New Cellular Neural Networks Template for Image Halftoning based on Bayesian Rough Sets

Elsayed Radwan[1,3*], **Basem Y. Alkazemi**[2,3], *and* **Ahmed I. Sharaf**[3]

[1]Faculty of Computer and Information Sciences, Mansoura University, Egypt.
[2]College of Computer and Information Systems Umm Al-Qura University, Makkah, Saudi Arabia
[3]Deanship of Scientific Research, Umm Al-Qura University, Makkah, KSA

Summary
Image halftoning is a technique for varying grayscale images into two-tone binary images. Unfortunately, the static representation of an image-half toning, wherever each pixel intensity is combined by its local neighbors only, causes missing subjective problem. Also, the existing noise causes an instability criterion. In this paper an image half-toning is represented as a dynamical system for recognizing the global representation. Also, noise is reduced based on a probabilistic model. Since image half-toning is considered as 2-D matrix with a full connected pass, this structure is recognized by the dynamical system of Cellular Neural Networks (CNNs) which is defined by its template. Bayesian Rough Sets is used in exploiting the ideal CNNs construction that synthesis its dynamic. Also, Bayesian rough sets contribute to enhance the quality of the halftone image by removing noise and discovering the effective parameters in the CNNs template. The novelty of this method lies in finding a probabilistic based technique to discover the term of CNNs template and define new learning rules for CNNs internal work. A numerical experiment is conducted on image half-toning corrupted by Gaussian noise.
Keywords:
Digital Image Half-toning, Bayesian Rough Sets, Noise Reduction, Cellular Neural Networks, CNNs Templates.

1. Introduction

Digital halftoning [1-3] is a dithering method for altering grayscale images into two-hued binary images. These half-tone images can resemble the original images when viewing from a remoteness by the low pass sieving in the Human Visual System (HVS). The half toning techniques increase the number of observed representation levels by capitalizing on the fact that the HVS uses averaging of neighboring intensities. Unfortunately, many previous approaches search on the local neighbors to detect the pattern which cause the missing in subjective. Thus, image half-toning should be concluded by searching on the core features and extracting pattern from global image. Hence, its computation should be characterized by simple processing elements, local interconnect and massive parallelism[4-6] . Thus, image half-toning can be described by a set of rules that illustrate its interconnection pattern and processed synchronously. The extracted pattern should be discovered and treated as a dynamical system. Therefore, image half-toning processing is well suited for the dynamical system of Cellular Neural Networks.

Cellular Neural Networks (CNNs), spatial arrangement of locally coupled cells, has a great contribution in image processing, especially halftoning [7, 8]. CNNs based halftoning methods operate the large image chunk by chunk with 3×3 templates or 5×5 templates to enhance the quality of halftone image. CNNs overcome the problem of noticeable boundaries of image by careful selection of boundary cell values. Also, CNNs discover the global pattern and save the processing time, because they are parallel analog architecture with local interconnection. The important point in the CNNs based halftoning methods is how to find suitable templates[9], which enhance the quality of halftone image. Unfortunately, CNNs dynamic is affected by the superfluous cell associated with noise in the input image. Thus, to neglect noise effect, a tolerance-based method is needed for understanding CNNs template. The optimal CNNs template should be revealed based on a probabilistic machine learning technique.

Since the existed noise in any data pattern degrades its classification accuracy and cause a perturbation in the CNNs dynamic as a consequence of missing its correct architecture, then discovering the optimal local rules that discern the noise is a fundamental demand [10, 11]. Thus, previous works are appeared to reduce them. Most of them depend on the spatial and frequency domain.

This paper depends on filtering the local image from noise and discovering the dynamical function that govern the pixel relation among its local neighbors based on hybrid Bayesian Reasoning [12, 13] then performing connectivity among all pixels based on missive parallel processing. Also, this paper relies on Bayesian Rough Sets concept to discover the optimal template CNNs architecture by removing the superfluous neighboring cells. These superfluous neighboring cells have no effect on classifying the cell's output. Another important concept of Bayesian Rough Sets is its ability to defining new sign measure for

Manuscript received April 5, 2023
Manuscript revised April 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.4.11

each neighbor cell. The sign measure infers the relation among the template parameters and give an expectation of CNNs behavior. Also, by Bayesian Rough Sets, similarities and regularities in the input data are discovered and excluded. Hence the learning time is reduced. This paper finds conditions on the elements of the CNNs and ensures a correct functioning of the CNNs for an image halftone application. Moreover, a new supervised learning rules for CNNs temples are discovered. In demonstrating the efficiency of the new technique an experiment is conducted on half toning digital images database. A comparative study for some successive techniques is mentioned.

This paper is organized as follows, section 2 discusses the preliminaries, section 3 mentions the paper methodology where the role of Bayesian Rough sets is declared in feature reduction and in learning the CNNs template. An experiment associated with a comparative study is appeared in section 4. Finally, the paper is concluded in Section 5.

2. Preliminaries

In this part a short survey is introduced for image halftoning, Cellular Neural Networks and Bayesian Rough sets.

2.1 Image Half-toning

Image halftoning is a technique that convert analog image into digital one.it should keep the subjective of an image and save toner [14]. Many techniques treated this problem based on dithering, error diffusion and iterative methods. In dithering, a gray scale image is considered as chunks, e.g. 3×3 block. Each chunk is processed by binary grid. Unfortunately, dithering method is so fast, but noise is sharing based on pixel's spatial coordinates. On the other hand, error diffusion method results are accurate because it depends on circular filter. A sign error measure is defined based on the neighbor of pixels for processing [2, 15]. Unfortunately, this technique is costly and time consuming. Finally, the iteration method processes the global image iteratively. It depends on reducing the difference between the toned image and the gray scale image by inspecting on the subjective of binary values on the resulted image. Also, it's time consuming although it yields high quality toned image. Thus, it is better to treat the global image for accurate result moreover search on the occluded objects in suitable time. Hence, a hybrid probabilistic method (based on posterior probability) in addition to a parallel processing technique can improve the image binarization.

