The Scientific Temper (2024) Vol. 15 (4): 3182-3186
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.27
https://scientifictemper.com/

RESEARCH ARTICLE

Deep learning driven image steganalysis approach with the impact of dilation rate using DDS_SE-net on diverse datasets

Hemamalini V.[1*], Victoria Priscilla C.[2]

Abstract

The challenge of effective and precise steganalysis is crucial in the field of digital steganography. Steganalysis is a constantly evolving field of study that looks for hidden data in digital media. With the recent developments in communication and information technology, as well as information law compliance, image Steganalysis has drawn a lot of attention. The methods for steganography that are now available make it harder to identify steganographic material. This study presents a comprehensive investigation of the DDS_SE-Net architecture based on convolution neural networks employing various datasets in steganalysis using key performance measures, including accuracy, recall, precision, and F1-score. Additionally, this study looks at how rate of change of dilation in DDS_SE-Net contributes to the improved outcomes. In this work dilation rate of 3 gave comparatively better accuracy of 92.9% against WOW, 89.2 and 89.8% against S-UNIWARD and HILL, respectively. The results show that the deep learning framework selected and the data used in training have a major impact on how well the model performs steganalysis.

Keywords: Steganalysis, Deep learning, Steganography, Dilation, Convolution neural network, Separable convolutions.

Introduction

Data security can be efficiently maintained by using steganography, a technology that hides secret data inside a carrier for clandestine transmission. Steganalysis, on the other hand, is an anti-steganography method that seeks to determine whether or not hidden data are buried into a carrier. This is crucial since it stops steganography from being abused. The most prevalent mediums in steganography and steganalysis are digital photographs. By minimizing the false positive (FP) and false negative (FN) rates, the steganalysis algorithms may detect the existence of secret data in the stego signal, which deviates from the statistical characteristics of the original signal, by employing DL approaches. A common scenario encountered by steganalysts transitioning from lab-based steganalysis to real-world applications is when their detector is trained on photos from one cover source and then applied to images from another. The majority of steganalysis research used pre-existing datasets, such as BOSSbase (Westfeld & Pfitzmann, 2000), BOWS2 (Ankita Gupta, 2023), Alaska(Yousfi *et al*., 2019), etc., for their experimentation. An effective design provides excellent accuracy and performs effectively with a variety of datasets. Using a real-time dataset, the CNN-based DDS_SE-Net (Dilated Depthwise Separable convolutions with Squeeze and Excitation blocks) (C Victoria Priscilla, 2024) architecture performed well. This research examines how DDS_SE-Net functions using a variety of current and real-time datasets. The analysis is also done on dilation rate changes in DDS_SE-Net. In order to create the minimal structure known as the dilated filter, the dilation rate operates by adding spaces between the convolutional filter's weights. By inserting these gaps, the filter can bypass some input values and concentrate on those that are divided by the specified gap. This helps in expanding the view without increasing the parameters.

Related Work

Numerous CNN architectures were created with the introduction of deep learning (DL) to handle various image

Research Department of Computer Science, Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Affiliated to University of Madras, Chennai, Tamil Nadu,India.

***Corresponding Author:** Hemamalini V., Research Department of Computer Science, Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Affiliated to University of Madras, Chennai, Tamil Nadu, India, E-Mail: hemaveera123@gmail.com

How to cite this article: Hemamalini, V., Priscilla, C. V. (2024). Deep learning driven image steganalysis approach with the impact of dilation rate using DDS_SE-net on diverse datasets. The Scientific Temper, **15**(4):3182-3186.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.27

Source of support: Nil

Conflict of interest: None.

