The DICOM CT Image Compression Based On Enhanced Lossless Prediction And Multilevel Thresholding Based Hybrid Cuckoo Search With Hill Climbing (CS-HC) Algorithm Based Segmentation

1Murugan K, 2Mothi R, 4Venkatesh T and 5Supriya M
1,4,5Assistant Professor, KPR Institute of Engineering and Technology, India
2Assistant Professor, Mahendra Engineering College, India
1murugan88.kumar@gmail.com

3Vasudevan D
SAP ABAP Consultant
Mitsubishi Electric Corporation, Japan.

Abstract - In computer vision applications, image segmentation is a common image processing step. It is used to separate pixels into different groups. The rise in the threshold count would hinder the segmentation phase of images. At the same time, in the field of threshold implementation in the image, it becomes an NT concern. This thesis suggests a multilevel threshold based on optimization techniques to remove ROI and uses enhanced lossless prediction algorithm to compress DICOM images in telemedicine applications. The hybrid Cuckoo search with hill climbing (CS-HC) algorithm strengthens the process used by the search agent to update the optimal solution. This algorithm calculates the threshold value. The superior results are produced by the proposed multilevel level thresholding based on CS-HC, as seen by the simulation results. Optimization is efficient and it has a high degree of convergence. Effective results are provided by the proposed lossless compression algorithm based on classification and blending estimation as compared with JPEG lossless and lossy compression techniques. With various threshold values, the algorithm's efficiency is checked. To apply this algorithm, Matlab2010a is used and DICOM photos are used to validate it.

Keywords - Image Compression, Segmentation, CT lung image, thresholding, hill climbing approach, Cuckoo search algorithm.

1. Introduction

In computer vision and image recognition, the central challenge is image segmentation. Via image segmentation with homogeneous properties, numerous sub-regions are created. It has a significant impact on bio-medical applications. An automatic segmentation technique can measure the changes in anatomical configurations where modifications lead to diseases in biological tissue. Minimum cross entropy based global thresholding techniques were introduced in for cuboidal cell nuclei in pictures. Photos are hematoxylin- and eosin-stained prostate tissue. An atlas-aided fuzzy c-means (FCM-Atlas) is introduced in for the segmentation of fibro glandular tissue. It estimates breast MRI volumetric density. The details on the multimodality picture are combined by following a sparse representation. Under structural limitations, brain tissue segmentation is integrated.

Image thresholding methods are most commonly used in image segmentation. It has two classifications. They are optimum thresholding and thresholding dependent on property. In the optimal thresholding strategy, the target function is optimised to look for an optimum threshold. Various properties are calculated from the histogram in order to find the threshold in the property-based threshold.

Based on parameters like cross entropy, entropy and class variance, the gray-level regions are segregated. They are used in optimal thresholding methods to measure the threshold value. In Otsu's unsupervised method of selecting a threshold, for example, the class variance is maximised.

Optimum thresholding strategies are both efficient and straightforward in bi-level thresholding. But it entails an unreliable, unsupervised Otsu process. In the multilevel thresholding query, Otsu's unsupervised method increases the computation time. To solve the problems of multilevel thresholding, algorithms like ant colony optimization, Darwinian particle swarm optimization (DPSO) bacterial foraging algorithm, fractional-order Darwinian particle swarm optimization (FODPSO) and particle swarm optimization are used.

Due to its efficacy, the PSO based algorithms and their extensions are most widely used in image segmentation. Darwinian particle swarm optimization DPSO and FODPSO are included in expansion. The DPSO and FODPSO algorithms proposed in solve Otsu problem. The n-level thresholding issue is to reduce an optimization problem that searches for thresholds that

@ IJAICT India Publications 2020
M.G. Sumithra et al.(eds.). Advances in Computing, Communication, Automation and Biomedical Technology, https://doi.org/10.46532/978-81-950008-1-4_058
maximise RGB (red-green-) objective (fitness) functions.

