Int. J. Advance Soft Compu. Appl, Vol. 11, No. 2, July 2019
ISSN 2074-8523

Enhanced Massive Training Artificial Immune Recognition System for False Positives Reduction in Lung Nodules Classification

Hang See Pheng[1], Siti Mariyam Shamsuddin[2] and Ong Kok Haur[2]

[1] Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia, UTM Skudai, 81310 Johor, Malaysia.
e-mail: sphang@utm.my
[2]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, UTM Skudai, 81310 Johor, Malaysia.
e-mail: mariyam@utm.my , ongkokhaur@gmail.com

Abstract

Massive Training Artificial Immune Recognition System (MTAIRS) had been implemented in the computerized system to classify lung nodules on Computed Tomography (CT) scans. In this algorithm, large training sub-regions are trained, and the classification algorithm shows promising results in the lung nodules classification. However, in the output images of non-nodule cases, some false positives are still identified in the MTAIRS. False positives are always considered as a common issue in most of the development of classification algorithms of lung nodules detection. The effort of reducing false positives in the output images from MTAIRS is presented where the enhancement is based on the affinity function in MTAIRS algorithms. The quantitative assessment on the classification results for detection of lung nodules will be presented in this research.

Keywords: *Artificial immune system, lung nodules classification, false positive reduction, affinity function.*

1. Introduction

The machine learning has been applied in several classification models for medical images namely CT scans (Erickson et al, 2017). One of the most prevalent implementation of machine learning to classify the lesions and nodules

on medical images for example, malignant or benign by computerized models. In the classification process, the informative and representative features that well describing the patterns inherent in data are important role to ensure the performance of the computational models (Shen et al, 2017). In the pixel-based machine learning for medical image processing which uses pixel values in the image directly are used directly as input information Suzuki (2017). The advantages of using pixel machine learning in classification models is to avoid the loss of information in the medical image data processing. In our previous research finding, the pixel-based learning classification algorithm Massive Training Artificial Immune Recognition System (MTAIRS) was concluded to be able to classify all lung nodules on testing images in the experiments (Hang et al., 2013 and Hang et al., 2015). However, the false positives, which are known as a limitation of MTAIRS had been found from the classification results, especially for the non-nodules cases. False positives are always considered as a common issue in most of the development of classification algorithms of lung nodules detection. These false positives may confuse the users when analyzing the output of automated system in lung nodules detection. The false positives in lung nodules detection normally exist because blood vessels have similar contrast as lung nodules on CT images.

In this research, re-designing the component of MTAIRS algorithm is aimed to improve the performance in the study case. The fundamentals of MTAIRS is developed from artificial immune system (AIS) where the similar mechanisms have been employed. Therefore, AIS algorithm is initially reviewed in the early stage of algorithm enhancement. There are three basic components in the common design of AIS, which are representations of algorithm, affinity measure and immune process. The modification of the representations of algorithm will not give significant impact on the performance of algorithm since the technical terms or artificial elements are designed to explain the components in immune process. experiments have been conducted to test the classification results by applying two different types of AIRS mechanisms in memory cell generation that mimic diverse immune processes. The classification results in the application domain are found to be almost the same in the qualitative evaluation, and their accuracies are the same in quantitative analysis. Based on the design of the AIS algorithm, the affinity function is found to be potentially influencing the accuracy of classification results. Basically, the choices of affinity function depend on the types of training data to ensure promising classification results. Therefore, the problem formulation is formed to seek for the appropriate affinity function that can measure the proximity of two training sub-regions in the training process.

In the MTAIRS algorithms, the affinity measure is inversely proportional to the distance measure. The affinity determines the stimulation of clones or mutates in the massive training process and affects the generation of training instances in the memory cell generation. The Euclidean distance is originally employed in the affinity measures of MTAIRS. Therefore, appropriateness of Euclidean distance

function in this application domain is investigated based on the nature of the training data. In this paper, the modification of the affinity function is examined to reduction the false positives on the output images, especially for non-nodule cases.

2 Review of Affinity Function

There were some research had conducted to investigate the appropriateness of Euclidean distance in AIRS algorithm (Freitas and Timmis, 2007, Hamaker and Boggess, 2004, Seeker and Freitas, 2007). The Euclidean distance was mostly general used as affinity measures in AIRS. The distance function was first implemented by Perelson and Oster in their early work that proposed the concept of shape space for quantifying the chemical interaction between cells or molecules (Perelson and Oster, 1979). However, the Euclidean distance function was sometime not appropriate to be used in AIS algorithm due to especially in the different application domain (Casro and Timmis, 2002). From the experiments done by Hameker and Boggess (Hamaker and Boggess, 2004), there were several distance functions such as Euclidean, Manhattan, Overlap, Value Difference Metric were applied AIRS classification algorithm in order to investigated their feasibility in affinity measures. From these experiments, they found that the accuracy of algorithms was increased by using other distance function instead of Euclidean distance function. Therefore, non-Euclidean distance function was proposed by Hamaker and Bogges (Hamaker and Boggess, 2004) in the AIRS algorithm. Moreover, they also concluded that the choice of distance function was depending on the nature and features of data were real-value, discrete or nominal. Besides, Freitas and Timmis (Freitas and Timmis, 2007) suggested that affinity function in AIRS algorithm should be tailored according to the type of data used, instead of just using standard affinity function. They also recommended that applying affinity function that considering attribute weights was more suitable when performing the classification tasks.

