Gaussian Low-pass Channel Attention Convolution Network for RF Fingerprinting

Shunjie Zhang,[1] Tianhao Wu,[1] Wang Wei,[1] Ronghui Zhan,[1] and Zhang Jun[1]

[1]*National Key Laboratory of Science and Technology on Automatic Target Recognition, College of Electronic Science, National University of Defense Technology, Changsha, China*

wwang@nudt.edu.cn

Radio frequency (RF) fingerprinting is a challenging and important technique in individual identification of wireless devices. Recent work has used deep learning-based classifiers on ADS-B signal without missing aircraft ID information. However, traditional methods are difficult to obtain well performance accuracy for classical deep learning methods to recognize RF signals. This letter proposes a Gaussian Low-pass Channel Attention Convolution Network (GLCA-Net), where a Gaussian Low-pass Channel Attention module (GLCAM) is designed to extract fingerprint features with low frequency. Particularly, in GLCAM, we design a Frequency-Convolutional Global Average Pooling (F-ConvGAP) module to help channel attention mechanism learn channel weights in frequency domain. Experimental results on the datasets of large-scale real-world ADS-B signals show that our method can achieve an accuracy of 92.08%, which is 6.21% higher than Convolutional Neural Networks.

Introduction: The Automatic Dependent Surveillance-Broadcast (ADS-B) system has been widely used in aviation field due to its low cost and high accuracy. However, ADS-B system is vulnerable to lawbreaking attacks due to the non-encrypted message mode. With the rapid development of computation and datasets, deep learning (DL) has made a significant forward in computer vision, voice and natural language processing. For the field of radio frequency (RF) identification, DL has made rapid progress. LA Yun et al. [1] collected a dataset of 426,613 ADS-B long signals of category 1,661 aircraft and 167,234 ADS-B short signals of category 1,713 aircraft, and verified the effect of networks on the identification rate for different signal-noise ratios, sampling rates, and frequency carrier offsets. Although they achieve good identification results, but move ID address was not taken into account. YAN Ke et al. [2] proposed an end-to-end DL method based on ADS-B raw I/Q data, in which the collected ADS-B signal was pre-processed by wavelet transform denoising and then input to the network. The proposed method obtains the identification accuracy of 99%, but the method did not remove the ID address from ADS-B. J. Robinson et al. [3] designed the ADCC (Augmented Dilated Causal Convolution) module for raw I/Q data and achieved an 85% accuracy for ADS-B with ID address removed. The deep neural network framework of the frequency principle (F-Principle) [4] demonstrated that deep learning tends to preferentially use low frequencies to fit the objective function. Xi Li et al. [5] extended the channel attention to frequency and experimentally demonstrated the effectiveness of using attention mechanism in the frequency domain.

Inspired by Frequency Channel Attention Networks [5]. This letter proposes the Gaussian Low-pass Channel Attention Convolution Network (GLCA-Net) to improve the accuracy of fingerprinting. GLCAM is able to capture the characteristics of ADS-B signal, and enhancie the low-frequency of the data while simplifying the computational work. In addition, we designed F-ConvGAP (Frequency-Convolutional Global Average Pooling) to replace traditional GAP (Global Average Pooling). The F-ConvGAP is learnable, and we can verify the characteristics of the network by showing the nature map of the module.

The remainder of this paper is organized as follows: In Section II, we describe the datasets and analyze the ADS-B signal. In Section III, we present the methodology, especially GLCAM and F-ConvGAP. Section IV, evaluation setup and experimental results are provided. Finally, Section V concludes.

ADS-B Encoding Format: The ADS-B signal uses Pulse Position Modulation (PPM) which has better anti-interference performance than Pulse Amplitude Modulation and Pulse Width Modulation. In this letter, we focus on the ADS-B signals of the long format S-model: standard output of 1090MHz with Extended Squitter (1090ES). ADS-B message structure as shown in fig.1.