Premature convergence to sub-optimal solutions is a typical problem in the Cuckoo Search (CS) algorithm inclusive optimization algorithm. The key explanation for this is the inability of CS optimization operators to retain the diversity of multi-generation solutions. By integrating CS with other strategies used for searching, this can be minimised by making hybrid algorithms.

In contrast with the original CS algorithm, the hybrid CS algorithms need more computational time. The hill climbing algorithm will achieve the best solution in a short time by combining it with different local search algorithms. Hence, a hybrid CS-HC solution is suggested in this article. The hybrid Cuckoo search with hill climbing (CS-HC) algorithm strengthens the process used by the search agent to update the optimal solution. This algorithm calculates the threshold value.

2. Proposed Methodology

The CT lung scan image is the dataset used by this suggested algorithm. By pre-processing, the noise in these images is removed [1]. Segmentation methods of multilevel thresholding are used to segment the pre-processed image. By producing several threshold levels, regions of interest and pixel values are extracted from the lung image as shown in Figure 1.

The value of the threshold is optimised. By using a hybrid swarm intelligence method, the best outcomes are achieved. It is a variation of the algorithm for cuckoo quest and hill climbing [2]. To compact DICOM images, an improved lossless prediction algorithm is suggested. It is used in applications for data transfer and storage [3].

Pre-processing

Pre-processing is a crucial step in picture handling. Noises are removed in pre-processing; contrast can be improved and weak boundaries and unrelated parts are decided. The medical image portion is impaired by this. Pre-processing techniques are used to analyse these problems. A complete filter is a Gaussian filter used in the time domain. It is a filter with a low pass with less group delay [5]. It is used in both low and high signals to reduce distortion.

Multilevel Thresholding

A similarity-based methodology is the threshold, and it is a fundamental segmentation algorithm. The threshold value defines how it functions [6]. The first class corresponds to pixels with a grey value higher than threshold value, while the second class corresponds to pixels with a grey value lower than the threshold value. Bi-level thresholding refers to the region of interest extracted by a single threshold value from an image.

Consider an image with K gray levels, bi-level thresholding is expressed as in Equation (1) and Equation (2),

\[ R_o = \{ i(x, y) \in I \mid 0 \leq i(x, y) \leq th - 1 \} \]

\[ R_i = \{ i(x, y) \in I \mid th \leq i(x, y) \leq th - 1 \} \]

Multiple thresholds are used in multilevel thresholding and image with multiple regions is created as in Equation (3), Equation (4), Equation (5) and Equation (6),

\[ R_o = \{ i(x, y) \in I \mid 0 \leq i(x, y) \leq th_1 - 1 \} \]

\[ R_i = \{ i(x, y) \in I \mid th_1 \leq i(x, y) \leq th_2 - 1 \} \]

\[ R_t = \{ i(x, y) \in I \mid th_i \leq i(x, y) \leq th_{i+1} - 1 \} \]

\[ R_m = \{ i(x, y) \in I \mid th_m \leq i(x, y) \leq k - 1 \} \]

Where, various thresholds are represented by \( th_1, th_2, th_3, \ldots, th_k \).

To compute the threshold values, parametric or non-parametric techniques are used. The probability density function is estimated by a parametric method for every class [7]. So, it is both time consuming and a complicated process. The non-parametric methods use variables such as error rate, entropy and class variance.

The classical methods used in bi-level thresholding to estimate a threshold value are Otsu and Kapur’s technique. The class variance is enhanced by the algorithm of Otsu, where Kapur’s algorithm maximises entropy. With the increase in computational complexity, multilevel thresholding techniques can use these techniques and are noise-sensitive and the threshold value needs to be defined by the user [8].