In the scope of our study domain, the research was conducted based on the classification of nodules on medical imaging. Therefore, the affinity function that considering the weights of attributes and suitable for image similarity measure were reviewed. In recent, there were several research had been done on surveying distance function to measure the similarity in the field of pattern recognition (Cha, 2007, Duda et al., 2012, Deza and Deza, 2006, Zezula et al., 2006). The common point of view from these surveys was the proximity of two objects should be measured by the suitable similarity function such as non-parameter similarity function(Cha, 2007, Zezula et al., 2006). Wu et al. (Wu et al., 2005) had proposed a non-linear function for calculating the distance that employed the theory of non-Mahalanobis method. The concept of non-linear distance function was also implemented by Frome et al.(Frome et al., 2007) for imaging processing in pattern recognition. Frome et al. claimed that this method was also known as image-to-

image distance function which could provide quantitative measure for the degree of similarity between images. Consequently, Bhattacharya et al. (Bhattacharya et al., 2012) had developed another non-affinity function that implemented to for similarity measures in classification algorithm. This recently developed affinity function considered the impacts of the other training points for any particular feature. Bhattacharya et al. (Bhattacharya et al., 2012) had applied the developed affinity function in the classification algorithm. They conducted the experiments to test the performance of the non-linear function compared to other popular distance functions. From the results analysis, their method had revealed highest average accuracy by 10-fold cross validation for eight standard dataset. Therefore, this non-linear affinity function that takes into account the influence of attribute wise with good performance in classification will be further assessed in the research domain.

3 Methodology of Enhanced MTAIRS (E-MTAIRS)

In the study of pattern recognition, the measure of similarity is normally computed by the distance function. From the concept of mathematics, the distance the measures of how far are two items. However, in the measurement of the proximity between two objects, similarity measure is more appropriate to describe the closeness of two objects. The training process in the MTAIRS algorithms involves the measurement of affinity that employs the Euclidean distance function. Therefore, a non-linear similarity function, developed by Bhattacharya et al. [1] is implemented in the training mechanism in proposed MTAIRS to obtain the measurement of similarity of two training sub-regions instead of distance measure by Euclidean distance. It is found that this similarity function shows a better performance in classification compared to other distance function namely, Euclidean distance and squared Euclidean. Generally, this similarity function is used to obtain the measure of closeness between both, testing and training points in the classification algorithm by considering the impact of vicinity training points. The proposed similarity function is formed by the multiplication of weight function with linear distance function (Equation 1).

$$d(v_t, v_i) = \sum_{j=1}^{d} w_j^i \delta_j^{ti} \tag{1}$$

where w_j^i is the weight function for the *i*th train point along *j*th feature and δ_j^{ti} is the measure of distance of test point, v_t and train point, v_i along *j*th feature. Equation 2 reveals the comprehensive formulation for the similarity function corresponding to Equation 1. The first term of the weight function represents the influence of all training instances to the respective test instance with respect to the effect of all training instances to both, the respective test and training instances. The next term corresponds to the affinity of entire training points to the test

instance regarding the affinity of all training instances to the respective training instance.

$$d(v_t, v_i) = \left(\left(\frac{\sum_{l=1}^{N} \left| v_{tj} - v_{lj} \right|}{\sum_{l=1}^{N} \left| v_{tj} - v_{lj} \right| + \sum_{m=1, m \neq i}^{N} \left| v_{ij} - v_{mj} \right|} \right)^{1/2} + \left(\frac{\sum_{l=1}^{N} \left| v_{tj} - v_{lj} \right|}{\sum_{m=1, m \neq i}^{N} \left| v_{ij} - v_{mj} \right|} \right)^{1/2} \right) \times \sum_{j=1}^{d} \left| v_{tj} - v_{ij} \right| \quad (2)$$

The non-linear function in affinity measure is employed in Enhanced-MTAIRS (E-MTAIRS) to replace the conventional Euclidean distance function. There are two main reasons for choosing the non-linear distance function in the classification algorithm. Firstly, the identification of the non-linear affinity function is found to provide better classification results compared to Euclidean distance. Secondly, the enhanced algorithms are applied in the training of image data involving the measure of similarity between images using massive sub-regions of two images. Thus, the non-linear affinity function which considers the influences of other sub-regions in the training processes is chosen. Fig. 1 shows the substitution of the identified non-linear affinity function in the overall training process of E-MTAIRS. The memory instances are randomly generated in the initialization phase from the input of training file. The memory instances are then ready to seek for the most similar training instances from training data by using the non-linear affinity formulation for new memory cell identification and ARB generation.

