www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 04, Issue 01, January 2024, pp : 379-384

e-ISSN : 2583-1062

Impact Factor : 5.725

A REVIEW ON REVIVING ROLE OF ATTENTION MECHANISM

Mohammad Rafi[1], Rakshitha C[2], Tejaswini Surve K S[3]

[1,2,3]Department of Computer Science and Engineering, UBDTCE, india.

ABSTRACT

In the realm of image processing, attention mechanisms have emerged as a powerful paradigm, mirroring the human cognitive process of selectively focusing on relevant information. However, amidst the recent surge in deep learning techniques, the attention mechanism's significance in image processing has somewhat diminished. This paper aims to revive and accentuate the crucial role of attention mechanisms in image-related tasks.The study begins by providing a comprehensive overview of attention mechanisms, elucidating their underlying principles and their evolution in the context of image processing. Subsequently, the paper delves into the challenges faced by traditional convolutional neural networks (CNNs) in handling complex and diverse visual data, underscoring the need for improved models that can effectively capture intricate relationships within images.

To revitalize the attention mechanism's prominence, the research explores novel architectures that seamlessly integrate attention modules into existing convolutional frameworks. By doing so, the study seeks to enhance the models' capability to discern and prioritize salient features, leading to improved accuracy and efficiency in image processing tasks. Furthermore, the paper investigates the application of attention mechanisms in various image processing domains, including image classification, object detection, and semantic segmentation. It discusses specific scenarios where attention mechanisms prove particularly beneficial, such as handling occlusions, scale variations, and fine-grained details.

1. INTRODUCTION

The concept of attention, a vital cognitive function in human perception, involves selectively focusing on the salient parts of a scene, enabling efficient processing of visual information [58]. This ability allows humans to filter relevant information with limited computational resources, enhancing both efficiency and accuracy in perception. In recent years, attention mechanisms have found applications in various computer science domains, including natural language processing and Computer Vision (CV) [59]. In these contexts, attention acts as a technique to emphasize specific parts of input data when generating output, essentially assigning importance weights to different input features. Attention mechanisms originated from the investigations of human vision. In cognitive science, only part of all visible information is noticed by human beings due to the bottlenecks of information

processing. Inspired by this visual attention mechanism, researchers have tried to find the model of visual selective attention to simulate the visual perception process of human beings, so as to model the distribution of human attention when observing images as well as videos and expand its applications. In recent years, important breakthroughs of attention mechanisms have been made in the fields of image and natural language processing. It has been proven that attention mechanisms can improve the performance of models, and they are also consistent with the perceptual mechanism of human brain and eyes. Taking the field of computer vision for example, most research combining deep learning and visual attention mechanisms concentrates on the use of mask. The principle of mask is that the key features in the image data are identified by another layer with new weight. By learning and training, deep neural network can learn the areas where attention needs to be paid in each new image, thereby forming attention. This idea further evolved into two different types of attention: soft attention and hard attention. The mechanism of soft attention is realized via gradient descent and is of differentiability and continuity. In neural networks, the weight of attention can be learned through forward propagation and backward feedback [2]. Hard attention mechanism, however, is not differentiative, which is often achieved by reinforced learning and motivated by the benefit function to make the model pay more attention to the details of some parts. This paper will introduce in three parts: the first part is the computational models of visual selective attention; the second part is the classification of the attention mechanism models of computer vision; the third part is the summary of the existing attention mechanisms and an outlook. In recent years, the field of image processing has undergone a transformative evolution fueled by advancements in deep learning techniques. Convolutional Neural Networks (CNNs) have become the cornerstone of image-related tasks, achieving remarkable success in tasks such as image classification, object detection, and semantic segmentation. However, as the complexity and diversity of visual data continue to grow, there is a pressing need to reevaluate the role of attention mechanisms within the realm of image processing. Attention mechanisms, inspired by human cognitive processes, have proven to be instrumental in capturing contextual information and selectively focusing on relevant features. Initially embraced in natural language processing tasks, attention mechanisms have somewhat taken a backseat in image processing, overshadowed by the success of traditional CNN architectures. This shift has raised questions about the potential benefits and untapped potential that attention mechanisms could bring to enhance the capabilities of image processing models. This paper seeks to revive the role of attention mechanisms in image processing, recognizing their unique ability to refine feature selection and improve the handling of intricate visual patterns. By reintroducing attention mechanisms into the forefront of image processing research, we aim to address the limitations of existing models and pave the way for more robust, accurate, and interpretable solutions.