The results of segmentation techniques based on automatic and iterative thresholding are not up to the expected level. By using an optimization algorithm, the optimum threshold value is selected [9]. A swarm optimization technique is used for multilevel

**Fig. 1: Proposed Methodology Diagram**

*Input Dataset*

The DICOM CT images of abdomen in real-time are used to evaluate the algorithm. The photos were obtained from the Metro Scans and Research Laboratories, Thiruvananthapuram. The images are captured using the 0.6 mm slice thickness of an Optima CT machine. For segmentation, 7 CT abdomen data sets are used for this work and 6 CT abdomen data sets are used for compression [4].
thesholding techniques. It generates high value for performance metrics such as speed, robustness and accuracy compared to traditional methods. In this work, using the Cuckoo search with hill climbing (CS-HC) algorithm, the threshold of a multilevel thresholding technique is estimated [10].

Cuckoo search with hill climbing (CS-HC) algorithm for estimation of the threshold for multi thresholding technique

In this work, proposed cuckoo search algorithm is used for selecting CH nodes with hill climbing (CS-HC). It has an objection function to compute a threshold for a multiple threshold technique. The aim to select the optimal threshold value.

Cuckoo Search Algorithm

CS algorithm is a Meta-heuristic search procedure. It utilizes subsequent idealized rules:
1. One egg is laid by each cuckoo at time, dumps it in haphazardly selected nest.
2. In next generation, execute a best nest having high quality of eggs.
3. Fix number of obtainable host nests. With a probability, an alien egg can be found by a host.

An egg can be thrown away or nest are dumped or new nest can be constructed in a new location by a host bird. The basic steps in cuckoo search algorithm are edited. Levy flight is computed to produce a novel solution for i \textsuperscript{th} cuckoo as shown in Equation (7).

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \bigoplus \text{Levy}(\lambda)$$

(7)

where, step size is represented by $\alpha$. This value depends on the problem. Entry wise multiplication is performed to compute a product. This procedure measures a Levy flight. Depends on the subsequent probability distribution, its step length is distributed as shown in Equation (8).

$$\text{Levy } u = t^{-\lambda}, 1 < \lambda \leq 3$$

(8)

It has infinite variance. A random walk procedure is formed by successive steps or jumps by cuckoo. This corresponds to a power-law step length distribution with important tail.

Algorithm: Cuckoo Search

Determine Objective function $f(x)$, $x = (x_1, x_2, \ldots, x_d)$
Initialize population of n host nests $x_i$ ($i=1,2,\ldots,d$)
While ($t < \text{MaxGeneration}$ or (stop criterion));
Attain cuckoo (say i) randomly and generate a novel solution by levy flights
Estimate its quality/ fitness $F_i$
Select nest between n (say j) randomly and produce novel solution by levy flights
End
Throw away fraction ($P_a$) of worse nests and construct novel ones at novel locations through levy flights
Maintain finest solutions or nests with quality solutions
Position solutions and acquire current best
End while

Post procedure outcomes and visualization

End

Hill Climbing (HC)

A technique for mathematical optimization is Hill Climbing (HC). This falls within the local search family. It calculates the current state in order to get the neighbourhood's best solution. If it reaches the goal state, it terminates the process. Otherwise, the current state is continuously updated. The updating is done by applying the current state to new operators. There is a two-step system. In the first step, a new operator that is not in the current state is selected and applied to the current state. In the second step, new states are computed.

It is recommended that the above statements move towards a better solution than the current solution. It is the fundamental idea of HC. The quality of the solution is enhanced by this.

1: $i = \text{initial solution}$
2: While $f(s) \leq f(i) \forall s \in \text{Neighbours}(i)$ do
3: Generates an s $\in \text{Neighbours}(i)$;
4: If fitness ($s$) > fitness ($i$) then
5: Replace s with the i;
6: End If

Based on the problem, HC can be adjusted. This is a major advantage of HC. We can change and customize any aspect of this algorithm. It is used in discrete and conversion domains.