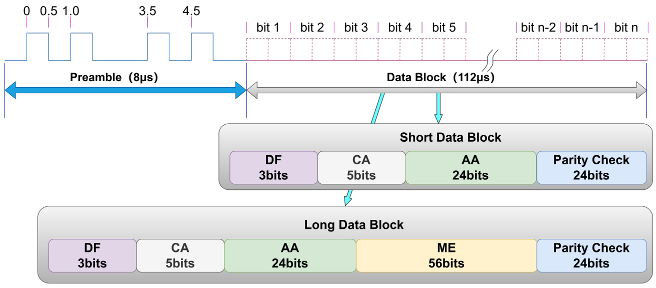

Fig 1 *Description of the ADS-B message structure.*

Each transmission contains 8 μs preamble and 112 μs data block. Taking a long data block as an example, the first 8 bits of a data block possess the downlink format (DF) and capability (CA), which remain constant in time and space for the same flight purpose. The next 24 bits are the aircraft address (AA) that uniquely identifies the aircraft all over the world. The 56 bits ME field consists of aircraft surveillance information (the short data block excludes this part). The last 24 bits parity check is used for receiver verification. The 24 bits aircraft address from each transmission is easy to learn by Deep Neural Networks (DNN), which affects the generalization ability of the algorithm for RF fingerprinting. Thus the aircraft address I/Q is removed during the network training process. The ID address is just extracted as a device label, in order to ensure that there is no direct semantic relationship between the label and the training data during the training of the neural network.

Each transmission of the long data module contains a total of 112bits of fields, which corresponds to a time length of 112 μs. Our equipment uses universal software radio peripheral with a sampling rate of 20MHz, thus 112 μs corresponds to 2240 sampling points containing 112 bits of information. Therefore, the ratio of information to sampled data is 112 / 2240 = 1 / 20.

ADS-B Raw I/Q Data Analysis: The ADS-B datasets are stored as a sequence of time-domain in-phase and quadrature (I/Q), it contains all the characteristics of the signal compare with constellation, fast Fourier transform and other format. Datasets consists of 33,647 transmissions from 100 devices. Each transmission contains 6000 fixed discrete complex sampling points. Analysis of raw signal is show in fig.2.

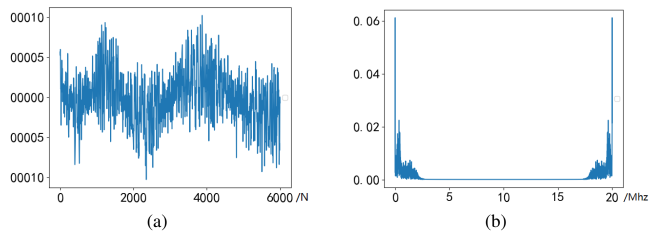

Fig 2 *Analysis of raw signals. (a) Time domain characteristics of signal. (b) Frequency domain characteristics of signal.*

As shown in fig.2. (b), the low frequency part of the signal has a rich spectral width which we are interested in.

Method: DL methods are able to extract fingerprint features from different data fragments. Gaussian Low-pass Channel Attention module (GLCAM) is proposed to enhance the features of the RF fingerprint. The operation of the channel attention is to learn a set of weighting parameters which is able to enhance the region of interest in the task. Global Average Pooling (GAP) is a special form of the two-dimensional discrete cosine transform. GAP would ignore many useful non-0-frequency components [5]. Through extensive mathematical derivations, deep neural networks are usually fitted from low to high frequencies at the beginning, middle and end phases of training.

