IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

An Accurate Ensemble Classifier for Medical Volume Analysis: Phantom and Clinical PET Study

Mhd Saeed Sharif [a], Maysam Abbod[b], Ali Al-Bayatti[c], Abbes Amira[c], Ahmed Alfakeeh[d], Bal Sanghera[e]

[a] School of Architecture, Computing and Engineering, ACE, UEL, University Way, London.
[b] College of Engineering, Design and Physical Sciences, Dep of EC, Brunel University London.
[c] School of Computer Science and Informatics, CEM, De Montfort University, Leicester, LE1 9BH, UK
[d] Department of Information Systems, Faculty of Computing & Information Systems, King Abdul Aziz University,
[e] Paul Strickland Scanner Centre, Mount Vernon Hospital, London.

Abstract

The predominant application of positron emission tomography (PET) in the field of oncology and radiotherapy and the significant amount of medical imaging research have led to an imminent need for effective approaches for PET volume analysis and the development of accurate and robust volume analysis techniques to support oncologists in their clinical practice, diagnosis procedure, arrangement of the right radiotherapy treatment and evaluation of patients' response to the provided therapy. This paper proposes an efficient optimised ensemble classifier to tackle the analysis problem of squamous cell carcinoma in patients' PET volumes. This optimised classifier is based on an artificial neural network (ANN), fuzzy C-means (FCM), an adaptive neuro-fuzzy inference system (ANFIS), K-means and a self-organising map (SOM). Four ensemble classifier machines are proposed in this study. The first three are built using a voting approach, an averaging technique and weighted averaging, respectively. The fourth novel ensemble classifier machine is based on the combination of a modified particle swarm optimisation (PSO) approach and a weighted averaging approach. Experimental national electrical manufacturers association and international electrotechnical commission (NEMA IEC) body phantom and clinical PET studies of participants with laryngeal squamous cell carcinoma are utilised for the evaluation of the proposed approach. Superior results are achieved using the new optimised ensemble classifier when compared with the results from the investigated classifiers and the non-optimised ensemble classifiers. The proposed approach can identify the region of interest class (tumour), with an average accuracy of 98.11% achieved in the studies of participants with laryngeal tumour. This system underpins the expertise of clinicians for PET tumour analysis.

Keywords: Medical Imaging; Tumour; Committee Machine; Particle Swarm Optimisation; Squamous Cell Carcinoma.

1. Introduction

The investigation and analysis of the volume of positron emission tomography (PET) is crucial for different clinical and diagnosis procedures, such as decreasing noise, artefact evacuation, tumour evaluation in the management stage and helping to plan the right radiotherapy treatment for patients [1]. PET is dynamically consolidated for the administration of patients. The outcomes of clinical investigations utilising fluorodeoxyglucose (FDG)-PET have shown its advantage in the analysis, organising and assessment of patient reactions to treatment [2–4]. The use of cutting-edge elite programming evaluation methodologies is valuable in helping clinicians with clinical findings and radiotherapy arrangement. In spite of the fact that the undertaking of therapeutic volume examination seems basic, insightful knowledge of the organs and information about physiology is necessary for achieving such an assignment on clinical restorative images. Basically, the clinical expert monitors each slice, decides outskirts among the image districts and characterises every area. This is generally finished image by image (2D) for a 3D volume and requires a re-slicing of information into the transaxial, sagittal and coronal planes. Also, recognition of thinner image features and difference changes is frequently required. Despite the fact that, for a commonplace 3D informational index, the comprehensive clinical expert manual investigation can take a few hours to finish, this methodology is possibly the most dependable and precise technique for restorative image examination. This is because of the monstrously multifaceted nature of

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

the human visual framework, a framework appropriate to this assignment [5–8].

The systems based on a combination of classifier frameworks can regularly provide viable answers for identifying different patterns and performing classification requirements. These can be identified by using distinctive terms, such as combined decisions, different expert combinations, blends of specialists, ensemble classifiers, fusions of different approaches, accord conglomerations, dynamic classifier determinations, composite classifier frameworks, half-and-half strategies, smart agents, frameworks supposition pools and board-of-committee ensemble machines [9–17]. The inspiration for such frameworks might be gotten from an exact perception that particular classifiers are prevalent in various circumstances, or it might follow from the idea of the application in question. In addition, endeavours may be concentrated on improving their speculation capacity and upgrading the precision of the classification.

Ensemble classifiers have different structures and are employed in a variety of applications [18]. Ensemble classifiers significantly outperformed the other approaches used for data of microwave breast screening acquired in the clinical trial presented in [9]. Another study presented a neural network ensemble design for image classification purposes, but this work only used neural networks; the ensemble we developed, in contrast, deals with the image as whole and use a combination of techniques [19]. Fusion of contextual information for the purpose of image recognition has been presented in [20], which shows that fusing the information can lead to better outcomes. Multicategory classification problems have been addressed using ensembles of binary classifiers in [21]. A gas-recognisable committee machine that joins different gas identification approaches to get a bound-together choice with improved precision was presented in [22]. The ensemble classifier was executed by amassing the yields of five gas-recognisable proof approaches through a cutting-edge casting ensemble with very good outcomes. A casting ensemble for spoken-influence classification was employed in [23], and the achieved committee precision was contrasted with the correctness of each separate classifier. In another study, a weighted casting committee machine was employed for recognising the human face and voice [24]. A hierarchical ensemble classifier was proposed in [25] based on multiple Fisher's linear discriminant classifiers, where each one embodied different facial evidence for face recognition.

An ensemble classifier was used in [26] to identify tissue in black and white. A group of multilayer perceptrons (an essential form of neural system) was set to learn data inputs comprising of different surface patterns, and the data outputs comprised of tissue type classes that were controlled by clinical specialists. At that point, an ensemble classifier was developed via preparing a Bayesian classifier to join the classification approaches of the neural systems. Results were contrasted among comparable AI based techniques, such as support vector machine and multiclass Bayesian ensemble classifier. The designed methodology was employed to identify the weight ulcer, which is a clinical pathology of restricted harm to the skin and basic tissue brought about by weight, shear or contact. It employed a mean move methodology and a locale-developing procedure for productive region division. Analysis, treatment and care of weight ulcers are exorbitantly expensive for wellbeing administrations. The right injury assessment is a basic errand for enhancing the adequacy of treatment and care. Physicians normally assess each weight ulcer by visual investigation of the harmed tissues, which is an inadequate methodology for assessing the dimension of the injury.

An ensemble classifier of neural systems was presented in a paper identifying the masses found in mammograms as dangerous or benign [27]. This ensemble was employed to group masses into two classes, malignant and benign, without arranging the images into the correct number of classes. This study used twenty areas of intrigue identified with harmful tumours and thirty-seven others belonging to benign tumours. A set of multilayer perceptrons was utilised as a complete ensemble of neural systems. The outcomes were achieved by consolidating the reactions of the individual classifiers. The research study proposed in [28] investigated several AI techniques for the purpose of identifying harmful and amiable bunched microcalcifications. This study's kernel-based approach methodology accomplished a performance of 85%.

A committee machine of neural systems intended to enhance the precision and vigour of identifying samples of gene information was developed in the research of [29]. Another committee machine, based on a voting framework for identifying multiclass protein creases, was presented in the research of [30]. These studies are crucial because identifying the protein structure is a critical for knowing the relationship and arrangement between sequence structures and conceivable practical transformative connections.

