Medical image registration and classification using smell agent rat swarm optimization based deep Maxout network

Sanjeevkumar Angadi[1], Mukesh Kumar Tripathi[2], Chudaman Devidasrao Sukte[3], Shivendra[4]
[1]Department of Computer Science and Engineering, Nutan College of Engineering and Research, Pune, India
[2]Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India
[3]Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, India
[4]Department of BCA, D.K College, Dumraon, India

Article Info

ABSTRACT

Medical image registration (MIR) is a crucial task in clinical image processing, involving the alignment of images from different modalities, such as magnetic resonance imaging (MRI) and computed tomography (CT), across various time points and subjects. Despite numerous advancements, no universal method caters to all MIR applications. This paper introduces the smell agent rat swarm optimization based deep Maxout network (SARSO-DMN) for MIR and classification. This work aims to enhance the accuracy and efficiency of medical image alignment and classification, addressing the challenges posed by diverse imaging modalities and temporal variations. The problem involves effectively registering CT and MRI images, followed by inhale and exhale classification. The proposed approach begins with feeding the input images into a convolutional neural network (CNN), followed by applying a deformation field to generate an intermediate output (output-1). This output, along with the input MRI images, is further processed by a CNN to produce output-2. Subsequently, output-2 and the input MRI image are subjected to another CNN, resulting in the final registered image. The classification phase utilizes a DMN optimized by the SARSO algorithm, which combines smell agent optimization (SAO) and rat swarm optimizer (RSO). The results demonstrate that SARSO-DMN achieves a maximum accuracy of 90.7%, a minimum false positive rate (FPR) of 11.3%, and a maximum true positive rate (TPR) of 91.2%. The SARSO-DMN approach provides a robust solution for MIR and classification, leveraging advanced optimization techniques to enhance performance.
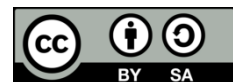
Corresponding Author:

Sanjeevkumar Angadi
Department of Computer Science and Engineering, Nutan College of Engineering and Research
Pune, India
Email: angadi.sanjeevkumar@gmail.com

1. INTRODUCTION

In clinical image assessment, deformable image registration (DIR) is the pre-processing method, which identifies non-linear spatial transformation to align an image. It is significant for several medical applications; wherein spatial aligning of anatomical structure is necessary. Such modalities include image-assisted process for diagnosing as well as patient organization, where images are obtained at diverse points in a time or utilizing various modalities [1]. Particularly in cardiac image assessment, DIR is utilized in image-assisted interferences, which need myocardial motion tracking or else in myocardial perfusion researches [2]. With quick progression of computer and clinical technologies, clinical images of diverse modalities can be

acquired through computed tomography (CT), nuclear magnetic resonance imaging (MRI) and other types of imaging tools [3]. In an area of clinical images, there are few general issues in deep learning-enabled techniques [4], like longer training period, higher necessity for hardware and worst portability of system. An indistinctness of conventional registration technique frequently emerges owing to enormous variation among reference and floating images [4]. A registration quality directly influences the consequence of its chore and follow-up tasks, thus the probe on MIR technology has extensive and reasonable importance. Depending upon types of spatial transformations, an image registration includes non-linear transformation as well as linear transformation. A linear transformation consists of affine registration and rigid registration. It is the total transformation for global image and is frequently utilized as pre-registration function for complicated multiple stage registration [5]. This work significantly contributes to the field of clinical image processing, offering a method that has the potential to improve diagnostic accuracy and patient outcomes in real-world clinical settings.