Fig. 2 presents the relative distance of memory instances with concerned training instance by applying the concept of non-linear distance function that takes into account the vicinity impact of other training instances in training file in the routine process of memory cell identification and ARB generation. Assuming that there are *N* training instances in the training file, the affinity measures are calculated for *N* times to measure the similarity between initial memory instances and each of the training instances, respectively. Based on the non-linear affinity function, the distance, *d* between memory instances (*Mc*) and concerned training instances point along attributes is calculated. Further, the distances between memory point and training instances are represented by d1, d1, ..., dN, where the total of these distance measures is the numerator of 1st term in weight function in Equation 2. Meanwhile, z1, z2, ..., zN are the distances between concerned training point and each training instance, where the total of these distance values is denominator of second term in weight function in Equation 2. Furthermore, the obtained affinity values will then determine the stimulation values in the continuing training process such as cloning and mutation. After the training process of memory cell identification and ARB generation, as well as competition for limited resources and memory cell introduction, the classification of sub-

regions is performed by E-MTAIRS based on the compilation of the best memory cells for each of the training classes.

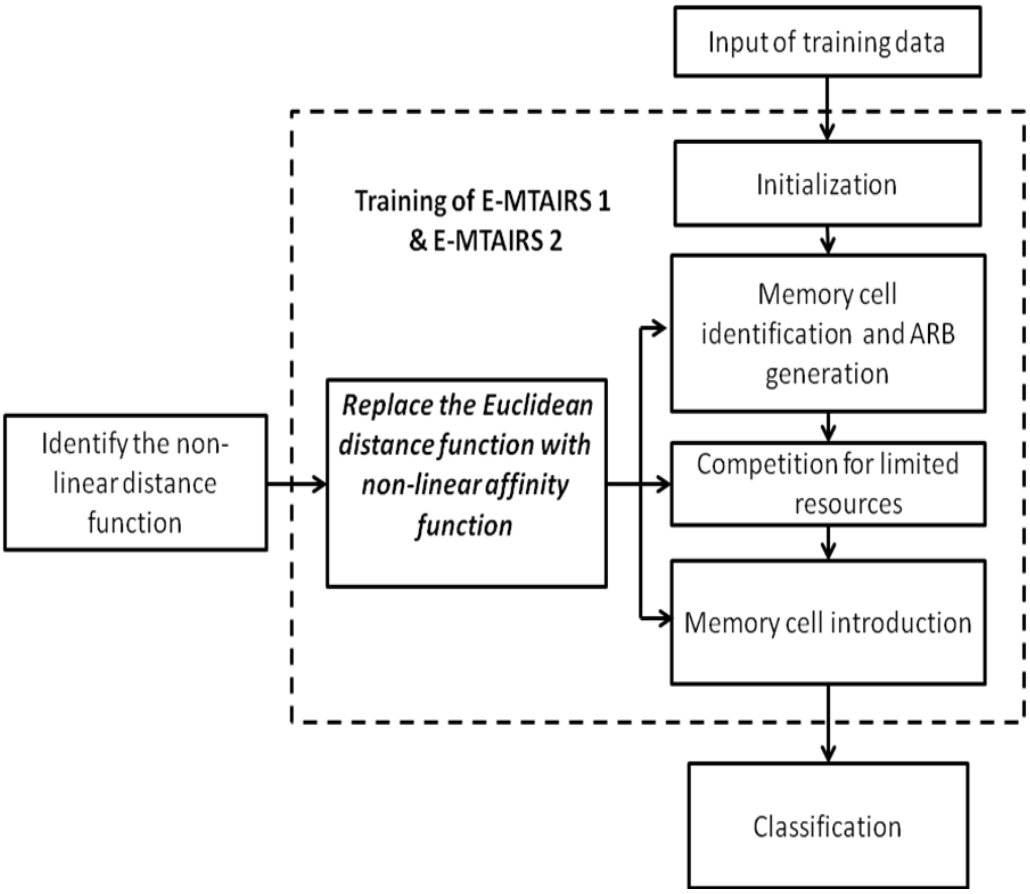


Fig. 1: The implementation of non-linear affinity function in E-MTAIRS

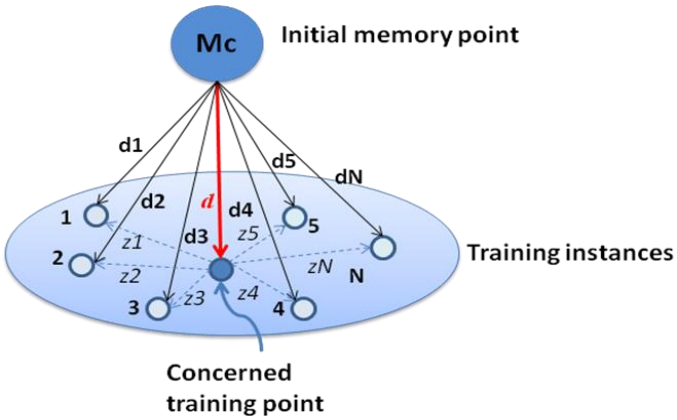


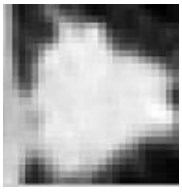
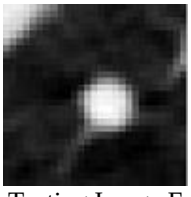
Fig. 2: The distance relation among both, testing and training points in non-linear distance function.

4 Experimental Results of E-MTAIRS

The experimental results of E-MTAIRS are presented by qualitative analysis to assess the quality of output images and quantitative analysis to calculate the percentage of difference with respect to the false positives reduction. The validation is done to assess the impact of affinity function in E-MTAIRS algorithm. Qualitative analysis is performed to examine the output images produced from the E-MTAIRS classification compared to MTAIRS. In addition, the pixel-based quantitative analysis is applied to calculate the accuracy of outcome from E-MTAIRS for both, nodule and non-nodule cases.

4.1 Qualitative Analysis for Nodule Cases

In the nodules visualization for testing image (Testing Images A-J), the output images of MTAIRS and E-MTAIRS are presented in Fig. 3. The visualization results show that MTAIRS and E-MTAIRS are able to identify all the nodule cases in testing images. Based on the image comparison, it is found that the output images produced by E-MTAIRS reveal more appropriate amount of bright pixels to represent lung nodules area compared to MTAIRS. This can be obviously seen from the output images that consist of larger size nodules (Testing Images A, C, I and J) with effective diameter of 9 mm to 18 mm. The gloomy noise in the output image produced by E-MTAIRS can be significantly reduced, especially for the latter. In this context, gloomy noise is defined as the unwanted grey pixel that exists on the nodules area of output results. Besides, for the overall visualization, the size of the area highlighted by E-MTAIRS are more fitted with the size of original testing images, compared to the output images produced by MTAIRS. This can be observed from the smaller size of lung nodules in Testing Images B, D, E, F, G and H ranging between 5 mm to 9 mm. The bright noise is significantly decreased in the classification results of E-MTAIRS, mainly at the surrounding area of lung nodules for those testing images. From the overall qualitative analysis, it can be concluded that the visualization of lung nodules can be better enhanced by E-MTAIRS in term of contrast and size of output images.

Original Images	**MTAIRS**	**E-MTAIRS**	**Original Images**	**MTAIRS**	**E-MTAIRS**
 Testing Image A			 Testing Image F		