The study presented in this paper investigates the efficiency of different committee machines and proposes an efficient optimised committee machine to tackle the analysis problem of squamous cell carcinoma in patients' PET volumes. This optimised classifier is based on an artificial neural network (ANN), fuzzy C-means (FCM), an adaptive neuro-fuzzy inference system (ANFIS), K-means and a self-organising map (SOM). This study includes four committee machines. The first is based on a voting approach, where every single participated technique should generate a specific outcome. The second is based on an averaging technique, where the class outcome revealing the biggest average weight will be selected as the most accurate one. The third is based on weighted averaging, where the generated outcomes from all the techniques are timed with the archived predicted weights. The fourth novel committee machine depends on the combination of a modified particle swarm optimisation (PSO) and weighted averaging approaches. The proposed optimised committee machine is evaluated using experimental national electrical manufacturers association and international electrotechnical commission (NEMA IEC) body phantom and clinical PET studies of 7 participants diagnosed with laryngeal squamous cell carcinoma. Very promising outcomes are achieved using the new optimised committee machine (CM4), as illustrated in the following sections.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

2. Theoretical Background

2.1 Voting Based Ensemble

Voting technique is one of the famous methodologies for consolidating the outcomes of different classifiers. In this method, every single classifier should generate a decision rather than a weight. The chosen class is the one predominantly selected by the various classifiers. Hence, the yield forecast (V_p) is resolved as pursues:

$$V_p = \begin{cases} x_1 \ when \ \sum_{i=1}^{K} x_i > T \\ x_2 \ when \ \sum_{i=1}^{K} x_i < T \\ tie \ when \ \sum_{i=1}^{K} x_i = T \end{cases} \quad (1)$$

where K is the quantity of classifiers and T is a turning limit. If 50% of the involved classifiers vote in favor of one class and the remaining 50% vote in favor of the other class; then a tie status is reached. This occurs when an evenly divisible number of classifiers is employed in the ensemble. Nevertheless, in the proposed ensemble, an odd number of classifiers is conveyed to stay away from this issue. Moreover, the most well-known technique among median, least and greatest techniques is the preponderance vote technique [31–33].

2.2 Averaging Based Ensemble

Averaging based ensemble carries out an averaging technique on the outcome of every single classifier for every representative class across the whole ensemble. The class outcome with the largest amount is then selected. The outcome is shown in Equation 2:

$$Q(x) = \underset{j=1\dots N}{\operatorname{argmax}} \left(\frac{1}{K} \sum_{i=1}^{K} y_{ij}(x) \right) \quad (2)$$

where N is the classes quantity, $y_{ij}(x)$ is the outcome value of the i^{th} classifier for the j^{th} class of the input x; K is the classifiers quantity employed in the whole ensemble [32].

2.3 Weighted Averaging Based Ensemble

Weighted average technique is similar to the past averaging technique; nevertheless, the classifiers' outcomes are timed with archived predicted weights. The outcome is shown in following equation [32]:

$$Q(x) = \underset{j=1\dots N}{argmax} \left(\frac{\sum_{i=1}^{K} w_i y_{ij}(x)}{\sum_{i=1}^{K} w_i} \right) \quad (3)$$

w_i are the weights where $i=1,\dots,K$, are inferred by reducing the classifiers' error for the training group. In the application in question "positron emission tomography", the predicted classifier's precision of every single relating class is employed as a practical weight for that individual class. Every ideal classifier's outcome y_{di} can be generated as the real outcome y_i in addition to a generated error e_i. This is illustrated in following equation:

$$y_{di}(x) = y_i(x) + e_i(x) \quad (4)$$

2.4 PSO Particle Swarm Optimisation

Particle swarm optimisation approach is set up dependent on the behaviour of social swarm. It can control convergence [34–36]. That behaviour in PSO can be be portrayed as pursues: In the event that a swarm of feathered creatures is considered, at that point their goal is to locate a warm spot to travel to. Having no earlier learning of that place, the winged creatures start in arbitrary ways with irregular speeds searching for the spot. Each feathered creature can recollect the discovered area, and by one way or another knows the headings where different winged animals found a legitimate spot. The reluctant flying creature, between the course found and those detailed by different winged creatures, quickens in the two headings modifying its bearing to fly somewhere close to the two known headings. Amid flying, a winged creature may locate a superior course (place) than the one found beforehand. It would then be attracted to this new area just as the other best area found by the entire swarm. Sporadically, one winged animal may fly preferable way over had been experienced by any flying creature in the swarm. The entire swarm would then be drawn toward that area in extra to their very own revelation. In the long run, the feathered creatures' flight guides them to the best spot they are searching for [37, 38].

There are numerous points of interest in PSO; for example, its algorithmic effortlessness. Moreover, it has one straightforward operator, which is the velocity. This element prompts the decrease of computational time and intricacy. In particle swarm optimisation, there are a definite group of elements which must be chosen controlled cautiously as per the application in question. These parameters (elements) are experimentally investigated and optimised based on the proposed application.

3. Methods and Materials

3.1 Phantom and clinical studies

3.1.1 Phantom Studies Data

The main informational index utilized in this research investigation is gathered by deploying the NEMA IEC image quality body phantom. This phantom comprises of a curved water filled cavity which contains 6 spherical inserts suspended by plastic rods of volumes 0.5, 1.2, 2.6, 5.6, 11.5 and 26.5 ml. The The internal measurements of these circles are 10, 13, 17, 22, 28 and 37 mm, respectively. The volume of the PET has a size of 168×168×66 voxels; each voxel has measurements of 4.07×4.07×5 mm^3, equal to a voxel

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

volume of 0.0828 ml. The deployed phantom was broadly utilized in the literature aiming for the evaluation of image quality and the quantitative processes approval [39–42]. In the deployed PET volumes, the PET emission data was reconstruct using a CT-based attenuation correction performed after the Fourier rebinning and model-based scatter correction. The PET volumes were reproduced utilizing two-dimensional iterative standardized normalized attenuation weighted ordered subsets expectation-maximisation (NAW-OSEM). In this experiment, the following default parameters were employed: ordered OSEM repetitive reconstruction, 4 recurrence with 8 subsets, pursued by a post-processing Gaussian filter (5 mm) [43]. The generated phantom volumes were obtained using a GE DST clinical PET-CT scanner.

3.1.2 Clinical Studies Data

The employed clinical data set in this research consisted of clinical data of PET images from 7 patients with T3–T4 laryngeal squamous cell carcinoma. *T3* represents a tumour in the larynx that made one of the vocal strings incaplable of movement, on the other hand the *T4* represents a tumour that reached out behind the larynx. Preceding the treatment, every patient experienced a FDG-PET examination. Patients were immobilized with a tweaked thermoplastic mask attached to a flat table top to avert complex neck motions. The procedures was as follows: a 10 minute transmission scan has been performed using the Siemens Exact HR camera (CTI, Knoxville, USA). Afterwards, a 1 hour dynamic 3D emission scan was performed instantly after intravenous infusion of 185-370 MBq (5-10 mCi) of FDG. This scan has 8 frames with variable span running from 90 to 600s. The acquired volumes have been redressed for dead time, arbitrary, disperse, lessening and rot and after that recreated utilizing a 3D OSEM approach. The ground truth of the tumour was based on the knowledge of expert clinicians who use their training and experience to identify suspected sites of disease through quantification measurement and visual assessment, while the histology (through biopsy) confirmed whether a disease was present at individual suspected sites and characterised its physical distribution and extent in biopsied tissue accordingly. The volume of this deployed data set was 128 × 128 × 47 voxels associated with every participant, where the voxel volume is 2.17 × 2.17 × 3.13 mm^3 [44, 45, 46].

3.2 The Optimisation Approach

This research has developed an optimisation algorithm built based on the PSO approach [34]. It can be summarised as pursues:

1. The initial step of actualizing the particle swarm optimisation approach is to choose the parameters and characterize the looking extent for every one of them.

2. The mean square error fitness function is chosen to show the goodness of the optimisation solution. The PSO particles are optimised by the fitness function, which is formulated as the objective function.

3. Each Every molecule starts at its very own irregular area with an arbitrary speed hunting down the ideal position in the arrangement space. As the underlying position of every molecule is the main area experienced toward the starting, this position ends up particular of every molecule (*pbest*). Every molecule has its very own *pbest* dictated by the way that it has experienced.

4. The principal worldwide best arrangement found by the remainder of the swarm (*gbest*) is then chosen from among these underlying positions. From that point onward, the methodology moves every molecule separately just barely through the whole swarm, and thinks about *gbest* and *pbest*.