In this introduction, we will provide an overview of the evolution of attention mechanisms, their initial successes, and subsequent underutilization in image processing. We will highlight the challenges faced by traditional CNNs in handling complex visual scenarios and emphasize the need for renewed exploration of attention mechanisms. Additionally, we will outline the objectives of this study, including the exploration of novel architectures that seamlessly integrate attention modules into existing frameworks, and the subsequent validation of their efficacy through extensive experimentation on benchmark datasets.

As we delve into the subsequent sections, it becomes evident that the revitalization of attention mechanisms holds the key to unlocking the full potential of image processing models, ushering in a new era of enhanced performance and adaptability in the face of evolving visual data challenges.

2. LITERATURE REVIEW

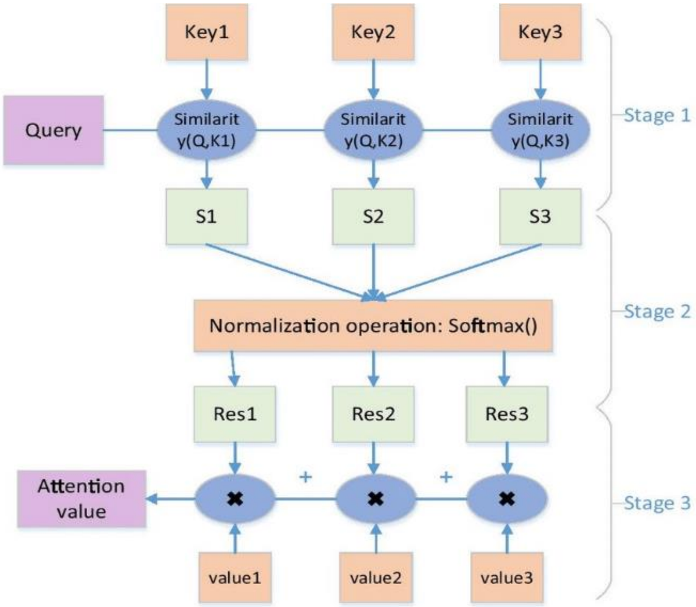
General Process of attention mechanism:

Attention mechanisms in image processing function as a mechanism for selectively focusing on specific regions or features within an image, allowing the model to weigh the importance of different parts of the input during processing. The general process of attention mechanisms in image processing can be outlined as follows:

Input Representation: Begin with an input image represented as a grid of pixels or as feature maps extracted from previous layers in a neural network.

Feature Extraction: Employ convolutional layers to extract hierarchical features from the input image. These features capture patterns and representations at different levels of abstraction.

Query, Key, and Value Computation: Transform the extracted features into three sets: queries (Q), keys (K), and values (V). These sets are computed through linear transformations to project the features into a shared space.



Attention Score Computation: Calculate the attention scores by measuring the similarity between the queries and keys. A common method is to use a dot product or other similarity metrics, followed by scaling and applying a softmax function to obtain normalized attention weights.

Weighted Sum of Values: Multiply the attention weights by the corresponding values to compute a weighted sum. This step emphasizes the more relevant features based on the attention scores.

Integration with Original Features: Combine the weighted sum with the original features to create an enriched representation that highlights important regions or features based on the attention mechanism.

Output Generation: Pass the integrated representation through subsequent layers of the neural network for further processing. The attention-enriched features contribute to the final output, whether it's classification, object detection, segmentation, or another image processing task.

Training and Optimization: During the training phase, the attention mechanism is optimized by adjusting the parameters of the model to minimize the error between predicted and actual outputs. This involves backpropagation and gradient descent optimization techniques.

Interpretability and Visualization: Attention mechanisms often offer interpretability benefits. Visualizations of attention maps can be generated to understand which parts of the input image the model is focusing on, providing insights into its decision-making process.

Application to Various Tasks: Attention mechanisms find applications in diverse image processing tasks, such as image classification, object detection, image captioning, and image segmentation. The ability to selectively attend to relevant information makes them versatile tools in handling complex visual data.