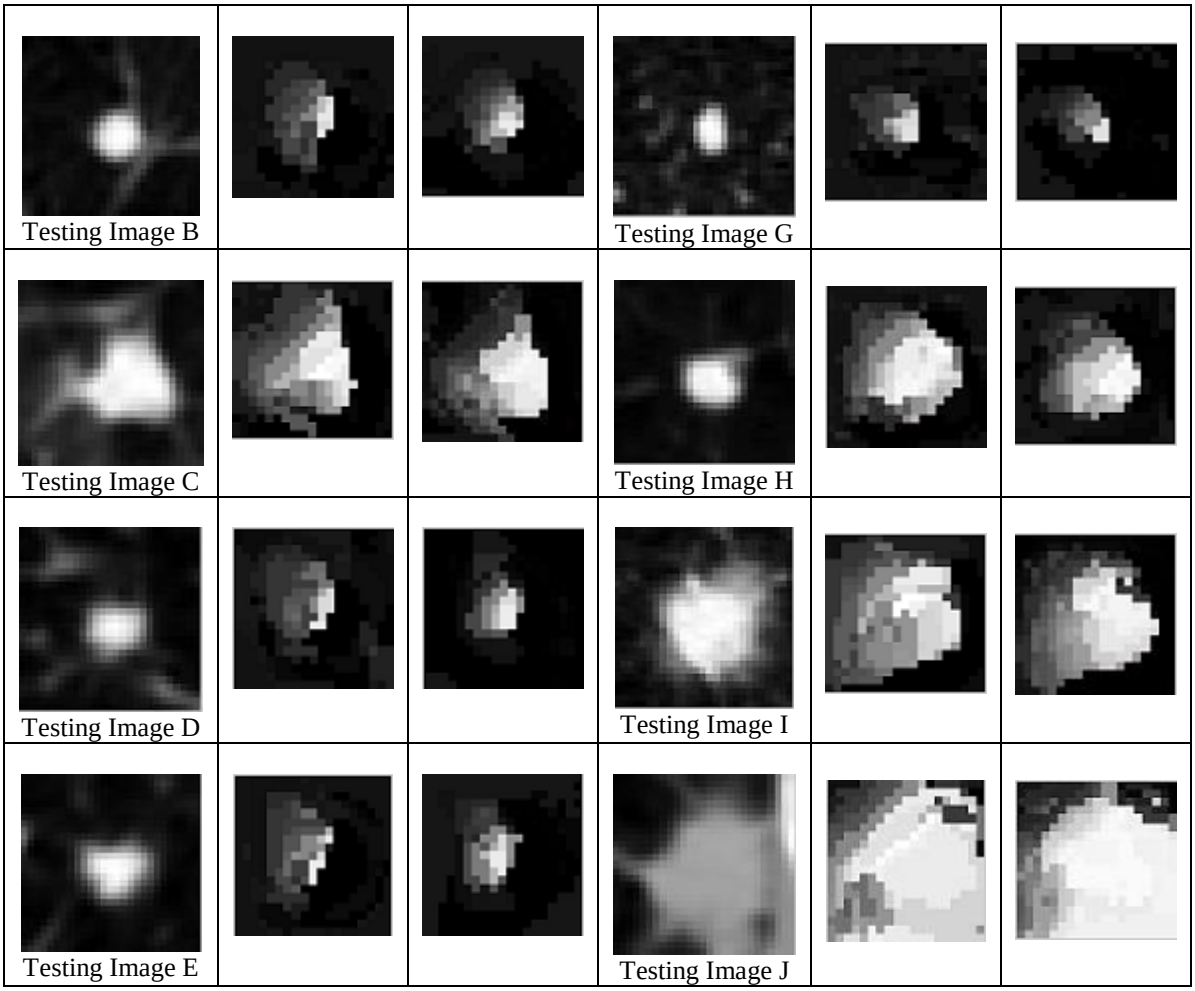


Fig. 3: Output images of nodule cases produced by MTAIRS and E-MTAIRS compared to original testing images

4.2 Qualitative Analysis for Non-nodule Cases

In the visualization of non-nodule cases, the output of MTAIRS and E-MTAIRS are presented in Fig. 4. The testing images for non-nodule cases are labeled as testing Images AN, BN, CN, DN, EN, FN, GN, HN, IN, JN. Based on the visualization results, E-MTAIRS is able to eliminate most of the blood vessels presented by bright pixels in the original testing images. Besides, the qualitative analysis also shows that less bright noise is identified in the non-nodule cases for the overall results. However, to further investigate the performance of E-MTAIRS, the quantitative analysis will be further discussed to provide objective judgments in results validation.

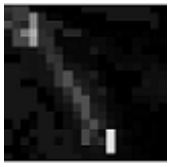
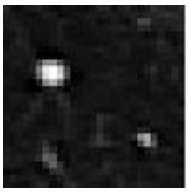
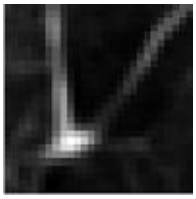
Original Images	MTAIRS	E-MTAIRS	Original Images	MTAIRS	E-MTAIRS
 Testing Image AN			 Testing Image FN		
 Testing Image BN			 Testing Image GN		
 Testing Image CN			 Testing Image HN		
 Testing Image DN			 Testing Image IN		
 Testing Image EN			 Testing Image JN		

Fig. 4: Output images of non-nodule cases produced by MTAIRS and E-MTAIRS 1 compared to original testing images

4.3 Quantitative Analysis for Non-Nodule Cases

Results validation based on quantitative evaluation is performed for non-nodule cases to investigate the classification results of E-MTAIRS compared to MTAIRS. The purpose of performing the quantitative analysis is to obtain the objective evaluation on the different results between classification algorithms. The evaluation is carried out by comparing the absolute difference between original testing images and output results to obtain the accuracy of pixel-based classification. Furthermore, the results of output images will be compared directly with the teaching images. The teaching image is formed by "dark" image which consists of all zeros as intensity. As a result, the pixel values in non-nodule output images should be approximately zero since the blood vessels and other high contrast healthy tissues should have been eliminated after the classification of pixels. In the pixel-based evaluation, based on the absolute difference between both, original testing and output images from classification will be performed to assess the accuracy of non-nodules classification by MTAIRS and E-MTAIRS. In the qualitative analysis, the accuracies of classification for both models are obtained by pixel-based evaluation, based on absolute different of two images. Fig. 5 shows the pixel based evaluation procedure, based on non-nodules classification.