Received: 02/11/2024 **Accepted:** 20/11/2024 **Published** : 20/12/2024

sources in steganalysis. With three fully connected layers that culminate with a Softmax layer and five convolutional layers coupled by Gaussian activation, the Qian-Net architecture (Qian *et al.*, 2015) uses high-pass filters to reduce the image's content. In Zhu-net architecture (Zhang *et al.*, 2020), separable convolutions are utilized to increase the signal-to-noise ratio by constricting the image content and utilizing the channel correlation of the residuals. SPP is used to gather the local features. Adopting data augmentation helps the network function better. GBRAS-NET(Reinel *et al.*, 2021) incorporates the most advantageous features from previous architectures. During the pre-processing stage, a total of 30 SRM filters are employed, utilizing a 3*TanH activation function. The feature extraction process involves the utilization of a sequence of convolution layers that incorporate depth-wise separable convolutions. This model incorporates a soft max layer along with global average pooling in the absence of a fully connected layer. The utilization of inverted residual blocks paired with a self-attention mechanism in CIRNet (Ankita Gupta, 2024) aims to minimize both the detection error rate and computing cost for steganalysis. The inverted residual blocks incorporate lightweight depth-wise and pointwise convolutions, as well as a self-attention module. The integration described in the study reduces the number of floating-point operations and network parameters while simultaneously improving the prominence of feature maps associated with the embedding regions. In the feature extraction phase of architecture in (Ntivuguruzwa & Ahmad, 2023), two-dimensional depthwise separable convolutions were utilised to enhance the signal-to-noise ratio, while conventional convolutions were employed to model local features. A revolutionary architecture for deep convolutional neural networks that Inception inspired was introduced by (Chollet, 2017). In this architecture, the modules of Inception have been replaced by depthwise separable convolutions.

Convolutional neural networks (CNNs) use dilated convolution, sometimes referred to as atrous convolution, a sort of convolution operation that allows the network to have a bigger receptive field without raising the number of parameters. The idea of dilated convolution originates with wavelet decomposition (Shensa, 1992). The dilated convolution operator, which performs wavelet decomposition, has been mentioned frequently. The conventional CNN's drawback is its enormous processing power consumption. To tackle this problem, a dilated CNN model is built. The hybrid dilated CNN (HDC) (Lei *et al.*, 2019) is built to overcome the detail loss problem in the dilated CNN model by piling dilated convolution kernels with different dilation rates one after the other.

Squeeze-and-excitation networks were developed by stacking the SE blocks together, and they demonstrated exceptional performance on challenging datasets (Hu, 2018). Extra features can be extracted from the digital image using the feature extraction along with the fusion layer. Therefore, memory utilization is increased while inference efficiency is boosted when the RepVgg block is used in SFR-Net (Xu *et al.*, 2021). The SE block increases the detection accuracy rate by learning feature weights to generate valid or ineffective feature maps with moderate weights or effective ones with massive weights. The benefits of convolutions and attention mechanisms are combined in the convolutional vision transducer CVTStego-Net (A *et al.*, 2024) to capture both regional and global dependencies in spatial domain image steganalysis. In the pre-processing stage, a bifurcation made up of 30 SRM filters is employed to enhance steganographic noise. The noise extraction and analysis stage uses SE-Block with residual operations to lessen the effect of redundant data and increase sensitivity in steganographic noise. During the classification phase, the local and global spatial associations of the steganographic noise are connected by combining SE-Block with a convolutional vision converter.

Materials and Methods

Three crucial elements make up the CNN-based DDS_SE-Net architecture used in this study's feature extraction phase. SE blocks, dilation, and depthwise separable convolution, as seen in Figure 1. The filter is dilated in a dilated convolution operation by adding gaps within the filter values. The dilation rate is a hyperparameter and can be changed to control the gap sizes. The dilated convolution decreases to a regular convolution at a dilation rate of 1. It contributes to a greater receptive field without raising parameters and lowers processing power usage. As part of a network architecture technique called depthwise separable convolution, a convolution operation is divided into two parts: depthwise convolution, which operates on individual input channels and pointwise convolution, which boosts the dimension of the feature map by incorporating information from different channels. This method lessens the overfitting issue and parameter count.