By incorporating attention mechanisms into image processing models, these steps collectively enhance the model's ability to discern and prioritize relevant information, leading to improved performance and adaptability in the face of varying and intricate visual patterns.

3. METHODOLOGY

Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint

Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolutionand Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism.

Image Water Ripple Detection Method Based on Constraint Convolution and Attention Mechanism

Design of image water ripple detection method

3.1 mage denoising based on attention mechanism

The removal of noise in images is a very important step in image processing, as noise can bring a lot of problems such as image quality degradation and loss of key information in the image, which will lead to incorrect judgments in subsequent work. Traditional methods for processing noisy images cannot effectively distinguish between complex noise and natural textures and do not effectively address issues such as loss of image edge information and excessive smoothness of details [6-8]. In response to the above issues, this article proposes an image-denoising algorithm based on an attention mechanism. We calculate the "attention map" from both channel and spatial aspects and then multiply the calculated "attention map" with the image feature map for adaptive feature learning, to mine deeper levels of noise in complex backgrounds and improve denoising effectiveness. c_1 represents the input characteristic diagram, G(x) represents the operation of the channel attention module, D(x) represents the operation of the spatial attention module, and U(x) represents the calculation result. The following equation is a mathematical representation of the dual attention mechanism, where the feature map is input to the channel attention module for the multiplication operation.

U G c =

$U_1=G(c_1)$

We multiply the feature map obtained from the previous step with the original feature map

$U_1=G(c_1)$

We input the multiplied result into the spatial attention module for operation.

□□32UDU□ (3)

We multiply the results obtained from the spatial attention module with the results obtained from the channel attention module to obtain the final result.

In this experiments, dividing manually forged test images into uniform overlapping blocks of size $B \times B$ pixels, where the B is varied from 6 to 36. The performance characteristics of the various techniques presented in terms of DA, have been presented in Fig. The results shown in the plots are the averages taken over all our test images. From the Fig, it is evident for all the algorithms, the DA increases with increasing forgery size. The maximum and average DA of all discussed algorithms are presented. Among all the techniques, the CWT-based method exhibits the best DA of 99.59% when the forgery size is 40%. This is due to the inherent properties of CWT exploited in the class of algorithms such as rotation invariance, robustness to noise and multi-level representation, which makes it an extremely efficient method for feature extraction. The characteristic FPR variation for region duplication detection for all the techniques versus unit block size and forgery size. It is the evident that for all the algorithms, the FPR decreases with increasing the forgery size. By wavelet-based copy– move forgery detection methods, several identical blocks get falsely detected at the boundaries of the images they contribute to the FPs which is not possible to be eliminated completely by adjusting the threshold. Presenting the false copy–move forgery detection results, for all the schemes, according to varying forgery sizes. Among all the schemes, the DCT-based techniques demonstrate the lowest rate of FPs. Similarly, the results for region duplication attacks falsely missed by the stateof-the-art techniques have been presented. Figure shows the plot of FNR versus unit block size and forgery size. From the Fig, it is evident that the FN detection rate diminishes with increasing forgery size, for all the techniques. It may be by the observation that the FN detection rate is trivial for all the techniques presented in this paper. The major challenge in this area of forensic research is to minimize the rate of FPs, as is evident. From the Figures, it may be observed that for any copy– move forgery detection technique, its DA and FN forgery detection characteristics are inversely proportional to each other. This is due to the fact that DA is directly computed depending on the number of correctly detected copy–moved pixels, while the FNR is determined by the number of undetected copy–moved pixels. The computational complexity of any block-based copy– move forgery technique increases as the unit detection block size is reduced. On the other hand, a smaller unit detection block size ensures higher DA. Hence, in such algorithms, it is desirable to obtain a correct trade-off between DA and computational complexity, by selecting an appropriate unit block size. In this regards, the experimental results may help a user to select the most suitable method and unit block size, according to the requirements.

