Svādhyāya - International Journal of Transdisciplinary Research and Development (SIJTRD) ISSN Online : 2583-1739

Vol 2 (2), Dec 2022, pp 13 - 23

ANALYSIS OF DEEP LEARNING ARCHITECTURES APPLIED IN STEGANALYSIS

V. HemaMalini[1] and C. Victoria Priscilla[2]

[1] Research Scholar, Research Department of Computer Science,
SDNB Vaishnav College for Women, University of Madras, Chennai, India
[2] Associate Professor & Head, PG Department of Computer Science,
SDNB Vaishnav College for Women, University of Madras, Chennai, India
Email : hemaveera123@gmail.com[1], aprofvictoria@gmail.com[2]

ABSTRACT

Steganalysis has achieved importance in security as the capture of hidden messages lead to the avoidance of ruinous security incidents. Many CNN architectures have proved good accuracy in detection with 3 phases such as pre-processing, feature extraction and classification. Analysing those studies would give us a better knowledge for further study. This work does an analysis based on the performance of the popular and strong architectures in Steganalysis using Deep learning. Detection accuracy has been improved evenly every year by various architectures. Applying the same payload and using the same database, from the year 2015, when checked with same Steganographic algorithm (S-UNIWARD), accuracy increased from 69% by QIAN-Net in the year 2015 to 87.1% in the year 2021 by GBRAS-Net. The increase in accuracy could be noticed when applying different payloads.

KEYWORDS: *STEGANALYSIS, DEEP LEARNING, CONVOLUTIONAL NEURAL NETWORK, MACHINE LEARNING*

INTRODUCTION

Steganalysis and steganography are the two pages of a same paper. While Steganography hides messages steganalysis tries to detect their existence or retrieve the embedded data. Many robust methods of steganalysis have been presented in the literature over the last few years. Traditional Steganalysis has two stages. First stage does extracting the features manually. The second stage contains binary classifiers using Ensemble Classifiers, Support Vector Machines; K-Means etc.(Arivazhagan Selvaraj, Amrutha Ezhilarasan, Sylvia Lilly Jebarani Wellington, no date)(Kodovský, Fridrich and Holub, 2012)(Cogranne and Fridrich, 2015)(Pevny, Fridrich and Ker, 2009). Detectors of steganography

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

developed on deep CNN have proved themselves superior to the previous detection classifiers. The most demanded source of cover objects are particularly images. To hide the information, steganography is applied using the algorithms HUGO(Pevný, Filler and Bas, 2010), WOW(Binghamton, 2012), HILL(Li *et al.*, 2014), S-UNIWARD (Holub, Fridrich and Denemark, 2014) and MiPOD (Sedighi, Cogranne and Fridrich, 2016) in the spatial domain. By the advances in DL, CNN construction began for spatial image Steganalysis. Qian-Net architecture (Qian *et al.*, 2015) was proposed in 2015. High-pass filters are used to minimize the content of the image with five convolutional layers connected with Gaussian activation and 3 fully-connected layers ending with a Softmax layer. The next updated architecture Xu-Net (Xu, Wu and Shi, 2016) employed an absolute value layer and the front layers had TanH activation. As a part of image pre-processing it also contained fixed High pass filters. The prominent Steganalysis architectures are presented below.

EXISTING DEEP CNN ARCHITECTURES IN STEGANALYSIS

BOSSbase 1.01 database and BOWS2 database have 10000 grayscale images of 512 *512 *1 size each. The architectures analysed in this study use both or any one of the above databases for their research.

YE-NET (Ye, Ni and Yi, 2017)

It was proposed in the year 2017. The first layer which serves as the pre-processing layer is begun with the basic HPF. It is used to calculate the residual maps in SRM. TLU is the activation function adopted to capture the embedding signal structure. The accomplishment of selection channel is done to improve the performance of the CNN.