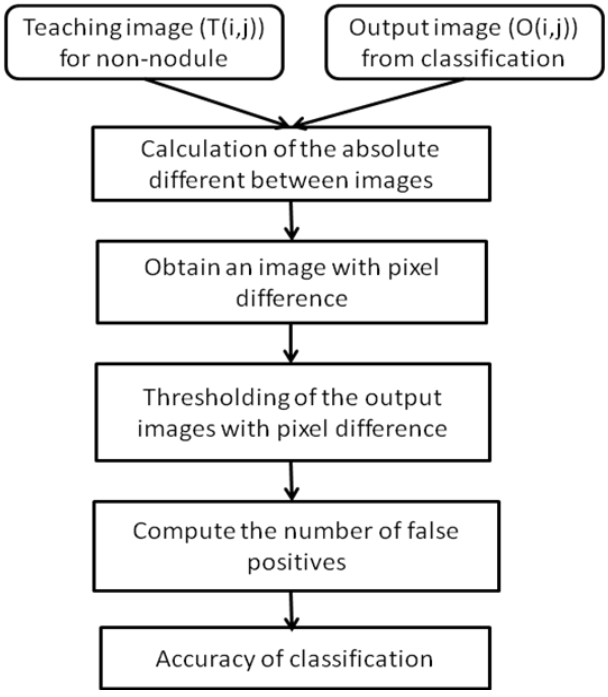


Fig. 5: Procedures of pixel-based evaluation of non-nodule cases

Firstly, the absolute difference between both, teaching and output images is calculated. The formulation to compute absolute differences between the teaching images $T(i, j)$ and output image $O(i, j)$ is shown in Equation 3.

$$Q(i, j) = |T(i, j) - O(i, j)| \tag{3}$$

Further, image $Q(i, j)$ is generated based on the pixel differences from Equation 3. The difference between true positives and false positives can be obviously revealed by thresholding the output of $Q(i, j)$, The threshold value of 0.2 has been determined in the thresholding of $Q(i, j)$. This threshold value is selected since the pixel values below 0.2 are shown darker in the images. Moreover, the number of false positives is obtained, and the accuracy of classification is calculated. The evaluation is repeated to calculate the accuracy between each of both, teaching and output images (Images AN, BN, CN, DN, EN, FN, GN, HN, IN and JN) produced by MTAIRS and E-MTAIRS. The pixel- evaluation has been carried out for E-MTAIRS, and the results are compared with MTAIRS. Equation 4 shows the formulation to calculate the percentage of false positives reduction.

$$\text{Percentage of FP reduction} = \frac{\text{FP of MTAIRS} - \text{FP of E-MTAIRS}}{\text{FP of MTAIRS}} \times 100\% \tag{4}$$

where FP represents the false positives. Based on these percentages of false positives reduction, the ability of E-MTAIRS in significantly reducing the false positives in most of the testing images is proven. The pixel evaluation is done to test the performance of E-MTAIRS compared to MTAIRS for non-nodule images.

Original Images	**MTAIRS**	**False positives**	**E-MTAIRS**	**FALSE POSTIVES**
 Image AN				
 Image BN				