4. CONCLUSION

In the last decade, there has been quite a lot of researches in the direction of image forgery detection. Specifically in the field of copy–move forgery or region duplication detection in images has gained a lot of research interest due to the fact that this form of forgery is one of the most primitive forms of attacks on digital images. However, it is not trivial to detect this form of forgery because the natural statistical properties of the images are not altered here. In this paper, by providing a detailed review of state-of-the-art copy–move forgery detection algorithms, their implementation, performance evaluation and comparison. In this paper, by introducing a set of standard parameters with respect to which by having performed the experiments for the performance evaluation and comparison.

The parameters introduced in this paper encompass three different dimensions of conventional forgery detection operations. The proposed parameterization would help the users select an appropriate forgery detection algorithm according to the requirements, and the expected forgery type. Future research in this direction would include incorporating more parameters into the proposed platform in order to optimize its efficiency in terms of image forgery detection evaluation and comparison. By proposing a novel training scheme for improving the robustness of the image forgery detection against various OSN based transmissions.

The proposed scheme is designed with the assistance of the modeling of a predictable noise $\boldsymbol{\tau}$ as well as an intentionally introduced unseen noise. Experimental results are provided to demonstrate the superiority of our scheme compared with several stateof-the-art methods. Further, by building an OSN transmitted forgery dataset for future research of the forensic community.

As the future work, may be by extending the proposed robust training scheme to deal with more complex. degradation scenarios, such as screen capturing, printing and re-photographing, etc. Additionally, investigating whether an image restoration network can be used to assist the forgery detection in severely degraded scenarios.

5. ROBUSTNESS EVALUATIONS

Although the proposed scheme is mainly designed to counter the lossy operations conducted by OSNs, and also like to evaluate its robustness under some more commonly used degradation scenarios, such as noise addition,cropping, resizing, blurring, and standalone JPEG compression. Such evaluation is very critical in real-world cases because these types of post-processing operations are often adopted to erase or conceal the forged artifacts. To this end, applying these post-processing operations to the original test set Columbia and report the quantitative comparisons. For the convenience of demonstration, utilizing a unified parameter p for controlling the magnitudes of different operations.

The origin of the horizontal axis (p = 0) corresponds to the case without any postprocessing. It can be observed, the competitors [12,27] cannot perform consistently with the increase of the perturbation intensity, while this method can generalize well to defeat these post processing operations.

6. REFERENCES

[1] P. Zhuang, H. Li, S. Tan, B. Li, and J. Huang, "Image tampering localization using a dense fully convolutional network," IEEE Trans. Inf. Forensics Security, vol. 16, pp. 2986–2999, 2021.

[2] S. Lyu, X. Pan, and X. Zhang, "Exposing region splicing forgeries with blind local noise estimation," Int. J. Comput. Vis., vol. 110, no. 2, pp. 202–221, Nov. 2014.

[3] Y. Li and J. Zhou, "Fast and effective image copymove forgery detection via hierarchical feature point matching," IEEE Trans. Inf. Forensics Security, vol. 14, no. 5, pp. 1307–1322, May 2019.

[4] X. Kang, M. C. Stamm, A. Peng, and K. J. R. Liu, "Robust median filtering forensics using an autoregressive model," IEEE Trans. Inf. Forensics Security, vol. 8, no. 9, pp. 1456–1468, Sep. 2013.

[5] H. Li, W. Luo, and J. Huang, "Localization of diffusion-based inpainting in digital images," IEEE Trans. Inf. Forensics Security, vol. 12, no. 12, pp. 3050–3064, Dec. 2017.

[6] M. Huh, A. Liu, A. Owens, and A. A. Efros, Fighting fake news: Image splice detection via learned self-consistency," in Proc. Eur. Conf. Comput. Vis., 2018, pp. 101–117.

[7] J.-L. Zhong and C.-M. Pun, "An end-to-end denseInceptionNet for image copy-move forgery detection," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 2134– 2146, 2020.

[8] J. Chen, X. Kang, Y. Liu, and Z. J. Wang, "Median filtering forensics based on convolutional neural networks," IEEE Signal Process. Lett., vol. 22, no. 11, pp. 1849–1853, Nov. 2015.

[9] H. Wu and J. Zhou, "IID-Net: Image inpainting detection network via neural architecture search and attention," IEEE Trans. Circuits Syst.Video Technol., early access, Apr. 22, 2021, doi: 10.1109/TCSVT. 2021.3075039.