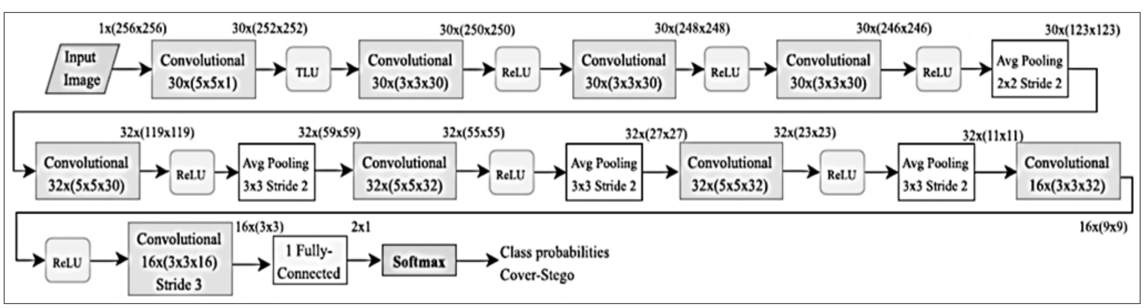

FIGURE 1 - YE-NET ARCHITECTURE

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Svādhyāya - International Journal of Transdisciplinary Research and Development (SIJTRD) ISSN Online : 2583-1739

Vol 2 (2), Dec 2022, pp 13 - 23

YEDROUDJ-NET(Yedroudj, Comby and Chaumont, 2018)

Proposed in the year 2018, this architecture combines the specialties of various papers. 30 filters for pre-processing, TLU activation, Batch Normalization and scale layer after Convolutions are notable points. It does not use Channel selection as in Ye-Net. There are three fully connected layers.

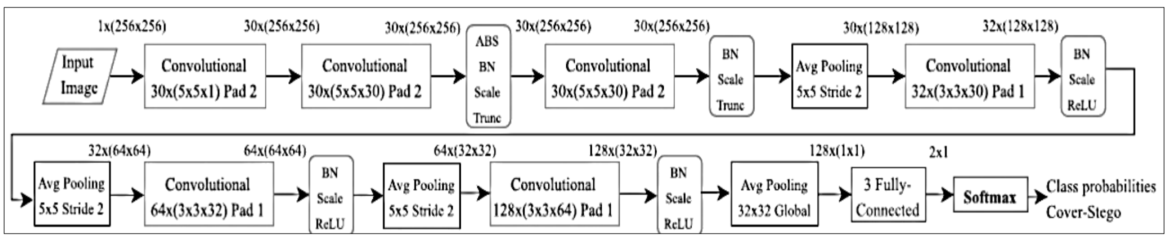

FIGURE 2 - YEDROUDJ-NET ARCHITECTURE

SR-NET(Boroumand, Chen and Fridrich, 2019)

This CNN was proposed in 2018. A deep residual architecture is designed. This cuts down the use of externally imposed elements and heuristics. It provides accuracy in diagnosing spatial domain as well as JPEG Steganography. Significantly expanded front part is the key part that helps to compute noise residuals. Four types of layers are designed with different convolution flow.