2.2 Cellular Neural Networks

CNNs is a dynamical computing grid of cells associated with a set of cell state transition rules that describe the relation among neighbor cells and real time monitor. It is completely described by a cellular behavior and Neural Network models [16, 17]. CNNs hold the benefit of Cellular Automata model [5, 18], where the transition rules label its state, and its production in the actual time, Neural Networks. CNNs [19] is (2D/3D) spatial arrangement of nonlinear analogue to digital dynamical processors called cells. Each cell cooperates synchronously within finite local neighbors that belong to the sphere of influence with radius r,

$$N_r\,(i,j) = \{c_{kl} : max(\,i\,-\,k\,, j\,-\,l\,) \leq \,r\} \tag{1}$$

CNNs is characterized by a nonlinear dynamical system. Each CNNs dynamical cell has a state x evolved in the real time according to some a set of dynamical rules in purpose of producing an output y, a piece-wise sigmoid function on the current state. The cell dynamics are functionally determined by a small set of parameters which control the cell interconnection strength called templates (A, B, z) where A is the feedback template, B is feedforward template. $Aand\ B$ are $(2r+1)\times(2r+1)$ real valued matrices, and z is the search bias's threshold. Thus, CNNs should be recognized based on its topology and its dynamics. The cell dynamics are described by following equations

$$\frac{dx_{ij}}{dt} = -x_{ij}(t) + \sum_{kl\in N_{ij}^r} A_{ij,kl}y_{kl} + \sum_{kl\in N_{ij}^r} B_{ij,kl}u_{kl} + z_{ij} \tag{2}$$

$$y_{ij} = f(x_{ij}) = \frac{1}{2}\ \left(|x_{ij}+1| - |x_{ij}-1|\right) \tag{3}$$

$$-1 \leq x_{ij}(0) \leq 1 \quad -1\ \leq u_{ij}(t) \leq 1 \quad |z| \leq z_{max}, 1 \leq i \leq M \quad 1 \leq j\ \leq N$$

2.3 Bayesian Rough Sets

Rough sets [20, 21] is an indiscernibility approach that reveal objects mechanism and process vagueness and uncertainty to extract the knowledge. It reduces the feature subset along with saving the semantics of the whole Information System (IS). In rough set approach, an approximation space is defined based on a 1-1 onto indiscernibility relation among objects of the universe U. Rough Set Model (RSM) generates two subsets by means of the equivalence relation E; namely, lower and upper approximations.

$$\mathfrak{L}(\mathrm{E(X)}) = \{\mathrm{u} \in \mathrm{U}: \mathrm{P}(\mathrm{X}|[\mathrm{u}]_\mathrm{E}) = 1\}, \tag{4}$$

$$\mathfrak{U}(\mathrm{E(X)}) = \{\mathrm{u} \in\ \mathrm{U}: \mathrm{P}(\mathrm{X}|\ [\mathrm{u}]_\mathrm{E}) > 0\} \tag{5}$$

The Pawlak approximation space (U, E) measures the approximation roughness by

$$\rho(X) = 1 - \alpha(X) = 1 - \frac{|\mathfrak{L}(E(X))|}{|\mathfrak{A}(E(X))|} \quad (6)$$

Unfortunately, variations in the dataset and missing attribute values disable the indiscernibility relation to extract equivalent classes. Therefore, some extensions of rough sets are presented like covering rough sets model [22], dominance rough sets and tolerance rough sets [23, 24]. These extensions depend on defining a major inclusion relation, like similarity relation, instead of equivalence relation. By majority relation some degree of noise data, misclassified objects, are allowed in the largely correct classification. As a result, all rough sets extensions currently in use depend on different kinds of heuristics where the resulted decision rules contain the crisp and probabilistic ones.

Bayesian Rough Sets [25, 26], a fashion of Bayesian reasoning and rough set, is introduced for discovering the latent knowledge in the uncertainty dataset. BRSM depends on increasing the prior probability of occurrence of the target event. It deals effectively with distorted data and discovers latent useful knowledge in boundary region. BRSM defined the positive and negative regions, certain knowledge, by increasing the prior probability of subset $X \subseteq U$ based on

$$Pos^*(X) = \{u \in U: P(X|[u]_E) > P(X)\}, \quad (7)$$
$$Neg^*(X) = \{u \in U: P(\neg X|\,[u]_E) > P(\neg X)\}$$

Thus, uncertain knowledge is defined as the set of all elementary knowledge that has no effect on the prior probability, i.e.

$$BND(E(X)) = \{u \in U: P(X|[u]_E) = P(X)\} \quad (8)$$

Thus, BRSM can search for the optimal local rules that discover the pattern and excluding noise from the most simplified construction. BRSM (almost) preserve consistency with data and classify unseen objects with the lowest error risk. Therefore, the capability of classifying more objects with high accuracy.

3. Designing CNNs Template by Bayesian Rough Sets

By this work, the assumption is representing the image halftoning operation as steadiness conditions of coupled CNNs. The learned CNNs template is measured depending on the data consistency, by means of BRSM. For any input image $U_{\{m\times n\}}$, the output of any cell y_{ij} (∞) is uniquely dominated by a slight array of U subjected to (2r +1)×(2r +1) revolve around cell c_{ij}, with a sphere of influence N_r (ij) .So, the output y_{ij} (∞) is a map from (2r +1)(2r +1) of response cells, as demonstrated in Figure 1, and the initial state x_0 into the set $\{-1,1\}$. The attribute values of CNNS should be in [-1,1], thus a conversation function $cf(.)$is used.

$$cf(c_{kl}) = -1 + 2\frac{c_{kl}}{255}$$

where $y_{ij} : f\left(x_0, c_1, \ldots, c_{(2r+1)(2r+1)}\right) \rightarrow \{-1,1\}$.