$Y(x)$ is an indicator function, i.e.,

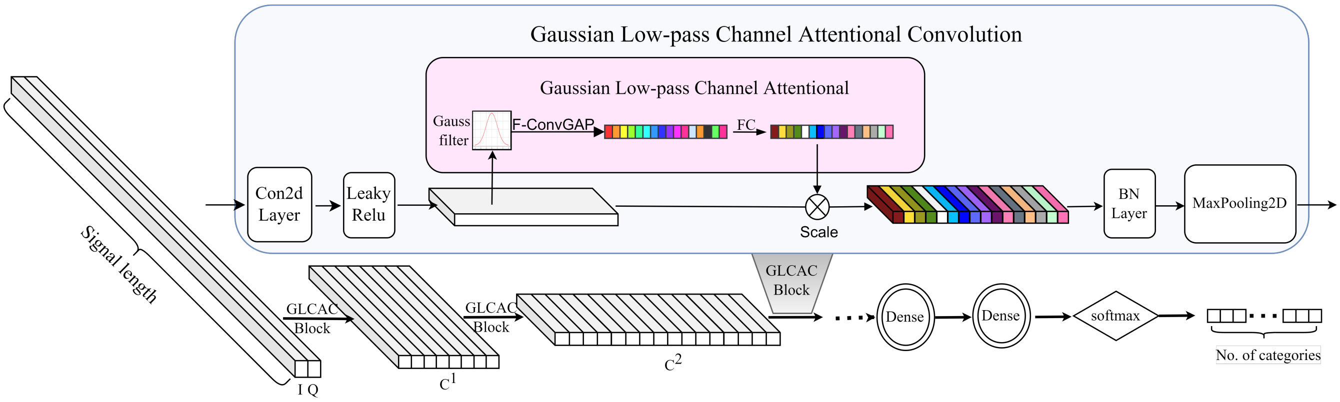

Fig 3 *The architecture of GLCAM. In this module, input data passes through a 2D convolution and LeakyRelu, then passes through a Gaussian low-pass filter, the weights of the channels are initialized by a Frequency Convolution Global Average Pooling (F-ConvGAP) and then passed through a fully connected and then through a batch normalization and max pooling, where the dark colors in the attention block represent enhancement, the light colors represent mitigation, and the white colors represent an assignment of 0. Best viewed in color.*

$$Y(x) = \begin{cases} 1, & |x| \le \tau, \\ 0, & |x| > \tau, \end{cases} \tag{1}$$

where τ is threshold. The frequency domain filter of the ideal window function corresponds to the time domain of the sinc function, since the time domain sinc function has negative values, which is not conducive to network convergence. Here is the convolution theorem:

$$\mathcal{F}^{-1}\left[g_1 \cdot g_2\right](x) = \mathcal{F}^{-1}(g_1) * \mathcal{F}^{-1}(g_2) \tag{2}$$

Where $\mathcal{F}^{-1}$ is inverse Fourier transformation,and * is indicates convolution operator. g_1 and g_2 represent frequency domain functions. Considering that Gaussian functions have similar characteristics in the time and frequency domains. The One-dimensional Gaussian:

$$G^{\delta}(x) = \frac{1}{\sqrt{2\pi}\,\delta} e^{-\frac{x^2}{2\delta^2}} \tag{3}$$

Where δ is the variance of Gaussian function G. For brevity of proof, formula (3) simplified as $G(x) = e^{-ax^2}$. $\hat{\mathrm{G}}(\omega)$ is Fourier transform of a Gaussian function:

$$\hat{\mathrm{G}}(\omega) = \int_{-\infty}^{+\infty} e^{-ax^2 - j\omega x} dx = e^{-\frac{\omega^2}{4a}} \int_{-\infty}^{+\infty} e^{-\left(\sqrt{a}x + \frac{j\omega}{2\sqrt{a}}\right)^2} dx \tag{4}$$

Respectively, where $\hat{}$ indicates Fourier transform. Let $u = \sqrt{a}x + \frac{j\omega}{2\sqrt{a}}$:

$$\hat{\mathrm{G}}(\omega) = \frac{2}{\sqrt{a}} e^{-\frac{\omega^2}{4a}} \int_{-\infty}^{+\infty} e^{-u^2} du = 2\sqrt{\frac{\pi}{a}} e^{-\frac{\omega^2}{4a}} \tag{5}$$

As shown in equation (3) and (5), the Fourier transform of a Gaussian is still a Gaussian, i.e., it can approximate $\mathrm{Y}(x)$ by $G^{\delta}(x)$ with a proper δ. We can equivalently perform the examination in the spatial domain to avoid the high-dimensional Fourier transform. The low frequency part $y^{low}(x)$ can be derived by

$$\mathcal{F}^{-1}\left(\hat{y}^{low}(x)\right) = \mathcal{F}^{-1}(\hat{y}(x)) * \mathcal{F}^{-1}\left(\hat{G}^{\sigma}(x)\right) = y * \hat{G}^{\sigma} \tag{6}$$

As shown in equation (6), the Gaussian function has the same low-pass property in the time and frequency domains. Then we can directly add a Gaussian low-pass filter to the data to reduce the number of fast Fourier transform operations to improve the learning time of the neural network.

