Digital Halftone Classification using Simplified CNN and Stochastic Statistics

Jing-Ming Guo* and Sankarasrinivasan S†
*†National Taiwan University of Science and Technology
* E-mail: jmguo@seed.net.tw Tel: +886-27303241
†E-mail: D10507805@mail.ntust.edu.tw Tel: +886-978374981

***Abstract*— ** Digital halftone technique is widely utilized in various printing applications to generate high quality printing output. The halftone classification is a fundamental requirement for the optimal reconstruction of binary printed images during inverse halftoning and image authentication. In this study, a new classification scheme based on convolutional neural networks (CCN) is proposed. For effective training, a new approach exploiting the properties of stochastic halftone parameters is proposed to obtain the favorable halftone patch. The CNN network has been optimized to perform classification with minimal layers, and the model is trained using cross entropy, and make use of softmax classification layer. During testing, the sub-halftone patches are classified individually and then majority voting strategy is performed to identify the actual halftone class. For performance evaluation, a deep learning halftone database has been developed for the 24 halftone varieties which also includes the latest multitone images. From experimental result, it has been validated that the proposed architecture achieves average correct classification rate (ACCR) of 100% for several halftone class, and showed superior performance in comparison with state-of-the-art methods.

Keywords: CNN, classifier, digital half-toning, stochastic statistics, printing.

I. Introduction

Digital halftoning involves generation of printable binary image from the grayscale or color image [1] [2] which is typically useful for multimedia printing. In practical scenario, two types of print marking engines such as laser and ink-jet are widely used in commercial and personal printing applications. The laser printers (electro-photography printers) are used for large scale printing and inherently suffer from producing unstable dot patterns. In relevant to this, the clustered dot halftone patterns are adopted, which pertains the binary patterns of clustered dot nature corresponding to green noise spectrum. Conversely, ink-jet printers are widely used in the personal printing application and it can deliver a stable dot patterns with good visual quality. In general, inkjet printer delivers homogeneously distributed dispersed pattern whose spectral distribution belongs to the blue noise region, and dispersed dot halftones are useful for this application [3]. Hence, the halftones are commonly classified as clustered dot and dispersed dot patterns. The digital halftones are mainly generated using four types of techniques such as ordered dithering (OD) [4], error diffusion (ED) [5], dot diffusion (DD) [6] and iterative approaches [7].

On the other hand, halftone classification is very important for many applications that involves printed image processing, document authentication, source printer identification, inverse halftoning and various reconstruction methods [8]. The first halftone classification involves three-layer back propagation network and performs classification based on the one dimensional correlation method [9]. The strategy considers five types of halftone images, and achieves an accuracy of 100%. Guo. et. al [10] proposed a new approach based on the least mean square filter to obtain the optimal feature and adopted naïve Bayes method to achieve it has limited classification accuracy on the dispersed dot halftones. Wen et. al [11] published a new technique which involves construction of statistical matrix descriptor for the halftone images which considers six types of error diffused halftones and achieved good classification. Similar approach for error diffusion classification is also proposed using the spectral characteristics [12]. The method involves spectral regression kernel discriminant to achieve dimensionality reduction and then nearest centroid classifier is adopted. As the approach is based on error diffusion kernels, the strategy cannot be extended to other halftone techniques.

Zhang. et. al, [13] proposed a deep neural network based halftone classifier considering 14 different types of halftones. The method used sparse auto-encoder and softmax regression for classification. The method has limited accuracy on the dot diffusion halftones, and achieved 100% towards the clustered dot patterns. Finally, Guo, et. al. [14] proposed a new strategy based on the stochastic halftone statistical parameters. The method used hand-crafted feature extraction and extreme learning machines for rapid estimation. The technique achieved a good classification accuracy on many halftone types and considers 18 types of halftone images for classification. Though this approach achieved a 100% classification on many halftone classes, it has limited performance on practically implemented halftone. In general, the error- and dot-diffusion halftones are widely used in practical applications and their classification accuracy are limited to 97% and 94%, respectively. In our study, we attempt to address this issue based on the popular convolutional neural networks (CNN's) based classifier. Moreover, the concept of stochastic halftone statistics is also exploited to enhance the performance.

The main contribution of this paper are as follows:

1) In comparison to existing approaches [9]-[14], the proposed work considers more halftone class for classification. The previous digital halftone database v1.0 [15] is upgraded to meet the enhanced deep learning and classification.

2) The stochastic and spatial statistical parameters are found to be highly effective to identify the optimal image patches for training and testing.