Cuckoo Search Algorithm with Hill Climbing Algorithm

The HC and CS techniques are combined to propose a CS-HC algorithm. The cuckoo species obligate brood parasitic behaviour is used by CSA. CSA combines Levy flight with behaviour of cuckoo species. Levy flight is random with power law step length distribution and it has heavy tail. Some fruit flies and bird’s behaviour are used in this. In order get efficient results in global exploration, Levy flight can be used.

The CSA is an efficient metaheuristic swarm-based algorithm. In search space, it maintains a balance between global wide exploration and local nearby exploitation efficiently. There is a chance for slow convergence with worst results. In order to avoid this, HC method for deep exploitation is combined with basic CSA to enhance the ability in search. It is termed as CS-HC algorithm. The benchmark functions are optimized using this.

Pseudo Code: Hybrid CS_HC

Start
Define Objective function $f(x)$, $x = (x_1, x_2, \ldots, x_d)$
Initial a population of n host nests $x_i$ ($i=1,2,\ldots,d$)
Define the cuckoo search parameters $P_a$
Begin CS
Evaluate its quality/ fitness $F_i$
Get a cuckoo (say i) randomly and generate a new solution by levy flights
Choose a nest among n (say j) randomly
If ($F_i > F_j$)
Replace j by the new solution
End
Then
Calculate the neighbouring nest
Find the maximum of neighbouring nest
If (Larger than the current one)
Local minimum is found
If not again calculate the neighbouring nest
If the local minimum is larger than the current local
minimum
Abandon a fraction \((P_a)\) of worse nests and build new
ones at new locations via levy flights
Keep the best solutions or nests with quality solutions
Rank the solutions and find the current best based on the
max iteration
End while
Final best population of nests
End

The number iterations are searched by using
standard cuckoo search algorithm in CS-HC. Search is
accelerated by passing the obtained best solution to the
HC. This rectifies the issue of slow convergence in
standard cuckoo search algorithm. HC is an iterative
type of algorithm. For a problem, it assigns an arbitrary
solution. By changing a single solution element
incrementally, it attempts to find the best solution. This
incremental changing process continuous until the
change produces better results. After algorithm’s
convergence, CSA is given with solution to check it by
using fraction probability \(P_a\).

Classification and Blending Prediction Based Lossless
Compression Algorithm

[11] proposed blending predictor. Based on this
predictor, this research work proposed a lossless
compression algorithm. The image with sharp
horizontal edges can be predicted well by the classical
predictor. In this work, perform the static predictor on
the estimated set of neighbouring pixels. The vector
quantization is also having a similar initial stage. The
following discussion presents the various stages of
lossless compression technique [12],[13].

Classification
The \(P^i\) represented the pixel that has to be predicted and
neighbouring pixel set is represented by \(\emptyset_C\).
Computation of M pixels with minimum distance vector
from pixel which has to be predicted in casual context
\(\emptyset_C\) is the major aim of classification.
The following expression shows the distance to the
pixel to be predicted as shown in Equation (9) and as
shown in Equation (10).
\[
D(i,j) = ||u(i,j) - u(x,y)|| \quad (9)
\]
\[
D(i,j) = \sum_{k=1}^{M} |u_k(i,j) - u_k(x,y)|^2 \quad (10)
\]
The blending context \(\emptyset_B\) is represented by minimum
distance vectors of set of M pixels. The cells are formed
by grouping pixels based on the Euclidean distance.
This process is similar to the vector quantizer.

Blending

The prediction is generated by a blending context \(\emptyset_B\)
which blends set of static predictors \(f = \{f_1, f_2, f_3, \ldots, f_n\}\). A penalty factor \(G_K\) is coupled with
all predictors in set \(f\). Based on the pixel prediction by
context \(\emptyset_B\), penalty factor defined as shown in Equation
(11).
The mean square error represents the penalty as
\[
P_k = \sum_{(i,j)\in\emptyset_B} |f_k(i,j) - f(i,j)|^2 \quad (11)
\]

Prediction and Error Correction
The weight of a predictor in a set \(f\) corresponds to the
inverse of penalty \(P_k\). Static prediction is computed by
weighted average of it and pixel prediction is done
based on this.