5. The particles speed is imperative key in the optimisation system. The speed of the molecule is changed by the overall areas of *pbest* and *gbest.* Every molecule is processed as a point in the *D* dimensional defined space. The i^{th} molecule can be shown as $X_i = (x_{i1}, x_{i2}, ..., x_{iD})$.

The best past position (the position which gave the best wellness esteem) of the i^{th} molecule is monitored and can be shown as $P_i = (p_{i1}, p_{i2}, ..., p_{iD})$. A list of the best molecule among every one of the particles in the populace is depicted using the symbol *g*. The range of velocity variation for particle *i* can be shown as $V_i = (v_{i1}, v_{i2}, ..., v_{iD})$. This velocity is updated according to the following equation:

$$v_{ip}(q+1) = \omega\, v_{ip}(q) + c_1.rand_{(1)}.\left(p_{ip}(q) - x_{ip}(q)\right) + c_2.rand_{(2)}.\left(p_{gp}(q) - x_{ip}(q)\right) \quad (5)$$

where $0 \le i \le (n-1)$, $1 \le p \le D$, *n* is the number of particles in a group and *D* is the dimension of the search space. For each particle, there are *D* number of parameters that are used to identify the particle location in the search space. For a specific particle, *q* is the repetition indictor, $v_{ip}(q)$ is the speed of particle *i* at repettion *q* and *rand(1)* and *rand(2)* are random values within [0..1]. ω is the inactivity weight factor, which determined to what degree the molecule stays along its unique course unaltered by the draw of *gbest* and *pbest*. c_1, c_2 values are the quickening constants, where c_1 is a factor choosing how much the molecule is impacted by the memory of its best area and c_2 is the value choosing how much the molecule is affected by the remainder of the particle groups.

6. In light of the past phase and after refreshing the speed value, the fresh location, to which the particle moved, is recognised based on the following:

$$x_{ip}(q+1) = x_{ip}(q) + v_{ip}(q+1) \quad (6)$$

where x_{ip} is the available location of a certain particle *i* at a repetition *q*.

7. Afterwards, the process from step four onward is repeated once more. The repeated procedure is iterated again and again until the until the end model condition is achieved.

3.3 The Ensemble Optimisation Approach

In light of the comprehension of the reasonable premise of the PSO, an algorithmic methodology is developed in this research as a novel enhancement approach method for PET

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

applications to improve the performance of the developed ensemble. The PET optimisation approach can be summarised as pursues:

1. Every particle among the group of particles speed up toward the best overall position and best settings and keep checking its own current available position.

To implement the developed approach, a specific procedure is followed: First, we need to select then optimise 5 different parameters—one value for every proposed classifier—and provide them a reasonable values, where the best solution is searched for and achieved. In view of the underlying experiments completed for the application of PET volumes, the best-chosen values for all the processed data sets is [0 ... 1.5]. The optimised values (R_1, R_2, R_3, R_4, R_5) were employed within the developed ensemble in order to obtain the optimal outcome in light of the following:

$$Q(x) = \underset{j=1\ldots N}{argmax}\left(\frac{\sum_{i=1}^{K} R_i Acc_i y_{ij}(x)}{\sum_{i=1}^{K} Acc_i}\right) \quad (7)$$

where Acc_i is the classification accuracy of a particular classifier i whereas N refers to how many classes are processed.

2. The second step is an essential one to evaluate the quality of the achieved solution. The calculation of the accuracy method was employed to provide a single figure referring to the effectiveness of the provided solution. The performance index of this function can be calculated as follows:

$$PI = 1 - \frac{1}{N}\sum_{i=1}^{K}\frac{tp_i}{tp_i + fp_i + fn_i} \quad (8)$$

where PI is the performance index, f_p is false positives, t_p is true positive and f_n is false negatives. These values are extracted from the confusion matrix.

3. Every molecule has its own *pbest* controlled by the way that it has experienced. The main worldwide best arrangement found by the remainder of the swarm (*gbest*) is then chosen from among these underlying positions.

4. Every molecule is moved exclusively just barely through the whole swarm, and the *pbest* and *gbest* are thought about. The exactness work restores an incentive to be relegated to the present area. In the event that that esteem is more noteworthy than the incentive at the separate *pbest* for that molecule, or the worldwide *gbest*, at that point the proper areas are supplanted with the present areas.

5. The molecule speed is calculated based on Eq. number 5. We have introduced the random parameter in Eq number 5 to imitate the lightly unforeseeable behaviour element of swarm in nature. Variant empirical parametric studies have been performed in the literature to determine the optimal amount of the c_1 with c_2. It has been determined before that the best choice for both c_1 combined with c_2 is 2.0 [47, 48]. But, the first round of experiments performed using every single data sets of PET shows that the optimal amount of w is 0.7298 and the best amount of c1, and c2 is 1.49618. The obtained experimental values are achieved throughout the following experimental approach, where compels conditions have been connected to these elements, for example, constraining the w within the values of [0..1]. The proposed approach has set the index performance to a value of 0.01. The proposed algorithm has saved, compared, and choose the best parameters to be deployed:

$$c_1 = c_2 = 1.49618$$

$$w = 0.7298$$

1: *Determine the constrains for each parameter*
2: The process of Initialising c_1
3: The process of Initialising c_2
4: The process of Initialising w
5: As long as PI > 0.01 do
6: $c_1 = c_1 + c_1 * rand(\pm 0.1)$
7: $c_2 = c_2 + c_2 * rand(\pm 0.1)$
*8: Calculate w = w + w * rand(0.1)*
9: Import PSO Approach
10: Save the achieved parameters in an array
11: Finish the cycle
12: Evaluate the parameters and deploy the best ones.
Where rand(0.1) was set in the range [0..0.1].

6. The new location for every single particle based on Eq. 6 is calculated.

7. The proposed approach has performed a significant number of experiments which determined the finishing paradigm standard at 100 epochs in respect for every single data set in the PET application. Where if the ultimate number of repetition is significantly big, the algorithm may become hanging tight waiting for an adjustment to be made for the consistent elements in the process; moreover, acquiring just a little number of repetition may lead to no adequate time in the swarm particles to comprehensively discover all the solution area and subsequently figure out what is the optimal answer.

8. A significant number of experiments were performed deploying all the phantom and clinical data sets to achieve the optimal correlation among the particles number (N_P) with the highest number of repition (N_{it}). Hence, it was determined that the complete analysis of such correlation maybe described by Eq. 9 (applicable for both phantom and clinical data sets. Moreover, choosing the right number of particles is very important for improving the performance of the approach in question. Therefore, these initial experiments have evaluated the most suitable number of particles for the data sets in question. The achieved results will be discussed in the results section.

$$N_{it} \cong \frac{3}{2}N_P \quad (9)$$

3.4 Description of the Optimised Proposed System

The proposed medical volume ensemble classifier for the analysis and classification of PET images is shown in Fig 1. Practical NEMA IEC phantom and pharyngolaryngeal squamous cell carcinoma data sets are used in this research to assess the performance of the developed optimised approach. The pre-processed PET images obtained from the scanner has gone through the initial 5 classifiers as follows:

1. Feedforward neural network (FFNN)

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Different parameters have been explored to achieve an appropriate architecture for the artificial neural network suitable for the PET application. These parameters include the training techniques, determining the number of the hidden layers employed in the proposed network architecture and determining the number of hidden neurons in every single layer [6]. Repeated experiments were performed 10 cycles associated with every architecture of the network. Afterwards, the optimal architecture was selected and trained.

2. ANFIS

Iterative test were performed to assess the most appropriate parameters for ANFIS approach [49]. The experiment shows that following selected elements can achieve and generate the optimal performance and results. The influence value is set to 0.1, the accept ratio is set to 0.1, the squash element is determined at 0.25, and the reject ratio is set at 0.0015.

3. Self-organising map (SOM) [50]: For PET application, the learning rate is chosen at a value of 0.6, and the number of training epoch is set to 1000 in the proposed experiments.
4. Fuzzy C-means (FCM) [51]: The experiments determined the pursue parameters to achieve the FCM convergence, which suits the proposed PET application: the number of epochs is determined at a value of 500, the value *m* is equal to 2 and the least value for improvement is set at a value of 1e-5.
5. *K*-means [52].