3) The CNN architecture is optimized with respect to halftones and the proposed architecture consists of minimal convolution layers and some standard CNN techniques such as max pooling and data augmentation are eliminated as it tremendously affected the classification.

II. CNN BASED HALFTONE CLASSIFICATION

The proposed approach adopts the CNN architecture, and optimizes it with respect to halftone images. Many classification problems in image processing area mainly focus on extracting the best feature of the object/scene such as its edges, corners and the abstract profile. During the training process, the algorithm slowly learns the most distinguishable features that can result in accurate classification. In contrast to this, the main objective in the classification of halftone images involves training the network to understand the pattern representation (distribution) of each class rather than the object or scene in an image. Thus, in the training phase, the distinctive halftone patterns for each class need to be introduced to the CNN model for effective learning and pattern understanding. Each halftone methods approximate the original images in a unique way, and the output pattern has different dot distribution.

From the previous research works [13] [14], it can be also validated that different halftones possess similar spatial properties such as mean and variance, and they require more complex parameters for analyses. The more effective way to characterize the halftone is through stochastic geometry which basically deals with mathematical modelling of random process [1].

A. *Identification of effective patch*

An effective patch is defined as the sub-halftone image of smaller size such as 32x32 which contains maximum distinctive information about a particular halftone class [13]. Usually the halftone patches with homogenous distribution of minority pixels (black or white dots) and containing the equal number of black and white pixels, is expected to meet this requirement. In the previous approaches, the local entropy obtained from the halftone patch of 30x30 is used to construct the histogram (termed as entropy distribution). Subsequently, the statistical mean and variance are used to threshold the effective halftone patches. Yet, the standard statistical parameters cannot accurately represent various aspects of halftones such as homogeneity, directional distribution, pattern alignment and artifacts. Thus, this study exploited the stochastic halftone analysis (SHA) parameters to distinguish and extract the effective patch. In this study, a new effective patch estimation is carried out by exploiting the stochastic statistical parameters such as directional distribution function (DDF), radially averaged power spectral density (RAPSD) and anisotropy (ANIS). Among them, DDF is computed in spatial domain, RAPSD and ANIS are calculated in the frequency domain. To begin with, the DDF is defined as

$$DDF_a(r1, r2) = \frac{E\{\Gamma_m^a | y \in \Phi\}/N(\Gamma_m^a)}{E\{\Gamma_m | y \in \Phi\}/N(\Gamma_m)} \tag{1}$$

where ϕ refers to the point process, and $r1$ and $r2$ are the inner and outer radius of the portioned portions. The expected number of points in a specific portion is provided by Γ_m^a. To summarize, the DDF effectively represents the expected number of points in a specific section. On the other hand, the spectral power is defined as

$$\hat{P}(f) = \frac{1}{K}\sum_{i=1}^{K}\frac{|DFT_{2D}(\varphi_i)|^2}{N(\varphi_i)} \tag{2}$$

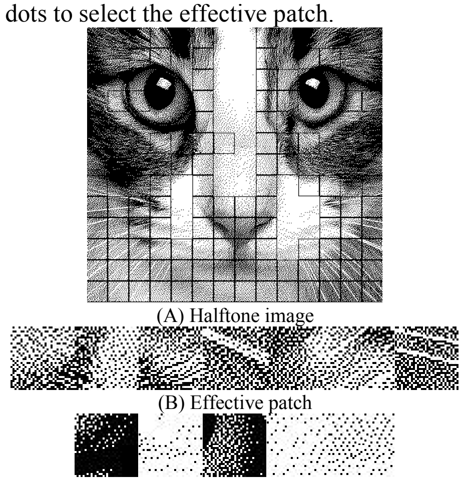
where DFT_{2D} refers to the 2-D fourier transform of the process φ_i; $N(\varphi)$ corresponds to total number of pixels; K means averaged periodograms. The spectral domain is partitioned into various annular rings $R(f_p)$ of width Δ_ρ to obtain RAPSD as follows.

$$RAPSD(f_\rho) = \frac{1}{N(R(f_p))}\sum_{f \epsilon R(f_p)} \hat{P}(f) \tag{3}$$

where $N\left(R(f_p)\right)$ is the number of frequency samples. Then the second parameter ANIS is measured by computing the noise to signal ratio of the frequency samples in the annular rings and it is a measure of isotropy.

$$ANIS(f_\rho) = \frac{1}{N(R(f_p))^{-1}}\sum_{f \epsilon R(f_p)} \frac{\left(\hat{P}(f) - P(f_\rho)\right)^2}{P^2(f_\rho)} \tag{4}$$

A standard effective patch is constructed by thresholding a middle gray scale value (=128) using the DBS approach. When the middle gray term undergoes iterative dithering, the resulted dispersed pattern comprises of equal number of black and white pixels. The different characteristics of the obtained standard patch is estimated using the mentioned stochastic parameters. Subsequently, the stochastic parameters obtained from the sub-halftone image patches and standard patch are correlated to identify an effective patch. The correlation coefficient greater than 0.7 is set for dispersed dots, and 0.4 for clustered dots to select the effective patch.



