International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 21st Nov 2020. Link :http://www.ijiemr.org/downloads.php?vol=Volume-09&issue=ISSUE-11

DOI: 10.48047/IJIEMR/V09/I11/47

Title **Survey on Picture-level Just Noticeable Difference (PJND) using Deep Learning (DL)**

Volume 09, Issue 11, Pages: 245-250

Paper Authors

S.S.Nandini, Dr.Jhansi Rani Singothu

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Survey on Picture-level Just Noticeable Difference (PJND) using Deep Learning (DL)

S.S.Nandini[1],Dr.Jhansi Rani Singothu[2]
[1]Research Scholar, Department of Computer Science and Systems Engineering,
Andhra University College of Engineering(A),Andhra University, Visakhapatnam, Andhra Pradesh, India
[2]Assistant Professor, Department of Computer Science and Systems Engineering,
Andhra University College of Engineering(A),Andhra University, Visakhapatnam, Andhra Pradesh, India

Abstract: Image distortion is one of the compression techniques which is used to find the modifications in the image. It is highly impossible to the human to notice the distortion within the image. Every pixel of the image is to be observed carefully weather the image is changed or not. Deep Learning (DL) is most widely used to solve various complex issues. Picture-level Just Noticeable Difference (PJND) is the method which is used to detect the noticeable change in the image. In the process leading to image correction and enhancement, the detection of artifacts and the quantification of their impact on the image-quality levels are of paramount importance. Image Quality Assessment (IQA) algorithms aim at consistently reflecting human quality perception in order to provide reliable estimates of the quality level of the images being processed for enhancement.

Keywords: *Deep Learning, Image Quality Assessment, Image distortion, Picture-level Just Noticeable Difference (PJND)*

Introduction

Images/Videos are commonly explored in various multimedia services and become an indispensable part in people's daily life. To provide a high quality of multimedia experience, there are many researches devoting to the development of image/video processing, image/video coding, and robust transmission technologies. Since human eyes are the ultimate receivers of images/videos in general, how to describe perceptual characteristics of human vision more precisely and efficiently has been drawing lots of attentions from both academic and industrial societies [1]–[4].

JND indicates the maximum distortion that is allowed to be injected with the constraint that no visual quality variation is perceived. In the literature, the conventional JND models can be mainly categorized into pixel-wise models which calculate the JND threshold for each pixel straightforwardly in the pixel domain and sub-band JND models which calculate the JND threshold for each sub-band coefficient

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

by transforming the image into frequency domain (e.g., DCT, wavelet) representation.

Literature Survey

Based on Chou and Li [5], Yang et al. [6] exploited a nonlinear additively model to reduce the overlapping effects between LA and CM. Since these two methods overestimated the masking effects in the edge regions, Liu et al. [7] decomposed one input image into two images, one is named structural image and the other is the textural image, followed by performing edge masking (EM) estimation and texture masking (TM) estimation, respectively. Considering that the CM effect is not comprehensively evaluated, Wu et al. [8] proposed the disorderly concealment effect based on free-energy principle for JND estimation. Motivated by the observation that the HVS is highly sensitive to the repeated pattern in visual signal, Wu et al. [9] introduced the concept of pattern complexity to decide the total masking effects. With image saliency information, Hadizadeh et al. [10] developed a saliency-guided JND model by the normalized Laplacian pyramid. the textural image proposed by Liu et al. [7] is further decomposed into two portions according to the regularity of the texture.

Then, Huang et al. [11] presented an architecture similar to the one proposed by Bondi et al. [12], and the authors were able to improve over the accuracy obtained by Freire-Obregón et al. [13] by using Batch Normalization and more convolutional layers. Following along the line of deeper CNN architectures, Yao et al. [14] (code available at https://github.com/grasses/Camera-Identification) put forward a 13-layer convolutional neural network.

Chen et al. [15] investigated the use of a residual neural network (ResNet) with 26 layers for source camera identification, and proved its effectiveness with multiple experiments: the accuracies obtained for brand-, model-, and device-level identification were 99.12%, 94.73%, and 45.81%, respectively. According to their paper, ResNet has better performance than AlexNet, GoogleNet, and the scheme from Bondi et al. [12]. Ding et al. [16] extended this last method by combining ResNet architecture with a multi-task learning strategy, further improving the performance. The three tasks (brand-level, model-level, and sensor-level classification) are integrated into one framework and share the weights of the CNN architecture.

Related work on Picture-level Just Noticeable Difference (PJND)

Hoffman and Stolitzka [18] proposed tests to determine if a compressed image differs from a reference image by more than one JND. The testing environment was according to ISO 3664. The monitor used had a 24.3-in diagonal and resolution 1920 × 1200. A reference (uncompressed) image and an image consisting of the alternating reference image and a distorted version were presented side by side. The observer had to identify which of the two images was non-flickering. A database of about 250,000 responses collected from 35 observers to 18 images was made available. The flicker method proposed in this paper was adopted as a standard [19].