These previously mentioned five classifiers provide the following classification predications (Pre): Pre1, Pre2, Pre3, Pre4 and Pre5, respectively.

The target class (class 5) is the tumour in the clinical data sets and the spheres (class 4) in the phantom data sets (simulated tumour). Each class refers to a different structure within the processed data sets. The selection of the most appropriate classes number (N) related to each data set in the analysed PET images, is made by experimenting and evaluating different values of N. The optimum value of N is determined based on the Bayesian information criterion (BIC) approach. BIC approach obtained a great fame as noteworthy methodology for model determination, it was employed in different applications e.g. image analysis sociological, biological research activities, etc. The BIC weights are determined gradually against the expanding estimations of *N*. The classes number *N* is selected from a range of 2 to 8, whereas in the proposed medical application of PET, any extra detachment is pointless, based on the analysis of the medical expert. BIC values will in general increment inconclusively as the quantity of parts increments in its model. An expansion in BIC esteem shows an improved model fit, be that as it may, these qualities ordinarily balance out on a surmised bend level, the start of which is typically taken to demonstrate the ideal *N* esteem for every datum set. The drawing of the BIC values against the classes number *N* of the phantom data set illustrates that the optimal *N* number is achieved at four. However, the optimal N value for the clinical data sets is five classes for each patients, where class number 5 refers to the region of interest (tumour), while the remaining classes are the different other structures presented in the image [53–56].

The generated outcomes from the deployed classifiers are sent to three ensemble classifiers at an initial stage. The first ensemble (CM1) gathers the five classifiers' outcomes in light of the voting technique. The second ensemble (CM2) gathers the five classifiers' outcomes in light of the averaging approach, while the third ensemble (CM3) gathers the five classifiers' outcomes in light of the weighted averaging technique. After the initial stage which deployed 3 different ensemble classifiers, a novel approach for optimisation built in light of the PSO approach is introduced and combined with the existing ensemble to improve the achievement and enhance the overall accuracy of the classification. The developed optimiser intensively looks for the most appropriate values for the optimisation procedure (R1, R2, R3, R4, R5) for every single classifier. After searching for the best optimisation values, they are sent and employed in the new ensemble, which in turns creates the optimised predication for the classification as well as the overall accuracy associated with every single data sets. The optimised ensemble/committee machine (CM4) benefits from the mixture of supervised and unsupervised classifiers as well as from the proposed optimisation approach. The most appropriate five parameters deployed in the new ensemble CM4 have enhance the overall performance as well as the accuracy achieved by the new CM4, as discussed in the results section. Every single value of the optimisation parameters has the ability to represent the most appropriate solution related to its own classifier, as long as it was given suitable range of values. At the end of the optimisation process, a performance pointer is created. This pointer is employed to determine the performance of the optimised ensemble CM4. The outputs are evaluated at the following stage, whereas a misclassification value (MCV), confusion matrix, and Acc are employed to assess the ensemble performance. The MCV represents the number of samples which are wrongly classified over the whole number of samples.

The introduced ensemble classifier has generated significant results for all processed data sets including the phantom and clinical ones. The developed ensemble generated an accuracy of 99.9% for some clinical data sets. The results section illustrates the results generated by the introduced ensemble system.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

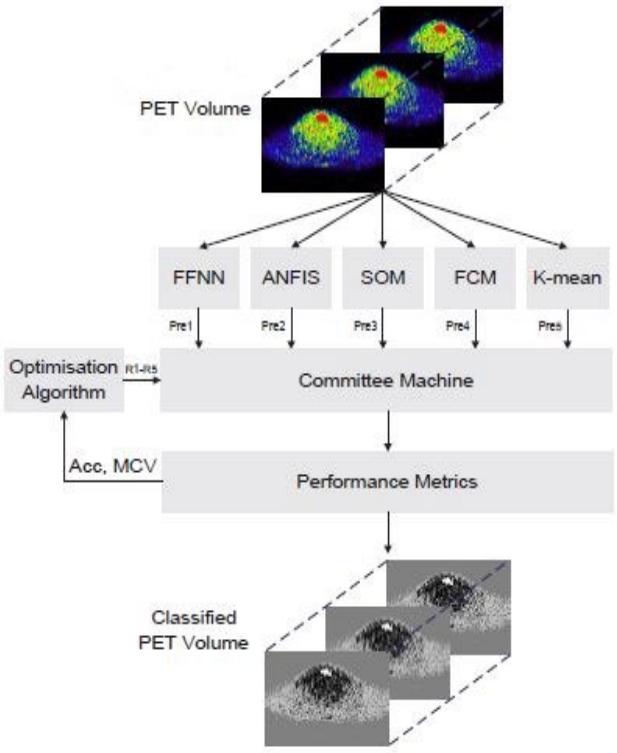

Fig. 1. Novel Optimised Ensemble Classifier for PET Volume Classification.

4. Results

The achieved results will be organised into two main sections for each type of data set (phantom and clinical). The first sub-section analyses the results from the first three committee machines, CM1, CM2 and CM3. The second sub-section discusses the results from the optimised committee machine, CM4, with a focus on the accuracy of region of interest class (tumour).

4.1 Phantom Data Set

4.1.1 The Committee Machines Results

The MCV, confusion matrix, and Acc, are employed to assess the developed ensembles performance. Table 1 shows the confusion matrix and Acc for the ensembles CM1, CM2 and CM3. The confusion matrix for the outputs of the first ensembles CM1 shows the following classification details: All class 1 voxels are precisely classified, there are 27 voxels related to class 2 are falsely classified into the first class. There are 85 voxels related to class 3 are falsely classified into class 1, while 25 voxels are falsely classified into the second class. Class number 4, which represents the simulated tumour, has 10 voxels falsely classified in the first class, while the remaining 42 voxels are falsely classified in the third class. For the region of interest there are 99 voxels precisely classified. In the second ensemble CM2 outputs, the confusion matrix illustrates the following results regarding the falsely classified numbers of voxels: 24 voxels related to class 2 falsely classified into class 1, 88 from class 3 falsely classified into class 2 and 52 from class 4 falsely classified into class 3, however all the first class has a precise classification.

Assessing the generated outcomes from the third ensemble CM3 in light of the confusion matrix demonstrates the following results for the falsely classified voxels: 23 voxels related to the second class are falsely classified into class 1, 141 voxels from class 3 are falsely classified into class 2 and finally there are 51 voxels related to class 4 are falsely classified in class 3, while voxels related to the first class is precisely analysed. Among the three ensembles CM1, CM2, and CM3; the highest accuracy is achieved through the CM3 ensemble.

CM	Class	Cl1	Cl2	Cl3	Cl4	Acc
CM1	Cl1	24939	-	-	-	0.9951
	Cl2	27	260	-	-	0.8333
	Cl3	85	25	2737	-	0.9473
	Cl4	10	-	42	99	0.6556
CM2	Cl1	27939	-	-	-	0.9990
	Cl2	24	263	-	-	0.7013
	Cl3	-	88	2759	-	0.9517
	Cl4	-	-	52	99	0.6556
CM3	Cl1	24939	-	-	-	0.9990
	Cl2	23	264	-	-	0.6168
	Cl3	-	141	2706	-	0.9337
	Cl4	-	-	51	100	0.6622

Table 1: The prediction analysis using a 4 ×4 confusion matrix for the CM1, CM2 and CM3 of phantom data set outcomes and the accuracy (Acc) related to every class.

4.1.2 The Optimised CM Results

Following the evaluation metrics results of CM1, CM2 and CM3 presented in the previous section, the best performance is achieved by CM3. However, higher classification accuracy and better performance are still required. Therefore, the developed optimised CM4 has been deployed within the system to process all the data sets. The initial experiments deploying the phantom data set show that the most appropriate particles number required for the phantom data 66, this number is associated with training epochs equals to 100. This practical value conforms with the one generated by Equation 9. Once the training is performed, the optimisation error/ performance index achieved is PI = 0.0061. The generated error for the optimisation procedure is illustrated in Figure 2, where the training epochs is set 100, which has a stabilised performance and is no longer reduced.