A network's multiple filters will first search each input channel for spatial properties, then combine the data over all possible output channels. When generating the output

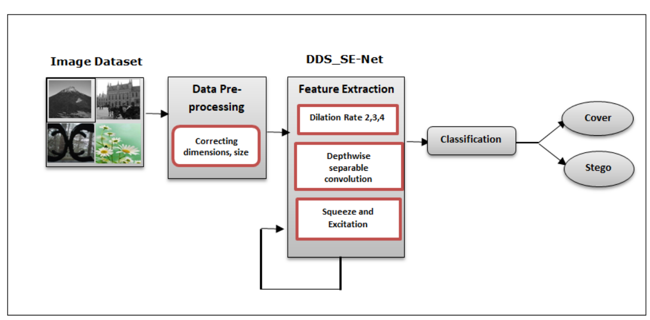

Figure 1: Process flow for varied datasets with DDE_SE-Net

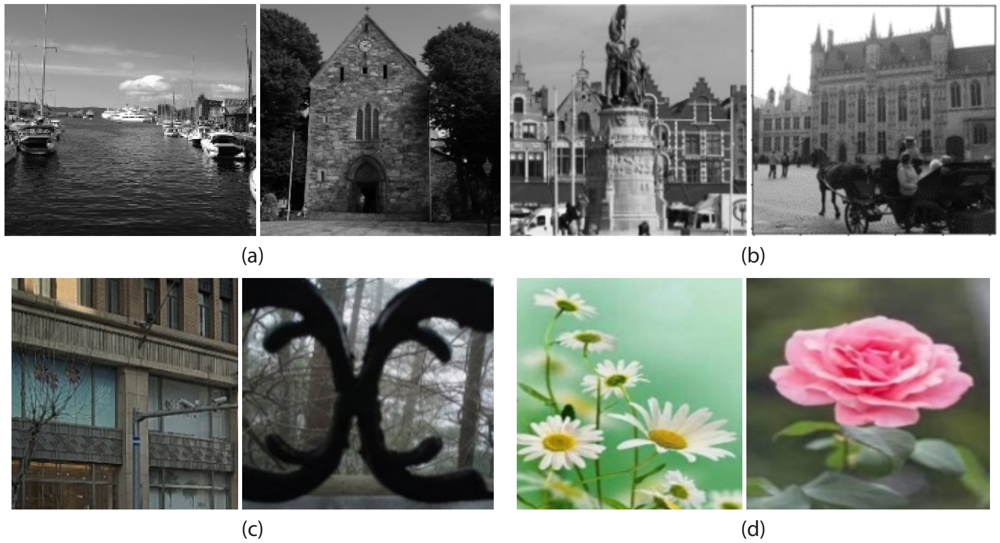

Figure 2: Sample dataset from (a) BOSSbase (b) BOWS2 (c) Alaska (d) Real-time

feature maps, the network gives each of its channels the same weight. By including a content-aware technique to adaptively weight each channel, SE Nets aim to change this. Better precision and robustness are aided by it.

Experiments

The experiments are conducted with three existing and one real time dataset to analyse the working of DDS_SE-Net architecture.

Data collection

BOSSbase includes an image steganalysis data set made up of 10,000 (512*512*1) grayscale images from 7 distinct cameras. About 11,536 color JPEG-formatted images are included in the BOWS2 database for image steganography. CNN applications employ ALASKA 8, a large dataset for image steganography that includes over 250,000 images divided into four different steganography categories. The real-time dataset consists of images collected online from Google, which are of different dimensions. There are differences in size and shape among the collected photographs. All the pictures were downsized to 480 by 640 pixels using an image editing application. From each dataset as in Figure 2, 5000 randomly chosen images are taken for this venture because of the processing complexity.

Parameter setting using dilation rate

The convolution operation's dilation rate parameter regulates how widely apart the kernel (filter) elements are separated. The kernel expands when there is dilation. The kernel can cover a greater portion of the input data without rising in size due to this stretching phenomenon. It provides more efficiency, aids in the creation of more complex feature maps, and detects larger patterns. Deep learning models may effectively identify patterns at many scales by varying the dilation rate, which enhances their effectiveness on tasks such as object recognition as well as image classification (Pandey, 2024). From Figure 3, it can be seen how the feature detector kernel is spaced when the dilation rate is 1,2 and 3. When the dilation rate is 1, it would be a normal convolution operation.

Results and Discussion

Once the data was collected and pre-processed, the three steganographic algorithms that were employed to produce the stego images were WOW (Binghamton, 2012), S-UNIWARD (Holub *et al.*, 2014), and HILL (Li *et al.*, 2014) with payload 0.4. Stego detection was accomplished by feeding the DDS_SE-Net model, with both the cover and the stego containing a dilation rate of 3. The results were tabulated in Table 1. Furthermore, the metrics for the real-time dataset were determined against the three steganographic methods, with changes in dilation rate of 2, 3, and 4, and the results are compared in Table 2.