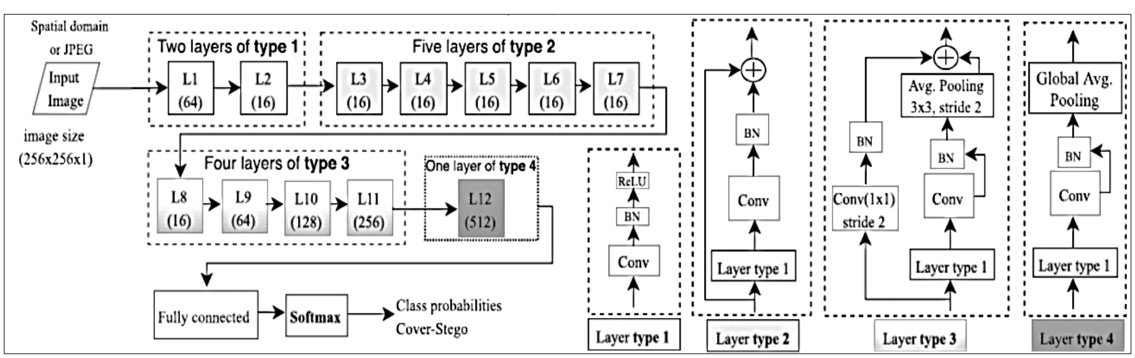

FIGURE 3 - SR-NET ARCHITECTURE

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Svādhyāya - International Journal of Transdisciplinary Research and Development (SIJTRD) ISSN Online : 2583-1739

Vol 2 (2), Dec 2022, pp 13 - 23

ZHU-NET(Zhang *et al.*, 2020)

To use the channel correlation of the residuals and restrict the image data, and to enhance the signal-to-noise ratio separable convolutions are used. To aggregate the local features, SPP is used. To improve the performance of the network, data augmentation is adopted. This CNN was proposed in 2019.

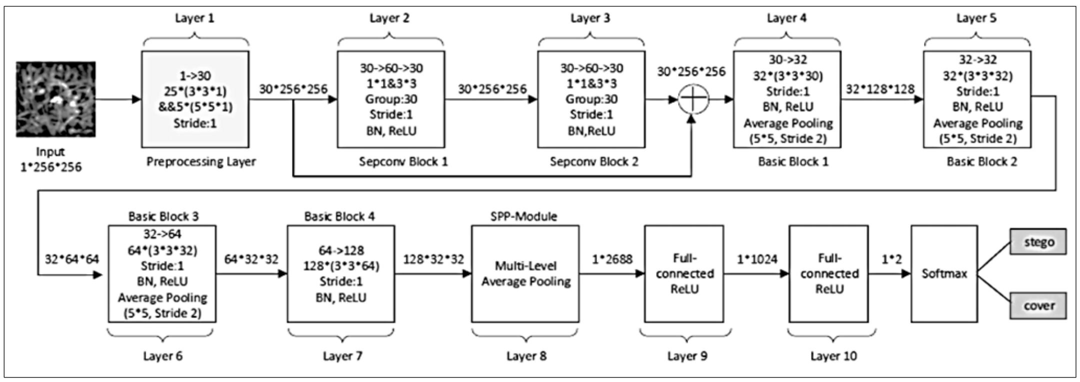

FIGURE 4 - ZHU-NET ARCHITECTURE

SIEMESE CNN(You, Zhang and Zhao, 2021)

Assuming that the noise of the natural images is similar between different subregions of images, this architecture is proposed in 2021. A Siamese-CNN based SiaStegNet is adopted, with two aligned subnets which consists of shared parameters.

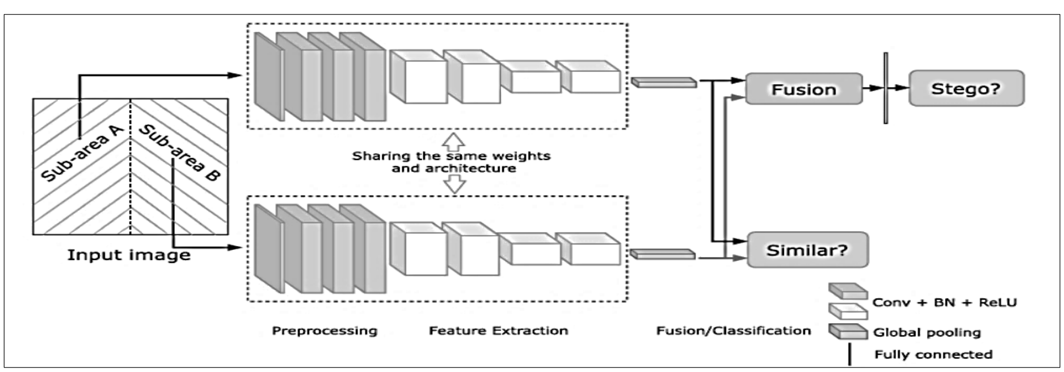

FIGURE 5 - SIAMESE-CNN ARCHITECTURE

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

GBRAS-NET(Reinel *et al.*, 2021)

This architecture was proposed in the year 2021. It uses all best features from previous architectures. 30 SRM filters in pre-processing stage are used with 3*TanH activation function. A sequence of convolution layers in feature extraction with separable convolutions and Depth wise convolutions. No fully connected layer. It has a Global average pooling with a soft max layer.