For feature reduction, decision table $DT = (U, C \cup Y(\infty), V, g)$ is constructed, C= $\{x_0, c_1, \ldots, c_{(2r+1)(2r+1)}\}$, V is the domain values of $|C|$ random variables, a value in [-1,1], and g is a random variable function where $P(g(c_i, x_{ij}), x_{ij} \cong N_r(i,j)) > 0$. Each row in DT associated with the pixel x_{ij} is characterized by a single record at location $i * m + j$.

$$y_{ij}(\infty){:}X_0 + \begin{pmatrix} c_{i-1,j-1} & c_{i-1,j} & c_{i-1,j+1} \\ c_{i,j-1} & c_{ij} & c_{i,j+1} \\ c_{i+1,j-1} & c_{i+1,j} & c_{i+1,j+1} \end{pmatrix} \longrightarrow \{-1,1\}$$

Figure 1: The CNNs effective box for classifying the output.

This work searches for a relation between the CNNs dynamic besides its interpretation and an image half-toning operation represented by a Decision Table. Since uncoupled CNNs is stable dynamical system, image half-toning operation is assumed as an uncoupled CNNs in purpose of discovering a new feedforward learning rule. Also, a superfluous template value is discovered. On the other hand, a sign measure is defined where the CNNs synaptic behavior is studied.

Proposition 1: The space invariant uncoupled CNNs can be represented well by consistent Decision Table (DT) and vice versa.

Proof (sufficient condition)

Consider an uncoupled CNNs then its dynamic is a stable, i.e. bipolar output. Also, its state equation is completely determined by the initial state X_0, the self-feedback value and the input of its neighbors' cells. Since the output for the cell ζ is determined by[27]

$$y_\zeta = sgn((a_{00} - 1)x_\zeta(0) + w_\zeta) \quad (9)$$
$$w_\zeta = \sum_{c_{kl} \in N_r(\zeta)} b_{kl} u_{kl} + z$$

w_ζ depends totally on the inputs of the neighbor cells, depicted in Figure 2.

$$\begin{pmatrix} u_{i-r,j-r} & u_{i-r,j} & u_{i-r,j+r} \\ u_{i,j-r} & u_{ij} & u_{i,j+r} \\ u_{i+r,j-r} & u_{i+r,j} & u_{i+r,j+r} \end{pmatrix} \Longrightarrow$$
$$\begin{pmatrix} u_{(2r+1)(2r+1)} & u_{2(2r+1)+r+1} & u_{2(2r+1)+1} \\ u_{2(2r+1)} & u_{((2r+1)(2r+1)+1)/2} & u_{(2r+1)+1} \\ u_{(2r+1)} & u_{(r+1)} & u_1 \end{pmatrix}$$

Figure 2: The transparent window for classifying the output.

then $y_\zeta = f(x_0, u_1, \ldots, u_{(2r+1)(2r+1)})$ has unique value, thus

$$P\big(Y = d \big| c_0, c_1, \dots, c_{(2r+1)(2r+1)}\big) = P(Y = d)$$

Hence $H(Y|C)$,

$$H(Y|C) = \frac{P(Y)-P(Y|C)}{\ln 2}\big(P(Y|C)\ln P(Y|C) + P(\neg Y|C)\ln P(\neg Y|C)\big) = 0 \quad (10)$$

The uncertainty measure equals zero, $Y = \{y_\zeta | y_\zeta = 1\}$ the event that cell's output lies in the positive saturation region. Thus, two Joint probability random variables are constructed

$X = \{c_0, \dots, c_{(2r+1)(2r+1)} | y_\zeta = 1\}$ and $\bar{X} = \{c_0, \dots, c_{(2r+1)(2r+1)} | y_\zeta = -1\}$ that yield disjoint and complement events, $P(X) + P(\bar{X}) = 1, X \cap \bar{X} = \phi$.

Hence, $DT = \big(U, X_0 \cup u_\zeta \cup y_\zeta\big)$, $u_\zeta = (u_1, u_2, \dots, u_{(2r+1)(2r+1)})$ is definable under $C = \{c_0, \dots, c_{(2r+1)(2r+1)}\}$, a consistent decision table.

(Necessity Condition) is proved in[28] . ■

In short, Uncoupled CNNs can be described as a Joint Probability Random Variable with $(2r+1)(2r+1)+1$ random variables associated with a measurable space if the decision table is consistent. Since image half-toning operation should save the subjective of an image, then a tolerance relation should be defined:

$$P\big(Y = d \big| c_0, c_1, \dots, c_{(2r+1)(2r+1)}\big) > P(Y = d)$$

i.e.

$$E\big(Y \big| c_0, c_1, \dots, c_{(2r+1)(2r+1)}\big) > E(Y)$$

Where Y is the bipolar classified output that represent a toned image, denoted as $\{X, \neg X\}$. In [29], authors have proved $P(Y = X) - P(Y = \neg X) = E(Y)$ for bipolar image. The expected value of the subjective point should be greater than the prior probability of the decision.

$$E\big(Y \big| c_0, c_1, \dots, c_{(2r+1)(2r+1)}\big) > P(X) - P(\neg X) \quad (11)$$

This means uncoupled CNNs is not a suitable operation for image halftoning, where the CNNs output should share in the subjective classification. Before computing the effective random variable based on BRSM that discern noise in the nearest positive cluster and save the subjective according to eq. (11), the maximum error caused in uncoupled CNNs by disconnecting a neighboring cell should be concluded.

Proposition 2: For Uncoupled CNNs, disconnecting a feedforward input cell, feedforward b_{kl} , cause and error limits of $abs(b_{kl}/(1-a_{00}))$, where a_{00} is the self-feedback.