[10] A. Li et al., "Noise doesn't lie: Towards universal detection of deep inpainting," in Proc. 13th Int. Joint Conf. Artif. Intell., Aug. 2021, pp. 1–7.

[11] D. Cozzolino, G. Poggi, and L. Verdoliva, "Splicebuster: A new blind image splicing detector," in Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS), Nov. 2015, pp. 1–6.

[12] L. Bondi, S. Lameri, D. Guera, P. Bestagini, E. J. Delp, and S. Tubaro, "Tampering detection and localization through clustering of camera based CNN features," in Proc. IEEE Conf. Comput. Vis. Pattern Recog nit. Workshops (CVPRW), Jul. 2017, pp. 1855– 1864.

[13] B. Bayar and M. C. Stamm, "Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection," IEEE Trans. Inf. Forensics Security, vol. 13, no. 11, pp. 2691– 2706, Nov. 2018.

[14] Y. Wu, W. Abdalmageed, and P. Natarajan, "ManTra-Net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 9543– 9552.

[15] O. Mayer and M. C. Stamm, "Forensic similarity for digital images," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 1331–1346, 2020.

[16] D. Cozzolino and L. Verdoliva, "Noiseprint: A CNN-based camera model fingerprint," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 114–159, 2020.

[17] W. Sun, J. Zhou, R. Lyu, and S. Zhu, "Processingaware privacy preserving photo sharing over online social networks," in Proc. 24th ACM Int. Conf. Multimedia, Oct. 2016, pp. 581–585.

[18] W. Sun, J. Zhou, Y. Li, M. Cheung, and J. She, "Robust high-capacity watermarking over online social network shared images," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 3, pp. 1208–1221, Mar. 2021.

[19] Security Ai Competition: Forgery Detection on Certificate Image. Accessed: Jan. 23, 2022. [Online]. Available:https://tianchi.aliyun.com/competition/entrance /531812/information

[20] C. Szegedy et al., "Intriguing properties of neural networks," in Proc. Int. Conf. Learn. Representat., 2014, pp. 1–10.

[21] Savchenko, V., Kojekine, N., Unno, H.: 'A practical image retouching method'. Proc. First Int. Symp. Cyber Worlds, 2002, pp. 480–487

[22] Redi, J.A., Taktak, W., Dugelay, J.: 'Image splicing detection using 2D phase congruency and statistical moments of characteristic function'. Society of Photooptical Instrumentation Engineers (SPIE) Conf. Series, 2007, vol. **6505**, p. 26

[23] Fridrich, A.J., Soukal, B.D., Lukáš, A.J.: 'Detection of copy–move forgery in digital images'. Proc. Digital Forensic Research Workshop, 2003

[24] Farid, A.P., Popescu, A.C.: ‘Exposing digital forgeries by detecting duplicated image region’. Technical Report, Hanover, Department of Computer Science, Dartmouth College, USA, 2004

[25] Kang, X., Wei, S.: ‘Identifying tampered regions using singular value decomposition in digital image forensics’. Int. Conf. Computer Science and Software Engineering, 2009, vol. **3**, pp. 926–930

[26] Zhang, J., Feng, Z., Su, Y.: ‘A new approach for detecting copy–move forgery in digital images’. 11th IEEE Singapore Int. Conf. Communication Systems, 2008, pp. 362–366

[27] Muhammad, G., Hussain, M., Bebisi, G.: ‘Passive copy–move image forgery detection using undecimated dyadic wavelet transform’, Digit. Invest., 2012, **9**, (1), pp. 49–57

[28] Yang, J., Ran, P., Tan, J.: ‘Digital image forgery forensics by using undecimated dyadic wavelet transform and Zernike moments’, J. Comput. Inf. Syst., 2013, **9**, (16), pp. 6399–6408

[29] Bayram, S., Sencar, H.T., Memon, T.N.: ‘An efficient and robust method for detecting copy–move forgery’. IEEE Int. Conf. Acoustics, Speech and Signal Processing, 2009, pp. 1053–1056

[30] Huang, Y., Lu, W., Sun, W., et al.: ‘Improved DCTbased detection of copy–move forgery in images’, Forensic Sci. Int., 2011, **206**, (1), pp. 178–184