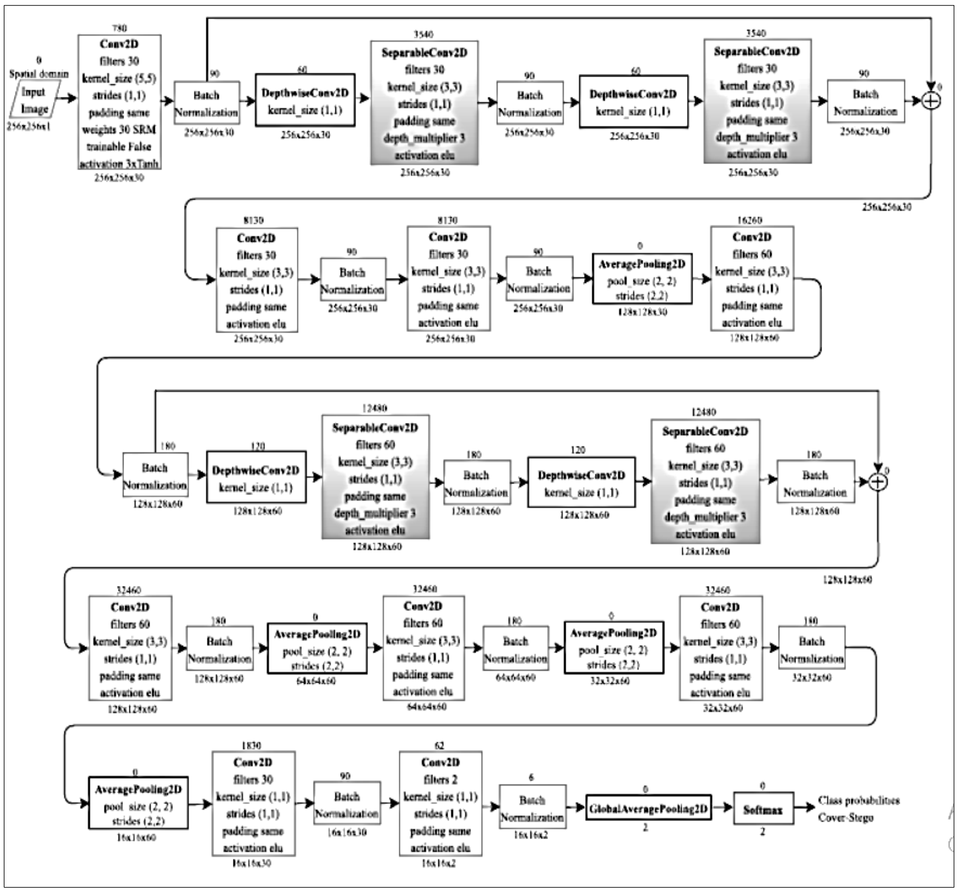

FIGURE 6 - GBRAS-NET ARCHITECTURE

ANALYSIS AND DISCUSSION ABOUT BUILDING BLOCKS

A systematic comparison of the architectures on steganalysis based on the building blocks of the Convolutional neural networks help in developing advanced networks to increase accuracy. Let's focus on them one by one.