(A) Halftone image
(B) Effective patch
(C) Non effective patch

Fig. 1. Extracted effective patches from a halftone image.

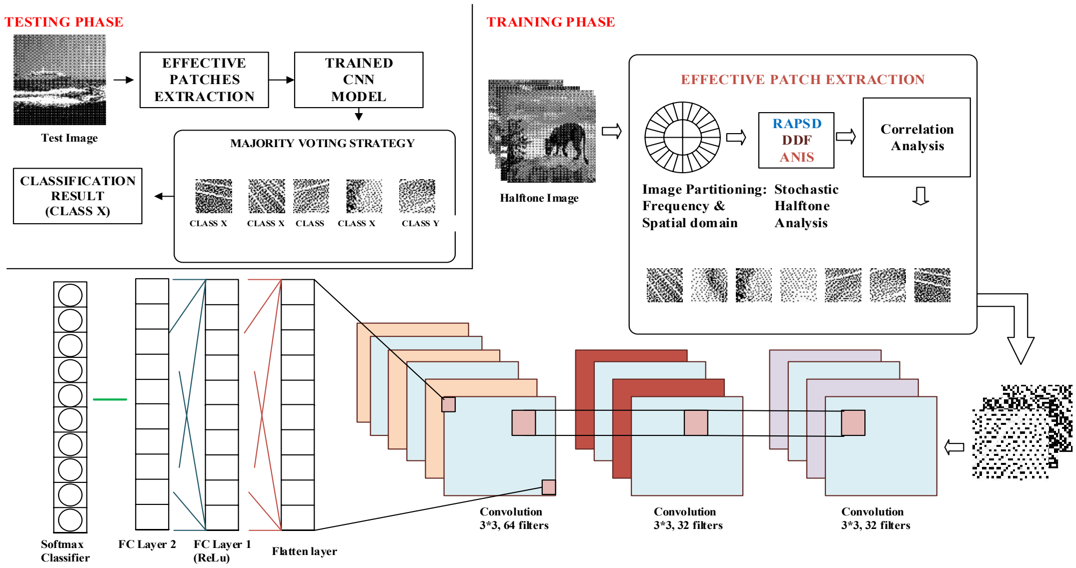


Fig. 2. Proposed Halftone Classification

Fig. 2. shows some of the extracted patches from the halftone image. It has been inferred from the extracted results, that the effective patches have equal distribution of black and white dots. This tremendously expedite the learning and the architecture can learn the unique pattern distribution of a particular halftone class more accurately. Moreover, it is visually evident that the non-effective patches do not contain maximal information and cannot be a precise indicator of any halftone class.

B. CNN Architecture

The proposed convolution network consists of three convolutional layers, followed by a flatten, and two fully connected layers as illustrated in Fig. 2. The softmax function is used as a final layer and the network is optimized by minimizing the cross entropy adjusting the variables of network layers. The first convolution layer filters take the image of size 32x32, and contains 32 filters of size 3x3 with the stride of 1 pixel. From the initial experiments, it has been found that lower filter sizes and stride values result in better performance. The second convolutional layer comprise of 32 filters of size 3x3 which is of the same configuration as the first layer. The third or final layer contains maximum number of filters (=64) and the filter size is maintained the same. All the convolutional layers comprising of rectified linear unit (ReLu) as activation function. Subsequently, the output of the convolution layers is required to pass through fully connected network, and hence a flatten layer is provided. The first fully connected layer is provided with the ReLu layer to learn some non-linearity. The next fully connected layer outputs vectors of the same number of halftone class and the estimates are difficult to interpret without normalization. Thus, its output is connected to the softmax function to perform normalization, and which represents each terms with their probability distributions. The cross-entropy function is used as a cost function which tends to zero when the model predicts the exact class. During training, the goal of optimization is to minimize the cross-entropy as close as zero to obtain the best model. The Adam optimizer [30] is adopted which uses squared gradients to scale the learning rate (based on root mean square prop) and moving average of the gradient (based on stochastic gradient descent) to perform adaptive moment estimation, and the learning rate of the optimizer is fixed at 0.0001. In traditional CNN model, the max-pooling layers are widely used as it reduces the complexity of upper layers by eliminating the non-maximal values. But in case of halftones, it has been found that max pooling (MP) of different halftone class may results in same output and hence the unique pattern distribution of the halftone class can be completely lost.