The effectiveness of DDS_SE-Net model has given good accuracy with four different datasets against three steganograpic algorithms with 0.4 bpp payload as in Table 1. With BOSSbase and real-time dataset, accuracy was more against WOW with 89.8 and 92.9% respectively. With BOWS2 and Alaska dataset, against S-UNIWARD, better accuracy of 90.2 and 90.07%, respectively was reached (Figure 4).

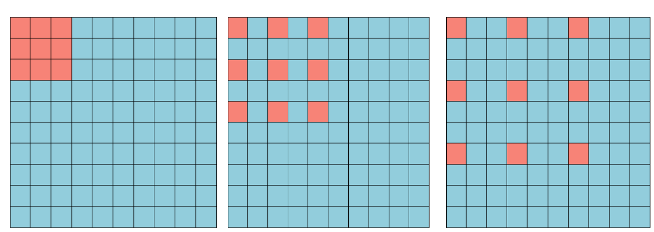

Figure 3: Sliding of feature detector kernel on the input feature map when dilation rate is (a)1 (b)2 and (c) 3

Table 1: Comparison of evaluation metrics of DDS_SE-Net against WOW, S-UNIWARD, and HILL with 0.4bpp using BOSSbase, BOWS2, Alaska and real-time datasets with dilation rate of 3

Dataset	*Steganographic algorithm*	*Accuracy*	*Precision*	*Recall*	*F1 Score*
	WOW	89.8	89.5	90.6	90.0
BOSSBASE	S-UNIWARD	89.3	87.4	92.0	89.6
	HILL	89.3	90.9	88.3	89.6
	WOW	87.2	87	87.7	87.4
BOWS2	S-UNIWARD	90.2	89	91.6	90.3
	HILL	88.9	88.9	90.4	89.6
	WOW	87.5	87.2	88.0	87.6
ALASKA	S-UNIWARD	90.07	90.2	90.3	90.2
	HILL	89.6	88.3	91.2	89.7
	WOW	92.9	91.5	94.0	92.7
REAL-TIME	S-UNIWARD	89.2	91.4	86.9	89.0
	HILL	89.8	91.5	87.9	89.6

Table 2: Comparison of evaluation metrics of DDS_SE-Net against WOW, S-UNIWARD, and HILL with 0.4 bpp using real-world datasets with dilation rate of 2, 3 and 4

Steganographic algorithm	*Dilation Rate*	*Accuracy*	*Precision*	*Recall*	*F1 Score*
WOW		87.17	87.35	87.28	87.31
S-UNIWARD	2	87.26	87.4	87.5	87.5
HILL		86.3	85.2	86.8	86.0
WOW		92.9	91.5	94.0	92.7
S-UNIWARD	3	89.2	91.4	86.9	89.0
HILL		89.8	91.5	87.9	89.6
WOW		87.5	87.08	87.64	87.36
S-UNIWARD	4	88.08	87.3	89.8	88.5
HILL		88.75	88.10	90.2	89.14

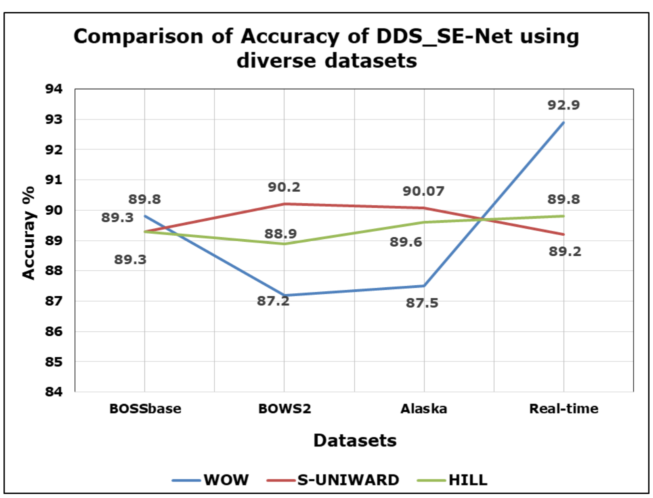

Figure 4: Comparison of accuracy of DDS_SE-Net against WOW, S-UNIWARD and HILL with 0.4 bpp using four different datasets