3. Result and Discussion
DICOM medical images are compressed efficiently by
the classification and blending prediction algorithm.
MATLAB 2013 a is used to implement this algorithm
and desktop computer is used to process it. The
specifications of the computer include 4 GB RAMS 64-
bit operating system with Intel Core i3 processor.
DICOM medical images are compressed efficiently by the classification and blending prediction algorithm. It is a lossless method. In initial stage of the CS-HC algorithm, value ‘0’ is assigned to the bird’s nest. The search space boundaries are used to set the values of neighbouring nest. The local minimum, local and neighbourhood values are assigned with least value based on the problem nature. The search is accelerated by passing the obtained best solution to the HC. The solutions single element is varied incrementally to get a best solution.

On new solution, perform incremental changes, if good results are produced because of this change. This process is repeated until the algorithm converges. The CSA solution is return after the convergence. Threshold and fraction probability $P_a$ are used to check the returned results.

The input image with its histograms is shown in Figure 2. For different threshold values, FODPSO outputs are shown in Figure 3.

The CS-HC outputs with different value of thresholding (2, 3, 4, 5) which are parallel to given CT images.

The experimental results the proved that the proposed algorithm for lossless compression is more efficient than the JPEG lossless and lossy compression algorithms. The JPEG lossy algorithm uses a Discrete Cosine Transform and compression is done using Huffman coder whereas in lossless algorithm run length encoder is used for compression and it used Adaptive predictor with Golomb. The experimental results are shown in Table 2.

The input image with its histograms is shown in Figure 2. For different threshold values, FODPSO outputs are shown in Figure 3.

The CS-HC outputs with different value of thresholding (2, 3, 4, 5) which are parallel to given CT images.

![Fig. 4: CS-HC results for various levels of thresholding (2, 3, 4, 5) parallel to input CT images (a, c, e, g).](image)

More details are generated by the images having high value of threshold as shown by the resultant images Figure 5 and Figure 6. Population based algorithms are used for optimization and experimental results are random as well as stochastic. The average value computed by executing each technique by 15times because of the stochastic nature. Table 1 shows the average standard objective values. The 7 data sets are represented using ID1 to ID 7. Threshold values used here are 2,3,4 and 5.

**Table 1: Pathological information of datasets for segmentation.**

| Dataset ID | Dimension | Pathological Information          |
|------------|-----------|-----------------------------------|
| 1          | 512*512   | Benign Cyst                       |
| 2          | 512*512   | Multiple Benign Cyst              |
| 3          | 512*512   | Benign tumor hemangioma           |
| 4          | 512*512   | Hepatocellular Carcinoma          |
| 5          | 512*512   | Liver Cirrhosis, Normal           |
| 6          | 512*512   | Malignant lesion with liver Cirrhosis |
| 7          | 512*512   | Liver Metastasis                  |

![Fig. 5: The first column represents the input images (ID1, ID2, and ID3), Second column represents the error image, the Third column represents the reconstructed image.](image)

![Fig. 6: The first column represents the input images (ID4, ID5, and ID6), Second column represents the error image, the Third column represents the reconstructed image.](image)
The current pixel is predicted using the predictions made by the static predictor. The predictions of it are assigned with a weight and sum of the predictions with its weights corresponds to a current pixel. Penalty factor is inversely proportional to it. The blending contexts efficiency is based on the penalty factor of predictor. The current prediction can be generated precisely, if the predictions made by predictors are correct. The predictors with high penalty factor correspond to inefficient predictor. They are blended out of the current blending context.

Average error $\bar{e}$ is computed based on the blending context $\emptyset_B$, 

$$\bar{e}(\emptyset_B) = \sum_{i,j\in \emptyset_B} \frac{[S(x,y)-S(i,j)]}{M}$$  \hspace{1cm} (13)

The final predictor of current pixel is defined based on the blending predictor error as expressed below, 

$$S(x,y) = S(x,y) + \bar{e}(\emptyset_B)$$  \hspace{1cm} (14)

The pixels local property is used to adjust the classification as well as the blending process. The neighbouring pixel’s casual set is used to predict current pixel. In local image context, higher order redundancy is removed by classification process.