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

FILTERS/KERNELS AND STRIDE

Kernels are matrix moving over the input data and calculating dot product with sub regions of input data and getting the output of dot products as the matrix. Kernel is a filter. In the other way we can say a filter as the collection of kernels. Filters help in removing the noise component residuals. It helps to expand the signal to noise ratio in the middle of the image signal and the weak stego signal. Kernel moves on the input data by the **stride** value left to right with one or more pixels column change in horizontal movement and top to bottom with one or more pixels row change in vertical movements. (Bai, 2019)

TABLE - 1(COMPARISON OF FILTERS)

ARCHITECTURE	FILTER APPLIED
Ye-Net	Residual maps in Spatial Rich Model computation done Kernel size is 5*5.
Yedroudj-Net	30 basic HPF are used from SRM. Kernel size set to 5*5.
SR-Net	The first two layers worked with 3*3 filters. No residual shortcuts or pooling are present.
Zhu-Net	Kernel size 3*3 to bring down the number of parameters in a tiny region of the image.
GBRAS-Net	To extract the noise residual map from the input image, a set of 30 HPF from SRM is used. Size is 5*5.
Siamese CNN	Initialization of weights to SRM filters (5*5) of a total count 30.

CONVOLUTION LAYERS AND FULLY CONNECTED LAYERS

Convolution layers are the constructing blocks of CNN. Each kernel convolves on the image and forms the activation map. The filters size is smaller than the image size. The last layers of the Convolution neural networks are the fully connected layers. It is the feed forward neural network. The output from the final convolution layer and the final pooling is flattened and fed to the classification layer.

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

TABLE- 2(COMPARISON OF LAYERS)

ARCHITECTURE	NUMBER OF LAYERS
Ye-Net	Totally 10 layers including one fully connected layer towards the end
Yedroudj-Net	9 layers including 1 pre-processing layer and 3 fully connected layers.
SR-Net	Total layers count to 13 of four types with some shortcut connections with a fully connected layer.
Zhu-Net	IT has a total of 10 layers -1 pre-processing layer, 2 separable convolution layers, 4 feature extraction layers and 2 fully connected layers. Uses 1 absolute value layer.
GBRAS-Net	9 convolution layers, 4 separable convolution and 4 depth wise separable convolution layers.
Siamese CNN	15 convolution layers of two sub regions and 1 fully connected layer. A dropout layer to prevent overfitting.

ACTIVATION FUNCTION AND NORMALIZATION

A node in between or at the end of the neural network that helps in deciding if the neuron will be activated or not is called the activation function. It says if the neuron is important to the network or not. Without an activation function, the neural network will just be like a linear regression model. Normalization is the pre-processing tool of data to standardize it. Data from different sources are brought inside the same range making the data suitable for training (Udofia, 2018).

TABLE – 3(COMPARISON OF ACTIVATION FUNCTION)

ARCHITECTURE	ACTIVATION FUNCTION
Ye-Net	TLU is used in first two layers and ReLU in the other layers.
Yedroudj-Net	Truncation function used in 1st and 2nd Blocks while blocks 3 through 5 use ReLU.
SR-Net	ReLU used after every convolution layer
Zhu-Net	ReLU is the activation function used.
GBRAS-Net	ELU activation function. 3TanH activation is used in Pre-processing layer.

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Siamese CNN	ReLU is the activation function used.

POOLING

To reduce the spatial size of the input image and to reduce the number of computations, we add a pooling layer. Pooling can be Max pooling or mean pooling. By reducing the size down sampling is performed. Thus, only important data is sent to the next layers in CNN.

TABLE - 4(COMPARISON OF POOLING)

ARCHITECTURE	POOLING
Ye-Net	Average pooling-layers 4 to 7
Yedroudj-Net	Average Pooling in 3 layers, A Global Average pooling layer
SR-Net	Average Pooling in 4 layers, A Global Average pooling layer
Zhu-Net	Average Pooling in 3 layers, A Multi-level Average pooling layer
GBRAS-Net	Average Pooling in 4 layers, A Global Average pooling layer
Siamese CNN	A global average pooling layer to reduce the feature dimensionality.