The experimental phantom data set is then processed through the new optimisation approach; CM4. The region of interest's voxels as well as voxels from class 2 are correctly classified in the corresponding class, there are 3 voxels related to class 1 are misclassified to class 2 and only 2 voxels related to class 3 are falsely analysed. The MCV = 0.0002, which is the closest to "0" among all the other

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

classifiers. The MCV values for all classifiers are presented in Table 4.

A significant accuracy improvement is achieved using the CM4 to detect and accurately classify the ROI (simulated tumour). Figure 3 shows a comparison between the new proposed ensemble classifier, CM4, and the remaining approaches (FFNN, ANFIS, SOM, FCM, K-means, CM1, CM2 and CM3) based on the Acc for the phantom data set. The simulated tumour (sphere) represented by class 4 (Cl4) is accurately classified using CM4.

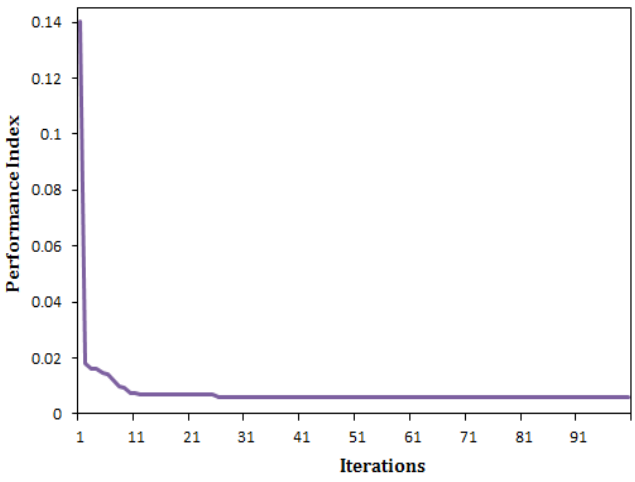

Fig. 2. The optimisation error model for the phantom data set.

Fig. 3. A comparison between the new proposed ensemble classifier, CM4, and the remaining approaches (FFNN, ANFIS, SOM, FCM, K-means, CM1, CM2 and CM3) based on the Acc for the phantom data set.

4.2 Clinical Data Set

4.2.1 The Committee Machines Results

The proposed ensemble classifier's performance is assessed through the deployment of different assessment and analysis metrics for the clinical data sets (Patients 1–7). The confusion matrix illustrates the following results related to the ensemble CM1 outcomes of the data related to Patient 1 (Pt 1) as an example of the clinical data sets for patients with pharyngolaryngeal squamous cell carcinoma. The class 2 voxels are precisely classified, there are 62 voxels related to class 1 are precisely classified. Moreover, there are 2790, 162 and 105 voxels related to the third class, fourth and fifth class, respectively are correctly analysed. Table 2 demonstrates the results obtained by the confusion matrix and Acc for every single class of Pt 1's data set.

The confusion matrix illustrates the following results related to the ensemble CM2 outcomes of the data related to Patient 1: There are 780 voxels related to class 1 falsely classified in class 2, 88 related to class 3 falsely classified in class 2, 1509 related to class 3 falsely classified in class 4, 495 related to classes 4 falsely classified in class 1 and 157 related to class 5 falsely classified in class 1. Class 2 is precisely classified, as illustrated in Table 2.

The ensemble CM3 precisely classified the voxels in class 2; as well as positively classified 230 voxels out of 278 voxels of the ROI, while CM1 and CM2 have classified just 105 and 121 voxels, respectively. The ensemble CM3 has also generated an accuracy for class 1, which is superior than the one obtained through the ensembles CM1 and CM2. The detailed of the accuracy and confusion matrix are illustrated in Table 2.

CM	Class	Cl1	Cl2	Cl3	Cl4	Cl5	Acc
CM1	Cl1	62	737	-	-	-	0.0172
	Cl2	0	9353	-	-	-	0.9204
	Cl3	1968	71	2790	303	-	0.5436
	Cl4	660	-	-	162	-	0.1440
	CL5	173	-	-	-	105	0.3776
CM2	Cl1	19	780	-	-	-	0.0130
	Cl2	-	9353	-	-	-	0.9150
	Cl3	-	88	3535	1509	-	0.6888
	Cl4	495	-	-	327	-	0.1402
	CL5	157	-	-	-	121	0.4352
CM3	Cl1	530	269	-	-	-	0.6633
	Cl2	-	9353	-	-	-	0.9644
	Cl3	-	76	4030	1026	-	0.7852
	Cl4	-	-	-	563	259	0.2969
	CL5	-	-	-	48	230	0.4283

Table 2: The prediction analysis using a 5 × 5 confusion matrix for the CM1, CM2 and CM3 generated outcomes for Pt 1 data set with laryngeal tumour and the Acc.

Similar results are achieved for Patient 2 (Pt 2). The details of the confusion matrix for CM (1–3) and the Acc of each class are presented in Table 3. In contrast, for Patient 3 (Pt 3), the accuracies achieved by ensemble CM3 for the first and fourth classes are 0.9225 and 87.66, respectively. These accuracy values are better than the ones obtained through CM1 and CM2. In the Patient 4 (Pt 4) data set, class 2 is detected correctly by CM1, CM2 and CM3. For

the Patient 5 (Pt 5) data set, CM3 has detected all the classes; however, CM1 and CM2 have only detected classes 2, 4 and 5. The CM 3 generated an accuracy of 1 (100%) for class 1 in Patient 6's (Pt 6) data set, on contrast, the accuracy obtained by CM1 and CM2 are 0.0267 and 0.0116, respectively. In addition, the CM3 accuracy of class 5 is the best among the first-stage committee machines (CM1, CM2 and CM3). For Patient 7 (Pt 7), CM 3 generated an accuracy of 1 (100%) for class 1. The voxels related to the second are all correctly classified through the ensembles CM1, CM2 and CM3; but only ensemble CM1 was not able to determine class 3 voxels.

CM	Class	Cl1	Cl2	Cl3	Cl4	Cl5	Acc
CM1	Cl1	53	2419	-	-	-	0.0123
	Cl2	0	8715	-	-	-	06661
	Cl3	53	1948	1213	-	-	0.3177
	Cl4	1022	-	-	5	-	0.0042
	CL5	129	-	-	147	77	0.2181
CM2	Cl1	16	2456	-	-	-	0.0059
	Cl2	-	8715	-	-	-	0.6392
	Cl3	-	2463	1354	-	-	0.2861
	Cl4	105	-	914	8	-	0.0068
	CL5	118	-	-	145	90	0.2549
CM3	Cl1	1931	541	-	-	-	0.7811
	Cl2	-	8715	-	-	-	0.7697
	Cl3	-	2066	1751	-	-	0.3735
	Cl4	-	-	871	156	-	0.1142
	CL5	-	-	-	339	14	0.0396

Table 3: The prediction analysis using a 5 × 5 confusion matrix for the CM1, CM2 and CM3 outcomes for Pt 2 with laryngeal tumour and the Acc related to every class.

4.2.2 The Optimised CM Results

The patients' data sets have also been processed using the new optimised approach, CM4. Comprehensive experiments were performed to make sure the most appropriate optimisation elements are employed to effectively classify the data sets in question. The optimum particle number required for all patients' data sets is 66 particles as well as 100 training iterations. Once the training procedure is completed, the error/performance indicator for the patients' data sets was $PI = 0.0005$. These optimisation parameters are used for the seven patients' data sets of pharyngolaryngeal squamous cell carcinoma. Figure 4 shows the performance index achieved for the optimisation procedure, which was associated with 100 epochs. The achieved model is generalised and validated to fit all the patients' data sets. This model illustrated a stable robust performance to analyse all the data sets with a stabilised index performance.