### 4. Conclusion

In this work, medical images are threshold in multilevel using a hill climbing algorithm which is based on cuckoo search. Also proposed a lossless compression algorithm for medical image classification and prediction of blending. The experimental results are better when compared to JPEG lossless and lossy algorithms. The ROI analysis of telemedicine images can be done using this proposed algorithm. The compression of images can also be done effectively for data transfer. In future, segmentation and compression algorithms can be implemented in hardware to make a portable system analysis and transfer of data.

### References

[1]. de Oliveira, D.L.L., do Nascimento M.Z., Neves, I.A., de Godoy, M.F., de Arruda, P.F.F., and de Santi Neto, D. Unsupervised segmentation method for cuboidal cell nuclei in histological prostate images based on minimum cross entropy. Expert Systems with Applications 40, 7331–7340, 2013.

[2]. Wu, S., Weinstein, S.P., Conant, E.F., and Kontos D. Automated histological tissue segmentation and volumetric density estimation in breast MRI using an atlased fuzzy C-means method. Med Phys. 40; 2013 DOI: 10.1118/1.4829496.

[3]. Wang, L., Shi, F., Gao, Y., Li, G., Gilmore, J.H., Lin, W., and Shen, D. Integration of sparse multi-modality representation and anatomical constraint for isointense infant brain MR image segmentation. NeuroImage 89 152–164, 2014.

[4]. H. Anandakumar and K. Umamaheswari, “Supervised machine learning techniques in cognitive radio networks during cooperative spectrum handovers,” Cluster Computing, vol. 20, no. 2, pp. 1505–1515, Mar. 2017.

[5]. H. Anandakumar and K. Umamaheswari, “A bio-inspired swarm intelligence technique for social aware cognitive radio handovers,” Computers & Electrical Engineering, vol. 71, pp. 925–937, Oct. 2018. doi:10.1016/j.compeleceng.2017.09.016

[6]. Otsu, N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, Cybernetics, SMC-9 62–66,1979.

[7]. Sathy, P.D., and Kayalvizhi, R. Optimal multilevel thresholding using bacterial foraging algorithm. Expert Systems with Applications 38,15549–15564,2011.

[8]. Maitra, M., and Chatterjee, A. A hybrid cooperative comprehensive learning-based algorithm for image segmentation using multilevel thresholding. Expert Systems with Applications 34,1341–1350,2008.

[9]. Zhiwei, Y., Zhaobao, Z., Xin, Y., and Xiaoang, N. Automatic threshold selection based on ant colony optimization algorithm. In Proceedings of the International Conference on Neural Networks and Brain, Beijing, China, Number 728–732, 2006.

[10]. Ghamsi, P., Couceiro, M.S., Benediktsson, and Ferreira, N.M.F. An efficient method for segmentation of images based on fractional calculus and natural selection. Expert Systems with Applications 39,12407–12417,2012.

[11]. Venkatesh, T., Shannugaraja, T., Maheswari, A., Dhivyadevi, R., Supriya, M., “A perspective approaches of clustering in wireless sensor networks-survey,” International Journal of Advanced Science and Technology, 29(7 Special Issue), pp. 1457-146,2020

[12]. T. Seemann, P. Tischer, and B. Meyer, ”History-based blending of image sub-predictors,” ITG Fachbericht, pp. 147_152, 1997.

[13]. Supriya, M., Dhivyadevi, R., Shannugaraja, T., Venkatesh, T., Jai Shankar, B., “Multi-ported memory on fpga for a high performance firt filters “International Journal of Advanced Science and Technology, 29(7 Special Issue), pp. 1464-1472,2020