G-SM-CAM: A Fast Visual Understanding of CNNs in SAR Images Interpretation

1st Jingyuan Fan
School of Electronic Engineering
Xidian University
Xi'an, China
jyfan@stu.xidian.edu.cn

2nd Zhenpeng Feng
School of Electronic Engineering
Xidian University
Xi'an, China
zpfeng_1@stu.xidian.edu.cn

3rd Hongbing Ji
School of Electronic Engineering
Xidian University
Xi'an, China
hbji@xidian.edu.cn

4th Mingzhe Zhu
School of Electronic Engineering
Xidian University
Xi'an, China
zhumz@mail.xidian.edu.cn

5th Bo Zang
School of Electronic Engineering
Xidian University
Xi'an, China
bzang@mail.xidian.edu.cn

6th Linlin Ding
School of Electronic Engineering
Xidian University
Xi'an, China
llding@stu.xidian.edu.cn

***Abstract*—Numerous automatic recognition methods based on convolutional neural networks (CNNs) can achieve high calculation efficiency owing to its end-to-end structure. However, the internal mechanism of CNNs is intransparent which is limiting or even disqualifying in SAR image interpretation. To provide a visual understanding of CNNs' mechanism, we propose a Group-Self-Matching class activation mapping (G-SM-CAM) inspired by the split strategy and Self-Matching CAM. In specific, feature maps are firstly split into several groups and then the sub-feature maps are merged as a new class activation. Finally, each class activation is "self-matched" with the input image as the renewed feature maps to generate saliency map. G-SM-CAM is efficient and effective on SAR images, which runs dramatically faster than Self-Matching CAM at minor cost of saliency map quality. Numerous experimental results demonstrate the validity and efficiency of G-SM-CAM based on a benchmark dataset MSTAR.**

***Index Terms*—synthetic aperture radar, target recognition, class activation mapping, interpretation of convolutional neural network**

I. Introduction

With numerous advanced synthetic aperture radar (SAR) imaging algorithms proposed, the quality of SAR imaging has reached a high level; however, SAR image interpretation develops far behind SAR imaging [1–3]. SAR image interpretation consists of image segmentation, object detection, and target recognition. Among these procedures, target recognition is usually regarded as the most challenging [2, 4]. Traditional target recognition mainly includes denoising, edge detection, region of interest (ROI) extraction, feature extraction and classification. It is these complex procedures that bring in huge computation burden which is an obstacle to realize real-time application [5, 6].

In contrast, some deep learning recognition algorithms, particularly combined with convolutional neural networks (CNNs), can achieve high speed of processing as well as satisfactory accuracy by merging the aforementioned procedures into an end-to-end framework [7]. However, the internal mechanism of CNNs is intransparent, which is a limiting or even disqualifying factor for SAR image interpretation [8]. Although many class activation mapping (CAM) algorithms are proposed to visualize CNNs' internal mechanism, like Grad-CAM [9], Grad-CAM++ [10], XGrad-CAM [11], Ablation-CAM [12], Score CAM [13], and Group CAM [14], they all show limited effects on SAR images as they are designed for optical images. Self-Matching CAM is the state-of-the-art CAM particularly designed for SAR images [15], whereas, Self-Matching CAM is considerably computing costly compared with other CAM methods.

In this paper, we propose an effective and efficient Group-Self-Matching class activation mapping (G-SM-CAM) for SAR images. The key advantages of G-SM-CAM are summarized as: (1) The computing efficiency improves 92.8% compared with Self-Matching CAM; (2) The saliency map of G-SM-CAM almost reaches the effect of Self-Matching CAM.

II. Related Work

As an important branch of CNN visualization algorithms, CAM methods [16] visualize CNN's decision by combining feature maps of deep layers linearly, as follows:

$$M_c(x,y) = \sum_k (\omega_c^k F^k(x,y)). \tag{1}$$

where ω_c^k is the weight of the k-th feature map F^k. Different definitions lead to different CAM methods. Selvaraju, et al. [9] proposed Grad-CAM to visualize an arbitrary CNN for classification by weighting the feature maps using gradients. Aditya et al. [10] proposed Grad-CAM++ by introducing higher-order derivatives in Grad-CAM to obtain a more precise highlighted region. These two gradient-based methods do not explain clearly why using the average of gradients to weight

This research was funded by the Fundamental Research Funds for the Central Universities, grant number: JB210206, JC2111.