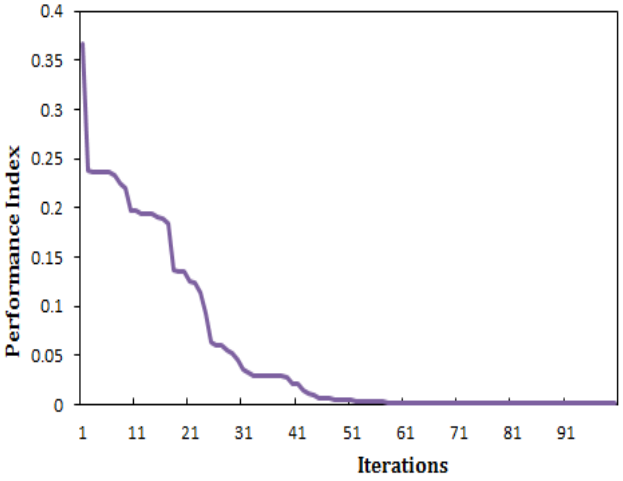

Fig. 4. The optimisation error model for clinical data sets.

In patient 1's data set, the processed voxels in classes 1, 2, 4 and 5 are precisely identified, only 6 voxels related to class 3 are falsely classified in class 4. In Pt 2's data set, in contrast, classes 1 and 4 are correctly classified; however, there are 2 voxels related to class 2 falsely classified in classes 1 and 4, three voxels related to class 3 falsely classified in class 1 and just 1 voxel related to class 5 falsely classified into class 1. In Pt 3's data set, the results illustrate the whole voxels in class 2 are accurately allocated, while only 3 voxels related to the region of interest are falsely classified to class 1. In Pt 4's data set, the results are the same, as 5 voxels out of the 1046 region of interest voxels are falsely classified in class four. In Pt 4's data set, only 1 voxel out of the 1174 region of interest's voxels are falsely classified to class 1. For the region of interest, classes 1 and 3 voxels in Pt 6's data set all of them are precisely allocated, and 3 voxels related to class 2 falsely classified into class 1 and two voxels related to class 4 are falsely classified into class 3.

The accuracy achieved for all the classes (Cl1–Cl5) using the developed CM4 is very satisfactory, as shown in Figure 5. CM4 can accurately classify not only the ROI (tumour) but also all the classes in the clinical data sets (Pt 1–7). This fact shows the robustness of the developed approach in identifying different classes in the data sets, which helps the radiation oncologist with handling the PET volumes and in the clinical diagnosis of squamous cell carcinoma patients.

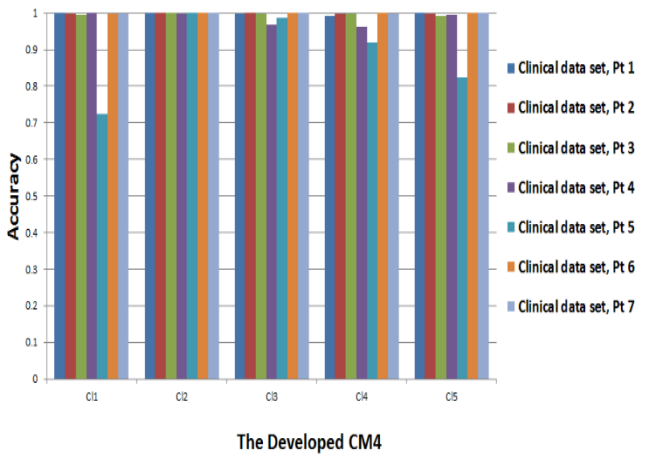

Fig. 5. The accuracy obtained by the developed ensemble (CM4) for all five classes in the patients' data sets.

Figure 6 shows the significant accuracy improvement achieved using the CM4 for classifying and detecting the ROI (tumour). An improvement of 100% has been achieved for class 5 (tumour), as FFNN was not able to detect any voxel of the tumour for all seven patients, while CM4 detected the tumour in all patients, with an average accuracy above 98%. CM4 outperformed all eight other approaches in detecting all the recommended classes, particularly the target ROI class (tumour).

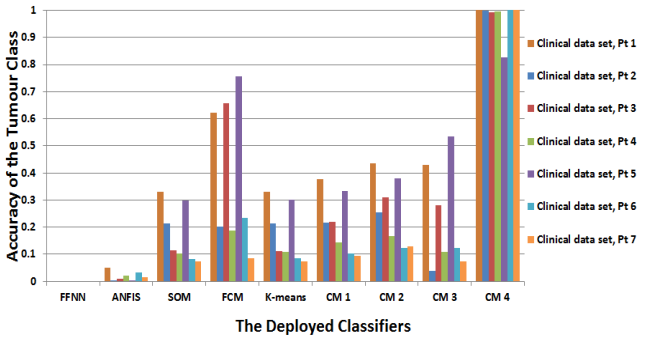

Fig. 6. A comparison between the new ensemble (CM4) and the remaining approaches (FFNN, ANFIS, SOM, FCM, K-means, CM1, CM2 and CM3) based on the Acc for the tumour class (Cl5).

On the other hand, for Pt 7, 1 voxel related to class 1 is falsely classified into class 3 and 4 voxels related to class 4 are falsely classified into class 1. Table 4 presents the MCV generated for the whole patients' data sets, where the least value of MCV=0.0003 is obtained in relation to the Pt 6 and Pt 7. The MCV values for all classifiers are presented in Table 4. Figure 7 shows representative segmentation results for Pt 1's clinical data set, where the black boundary represents the clinical expert estimation. The best match with this boundary is achieved by CM4, where the light blue boundary is almost overlapping the clinical expert boundary.

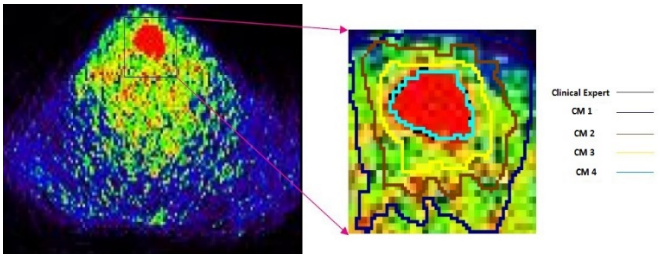

Fig. 7. Representative segmentation results for Pt 1's clinical data set, where the black boundary represents the clinical expert estimation. The light blue boundary is almost overlapping the clinical expert boundary.

5. CM Discussion

To evaluate the performance of CM4 in general for all the classes, not just for the ROI, an average accuracy (AAcc) metric has been introduced. This metric considers the average of the accuracies of all the subject classes. This value provides a general assessment about each one of the considered classifiers. The AAcc for the experimental phantom data set achieved by CM4 (0.9939) is the best among all the classifiers, CM1, CM2 and CM3. The AAcc obtained by CM4 for Pt 1's data set is 99.83%, as a significant improvement level of 78% is obtained in comparison to the lowest accuracy of 21.71% gained through the FFNN approach. The obtained AAccs are 99.94% and 99.93% for Pt 6 and Pt 7, respectively. The AAcc gained in relation to the Pt 6 and Pt 7 data sets are higher than the ones obtained for Pt 1, Pt 2, Pt 3, Pt 4 and Pt 5. This robust result indicates that the developed system has a stabilised robust performance with a higher accuracy than the other approaches.

Figure 8 illustrates a complete comparison in light of the accuracy of the new ensemble (CM4) and the remaining proposed approaches applied for all the analysed clinical data sets (Pts 1–7) as well as the phantom one. In addition to the previously discussed objective performance evaluations for the achieved results by the new system, a comprehensive subjective evaluation in light of the clinician's expertise has been carried out to validate the performance of this approach.