During the testing, the halftone image is divided into a sub-halftone images of size 32x32. Subsequently, the effective patches are extracted using the similar technique used for testing and classified individually. Further, the majority voting (MV) strategy [10] is used, which estimate final halftone class of the test image based on the maximum vote.

III. RESULTS AND PERFORMANCE EVALUATION

For performance evaluation, comprehensive experiments are conducted, including obtaining the optimization of CNN model with respect to number of layers, filter size, dataset splitting and the comparison studies over the classification accuracy with existing methods. The deep learning digital halftone database is generated which includes 24 halftone varieties [16]-[27]. The experiment is conducted in a system configuration of i5 processor, 3.40 GHz, 16 GB RAM and Nvidia GTX 1060 (6 GB RAM). The average correct classification rate (ACCR) is used as a performance index and is defined in Eq. 5.

$$ACCR = \frac{1}{N_c}\sum_{j=1}^{N_c} CCR_j. \quad (5)$$

It is the average of all the considered halftone varieties. During experimentation, for training 10,000 sub-halftone patches of size 32x32 is used for each class. The training, validation and testing set is divided in the ratio of 3:1:1 with the batch size of 200 and the experiment is performed for 100 epochs.

TABLE I: ACCR COMPARISON OF DIFFERENT HALFTONE CLASSIFICATION ALGORITHMS.

Halftone Types	**Chang, et. al [9]**	**Liu, et. al [10]**	**Zeng, et. al [12]**	**Zhang, et.al [13]**	**Guo. et. al [14]**	**Proposed**
	Maximum halftone class/CCR					
Ordered Dithering	2 /100%	2/ 100%	-	4/100%	4 / 99.6%	7/99.3%
Error Diff.	1/100%	3/ 98.2%	6 /97.9%	6/97.3%	8/ 97.2%	**11/98.74%**
Dot Diffusion	-	3/ 92.33%	-	3/ 89.3%	3/ 94.7%	3/**98.63%**
Iterative	-	1/ 100%	-	1 /100%	1/ 100%	1/**100%**
Multitoning	-	-	-	-	2 /99.4%	2/98.8%
ACCR	3/100%	9/ 97.6 %	6 /97.9%	14/ 96.65%	18/ 97%	**24/98.3%**

TABLE II. ACCR OF DIFFERENT MODEL DURING TRAINING.

S. No	No of Conv. Layer	Filter Size	No of filters	ACCR
Model-1	2	3x3	32/64	83%
Model-2	2	5x5	32/64	85.4%
Model-3	3 (MP)	3x3	32/32/64	74%
Model-4	3	3x3	32/32/64	99.8%
Model-5	3	5x5	32/32/64	100%

From the analysis, it has been concluded that the CNN requires minimum of three layers to obtain the better performance. From Model-3 it has been also validated that the addition of max-pooling (MP) results in poor classification accuracy. Among the Model-4 and 5, the filter of size 5x5 requires longer training time and performs only slightly better than 3x3. As there is no max pooling involved, the 3x3 filter is considered as the better choice to reduce computational demand. In Table 2, the comparison of the various existing methods with the proposed technique is tabulated. In comparison to the existing methods, the present work considered more halftone classes in ordered dithering and error diffusion types. From Table 1, it is also inferred that the existing approaches has limited performance particularly in error- and dot-diffusion halftones. It has been addressed in the previous evaluations that some of the halftone pairs are very similar [14] and have very close statistical properties. It has been also found that many halftone models are an improvisation version of one another and generate visually alike patterns which makes it very difficult for classification. The proposed CNN based deep learning architecture have overcome this issue and recognize the halftone pattern more precisely and achieved better performance.

In Fig. 3, the convolution layer output of clustered- and dispersed-dot for the first and third layers of convolution has been shown. It is inferred from the output that in the first layer the model learns on the distribution of lower level gray scale patches, and in the final convolutional layer the distribution towards the middle gray level value. Notably, the objective of our training is achieved as the deep learning feature contains equal distribution of black and white pixels. It has been already mentioned that this specific binary distribution contains maximal distinctive information and as the proposed architecture performs the classification using this distinctive feature, it results in the best classification accuracy.