The flow chart for individual identification of signals based on GLCAM is shown in fig.3. During the Gaussian Low-pass Channel Attention Convolution network (GLCA-Net) module the input data is shortened and the number of channels is increased. The equivalent sampling rate is halved every time passing the GLCA-Net which is related to the pooling size we set. Combined with Section II, we want the network to focus more on the first 1/20 of the frequency domain of the data, so the δ variance of Gaussian in the GLCAM is set to 1/20.

Frequency Convolution Global Average Pooling: We replace the GAP with a convolution kernel of the same size as the input network data, as we all know the GAP parameters are fixed in conventional Squeeze-and-Excitation modules. Our module is autonomously learnable and can adjust to the best filter by dynamic changing in the network. The design of F-ConvGAP is shown in fig.4.

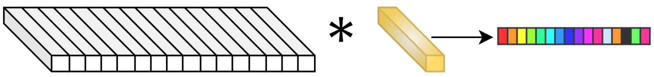

Fig 4 *F-ConvGAP Block.*

Experimental Design: We obtain the ID address of ADS-B signal by decoding, and threat them as the label of the network. The ID address of the aircraft is removed to ensure that feature extraction and identification is base on the residual signal. The overall flow of the experiments is shown in fig.5.

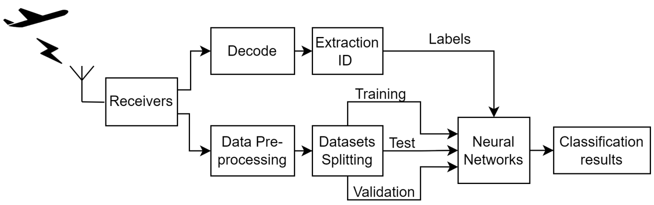

Fig 5 *Deep learning overall framework.*

Parameter Settings: The training and testing process of the algorithm were deployed on a Linux server using the TensorFlow framework. The model was trained and tested on an RTX1080Ti GPU with the support of a GPU acceleration library.

Our datasets contains 33,647 ADS-B transmissions with 100 devices. We randomly divided the entire datasets into three non-overlapping parts, namely the training set (80% of the datasets), the validation set (10%), and the test set (10%). We optimize the LSTM and Resnet as a comparison network that can achieve high accuracy. We reproduced CNN as one of the baseline models, the CNN has 8 Conv units and two dense layers. GLCA-Net consists of several Conv Units with the combination of GLCAM.

Comparison results of different classifiers: The validation accuracy during training is shown in fig.6. The GLCA-Net converges the fastest and has the highest recognition accuracy. Two metrics are chosen for evaluation, namely accuracy and model parameter size. The test accuracy is

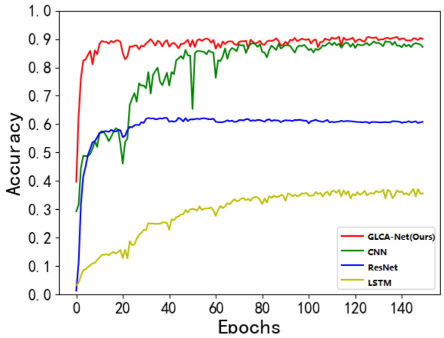

Fig 6 *Validation Accuracy During Training.*

measured using the average of 5 training sessions. The results are shown in Table 1.

Table 1. Comparison of identification performance on the test datasets.

Network	Parameters (M)	Accuracy (%)
LSTM	0.37	39.2
ResNet	0.49	60.49
CNN [3]	0.87	85.88
GLCA-Net(Ours)	1.06	**92.09**

As shown in Table 1, GLCA-Net achieve highest Accuracy.