RESULTS AND METRICS

Every year approximately a new Steganalysis architecture based on CNN is published with improvement in detection accuracy. Taking WOW Steganographic algorithm as reference with a payload of 0.4bpp, a QIAN-Net and XU-Net architectures proposed in years 2015 and 2016 respectively gave accuracy of 70.7% and 79.3%. Ye-Net architecture with Truncated Linear Unit (TLU) and the knowledge of selection channel gave accuracy of 76.7%. A clever fusion of Xu-Net and Ye-Net is applied in Yedroudj-Net architecture which gave the accuracy of 85.1%. A deep residual architecture is modelled to provide detection accuracy is proposed by SRNet obtaining accuracy of 86.4%. Use of smaller convolution kernels, separable convolutions spatial pyramid pooling is remarkable in the Zhu-Net architecture

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Svādhyāya - International Journal of Transdisciplinary Research and Development (SIJTRD) ISSN Online : 2583-1739

Vol 2 (2), Dec 2022, pp 13 - 23

acquiring 88.1% accuracy. Use of filter banks in Pre-processing stage and depth wise and separable convolutions in feature extraction enhanced the work in GBRAS-Net reaching an accuracy of 89.8%. The relationship between sub-regions of images before and after steganography is concentrated in Siamese CNN named SiaStegNet with accuracy 92.09%.

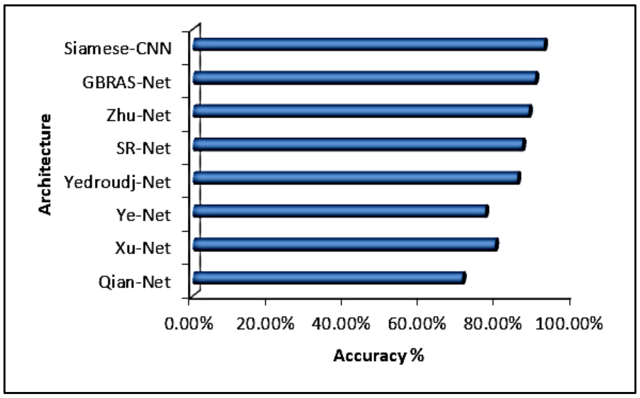

FIGURE 7 - COMPARISON OF ACCURACY % OF STEGANALYSIS BASED ON CNN

CONCLUSION

Machine learning algorithms were used in Classification of traditional Steganalysis. Deep learning has taken its place presently. This article does a performance analysis of the most robust Steganalysis architectures by comparing various building blocks of Deep learning network proposed in recent years. This helps to better understand what advancements and novelties are introduced in every stage of development in Steganalysis. This study guides the researchers to look for new methodologies and incorporate new techniques to further increase the accuracy in detecting hidden messages.

REFERENCES

Arivazhagan Selvaraj, Amrutha Ezhilarasan, Sylvia Lilly Jebarani Wellington, A.R.S. (no date) ‘Digital image steganalysis: A survey on paradigm shift from machine learning to deep learning based techniques’, IET Image Processing, 15(2), pp. 504–522. Available at: https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/ipr2.12043.

Bai, K. (2019) A Comprehensive Introduction to Different Types of Convolutions in Deep Learning, Towards Data Science. Available at: https://towardsdatascience.com/a-comprehensive-introduction-to-different-types-of-convolutions-in-deep-learning-669281e58215.

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Binghamton, S. (2012) 'Designing steganographic distortion using directional filters Vojtˇ ech Holub and Jessica Fridrich', Ieeexplore.Ieee.Org, pp. 234–239. Available at: http://ieeexplore.ieee.org/abstract/document/6412655/.

Boroumand, M., Chen, M. and Fridrich, J. (2019) 'Deep residual network for steganalysis of digital images', IEEE Transactions on Information Forensics and Security, 14(5), pp. 1181–1193. Available at: https://doi.org/10.1109/TIFS.2018.2871749.