Proof:

Since the output function, eq. (3), is a linearly increasing piecewise function of a trajectory, the changes emitted by neglecting some feedforward weight is like the variations in its trajectory. To define the perturbation, the output function should satisfy the Lipschitz condition of stability [27], i.e.

The output function is continuous, Let the model $y(x_{ij}) = f(x_{ij}) = \tanh(x_{ij})$

$|f(x_i) - f(x_j)| \le L|x_i - x_j|$ satisfied

For differentiation assumption, the output function is simulated by $y(x_{ij}) = f(x_{ij}) = tanh(x_{ij})$. Substituting in Mean Value Theorem the left-hand side becomes.

$$\big(f(x(b_{kl},t)) - f(x(0,t))\big) = f^{`}(x(\zeta b_{kl},t)) - x(0,t)) \quad , \zeta \in [0,1]$$

Since the trajectory of uncoupled CNNs is given by [27]

$$x(b_{kl},t) = \frac{b_{kl}}{1-a_{00}} + (x(0) - b_{kl} - a_{00})e^{(a_{00}-1)t}, t>0$$

Thus $\frac{df(x(b_{kl},0)}{db_{kl}} = \frac{df}{dx}\cdot\frac{dx}{db_{kl}}$

$$= sech^2(x)\Big(\frac{1}{1-a_{00}} - e^{-(1-a_{00})t}\Big),$$

$$= \big(1 - tanh^2(x)\big)\Big(\frac{1}{1-a_{00}} - e^{-(1-a_{00})t}\Big), t > 0$$

For the worst case, weak positive uncoupled CNNs

$$\frac{df(x(b_{kl},0)}{db_{kl}} = \frac{\big(1-tanh^2(x)\big)}{1-a_{00}} < \frac{1}{1-a_{00}}$$

$$\big(f(x(b_{kl},t)) - f(x(0,t)\big) < \frac{1}{1-a_{00}}(x(b_{kl},t) - x(0,t))$$

From Lipschitz condition 2, $\frac{1}{1-a_{00}} = L$

For uncoupled CNNs, the transient equation is well posed differential equation with Lipschitz constant

$$\hat{L} = 1 + L\|A\| = 1 + \frac{a_{00}}{1-a_{00}} = L,$$

Also, for uncoupled CNNs

$$x(\zeta b_{kl},t)) - x(0,t) < d\varepsilon, and\ \varepsilon \ll b_{kl}$$

Hence, for $d = 1$

$\big|f(x(b_{kl},t)) - f(x(0,t))\big| \le \Big|\frac{b_{kl}}{1-a_{00}}\Big|$ which complete the proof. ■

Corollary 3: For uncoupled CNNs, in the learning process, the changes in the feed-forward template is recognized by equation

$$b_{kl}(t+\Delta t) = b_{kl}(t) + \eta \sum_{i=1}^{M}\sum_{j=1}^{N}\big(y_{ij}^d - y_{ij}\big)\Big(1 - tanh^2\Big(\frac{b_{ij}}{1-a_{00}}\Big)\Big)\cdot\frac{u_{b_{kl}}}{1-a_{00}} \quad (12)$$

Where η is the learning rate, $u_{b_{kl}}$ is the corresponding input of b_{kl} , y_{ij}^d is the desired output, and y_{ij} is the actual output.

Proof:

For any Neural Networks, the learning rule for the back-propagation model is determined in accordance to the derivative of its error function. Thus, $w = w_{old} + \eta\frac{\partial E(w)}{\partial w}$

Where η is the learning rate and $E(w)$ is the error caused in the whole CNNs output. By substituting from proposition 2 the proof result directly. ■

Although Rough Sets yield the core knowledge with reduced set of features, called reduct, that is representative for the whole knowledge, it faces some difficulties with noisy and dynamical data. The rational objective function is missing because of noise and global behavior. Thus, some tolerance relation is needed to increase the probability of detecting the subjective of toned image. In bipolar decision, BRSM guarantee a symmetry and equivalence relation wherever the self- duality is satisfied [30]. BRSM depends on measuring relative gain.

$$R(Y|C) = \sum_{E\in U/C} P(E)\left(\max\left(g(X|E), g(\neg X, E)\right)\right) \quad (13)$$

$g(D|E) = \frac{P(D|E)}{P(D)} - 1$, E is corresponding to some domain values of Joint Probability r.v. $\mathcal{g}$. Thus, the reduced set of attributes $B \subseteq C$ for a decision System $DS = (U, C \cup D, V, \mathcal{g})$, where $U|D = \{X, \neg X\}$ should satisfy

$R(D|C) = R(D|B)$ and $H(D|B) - H(D|C) \leq \varepsilon$

2. B is independent subset of C, i.e. for each $a_i \in B, H(D|B - \{a_i\}) > H(D|B)$ where strict inequality holds. Because uncoupled CNNs depends totally on static input feature of the cells, thus the ***reduct*** algorithm for decision table describe its dynamic without loss on its semantics. Apparently, the feature with less uncertainty score, eq. (10), would be selected. Also, the boundary region, eq. (8), is reduced where the relative certainty gain is increased, eq. (13). The reduction algorithm, Algorithm (1), estimates first the relative certainty gain of every feature. Then, all features' uncertainty measure is considered. The consequence caused by disturbance in the decision system of uncertainty measure will decrease where a tiny perturbation of value will not result a great changes of uncertainty measure. The reduction algorithm considers a symmetric feedforward template, i.e. if c_i is a superfluous attribute, then $c_{(2r+1)(2r+1)+1-i}$ is considered a superfluous.

Algorithm (1) The computation of the reduced set of features that represent Uncoupled CNNs

Require: DT, C the set of all conditional feature, $U|Y = \{X, \neg X\}$ the decision class $X = \{y: y = 1\}$,
Ensure: B -a reduct of feature subset.