Ablation Study: We also design 8 ablation experiments of GLCA-Net. Frequency Channel Attention (FCA) [5] is reproduced as an ablation comparison. Compared with FCA, GLCA-Net adds a Gaussian low-pass filter to the attention of the frequency domain channel. It should be noted that GLCAM (layer 1) represents Conv Unit 1 and it joins the GLCA-Net. All the remaining seven layers have been replaced by FCA modules.

Table 2. Result of Ablation Study.

Network	Parameters (M)	Accuracy (%)
CNN(Backbone)	0.87	85.88
FCA[5]	1.06	89.95
GLCAM(1 layer)	1.06	89.71
GLCAM(1-2 layers)	1.06	90.19
GLCAM(1-3 layers)	1.06	89.62
GLCAM(1-4 layers)	1.06	90.67
GLCAM(1-5 layers)	1.06	90.93
GLCAM(1-6 layers)	1.06	92.00
GLCAM(1-7 layers)	1.06	91.02
GLCAM(1-8 layers)	1.06	**92.09**

As shown in Table 2. FCA obtains higher accuracy than CNN by 4.07 %, where FCA contains 8 layers frequency domain channel attention. The accuracy of our method is 2.14 % higher than FCA In terms of correctness. On the whole, as the GLCAM is added in more layers of the network, the recognition accuracy has increased, and the addition of GLCAM do not increase the number of parameters.

An Interesting Discovery About The Feature-Map: Finally, we visualize the convolution parameters in the first layer of GLCA-Net, as shown in Fig.7. Considering the fact in Fig.2 (b), we found that the network autonomously and consciously reduces the impact of 0-frequency by training. The first layer of this module is an autocorrelation filter in the frequency domain that focuses more on near-low frequency information than 0-frequency. Thus the experiments demonstrate that using GAP alone does not fully exploit the properties of the data.

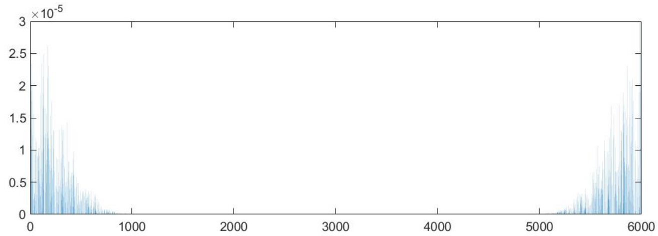

Fig 7 *F-ConvGAP Feature Map(1 layer).*

Conclusion: In this letter, a new attention module named GLCAM is proposed to explore the fingerprint of ADS-B signals. Firstly, adding attention module to frequency domain shifts the focus of feature enhancement, which is interest to signal identification. Secondly, we design a new GAP called F-ConvGAP in the block to train more efficiently. The improved deep learning network achieves good identification results in the experiments. In addition, this lightweight model can be better applied to engineering applications.

Acknowledgments: This work was supported in part by the National Natural Science Foundation of China (No. 61901500, No. 62001486), the National Natural Science Foundation of Hunan Province through the China Postdoctoral Science(2020TQ0082).

Received: DD MMMM YYYY *Accepted:* DD MMMM YYYY
doi: 10.1049/ell.10001

References

1. Yt, A.,Yun, L.A.,Hz, A.(et al.): Large-scale real-world radio signal recognition with deep learning. (2021)
2. Yan, K.,Xie, Y.: ADS-B signal radiation source recognition based on deep learning. Information Technology 45(2), 6 (2021)
3. Robinson, J., Kuzdeba, S., Stankowicz, J., Carmack, J.M.: Dilated Causal Convolutional Model For RF Fingerprinting. 2020 10th Annual Computing and Communication Workshop and Conference (CCWC) 0157–0162 (2020)
4. Luo, T., Ma, Z.,Xu, Z.J., Zhang, Y.: Theory of the Frequency Principle for General Deep Neural Networks. Statistics 2, (2019)
5. Qin, Z.,Zhang, P.,Wu, F.,Li, X.: FcaNet: Frequency Channel Attention Networks. IEEE/CVF International Conference on Computer Vision (ICCV) (2021)