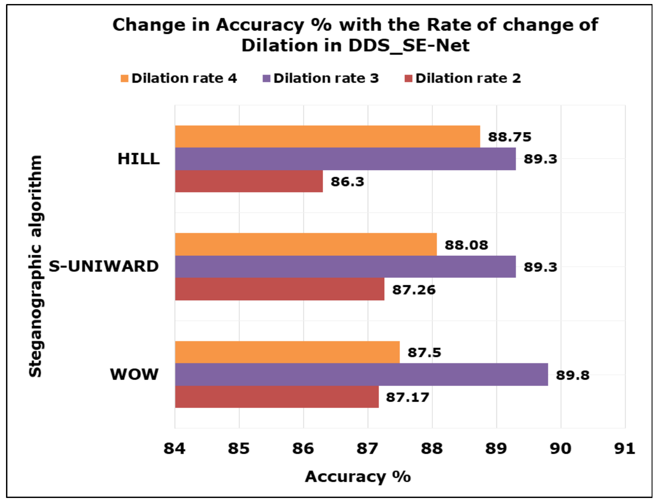

Figure 5: Comparison of accuracy of DDS_SE-Net against WOW, S-UNIWARD and HILL with 0.4 bpp using real world dataset with varying dilation rates

Table 2 gives the results of classification metrics of DDS_SE-Net model with varying dilation rates using real-time dataset against three popular steganographic algorithms with payload 0.4 bpp. It could be noted that when the dilation rate is 3, the accuracy and F1-score are more than the dilation rates 2 and 4 in this scenario. In case the dilation rate is 2, the coverage area by the filter on the input data is not very larger than the normal convolution. When dilation rate is 4, the network might have missed the essential fine details due to oversimplification, thus, resulting in less accuracy when compared to the dilation rate of 3 as in Figure 5, which has larger coverage and has not missed any important details too. This might change with other datasets and in other network models.

Conclusion

Using a variety of datasets for steganalysis, this work provides a thorough study of the DDS_SE-Net architecture based on convolution neural networks. The datasets used are BOSSbase, BOWS2, Alaska and real-time dataset. The results show that DDS_SE-Net architecture gives considerably better results using varying datasets against three prevailing Steganographic algorithms. Using BOSSbase (89.8 and 90%) and Real-time dataset (92.9 and 92.7%), accuracy and F1-score were more against the WOW algorithm. Accuracy and F1-score were high against S-UNIWARD with BOWS2 (90.2 and 90.3%) and Alaska (90.07 and 90.2%) datasets. Moreover, experiments were performed with varying dilation rates of 2, 3 and 4, where better results were achieved with dilation rate 3 using a real-world dataset with DDS_SE-Net model with 92.9% against WOW, 89.2 and 89.8% against S-UNIWARD and HILL, respectively. In the future, a framework will be created with best classification model with the best optimizer and dataset for steganalysis, which could be applied for any other application.

Acknowledgment

I want to sincerely thank Dr. C. Victoria Priscilla, my supervisor, for her invaluable guidance and support. I would also like to express my gratitude to my college Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women for encouraging me in taking forward my research work.

References

Bravo-Ortiz, M. A., Mercado-Ruiz, E., Villa-Pulgarin, J. P., Hormaza-Cardona, C. A., Quiñones-Arredondo, S., Arteaga-Arteaga, H. B., & Tabares-Soto, R. (2024). CVTStego-Net: A convolutional vision transformer architecture for spatial image steganalysis. Journal of Information Security and Applications, 81, 103695. https://doi.org/https://doi.org/10.1016/j.jisa.2023.103695Get rights and content

Ankita Gupta. (2023). *BOWS2*. https://doi.org/10.17632/kb3ngxfmjw.1

Ankita Gupta, R. C. & P. S. (2024). CIRNet: An Improved Lightweight Convolution Neural Network Architecture with Inverted Residuals for Universal Steganalysis. *Arabian Journal for Science and Engineering*. https://link.springer.com/article/10.1007/s13369-023-08630-x