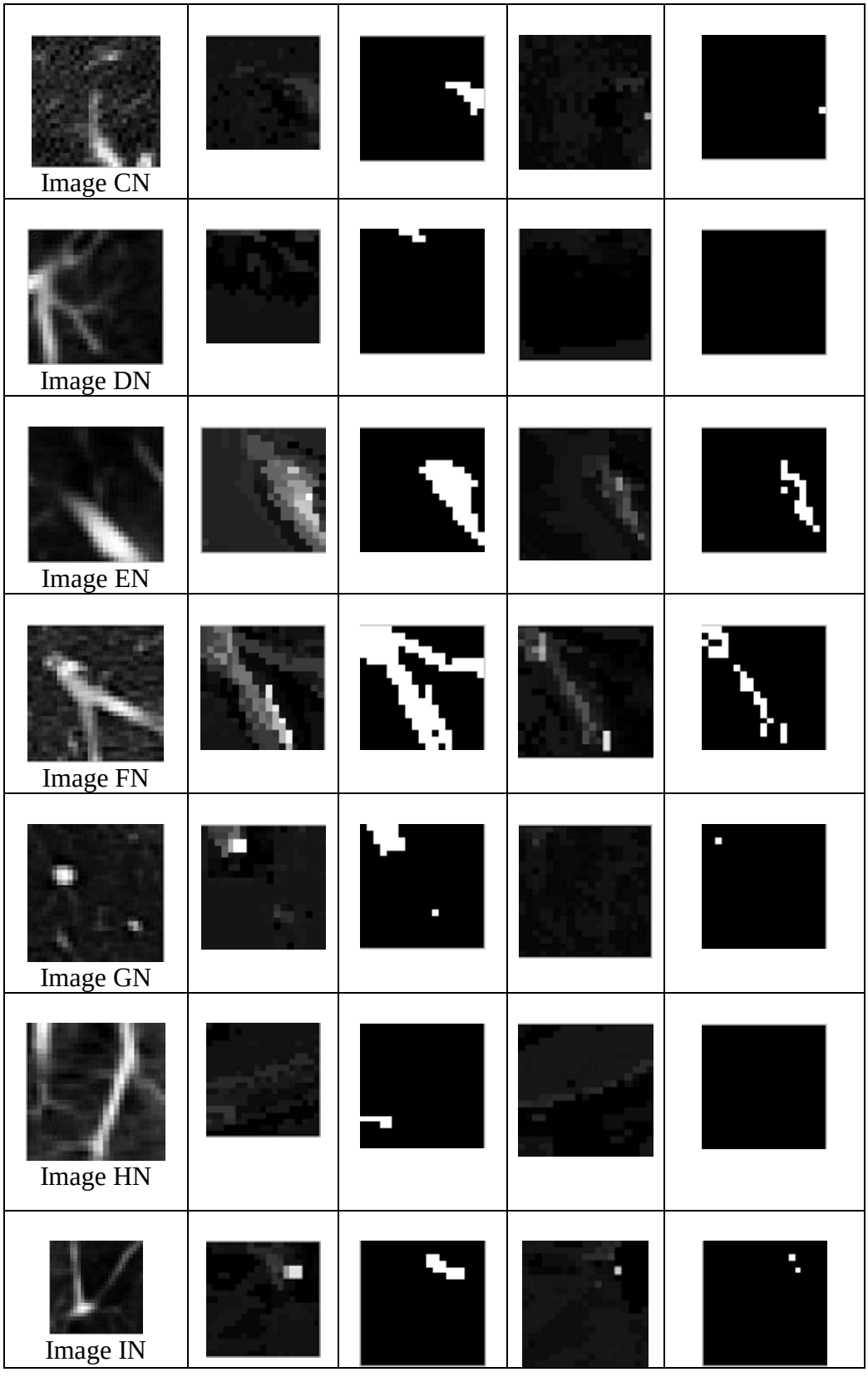


Image CN
Image DN
Image EN
Image FN
Image GN
Image HN
Image IN

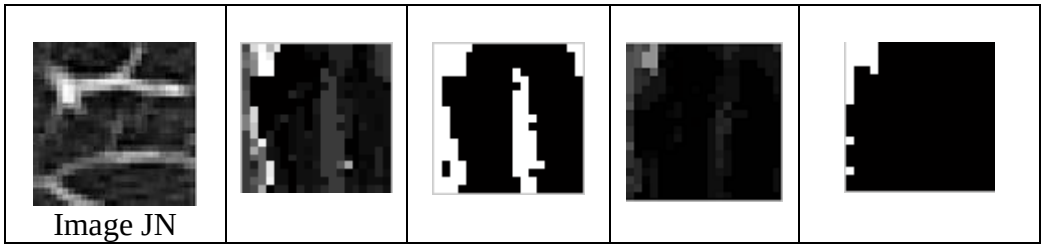


Fig 6: Comparison of absolute difference of output images compared to the original images by MTAIRS and E-MTAIRS

Fig. 6 illustrates the comparison of output images of E-MTAIRS and MTAIRS, and the images of their false positives evaluation. From these output images, it can be observed that most of the false positives have also been reduced by E-MTAIRS. Besides, all of the false positives in the output of E-MTAIRS are totally eliminated in testing images DN and HN. Based on the outcome of the evaluation, the percentage of false positives in output images of MTAIRS and E-MTAIRS is computed. Further, the percentages of false positives reduction of E-MTAIRS are calculated with Equation 4 and these outputs are revealed in Table 1. By referring to the percentages of reduction, E-MTAIRS capability of diminishing the false positives in the testing images is proven.

Table 1: Percentages of false positive and reduction of MTAIRS and E-MTAIRS

Image	Percentage of false positives in MTAIRS (%)	Percentage of false positives in E-MTAIRS (%)	Percentage of reduction (%)
A1	9.14	4.43	51.53
B1	6.09	4.71	22.66
C1	3.88	0.28	92.78
D1	1.39	0.00	100.00
E1	15.24	5.26	65.49
F1	32.41	8.59	73.50
G1	5.54	0.28	94.95
H1	1.94	0.00	100.00
I1	3.60	0.55	84.72
J1	27.15	5.54	79.59

5 Discussion

From the experimental results, E-MTAIRS with enhancement of affinity function are compared to the original version of MTAIRS. The performances of these two algorithms are tested in terms of nodules detection and non-nodules classification.

In the results evaluation for lung nodule cases, the performances of E-MTAIRS in qualitative analysis are slightly different in visualization compared to MTAIRS. The contrast in nodules and the healthy tissues surrounding the nodules are higher and this shows that the classified pixel values have fulfilled the expected results. In the training algorithm, the teaching image reveals that the pixel values which are close to 1 represent nodules area with bright color, while pixel values close to 0 represent lung area without nodules with dark color. Therefore, the use of E-MTAIRS is beneficial in generating the output of detected lung nodules. The qualitative analysis shows that the performance of E-MTAIRS is more optimized between the algorithms in the visualization of lung nodules in output images. Based on the results produced by MTAIRS, it is noticed that the output images consisting of bright noise may be confused with small likelihood of nodular. The distributed bright noise on output images are reduced by the enhanced algorithm in the classification results. Furthermore, the quantitative analysis has also been conducted for non-nodule cases. From the comparison of output images with the teaching image, the pixel differences of classification results can be visualized, and the false positive rate is computed according to the number of pixels that are recognized as false positives. In this results validation, it is found that E-MTAIRS better reduce the false positive in the output images compared to MTAIRS. The E-MTAIRS is able to reduce the false positives effectively by comparing the percentage of false positives reduction in overall analysis.