1. $B \leftarrow C, T \leftarrow C$
2. While $(T \neq \phi)$
3. Foreach $(c_i \in B)$
    a. compute $R(Y|c_i)$
4. compute $\min_i R(Y|c_i)$
5. If $(R(X|C) = R(X|B - \{c_i\})$ # Superfluous Cell
    a. $Delet(B, c_i)$ [31]
    b. $if\ H(D|B) - H\left(D|B - \{c_{(2r+1)(2r+1)+1-i}\}\right) \leq \varepsilon$)
        i. $Delet(B, c_{(2r+1)(2r+1)+1-i})$
        ii. $T \leftarrow T - \{c_{(2r+1)(2r+1)+1-i}\}$
6. Else
    a. Break
7. End if
8. $T \leftarrow T - \{c_i\}$
9. End while
10. Return B

According to the reduction algorithm, neglecting the superfluous attribute, the subsequent feedforward template should be removed too. Since CNNs is an analog to digital processor with dynamically stable CNNs, then $P(X|E) > P(X)$ where $X\ and\ E$ indicates the output $y = 1$ and the class caused by the input cells respectively. Thus,

$$P\left((a_{00} - 1)x_\xi(0) + \omega_\xi > 0\right) > \frac{N_{X^+}}{N} \quad (15)$$

Where $\omega_\xi = \sum_{c_j\in B} b_{Bj}u_j + z$, b_{Bj} is the feedforward template corresponding to the j[th] cell that form the reduced pattern B whereas z is the bias. If a feedforward connection is reduced from some effective cells, some decision subsets become inconsistent i.e. $P(\bar{X}|drop\ c_i) = P(\bar{X})$. Thus

$$P\left((a_{00} - 1)x_\xi(0) + \sum_{c_j\in B} b_{Bj}u_j + z - b_{Bi}u_i > 0\right) = \frac{N_{\bar{X}^+}^{c_i}}{N_{\bar{X}}^{c_i}} \quad (16)$$

Where $N_{\bar{X}^+}^{c_i}$ and $N_{\bar{X}}{}^{c_i}$ are the number of positive output and total number output by dropping the cell c_i. From eq. (15)

$$P(\bar{X} + b_{Bi}u_i > 0) > \frac{N_{X^+}}{N}$$

Since

$P(\bar{X} + b_{Bi}u_i > 0) = P(\bar{X} > 0) + P(-b_{Bi}u_i < \bar{X} < 0)$ where $\{\bar{X} > 0\}\ and\ \{(-b_{Bi}u_i < \bar{X} < 0)\}$ are mutually exclusive events, Thus

$P(\bar{X} + b_{Bi}u_i > 0) - P(\bar{X} > 0) = P(0 < \bar{X} + b_{Bi}u_i < b_{Bi}u_i)$

Hence

$$P(0 < X < b_{Bi}u_i) \geq \frac{N_{X^+}}{N} - \frac{N_{\bar{X}^+}^{c_i}}{N_{\bar{X}}{}^{c_i}} \quad (17)$$

Since the input vector of CNNs are in [-1,1], thus

$$P(0 < X < b_{Bi}u_i < b_{Bi}) \geq \frac{N_{\bar{X}}{}^{c_i} * N_{X^+} - N * N_{\bar{X}^+}^{c_i}}{N * N_{\bar{X}}{}^{c_i}} \quad (18)$$

Where, $N = N_{X^+} + N_{\neg X^+}$ and $N_{\bar{X}}{}^{c_i} = N_{\bar{X}^+}^{c_i} + N_{\neg\bar{X}^+}^{c_i}$, where $N_{\neg X^+}$ is the total number of negative output without dropping any cell attribute and $N_{\neg\bar{X}^+}^{c_i}$ is the total number of negative output with dropping cell attribute . by substitution in eq. (18)

$$P(0 < X < b_{Bi}u_i < b_{Bi}) \geq \frac{(N_{\neg\bar{X}}{}^{c_i} * N_{X^+}) - (N_{\neg X^+} * N_{\bar{X}^+}^{c_i})}{N * N_{\bar{X}}{}^{c_i}} \quad (19)$$

From probability axioms the right-hand side of eq (19) should be greater than zero, hence the nominator should be greater than zero, i.e.

$$(N_{\neg\bar{X}}{}^{c_i} * N_{X^+}) - (N_{\neg X^+} * N_{\bar{X}^+}^{c_i}) > 0 \quad (20)$$

Corollary 4; For CNNs architecture, the feedforward weight b_{Bi} related to the cell c_i exhibits excitatory or

inhibitory based on the sign of $(N_{\neg\bar{X}}{}^{c_i} * N_{X^+}) - (N_{\neg X^+} * N_{\bar{X}^+}^{c_i})$.
Proof: The proof is a direct result of Eq (20).

On behave of the image half-toning, the processed pixel is related to the input and the output neighbors synchronously to save the subjective of an image. Thus, the data consistency cannot be recognized by uncoupled CNNs. Hence, in purpose of saving the subjective of an image and enhance knowledge classification, new features should be added depending on the output neighbors. This means the output of image toned cells depends on the reduced set of input feature in addition to the output of the neighboring cells.

$$y_{ij} = f(x_0, B_u, y_{i-r,j-r}, y_{i-r+1,j-r}, \dots y_{i+r,j+r}) \tag{21}$$

By eq. (21) the image half-toning process is determined by coupled CNNs. For coupled CNNs, the CNNs dynamics is stated by dynamical rules with feedback weights representation. Thus, stability criteria of CNNs should be expanded to handle the state transient for the neighboring cells. Hence, the global representation of image half-toning is recognized, if the cell output is increased by contributing the state transient of the neighboring cells, the condition by eq (22) should be satisfied.