(A)

(B)

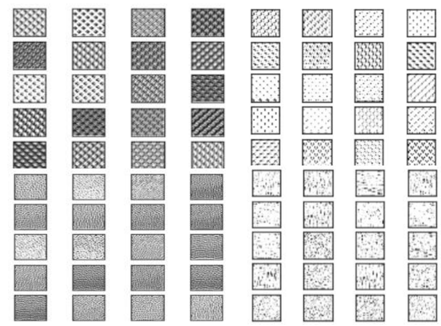


Fig. 3. Clustered (A) and Dispersed-dot (B) of first and third convolution layer for a sample test image.

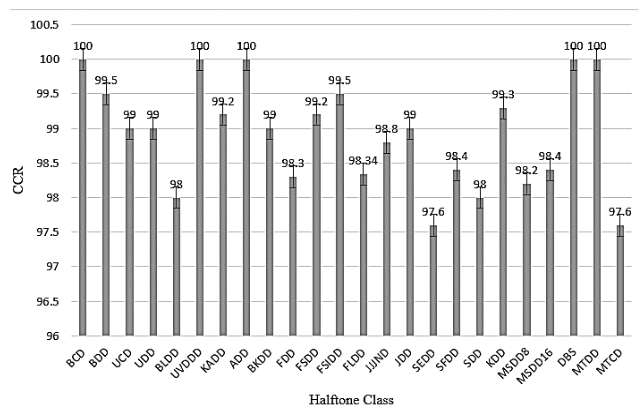


Fig. 4. Correct Classification Rate (CCR)) for different halftone class.

Fig. 4, shows the CCR of the 24 different halftones considered in this study. In comparison to the previous work, the proposed model achieves an improved classification over all the halftone, in particular for the error- and dot-diffusion halftones.

IV. CONCLUSION

A simple effective classification of halftone images based on CNN model is proposed. The proposed work considers over 24 different varieties of halftone images which are widely used in many multimedia printing applications. With this wide prospects, the inverse halftoning is a difficult task as it requires complete knowledge on the inherent halftone technique. To enhance the training and to identify the effective patches, the stochastic halftone parameters such as DDF, RAPSD and ANIS are adopted. In addition, the conventional CNN architecture is reconfigured to cope with halftone properties, and a simple three-layer architecture without max-pooling layer is developed. The Softmax classifier layer and adaptive moment estimator are utilized with the cross-entropy as the cost function. From CNN weightings, it has been validated that the model learns the distinctive characteristics more precisely, and results in the best classification accuracy in comparison with the state-of-the-art methods. The proposed approach achieved a consistent performance over all the 24 halftone varieties and achieved nearly 100% classification rate on several class. In future, the classification model can be extended to evaluate the quality assessment of halftones. The digital halftone database v2.0 is made open source and it will be useful for other halftone related researches such as restoration, retrieval and printed images processing [31].

REFERENCE

[1] R. Ulichney, *Digital halftoning*. MIT press, 1987.

[2] D. L. Lau, and G. R. Arce, Modern digital halftoning. CRC Press, 2008.

[3] D. L. Lau, R. Ulichney and A. R. Gonzalo "Blue and green noise halftoning models." *IEEE Signal Processing Magazine*, vol. 20, no. 4, pp. 28-38, Jul. 2003.

[4] B. E. Bayer, "An optimum method for two-level rendition of continuous tone pictures," *IEEE International Conference on Communications,* vol. 26, June 1973.

[5] R. W. Steinberg. and L. Floyd, "An adaptive algorithm for spatial greyscale," *Proceedings of the Society of information display*, vol. 17, pp. 75-77, 1976.

[6] D. E. Knuth, "Digital halftones by dot diffusion," *ACM Trans Graph (TOG),* vol. 6, no. 4, pp. 245–273, 1987.

[7] F.A. Baqai, and J. P.Allebach, "Halftoning via direct binary search using analytical and stochastic printer models," *IEEE Transactions on Image Processing,* vol. 12, no. 1, pp.1-15, 2003.

[8] M. J. Tsai et. al, "Source identification for printed documents," *International Conference on Collaboration and Internet Computing (CIC),* vol. 1, pp. 54-58, Dec. 2017.