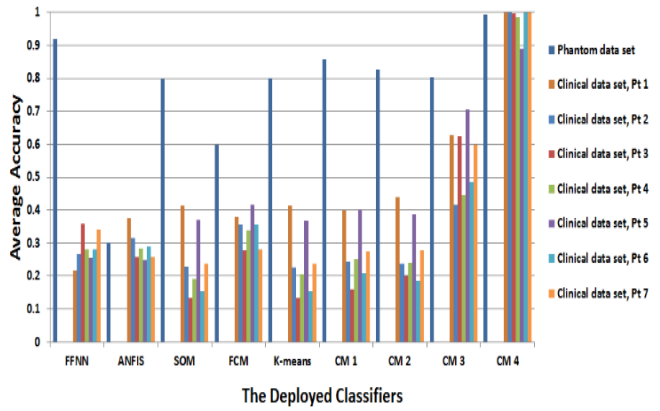

Fig. 8. A comparison between the new ensemble (CM4) and the remaining approaches (FFNN, ANFIS, SOM, FCM, K-means, CM1, CM2 and CM3) based on the AAcc.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Data Set	Phantom	Pt 1	Pt 2	Pt 3	Pt 4	Pt 5	Pt 6	Pt 7
Classifier	MCV							
FFNN	0.0024	0.5692	0.5000	0.4169	0.4662	0.4037	0.5050	0.4741
ANFIS	0.1384	0.3302	0.3073	0.5333	0.4421	0.5104	0.5539	0.5316
SOM	0.0080	0.1887	0.3807	0.7536	0.4680	0.2529	0.6279	0.4867
FCM	0.0139	0.3042	0.2692	0.7891	0.3474	0.2811	0.4064	0.4542
K-means	0.0076	0.1881	0.3808	0.7548	0.4534	0.2535	0.6271	0.4867
CM1	0.0066	0.2387	0.3858	0.7621	0.4331	0.2550	0.5993	0.4702
CM2	0.0058	0.1848	0.3784	0.7174	0.4320	0.2453	0.5915	0.4511
CM3	0.0076	0.1024	0.2329	0.1611	0.2918	0.1111	0.3535	0.2177
CM4	0.0002	0.0004	0.0004	0.0013	0.0062	0.0221	0.0003	0.0003

Table 4: The MCV values for all the approaches and ensembles (CM1, CM2, CM3 and CM4) of the phantom data set and clinical data sets (Pts 1–7).

6. Conclusions

This study proposed an efficient PET volume handling approach for a robust PET volume analysis of patients' squamous cell carcinoma. This approach was based on FFNN, ANFIS, SOM, FCM and K-means. After the initial evaluation of these five classifiers, three ensemble classifiers (CM1, CM2 and CM3) were built in light of different methodologies such as weighted averaging, voting and averaging techniques. As the performance evaluation of these three ensemble classifiers did not reveal a significant level of accuracy for classifying the ROI (especially for the patients' data sets), an optimised novel approach (CM4) based on the combination of the modified particle swarm optimisation (PSO) and weighted averaging approaches was developed for PET volume analysis. This approach overcame the misclassification problem associated with the previous approaches (CM1, CM2 and CM3). All the initial and developed approaches were evaluated using experimental NEMA IEC body phantom and clinical PET studies for laryngeal squamous cell carcinoma patients. Superior results were obtained through the new optimised ensemble/committee machine (CM4) when compared to the results from the other explored approaches and the non-optimised ensembles. The proposed approach can identify the region of interest class (tumour) precisely. The average accuracy obtained for the clinical studies of all patients (Pts 1–7) is 98.11%.

Promising results were achieved in particular for the patients clinical studies. Regarding the NEMA body phantom data set; the proposed approach gained and an overall accuracy of 99.39%. This accuracy was the highest in comparison with the accuracy of the other explored approaches and the non-optimised ensembles. The best improvement achieved by the ensemble CM4 was for the patients' data sets, where the overall voxels of the region of interest, tumour, were accurately allocated in the right class.

Different performance metrics were employed to validate the achieved performance. For example, an MCV value of 0.0002 was generated in regards to the NEMA phantom, and an MCV = 0.0003 was achieved for the pharyngolaryngeal clinical studies. The promising results achieved using different types of PET data, in particular for the patients clinical data, indicated the stability and robustness of the proposed approach. On the other hand, achieving an average accuracy of around 98% and matching the gold standard results fulfilled the aim of this paper, which was to assist clinicians in analysing the significant volumes of PET images precisely.

Acknowledgements

The authors would like to thank John Lee of the Centre of Molecular Imaging, Radiotherapy and Oncology at UCL for the medical data support. The authors gratefully acknowledge the support provided by the deanship of scientific research (DSR), King Abdulaziz University, under grant No. DF 3196111441.

7. References

[1] I. Ketata, L. Sallemi, F. Morain-Nicolier, M. Ben Slima, A. Cochet, K. Chtourou, S. Ruan, A. Ben Hamida, Factor analysis-based approach for early uptake automatic quantification of breast cancer by 18F-FDG PET images sequence, Biomedical Signal Processing and Control, Volume 9, Pages 19-31, 2014.

[2] S. Vauclin, C. Michel, I. Buvat, K. Doyeux, A. Edet-Sanson, P. Vera, I. Gardin, S. Hapdey, Monte-Carlo simulations of clinically realistic respiratory gated 18F-FDG PET: Application to lesion detectability and volume measurements, Computer Methods and Programs in Biomedicine, Volume 118, Issue 1, Pages 84-93, ISSN 0169-2607, 2015.

[3] G. Roman-Jimenez, R. De Crevoisier, J. Leseur, A. Devillers, J. D. Ospina, A. Simon, P. Terve, O. Acosta. Detection of bladder metabolic artifacts in 18F-FDG PET imaging, Computers in Biology and Medicine, Volume 71, Pages 77-85, ISSN 0010-4825, 2016.

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

[4] J. M. Mateos-Pérez, M. L. Soto-Montenegro, S. Peña-Zalbidea, M. Desco, J. J. Vaquero. Functional segmentation of dynamic PET studies: Open source implementation and validation of a leader-follower-based algorithm. Computers in Biology and Medicine, Volume 69, Pages 181-188, ISSN 0010-4825, 2016.

[5] M. Aristophanous, B.C. Penney, and C.A. Pelizzari. The development and testing of a digital pet phantom for the evaluation of tumor volume segmentation techniques. Medical physics, 35(7):3331–3342, 2008.

[6] M. S. Sharif and A. Amira. An intelligent system for PET tumour detection and quantification. Proceedings of the IEEE International Conference on Image Processing (ICIP), November 2009.

[7] T. Doshi, J. Soraghan, L. Petropoulakis, G. Di Caterina, D. Grose, K. MacKenzie, C. Wilson, Automatic pharynx and larynx cancer segmentation framework (PLCSF) on contrast enhanced MR images, Biomedical Signal Processing and Control, Volume 33, Pages 178-188, 2017.

[8] C. Ballangan, X. Wang, M. Fulham, S. Eberl, D. Dagan Feng, Lung tumor segmentation in PET images using graph cuts, Computer Methods and Programs in Biomedicine, Volume 109, Issue 3, Pages 260-268, ISSN 0169-2607, 2013.

[9] Y. Li, E. Porter, A. Santorelli, M. Popović, M.k Coates, Microwave breast cancer detection via cost-sensitive ensemble classifiers: Phantom and patient investigation, Biomedical Signal Processing and Control, Volume 31, Pages 366-376, 2017.

[10] M. S. Sharif and M. H. Alsibai. Medical Data Analysis Based on Nao Robot: An Automated Approach Towards Robotic Real-Time Interaction with Human Body. 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2017.

[11] Guoqiang Peter Zhang. Neural networks for classification: A survey. IEEE transactions on systems, man, and cybernetics-part c: applications and reviews, 30(4):451–462, 2000.

[12] F. Roli and J. Kittler. Multiple classifier systems. Third International Workshop, MCS 2002, Cagliari, Italy, 2364, 2002.

[13] R. Duin. The combining classifier: To train or not to train. International Conference on Pattern Recognition, Canada, 2002.

[14] T. Windeatt and F. Roli (Eds). Multiple classifier systems. Fourth International Workshop, MCS, Guilford, UK, 2709, 2003.

[15] M. S. Sharif, M. Abbod, and A. Amira. PET volume analysis based on committee machine for tumour detection and quantification. Proceedings of the IEEE International Conference on Developments in eSystems Engineering, DeSE2011-Special Session: Intelligent Techniques in Cancer Research,December 2011.