$$P(y_{ij}(t) | x_{kl}(\infty),\ c_{kl} \in N_r(ij)) > P(y_{ij}(t)) \tag{22}$$

Where $x_{kl}(\infty)$, *state transient*, is a function that depends on the input pattern, the initial state $x_{kl}(0)$ and the output of the neighbor cells. Based on the comparison principle the feedback template [32] is calculated as

$$a_{kl}(t + \Delta t) = a_{kl}(t) + \eta \sum_{i=1}^{M} \sum_{j=1}^{N} \left(2 - (y_{ij}^d - y_{ij})\right)(1 - y_{ij}^2)^2 y_{a_{kl}} \tag{23}$$

A summary of BRSM in learning image half-toning process based on CNNs is introduced by Algorithm (2), which illustrates the pseudo code of this method. BRSM deduce the Global Expected Certainty Gain, $\lambda(D|C) = \lambda(D|B), B \subseteq C$, which is normalized to measure the certainty factor.

$$\lambda(D|C) = \frac{egabs(D|C)}{2P(D)P(\neg D)} \tag{24}$$

and

$$egabs(D|C) = \sum_{E \in U|C} P(E) abs(P(D|E) - P(D))$$

If $\lambda(X|B) = 1$ then a consistent DT. This means the reduced feature set B is representative for image half-toning operation.

Algorithm (2) Learning CNNs for image half-toning using BRSM

Require: $DT = (U, X_0 \cup u_\zeta \cup y_\zeta), u_\zeta = (u_1, u_2, \dots, u_{(2r+1)(2r+1)})$ and r – radius of influence sphere

Ensure: $A, B, and\ z$ CNNs learning template

1. If $r > 3$ then
   a. Print $\alpha(X^+)$ eq (6)
   b. Return $A, B\ and\ z$
2. End if
3. 3. Reduce DT for Uncoupled CNNs using Algorithm (1) (reduct $= B_u$). Compute the behavior of each feature from eq (20)
4. If $\lambda(D|B_u) = 1\ or\ |POS^B(X^+) - X^+| < n\ then$
   a. Construct uncoupled CNNs feedforward template as the reduct of DT.
   b. Learn self-feedforward template using eq (12) where $a_{00} > 1\ and\ z = z + \eta\sum\sum(sech^2(x_{ij}) + y(\infty) - y^d)$
   c. Return $a_{00}, B\ and\ z$
5. Else
   a. Print $\alpha(X^+)$
   b. Reconstruct $\overline{DT} = (U, x_0, \bar{C})$, $\bar{C} = (B, y_{i-r,j-r}, y_{i-r+1,j-r}, \dots y_{i+r,j+r})$
   c. Construct decision rules as a probabilistic CA model
   d. If $\lambda(D|\bar{C}) = 1$ then
      i. Learn Feedback template A using eq(23) and feedforward template using eq(12) , $z = constant$
   e. Else
      i. Increase the radius $r = r + 1$
      ii. Reconstruct the DT.
      iii. Call Algorithm (2)
   f. End if
6. End if

4. Experiment (IMAGE HALF-TONING)

Image half toning is selected to demonstrate the ability of proposed method to recognize the propagating type template. According to our method this template cannot be recognized by a single layer with 3×3 but it can be recognized by 5×5 as shown below and in Figure 3

At first stage, 3×3 *template*, the algorithm is inconsistent as $\alpha = 0.8$, applying Bayesian rough sets algorithm and pruning the noise, we found there is no superfluous cells. The set of effective cells are all neighboring cells

(a) Gaussian corrupted Image

(b) The source image

The 3×3 image

The 5×5 image

Fig.3 Learning CNNs for image Half-toning

Concatenating the output corresponding to the set of effective cells as additional attributes, i.e. the new attributes became {u0, u1, u2, u3, u4, u5, u6, u7, u8, u9,y1,y2,y3,y4,y6,y7,y8,y9} where the output of the cell itself is removed , and check the consistency.

Now the decision algorithm is inconsistent, α =0.88, with robustness measure $\Delta = 0.013$

429 positive rules and 388 negative rules and the effective cells are $\{u_0, u_1\, u_3,\ u_7, u_9, y_1, y_2, y_3, y_4, y_6, y_7, y_8, y_9\}$ 1 with sign measure $\{(\frac{419}{381}),(\frac{415}{384}),(\frac{417}{380}),(\frac{414}{382}),(\frac{420}{385}),(\frac{405}{360}),(\frac{405}{365}),(\frac{407}{362}),(\frac{410}{363}),(\frac{402}{360}),(\frac{400}{361}),(\frac{403}{359}),(\frac{409}{366})\}$, i.e. {+,+,+,+,+,-,-,-,-,-,-,-,-}

Our template is considered as

$$A = \begin{bmatrix} - & - & - \\ - & a_5 & - \\ - & - & - \end{bmatrix} \quad B = \begin{bmatrix} + & 0 & + \\ 0 & + & 0 \\ + & 0 & + \end{bmatrix} \quad z = R_{real}$$

since the initial state is considered as the image itself

By applying learning rules, with epochs =50, the template parameters are as follows where the similarity among the template parameters is considered.

$$A = \begin{bmatrix} -0.291 & -0.291 & -0.291 \\ -0.291 & 1.426 & -0.291 \\ -0.291 & -0.291 & -0.291 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.392 & 0 & 0.392 \\ 0 & 2.804 & 0 \\ 0.392 & 0 & 0.392 \end{bmatrix}, \text{z=-0.166}$$

By increasing the radius of the sphere of influence to be two, and then checking the consistency , it is inconsistent algorithm with α =0.972 and applying Bayesian rough set algorithm for decision is applied, there is no superfluous cells and the set of effective cells is $\{C_0, C_1, C_2, C_3, \dots, C_{23}, C_{24}, C_{25}\}$