Cogranne, R. and Fridrich, J. (2015) 'Modeling and Extending the Ensemble Classifier for Steganalysis of Digital Images Using Hypothesis Testing Theory', IEEE Transactions on Information Forensics and Security, 10(12), pp. 2627–2642. Available at: https://doi.org/10.1109/TIFS.2015.2470220.

Holub, V., Fridrich, J. and Denemark, T. (2014) 'Universal distortion function for steganography in an arbitrary domain', Eurasip Journal on Information Security, 2014, pp. 1–13. Available at: https://doi.org/10.1186/1687-417X-2014-1.

Kodovský, J., Fridrich, J. and Holub, V. (2012) 'Ensemble classifiers for steganalysis of digital media', IEEE Transactions on Information Forensics and Security, 7(2), pp. 432–444. Available at: https://doi.org/10.1109/TIFS.2011.2175919.

Li, B. et al. (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), pp. 4206–4210.

Pevný, T., Filler, T. and Bas, P. (2010) 'Using high-dimensional image models to perform highly undetectable steganography', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6387 LNCS, pp. 161–177. Available at: https://doi.org/10.1007/978-3-642-16435-4_13.

Pevny, T., Fridrich, J. and Ker, A.D. (2009) 'From blind to quantitative steganalysis', Media Forensics and Security, 7254, p. 72540C. Available at: https://doi.org/10.1117/12.805601.

Qian, Y. et al. (2015) 'Deep learning for steganalysis via convolutional neural networks', Media Watermarking, Security, and Forensics 2015, 9409, p. 94090J. Available at: https://doi.org/10.1117/12.2083479.

Reinel, T.S. et al. (2021) 'GBRAS-Net: A Convolutional Neural Network Architecture for Spatial Image Steganalysis', IEEE Access, 9, pp. 14340–14350. Available at: https://doi.org/10.1109/ACCESS.2021.3052494.

Sedighi, V., Cogranne, R. and Fridrich, J. (2016) 'Content-adaptive steganography by minimizing statistical detectability', IEEE Transactions on Information Forensics and Security, 11(2), pp. 221–234. Available at: https://doi.org/10.1109/TIFS.2015.2486744.

Udofia, U. (2018) Basic Overview of Convolutional Neural Network (CNN), DataSeries. Available at: https://medium.com/dataseries/basic-overview-of-convolutional-neural-network-cnn-4fcc7dbb4f17#:~:text=The activation function is a,neuron would fire or not.&text=We have different types of,Rectified Linear Unit (ReLU).

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License

Xu, G., Wu, H.Z. and Shi, Y.Q. (2016) 'Structural design of convolutional neural networks for steganalysis', IEEE Signal Processing Letters, 23(5), pp. 708–712. Available at: https://doi.org/10.1109/LSP.2016.2548421.

Ye, J., Ni, J. and Yi, Y. (2017) 'Deep Learning Hierarchical Representations for Image Steganalysis', IEEE Transactions on Information Forensics and Security, 12(11), pp. 2545–2557. Available at: https://doi.org/10.1109/TIFS.2017.2710946.

Yedroudj, M., Comby, F. and Chaumont, M. (2018) 'Yedroudj-Net: An Efficient CNN for Spatial Steganalysis', ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2018-April(1), pp. 2092–2096. Available at: https://doi.org/10.1109/ICASSP.2018.8461438.

You, W., Zhang, H. and Zhao, X. (2021) 'A Siamese CNN for Image Steganalysis', IEEE Transactions on Information Forensics and Security, 16, pp. 291–306. Available at: https://doi.org/10.1109/TIFS.2020.3013204.

Zhang, R. et al. (2020) 'Depth-Wise Separable Convolutions and Multi-Level Pooling for an Efficient Spatial CNN-Based Steganalysis', IEEE Transactions on Information Forensics and Security, 15, pp. 1138–1150. Available at: https://doi.org/10.1109/TIFS.2019.2936913.

© 2022, SIJTRD, https://www.sdnbvc.edu.in/sijtrd/
This work is licensed under a Creative Commons Attribution 4.0 International License