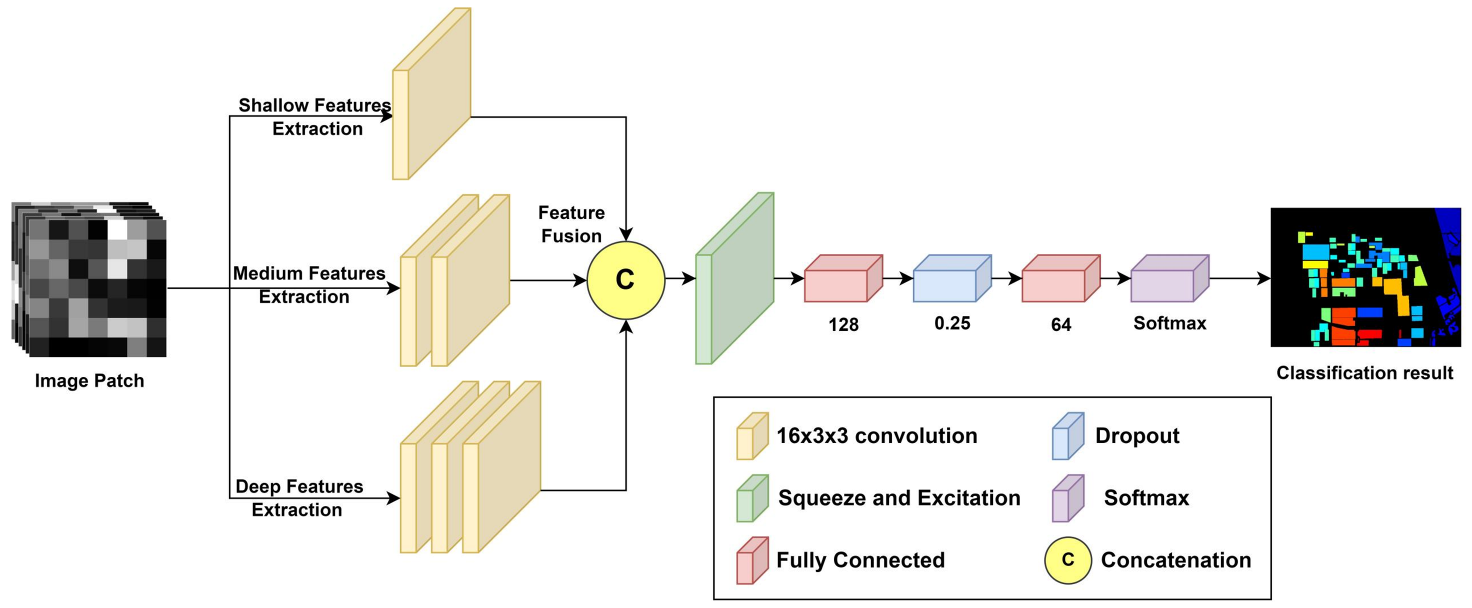

Fig. 2 Proposed Model

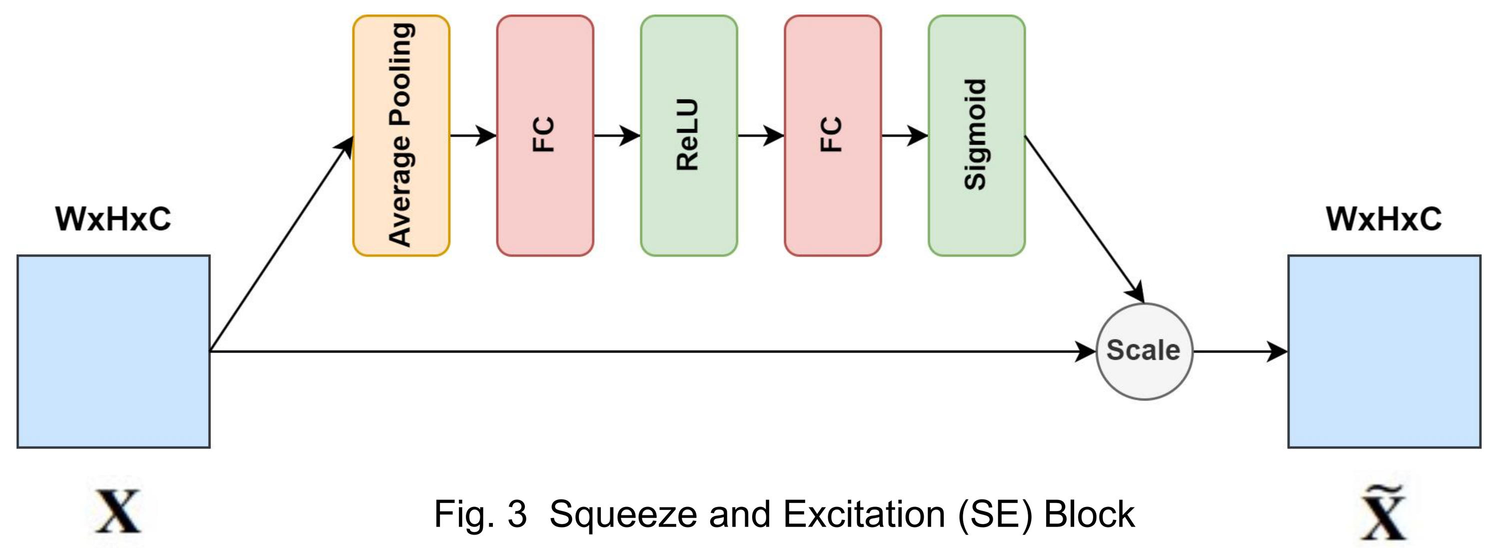

Fig. 3 Squeeze and Excitation (SE) Block

EXPERIMENTS

- Three experiments were conducted
  1. RV-CNN : Only real-valued layers were used
  2. CV-CNN : Complex-valued layers were used but with SE
  3. CV-CNN-SE: Using complex-values layers along with SE
- The image data was first split into patches of size 11 × 11 × 6
- Image patches are randomly divided into **1%** for training and **99%** for testing and evaluating the model performance.