Complete the decision algorithm by concatenating the output corresponding to the set of effective cells, the new attributes became $\{u_0, u_1, u_2, u_3, \dots, u_{23}, u_{24}, u_{25}, y_1, y_2, \dots, y_{14}, \dots, y_{24}, y_{25}\}$ and then check the consistency of modified decision algorithm. The modified decision algorithm is consistent with α =1, and after applying the algorithm the structure of optimal template is constructed:

$$A = \begin{bmatrix} -0.091 & -0.107 & -0.138 & -0.107 & -0.091 \\ -0.107 & -0.328 & -0.556 & -0.328 & -0.107 \\ -0.138 & -0.594 & 1.229 & -0.594 & -0.138 \\ -01.07 & -0.328 & -0.594 & -0.328 & -0.107 \\ -0.091 & -0.107 & -0.138 & -0.107 & -0.091 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.042 & 0 & 0.475 & 0 & 0.042 \\ 0 & 0.419 & 0 & 0.419 & 0 \\ 0.475 & 0 & 2.391 & 0 & 0.475 \\ 0 & 0.419 & 0 & 0.419 & 0 \\ 0.042 & 0 & 0.475 & 0 & 0.042 \end{bmatrix}, \quad \text{z=-0.06}$$

We defined the comparison to be the percent of error occurred as the result of the robustness changes on the template parameters and the number of iterations that are needed for each cell to enter the saturation region. Also, we extended the comparison to handle the ability to discover the optimal template structure. According to our comparison, we can say that GA is always in need of other methods to complete its shortcutting. Also, the truncation learning rules perform the same as GA.

a. The source image

b. 3 × 3image halftoning

5 × 5 CNNs image halftoning

Figure 4: image halftoning processing based on BRSM

Table 1: an image halftoning Comparative table using different learning methods for CNNs

	Error %	Robustness%	N-Experiment	N-epochs	Optimal Structure
Rough Sets +GA	0 0.0048 0.053	4 5 10	10	200 200 220	Yes
Bayesian Rough set	0.004	3	2	230	Yes
GA	0 0.012 0.071	2 5 10	10	250 250 327	No
Truncation Learning	0.006 0.187	0.5 10	5	300 300	No

A comparison among Bayesian Rough Sets in designing CNNs template for image-halftoning and other previous methods such as ordered dithering, error diffusion, table halftoning and GA for template learning is summarized in Figure 7. The evaluation process of the proposed system requires measuring the deviation of pixel output intensity with respect to the average of its neighbor to ensure the desired performance of the proposed system. Unfortunately, the Mean Square Error is not significant because of its violating value. So, the Peak Signal to Noise Ratio (PSNR) is used to declare the effectiveness of the proposed system and alter the MSE to be maximize with amplification.

$$MSE = \frac{1}{MN}\sum \sum_{for\ all\ M,N} \left(y_{ij}(\infty) - \overline{y_{ij}}\right)^2$$

Where $\overline{y_{ij}}$ is the average output taken over the cell neighborhood $N_r(i,j)$ and y_{ij}(\infty) is the settled output

$$PSNR = 10\log_{10}\frac{255^2}{MSE}$$

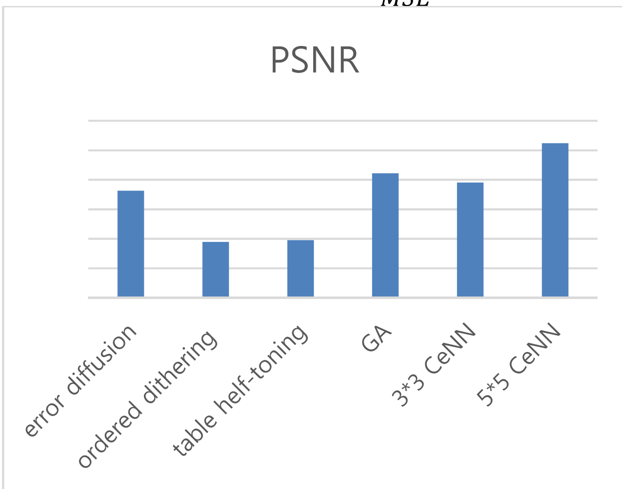

Figure 5: Chart of PSNR percent of halftone image by different methods compared with Bayesian rough set

According to the results, BRSM decreases the error with suitable robustness. It can also create an optimal CNNs template structure in learning the image half-toning. Moreover, the learning time is saved because of the dynamicity of CNNs. Unfortunately, the robustness of processing image half-toning depends totally on the learning rate that should be small enough. Thus, an evolutionary computing technique may increase the robustness of our process.

5. CONCLUSION

Since image half-toning processing is suffering from local representation and noise that corrupts the input image. This paper proposes a probabilistic machine learning algorithm to discover the global pattern concerned by an image and to process the optimal image half-toning template of CNNs. The problems of image half-toning, local representation of pixels neighbors and the corrupted region by gaussian noise, have been solved based on the CNNs dynamics and Bayesian rough sets. Bayesian Rough Sets discover the optimal and effective cells by generating a decision table and reduces its features without loss of its semantics. Moreover, the hidden pattern by noise is discovered. The optimal local rules are extracted and represented as an image half-toning process. New CNNs learning rules are generated, and the values of optimal template are determined to generate the robustness of a CNNs template. The experiment has been run on Lina image. The CNNs template has been constructed on 3×3 templates and 5×5 templates. The optimal template structure has been discovered. It has been proved that the quality of halftone image with 5×5 templates is higher than the corresponding one with 3×3 templates and other known methods.