[9] P. C. Chang, and C. S. Yu. "Neural net classification and LMS reconstruction to halftone images." *In Visual Communications and Image Processing*, vol. 3309, pp. 592-603, Jan. 1998.

[10] Y. F. Liu, J. M. Guo, and J. D. Lee. "Halftone image classification using LMS algorithm and naive Bayes," *IEEE Transactions on Image Processing*, no. 10, pp. 2837-2847, Oct. 2011.

[11] Z. Q. Wen, Y. X. Hu, and W. Q. Zhu, "A novel classification method of halftone image via statistics matrices," *IEEE Transactions on Image Processing*, no. 11, pp. 4724-4736, Aug. 2014.

[12] Z. Zeng, et. al, "Classification of error-diffused halftone images based on spectral regression kernel discriminant analysis," *Advances in Multimedia*, vol. 1, pp. 1-10, May 2016.

[13] Y. Zhang, E. Zhang and W. Chen, "Deep neural network for halftone image classification based on sparse auto-encoder," *Engineering Applications of Artificial Intelligence*, vol. 50, pp. 245-255, Apr. 2016.

[14] J. M. Guo and S. Sankarasrinivasan, "Generalized halftone classification approach using stochastic analysis," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-13, May 2018.

[15] J. M. Guo and S. Sankarasrinivasan, "Digital Halftone Database (DHD): A Comprehensive Analysis on Halftone Types," *Proceedings of APSIPA Annual Summit and Conference*, vol. 1, pp. 1091-1099, Nov. 2018.

[16] B. E. Bayer, "An optimum method for two-level rendition of continuous tone pictures," *IEEE International Conference on Communications,* vol. 26, June 1973.

[17] O. Bryngdahl, "Halftone images: Spatial resolution and tone reproduction," *Journal of Optical Society of America*, vol. 68, no. 3, pp. 416-422, 1978

[18] J. M. Guo, and Y. F. Liu, "Clustered-Dot Screen Design for Digital Multitoning," *IEEE Transactions on Image Processing,* vol. 25, no. 7, pp. 2971-2982, Jul. 2016.

[19] Kacker, D., Camis, T., & Allebach, J. P. Electrophotographic process embedded in direct binary search. *IEEE Transactions on Image Processing*, *11*(3), 243-257. Mar. 2002.

[20] R. A.Ulichney, "Void-and-cluster method for dither array generation," In *Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, pp. 332-344, Sep. 1993.

[21] Chandu, Kartheek, Mikel Stanich, Chai Wah Wu, and Barry Trager, "Direct multi-bit search (DMS) screen algorithm," *IEEE International Conference in Image Processing*, pp. 817-820, Sep. 2012.

[22] J. P. Seckar, "Relation of statistical information and visual quality in halftone images," *In Proceedings of the 21st International DAAAM Symposium*, vol. 21, no. 1, 2010

[23] D. Burkes, "Presentation of the Burkes error filter for use in preparing continuous-tone images for presentation on bi-level devices," *The file BURKES. ARC,* 1988

[24] Libaca study. (n.d.). Libcaca study: the science behind colour ASCII art. [online] Available at: http://caca.zoy.org/study/index.html [Accessed 30 Jan. 2019].

[25] J. F. Jarvis, C. N. Judice and W. H. Ninke, "A survey of techniques for the display of continuous tone pictures on bilevel displays," *Computer graphics and image processing*, vol. *5*, no. 1, pp.13-40, 1976.

[26] J. F. Jarvis J, C. Roberts, "A new technique for displaying continuous tone images on a bi-level display," *IEEE Trans Commun,* vol. 24, no. 8, pp. 891–898, 1976.

[27] M. Mese P. P. Vaidyanathan, "Optimized halftoning using dot diffusion and methods for inverse halftoning," *IEEE Transaction of Image Processing,* vol. 9, no. 4, pp. 691–709, 2000.

[28] K. Alex, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *In Advances in neural information processing systems*, pp. 1097-1105, 2012.

[29] P. N. Druzhkov, and V. D. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," *Pattern Recognition and Image Analysis*, vol. 26, no. 1, pp. 9-15, Jan 2016.

[30] P. K. Diederik, and Ba. Jimmy, "Adam: A method for stochastic optimization," *arXiv preprint*:1412.6980, 2014.

[31] J. M. Guo and S. Sankarasrinivasan, Digital Halftone Database v2.0 [online] Available at https://drive.google.com/file/d/1an1RIS3f63wArpjAyn1UwCCw9s8KMop1/view?usp=sharing [Accessed 30 Jan 2019].