[16] J. Kitler and F. Roli. Multiple classifier systems. Second International Workshop, MCS 2001 Cambridge, UK, 2096, 2001.

[17] H. Kim, H. Kim, H. Moon, and H. Ahn. A weight-adjusted voting algorithm for ensembles of classifiers. Journal of the Korean Statistical Society, 40(4):437–449, 2011.

[18] M. Termenon and M. Graa. A two stage sequential ensemble applied to the classification of alzheimers disease based on MRI features. Neural Process Letter, 35:1–12, 2012.

[19] G. Giacinto and F. Roli. Design of effective neural network ensembles for image classification purposes. Image and Vision Computing, 19(9-10):699– 707, 2001.

[20] J. Sill H. Kasdan X. Song, Y. AbuMostafa and M. Pavel. Robust image recognition by fusion of contextual information. Information Fusion, 3:277–287, 2002.

[21] F. Bovolo, L. Bruzzone, and L. Carlin. A novel technique for subpixel image classification based on support vector machine. IEEE Transactions on Image Processing, 19(11):2983–2999, Nov. 2010.

[22] M. Shi, A. Bermak, S. Member, B. Belhouari, and P. C. H. Chan. Gas identification based on committee machine for microelectronic gas sensor. IEEE Transactions On Instrumentation And Measurement, 55(5):1786–1793, 2006.

[23] D. Morrison and L Desilva. Voting ensembles for spoken affect classification. Journal of Network and Computer Applications, 30(4):1356–1365, 2007.

[24] X. Mu, J. Lu, P.Watta, and M. H. Hassoun. Weighted voting-based ensemble classifiers with application to human face recognition and voice recognition. Proceedings of International Joint Conference on Neural Networks, June 2009.

[25] Yu Su, S. Shan, X. Chen, and W. Gao. Hierarchical ensemble of global and local classifiers for face recognition. IEEE Transactions on Image Processing, 18(8):1885–1896, Aug 2009.

[26] F. Veredas, H. Mesa, and L. Morente. Binary tissue classification on wound images with neural networks and bayesian classifiers. IEEE Transactions On Medical Imaging, 29(2):410–427, 2010.

[27] L. Augusto da Silva, E. Del Moral Hernandez, and R. M. Rangayyan. Classification of breast masses using a committee machine of artificial neural networks. Journal of Electronic Imaging, 17(1):1–10, 2008.

[28] L. Wei, Y.i Yang, R. M. Nishikawa, and Y. Jiang. A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. IEEE Transactions On Medical Imaging, 24(3):371–380, 2005.

[29] B. Liu, Q. Cui, T. Jiang, and S. Ma. A combinational feature selection and ensemble neural network method for classification of gene expression data. BMC Bioinformatics, 5(1):136, 2004.

[30] Y. Chen, F. Chen, J. Y. Yang, and M. Q. Yang. Ensemble voting system for multiclass protein fold recognition. International Journal of Pattern Recognition and Artificial Intelligence, 22(4):747–763, 2008.

[31] C. Ji and S. Ma. Combinations of weak classifiers. IEEE Transactions on Medical Imaging, 8(1):32–42, January 1997.

[32] D. M.J. Tax, M. V. Breukelen, R. P.W. Duin, and J. Kittler. Combining multiple classifiers by averaging or by multiplying? Pattern Recognition, 33(9):1475–1485, September 2000.

[33] L.I. Kuncheva. A theoretical study on six classifier fusion strategies. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(2):281–286, February 2002.

[34] M. Clerc. Particle swarm optimization. ISTE, 2006.

[35] M. Clerc and J. Kennedy. The particle swarmexplosion, stability, and convergence in a multidimensional complex space. IEEE transactions on evolutionary computation, 6(1):58–73, 2002.

[36] Z. Bao, F. Wang, X. Zhao, and T. Watanabe. Fault-tolerant image fi lter design using particle swarm optimization. Artificial Life Robotics, 16:333–337, 2011.

[37] J. Robinson, S. Sinton, and Y. Rahmat-Samii. Particle swarm, genetic algorithm, and their hybrids: optimization of a profiled corrugated horn antenna. IEEE International Symposium on Antennas and Propagation Society, 1:314–317, 2002.

[38] J. Kennedy and W. M. Spears. Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator. IEEE World Congress on Computational Intelligence, pages 78–83, May 1998.

[39] C. Jonsson, R. Odh, P. O. Schnell, and S. A. Larsson. A comparison of the imaging properties of a 3- and 4-ring biograph PET scanner using a novel extended NEMA phantom. IEEE Nuclear Science Symposium Conference, 4:2865–2867, 2007.

[40] M. D. R. Thomas, D. L. Bailey, and L. Livieratos. A dual modality approach to quantitative quality control in emission tomography. Physics in Medicine and Biology, 50(15):187–194, 2005.

[41] H. Bergmann, G. Dobrozemsky, G. Minear, R. Nicoletti, and M. Samal. An inter-laboratory comparison study of image quality of PET scanners using the NEMA NU 2-2001 procedure for assessment of image quality. Physics in Medicine and Biology, 50(10):2193–2207, 2005.

[42] J. M. Wilson and T. G. Turkington. Multisphere phantom and analysis algorithm for PET image quality assessment. Physics in Medicine and Biology, 53(12):3267–3278, 2008.

[43] Data Spectrum Corporation.http://www.spect.com/products-pet.html, 2017.

[44] J. F. Daisne, M. Sibomana, A. Bol, G. Cosnard, M. Lonneux, and V. Gregoire. Evaluation of a multimodality image (CT, MRI, and PET) coregistration procedure on phantom and head and neck cancer patients: accuracy, reproducibility and consistency. Radiotherapy and Oncology, 69:237–245, 2003.

[45] J. F. Daisne, T. Duprez, B. Weynand, M. Lonneux, M. Hamoir, H. Reychler, and V. Gregoire. Tumor volume in pharyngolaryngeal

squamous cell carcinoma: Comparison between CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology, 233:93–100, 2004.

[46] X. Geets, J. A. Lee, A. Bol, M. Lonneux, and V. Gregoire. A gradient-based method for segmenting FDG-PET images: methodology and validation. European Journal of Nuclear Medicine and Molecular Imaging, 34(9):1427–1438, 2007.

[47] R. C. Eberhart and Y. Shi. Particle swarm optimization: developments, applications and resources. Proceedings of the 2001 Congress on Evolutionary Computation, 1:81–86, 2001.

[48] A. Carlisle and G. Doizier. An off-the-shelf PSO. Proceedings of the Particle Swarm Optimization Workshop, 2001.

[49] M. S. Sharif, M. Abbod, and A. Amira. Neuro-fuzzy based approach for analysing 3D PET volume. Proceedings of the IEEE International Conference on Developments in eSystems Engineering, DeSE2011-Special Session: Intelligent Techniques in Cancer Research,, December 2011.

[50] T. Kohonen. Self-organizing maps. Springer, 2001.

[51] Sadaaki Miyamoto, Hidetomo Ichihashi, and Katsuhiro Honda. Algorithms for fuzzy clustering: methods in c-means clustering with applications. Springer, 2008.

[52] A. K. Jain and R. C. Dubes. Algorithms for Clustering Data. Prentice Hall, 1988.

[53] S. H. Ebenuwa, M. S. Sharif, M. Alazab and A. Al-Nemrat. Variance Ranking Attributes Selection Techniques for Binary Classification Problem in Imbalance Data. *IEEE Access*, vol. 7, pp. 24649-24666, 2019.

[54] D. L. Weakliem. A criteque of the bayesian information criterion for model selection. Sociological Methods and Research, 27(3):359-397, 1999.

[55] C. T. Volinsky and A. E. Raftery. Bayesian information criterion for censored survival models. Biometrics, 56:256-262, 2000.

[56] J. T. Bushberg et al. The Essential Physics of Medical Imaging. Philadelphia: Lippincott Williams and Wilkins, 2006.