Zhang et al. [20] collected a large-scale dataset of perceptual judgements, which included asking subjects whether one reference patch and one distorted patch are identical. They used 20 types of distortions (e.g., photometric distortions, noise, blurring, and compression artifacts) and sequentially composed pairs of distortions. The two patches had a resolution of 64 × 64 and were shown for 1 s each, with a 250 ms gap in between.

Redi et al. [21] compared the performance of absolute category rating obtained by the single stimulus (SS) method with that of the quality ruler (QR) method. The QR consists of a series of reference images varying in a single attribute (sharpness), with known and fixed quality differences between the samples, given by a certain number of JND units [22]. For the QR method, the quality of an input image is compared to the image qualities on the ruler. The study showed that QR scores have narrower confidence intervals than SS scores.

Deep Learning Algorithms for Picture-level Just Noticeable Difference (PJND)

Deep learning is a subfield of machine learning that has evolved out of the traditional approaches to artificial neural networks. Various deep learning algorithms are discussed in this section to notice the Picture-level Just Noticeable Difference. With the advent of back propagation, neural networks began to see renewed interest and significant theoretical advancement in the form of recurrent neural networks (RNN), Convolutional neural networks (CNN), Long Short-Term Memory networks (LSTM) and Gradient decent (SGD).

1) CNN: Convolutional neural networks learn filter banks that are convolved with the original data. The filters can also be represented as a fully connected layer where the weights of the edges are tied together in way that replicates the convolution operation. This weight sharing structure allows for fewer parameters than having each weight be unique, and directly accounts for structure in the data. Each filter creates a new, processed version of the image.

2) RNN: Recurrent neural networks contributed a lot to the success of deep learning in various fields. It outperforms for time series classification because of its intra layer connections and inters layer connections. Training an RNN is a difficult task. Back-propagation must be modified to function in RNNs, since there are cycles in the graph. This is frequently handled through a technique known as back-propagation through time, wherein the network is "unrolled" for a discrete number of steps. This process creates a network with only forward connections, allowing back-propagation to work as normal, at the cost of limiting the impact of the recurrent connections.

3) LSTM: Long Short-Term Memory networks have proven successful for sequential data classification because they are specifically designed to classify for each time stamp. At each time point data is processed by individual LSTM unit. The processed result is fed to the next layer and

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

as well as well as remains within the same layer for the processing of next time stamp. The information passed through to the next layer is passed through an activation function. However, the recurrent connection within the layer is not subjected to an activation function. One of the vibrant problems in LSTM is the lack of activation function in recurrent layers to avoid the vanishing gradient problem. Due to this gradient problem it gives gradient a constant value of one. Each LSTM unit has a number of gates which control the flow of information. A gate is a combination of a sigmoidal activation unit and pointwise multiplication. The flow of information from one time stamp to another time stamp is controlled by these gates.

4) SGD: It is a derivative of traditional gradient descent, differing in that the error function is calculated using only a subsample of the available data. This is both easier to use and more efficient for training datasets that do not fit in memory. Furthermore, adding randomness to the optimization can aid in breaking through plateaus and avoiding local minima. The addition of momentum terms, which biases the gradient in the direction of recently calculated gradients, greatly improved the ability to train deep models by further increasing speed of convergence. Most studies focus on using machine learning to answer a neuroscience question, rather than on improving the machine learning techniques themselves. Generally only a single technique is applied to the classification of a dataset and Individual datasets have highly varying characteristics, and some are simply harder to classify than others, so it is difficult to draw conclusions on the effectiveness classification system as compared to another.