6 Conclusion

In the enhancement of both proposed classification, non-linear affinity function is chosen to replace Euclidean distance function since it has been proven to be more appropriate in computing the similarity measures between two sub-regions in the training process. In the training process of E-MTAIRS algorithm, the non-linear affinity function plays an important role to generate the best memory instances for each class before the classification procedure. The best memory instances in the outcome of the training process will directly impact the classification accuracy Due to the complexity of non-linear function, the efficiency of E-MTAIRS is lower compared to the MTAIRS based on the time factor analysis. However, their overall performance is still higher. The results comparison with MTAIRS and E-MTAIRS have shown better classification results in the nodule and non-nodule cases classification based on the discussed qualitative and quantitative analysis. The false positives on the outcome of non-nodules testing image are effectively reduced by the enhanced algorithms. Therefore, it can be concluded that replacing Euclidean distance with the non-linear affinity function can improve the overall performance of classification algorithm in terms of visualization and accuracy.

ACKNOWLEDGEMENTS
The authors wish to thank the Malaysian Ministry of Education (MOE), Universiti Teknologi Malaysia (UTM) and Research Management Centre (RMC) for financial sponsorship to this work through grant funding number Q.J130000.2754.03K40.

References

[1] Bhattacharya, G., Ghosh, K. & Chowdhury, A. S. (2012). An affinity-based new local distance function and similarity measure for k-NN algorithm. *Pattern Recognition Letters,* 33**,** 356-363.
[2] Casro, L. N. D. & Timmis, J. (eds.) (2002). *Artificial Immune Systems: A New Computational Intelligence Approach.*
[3] Cha, S.-H. 2007. Comprehensive survey on distance/similarity measures between probability density functions. *City,* 1**,** 1.
[4] Deza, M.-M. & Deza, E. (2006). *Dictionary of distances*, Elsevier.
[5] Duda, R. O., Hart, P. E. & Stork, D. G. (2012). *Pattern classification*, John Wiley & Sons.
[6] Erickson B. J. , Korfiatis P., Akkus Z., Kline T. L. (2017). Machine learning for medical imaging. Radiographics. 37(2), 505-15.
[7] Freitas, A. A. & Timmis, J. (2007). Revisiting the foundations of artificial immune systems for data mining. *Evolutionary Computation, IEEE Transactions on,* 11**,** 521-540.
[8] Hamaker, J. S. & Boggess, L. (2004). Non-Euclidean distance measures in AIRS, an artificial immune classification system. *In: Evolutionary Computation, 2004. CEC2004. Congress on*, 19-23 June 2004. (Vol.1, pp:1067-1073)
[9] Hang, S. P. and Shamsuddin, S.M, Ralescu, A. (2013). Massive Training in Artificial Immune Recognition Algorithm for Enhancement of Lung CT Scans. *In the proceeding of Fifth International Conference of Soft Computing and Pattern Recognition (SoCPaR 2013).* 15-18 December 2013, Hanoi, Vietnam.
[10] Hang, S. P. and Shamsuddin, S.M, Alwee, R. Kullback Leibler Divergence for Image Quantitative Evaluation. (2015). *In the proceeding of Simposium Kebangsaan Sains Matematik (SKSM23)* 24-26 November 2015, Universiti Teknologi Malaysia, Johor.
[11] Perelson, A. S. & Oster, G. F. (1979). Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, 81, 645-670.
[12] Seeker, A. & Freitas, A. A. (2007). WAIRS: improving classification accuracy by weighting attributes in the AIRS classifier. *In: Evolutionary Computation, 2007. CEC 2007*. IEEE Congress on, 25-28 Sept. 2007. (pp:3759-3765).
[13] Shen D., Wu G., Suk H-I. (2017). Deep Learning in Medical Image Analysis. Annual Review of Biomedical Engineering. 19(1), 221-48.

[14] Suzuki K. (2017). Overview of deep learning in medical imaging. Radiological physics and technology. 10(3), 257-73.
[15] Wu, G., Chang, E. Y. & Panda, N. (2005). Formulating distance functions via the kernel trick. *In: Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Chicago, Illinois, USA. (pp:703–709).
[16] Zezula, P., Amato, G., Dohnal, V. & Batko, M. (2006). *Similarity search: the metric space approach*, Springer.