RESULTS

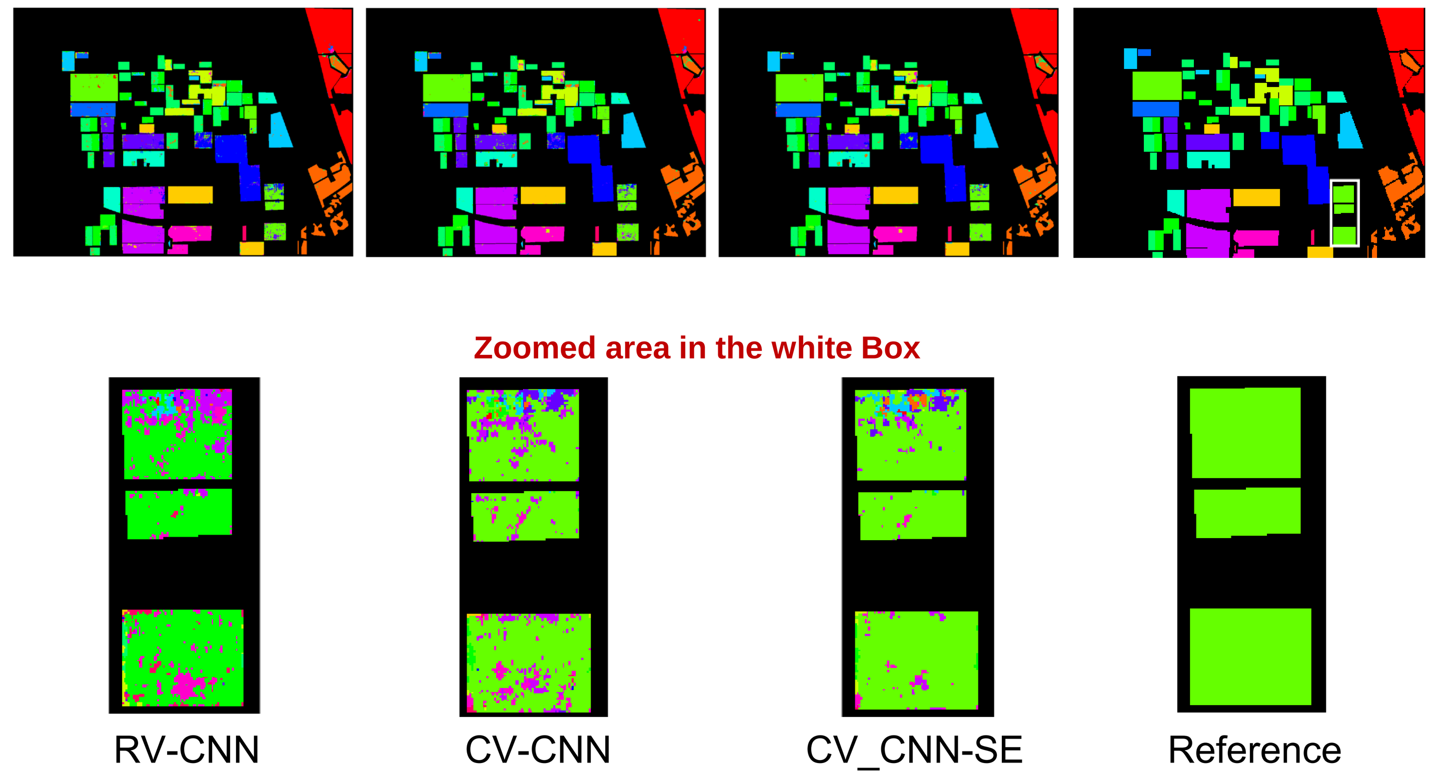

Fig. 3 Classification results.

Class #	Train	Test	RV-CNN	CV-CNN	CV-CNN-SE
Water	293	28956	98.28%	98.56%	**99.52%**
Forest	159	15696	**98.12%**	97.61%	95.88%
Lucerne	112	11088	95.17%	**95.22%**	94.46%
Grass	103	10098	86.39%	81.73%	**89.39%**
Rapeseed	219	21636	90.45%	91.96%	**95.11%**
Beet	148	14559	91.84%	**92.77%**	90.85%
Potatoes	214	21130	89.12%	92.40%	**94.52%**
Peas	104	10292	95.05%	94.90%	**95.22%**
Stem Beans	85	8386	94.16%	94.61%	**96.42%**
Bare Soil	64	6253	93.67%	94.08%	**94.20%**
Wheat	177	17462	91.92%	94.24%	**96.57%**
Wheat 2	107	10522	93.25%	93.31%	**95.88%**
Wheat 3	221	21801	95.89%	98.41%	**99.31%**
Barley	74	7295	96.16%	96.93%	**97.56%**
Buildings	6	572	85.31%	86.19%	**89.66%**
AA			93.01 ± 1.06%	93.53 ± 0.53%	**94.44 ± 0.37%**
OA			93.68 ± 0.93%	94.52 ± 0.50%	**95.65 ± 0.36%**
Kappa			93.10 ± 1.01%	94.02 ± 0.57%	**94.98 ± 0,41%**

Fig. 4 Classification performance of different methods used in this study.

CONCLUSIONS

- The proposed approach shows superiority in terms of Kappa, AA and OA.
- CV-CNNs perform better than RV-CNNs, results are furthered enhanced by using attention mechanism
- Visual inspection shows that the classification map obtained from the proposed model is visually similar to the reference image.

REFERENCES

1. GitHub - fudanxu/CV-CNN. GitHub. https://github.com/fudanxu/CV-CNN, Accessed July-9-2023
2. Hänsch, Ronny, and Olaf Hellwich. "Complex-valued convolutional neural networks for object detection in PolSAR data." In 8th European Conference on Synthetic Aperture Radar, pp. 1-4. VDE, 2010.
3. Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132-7141. 2018.

Email: mqalkhatib@ieee.org