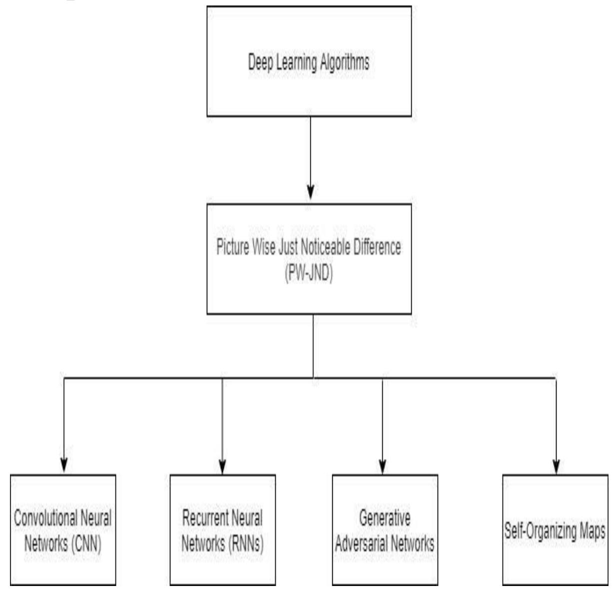

Figure 1: Deep Learning for PW-JND

Conclusion

Various JND models are discussed and the integration of deep learning algorithms is also plays a major role. JND models are usually employed in image/video coding frameworks to improve coding efficiency by removing perceptual redundancy of the visual content. The study revealed that, for the purpose of building publicly available data of subjective quality, the Single Stimulus method presents several drawbacks such as low confidence in the scores and susceptibility to range effects. Conversely, the Quality Ruler method is worth the implementation effort from a point of view of consistency and repeatability of the scores.

References

[1] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C.-C. J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," IEEE Trans. Image Process., vol. 26, no. 6, pp. 2682–2693, Jun. 2017.

[2] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content images," IEEE Trans. Image Process., vol. 25, no. 8, pp. 3838–3851, May 2016.

[3] L. Yu, H. Su, and C. Jung, "Perceptually optimized enhancement of contrast and color in images," IEEE Access, vol. 6, pp. 36132–36142,
2018.

[4] A. Yang, H. Zeng, J. Chen, J. Zhu, and C. Cai, "Perceptual feature guided rate distortion optimization for high efficiency video coding," Multidimensional Syst. Signal Process., vol. 28, no. 4, pp. 1249–1266, 2017.

[5] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," IEEE Trans. Circuits Syst. Video Technol., vol. 5, no. 6, pp. 467–476, Dec. 1995.

[6] X. K. Yang, W. S. Ling, Z. K. Lu, E. P. Ong, and S. S. Yao, "Just noticeable distortion model and its applications in video coding," Signal Process., Image Commun., vol. 20, no. 7, pp. 662–680, Aug. 2005.

[7] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 11, pp. 1648–1652, Nov. 2010.

[8] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," IEEE Trans. Multimedia, vol. 15, no. 7, pp. 1705–1710, Nov. 2013.

[9] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C.-C. J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," IEEE Trans. Image Process., vol. 26, no. 6, pp. 2682–2693, Jun. 2017.

[10] H. Hadizadeh, A. Rajati, and I. V. Bajic, "Saliency-guided just noticeable distortion estimation using the normalized laplacian pyramid," IEEE Signal Process. Lett., vol. 24, no. 8, pp. 1218–1222, Aug. 2017.

[11] Huang, N.; He, J.; Zhu, N.; Xuan, X.; Liu, G.; Chang, C. "Identification of the source camera of images based on convolutional neural network." Digit. Investig. 2018, 26, 72–80.

[12] Bondi, L.; Baroffio, L.; Guera, D.; Bestagini, P.; Delp, E.J.; Tubaro, S. "Camera identification with deep convolutional networks." arXiv 2016, arXiv:1603.01068.

[13] Freire-Obregón, D.; Narducci, F.; Barra, S.; Castrillón-Santana, M. "Deep learning for source camera identification on mobile devices." Pattern Recognit. Lett. 2019, 126, 86–91.

[14] Yao, H.; Qiao, T.; Xu, M.; Zheng, N. "Robust multi-classifier for camera model identification based on convolution neural network." IEEE Access 2018, 6, 24973–24982.

[15] Chen, Y.; Huang, Y.; Ding, X. "Camera model identification with residual neural network." In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017;pp. 4337–4341.

[16] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. "Going deeper with convolutions." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[18] D. M. Hoffman and D. Stolitzka, "A new standard method of subjective assessment of barely visible image artifacts and a new public database," Journal of the Society for Information Display, vol. 22, no. 12, pp. 631–643, 2014.

[19] ISO/IEC JTC 1/SC 29/WG 1, "29170-2 Information technology – Advanced image coding and evaluation – Part 2: Evaluation procedure for nearly lossless coding," Tech. Rep. JTC 1 / SC 29 N 14760, ISO/IEC, Geneva,February 2015.

[20] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in Proceedings of the IEEE
Conference on Computer Vision and Pattern Recognition (CVPR), pp. 586–595, 2018.

[21] J. Redi, H. Liu, H. Alers, R. Zunino, and I. Heynderickx, "Comparing subjective image quality measurement methods for the creation of public databases," in
The Proceeding of SPIE-IS&T Electronic Imaging, Image Quality and System Performance VII, vol. 7529,p. 752903, International Society for Optics and Photonics, 2010.

[22] B. W. Keelan and H. Urabe, "ISO 20462: A psychophysical image quality measurement standard," in The Proceeding of SPIE-IS&T Electronic Imaging,Image Quality and System Performance, vol. 5294,pp. 181–189, International Society for Optics and Photonics, 2003.