Binghamton, S. (2012). DESIGNING STEGANOGRAPHIC DISTORTION USING DIRECTIONAL FILTERS. *Ieeexplore.Ieee.Org*, 234–239. http://ieeexplore.ieee.org/abstract/document/6412655/

C Victoria Priscilla, V. H. (2024). A Three-Component Feature Extraction Using DDS _ SE-NET for Efficient Deep Learning-Based Image Steganalysis for Real-World Images. *INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY*, 3335–3343.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 1800–1807. https://doi.org/10.1109/CVPR.2017.195

Holub, V., Fridrich, J., & Denemark, T. (2014). Universal distortion function for steganography in an arbitrary domain. *Eurasip Journal on Information Security*, *2014*, 1–13. https://doi.org/10.1186/1687-417X-2014-1

Hu, J. (2018). Squeeze-and-Excitation_Networks. *Cvpr*, 7132–7141. http://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html

Lei, X., Pan, H., & Huang, X. (2019). A dilated cnn model for image classification. *IEEE Access*, *7*, 124087–124095. https://doi.org/10.1109/ACCESS.2019.2927169

Li, B., Wang, M., Huang, J., & Li, X. (2014). A NEW COST FUNCTION FOR SPATIAL IMAGE STEGANOGRAPHY College of Information Engineering , Shenzhen University , Shenzhen , GD 518060 , China Institute of Computer Science and Technology , Peking University , Beijing 100871 , China. *International Conference on Image Processing(ICIP)*, 4206–4210.

Ntivuguruzwa, J. D. L. C., & Ahmad, T. (2023). A convolutional neural network to detect possible hidden data in spatial domain images. *Cybersecurity*, *6*(1). https://doi.org/10.1186/s42400-023-00156-x

Pandey, A. K. (2024). *Dilation Rate in a Convolution Operation*. https://medium.com/@akp83540/dilation-rate-in-a-convolution-operation-a7143e437654

Qian, Y., Dong, J., Wang, W., & Tan, T. (2015). Deep learning for steganalysis via convolutional neural networks. *Media Watermarking, Security, and Forensics 2015*, *9409*, 94090J. https://doi.org/10.1117/12.2083479

Reinel, T. S., Brayan, A. A. H., Alejandro, B. O. M., Alejandro, M. R., Daniel, A. G., Alejandro, A. G. J., Buenaventura, B. J. A., Simon, O. A., Gustavo, I., & Raul, R. P. (2021). GBRAS-Net: A Convolutional Neural Network Architecture for Spatial Image Steganalysis. *IEEE Access*, *9*, 14340–14350. https://doi.org/10.1109/ACCESS.2021.3052494

Shensa, M. J. (1992). The Discrete Wavelet Transform: Wedding the À Trous and Mallat Algorithms. *IEEE Transactions on Signal Processing*, *40*(10), 2464–2482. https://doi.org/10.1109/78.157290

Westfeld, A., & Pfitzmann, A. (2000). Attacks on steganographic systems breaking the steganographic utilities ezstego, jsteg, steganos, and s-tools–and some lessons learned. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1768*, 61–76. https://doi.org/10.1007/10719724_5

Xu, G., Xu, Y., Zhang, S., & Xie, X. (2021). SFRNet: Feature Extraction-Fusion Steganalysis Network Based on Squeeze-and-Excitation Block and RepVgg Block. *Security and Communication Networks*, *2021*. https://doi.org/10.1155/2021/3676720

Yousfi, Y., Butora, J., Fridrich, J., & Giboulot, Q. (2019). Breaking Alaska: Color separation for steganalysis in JPEG domain. *IH and MMSec 2019 - Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, 138–149. https://doi.org/10.1145/3335203.3335727

Zhang, R., Zhu, F., Liu, J., & Liu, G. (2020). Depth-Wise Separable Convolutions and Multi-Level Pooling for an Efficient Spatial CNN-Based Steganalysis. *IEEE Transactions on Information Forensics and Security*, *15*, 1138–1150. https://doi.org/10.1109/TIFS.2019.2936913