Acknowledgements*:*
The Author would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work under the Grant code (20UQU0074DSR)

References

1. Frank, T., et al., *A Study on Halftoning Improvement for Low-Resolution Digital Print Engines With Machine Learning Methods.* IEEE Access, 2022. **10**: p. 19780-19795.
2. Zhang, F. and X. Zhang, *Image inverse halftoning and descreening: a review.* Multimedia Tools and Applications, 2019. **78**(15): p. 21021-21039.
3. Liu, T., *Three problems in imaging systems: texture re-rendering in online decoration design, a novel monochrome halftoning algorithm, and face set recognition with convolutional neural networks*. 2020, Purdue University Graduate School.
4. Liu, Z., et al., *A multi-level-optimization framework for fpga-based cellular neural network implementation.* ACM Journal on Emerging Technologies in Computing Systems (JETC), 2018. **14**(4): p. 1-17.
5. Qadir, F. and I.Q. Shoosha, *Cellular automata-based efficient method for the removal of high-density impulsive noise from digital images.* International Journal of Information Technology, 2018. **10**(4): p. 529-536.
6. Enescu, A., A. Andreica, and L. Diosan. *Evolved cellular automata for edge detection in grayscale images*. in *2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*. 2019. IEEE.
7. Zarándy, Á., A. Horváth, and P. Szolgay, *Cnn technology-tools and applications.* IEEE Circuits and Systems Magazine, 2018. **18**(2): p. 77-89.
8. Kung, J., D. Kim, and S. Mukhopadhyay, *Adaptive precision cellular nonlinear network.* IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2018. **26**(5): p. 841-854.
9. Hosokawa, Y. and Y. Nishio, *Applications and Oscillatory Phenomena of Cellular Neural Network Using Two Kinds of Cloning Templates.* Journal of Signal Processing, 2020. **24**(1): p. 1-9.
10. Priego, B., et al., *A cellular automata-based filtering approach to multi-temporal image denoising.* Expert Systems, 2018. **35**(2).
11. Jeelani, Z. and F. Qadir, *Cellular automata-based approach for salt-and-pepper noise filtration.* Journal of King Saud University - Computer and Information Sciences, 2022. **34**(2): p. 365-374.
12. Murphy, K.P., *Probabilistic machine learning: an introduction*. 2022: MIT press.
13. Xie, J., B.Q. Hu, and H. Jiang, *A novel method to attribute reduction based on weighted neighborhood probabilistic rough sets.* International Journal of Approximate Reasoning, 2022. **144**: p. 1-17.
14. Luo, J., et al. *Halftone Image Steganalysis by Reconstructing Grayscale Image*. 2021. Cham: Springer International Publishing.
15. Guo, J.-M. and S. Sankarasrinivasan. *Digital Halftone Classification using Simplified CNN and Stochastic Statistics*. in *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. 2021. IEEE.
16. Chua, L.O. and T. Roska, *Cellular Neural Networks and Visual Computing : Foundations and Applications*. 2002, Cambridge, UK: Cambridge University Press.
17. Zambrano-Luna, B. and W.A. Zuniga-Galindo, *p-adic Cellular Neural Networks.* arXiv preprint arXiv:2107.07980, 2021.
18. Elsayed, W.M., A.H. El-bassiouny, and E.F. Radwan, *Applying Inhomogeneous Probabilistic Cellular Au-tomata Rules on Epidemic Model.* International Journal of

Advanced Research in Artificial Intelligence (IJARAI), 2013. **2**(4): p. 39-47.

19. Sun, Y., et al., *Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification.* IEEE Trans Cybern, 2020. **50**(9): p. 3840-3854.
20. Pawlak, Z., *Rough sets, decision algorithms and Bayes' theorem.* European Journal of Operational Research, 2002. **136**(1): p. 181-189.
21. Rahman, A., H. Muhammah, and L. Sungyoung, *Rough set-based approaches for discretization: a compact reviews.* Artificial Intelligence Review, 2015. **44**(2): p. 235-263.
22. Yao, J., et al., *Formal concept analysis, rough sets, and three-way decisions*. 2022, Elsevier. p. 1-6.
23. Pięta, P. and T. Szmuc, *Applications of rough sets in big data analysis: An overview.* International Journal of Applied Mathematics and Computer Science, 2021. **31**(4): p. 659-683.
24. Radwan, E. and B. Alkazemi, *Extracting Knowledge from Incomplete Dataset by Tolerance Rough Sets with Consistency Measure.* International Journal of Soft Computing, 2018. **13**: p. 149-157.
25. Zhang, H., et al., *Bayesian rough set model: A further investigation.* International Journal of Approximate Reasoning, 2012. **53**(4): p. 541-557.
26. Ma, J., C. Zou, and X. Pan, *Structured probabilistic rough set approximations.* International Journal of Approximate Reasoning, 2017. **90**: p. 319-332.
27. Roska, L.O.C.a.T., *Cellular Neural Networks and Visual Computing*. 2004, Germany: Cambridge University Press.
28. Radwan, E., O. Nomir, and E. Tazaki, *A New Learning Method for Cellular Neural Networks Templates based on Hybrid of Rough Sets and Genetic Algorithms.* International Journal of Computer Science and Information Security, 2010. **8**(3): p. 155-164.
29. Zhang, T.-Y., *Bayesian Theorem In Rough Set Background And Its Application To Recovering Image*, in *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*. 2018, IEEE: Chengdu, China. p. 306-310.
30. Yang, C. and G. Li, *Incremental Attribute Reduction Under Variations of the Attribute Set Based on Conflict Region.* IEEE Access, 2021. **9**: p. 160195-160215.
31. Liu, D., T. Li, and J. Zhang, *Incremental updating approximations in probabilistic rough sets under the variation of attributes.* Knowledge-based systems, 2015. **73**: p. 81-96.
32. Elhady, A.M., et al., *Enhancing the Performance of CNNs using Evolutionary Programming.* International Journal of Computer Engineering and Information Technology, 2017. **9**(5): p. 87.