Deep learning [6] can learn hierarchical features by assuming simpler image features initially and thereafter increasingly constructing up many deep ones from prior levels. Few deep learning techniques like stacked auto-encoder (SAE) [7], deep belief network (DBN) and CNN have revealed the advantage in creating the complicate feature information in higher-dimensional data and have been used in image processing as well as assessment chores like image denoising, classification and segmentation [8]. Recently, deep learning, particularly CNN has subversive impact on a domain of computation vision that includes image segmentation, classification as well as target detection. A deep learning-enabled registration technique substitutes an iterative optimization procedure of convolutional approach utilizing the trained system for achieving quick registration of unobserved pair of image volumes. This has been generated quickly and has become a major direction of MRI probe nowadays by good quality of its outstanding computation speed as well as accuracy [9]. Deep learning techniques that indicate a promising domain of probe have developed for image registration lately. Many of these approaches concentrate on regions of interest (ROIs) in highly "static" parts like prostate, liver or brain. The deep learning image registration (DLIR) model [10] executed unsupervised affine as well as deformable 3D image registration [11]. In present years, deep learning and machine learning have been paved their path into image registration and other associated applications namely computer-aided diagnosis [12], obtaining highly precise and constant solutions. There is the shortage of consensus on categorical method, which benefits from robustness of deep learning against offering high accuracy registrations in spite of obtained image pair criterion [13].

Boveiri *et al.* [14] introduced deep learning for the non-rigid multi-model MRI that could dig the intrinsic features of clinical images more efficiently than hand-devised feature extraction techniques. Even though, this method was not capable to offer acceptable registered outcomes. Yang *et al.* [15] presented ⊥-loss for MRI image reconstruction and registration. It minimized end-point mistake with lesser residual errors, but phase error was increased slowly while increasing extra noise. Zhu *et al.* [16] designed progressive anatomically constrained deep neural network (PACN) for 3D- deformable MIR. This technique generated higher quality of deformation field, though it did not include fine-tuned hyperparameters. Tripathi and Shivendra [17] developed multiscale hierarchical deformable registration network (MHNet) for 3D brain MR image registration. It had probable for registering other organs as well as modal images, but still it was only utilized for brain MRI registration. It was able to execute intra-patient abdominal CT-MRI registration successfully, but it failed to predict helpful synthetic transformations for improving detailed alignment.

Medical image registration (MIR) is a critical task in clinical image processing, involving the alignment of images from various modalities, such as MRI and CT, across different time points and subjects. Despite the existing methods, more universal solutions must be available for all MIR tasks. This paper introduces a novel approach, the smell agent rat swarm optimization deep Maxout network (SARSO-DMN), to improve the accuracy and efficiency of MIR and classification. This study aims to address the limitations of current MIR techniques by developing a more robust and effective method. The problem statement focuses on the challenges of registering CT and MRI images, followed by accurate inhale and exhale classification. The proposed approach involves a multi-stage process starting with the input images fed into a convolutional neural network (CNN). A deformation field is then applied to produce an intermediate output (output-1). This output and the input MRI images are processed through another CNN to generate output-2. Subsequently, output-2 and the input MRI image are subjected to a final CNN stage, resulting in the registered image.

2. PROPOSED METHOD

The study introduces a novel approach for MIR and classification by combining SARSO with a DMN. The key findings highlight that this approach significantly improves the accuracy of image registration and classification compared to traditional methods. The integration of SARSO optimizes the alignment process effectively, while the DMN enhances classification accuracy [18]. Additionally, the method is computationally efficient, making it suitable for large-scale medical imaging tasks. The approach also

demonstrates robustness against noise and image variations, indicating its potential for reliable clinical applications in diagnosis and treatment planning. MIR transforms two or several groups of an imaging data into single coordinate scheme. It has vital part in diagnosis, surgical preparation to real-time assistance and post-procedural analysis. In this work, SARSO-DMN is designed for MIR and classification. The two input images considered are CT and MRI images. These two inputs perform image registration using CNN and then deformation field is applied, thus output-1 is acquired. Thereafter, output-1 and input MRI image is fed to CNN to obtain output-2. Then, output-2 and input MRI image are given to CNN. Figure 1 presents diagrammatical illustration of SARSO-DMN for MIR and classification.
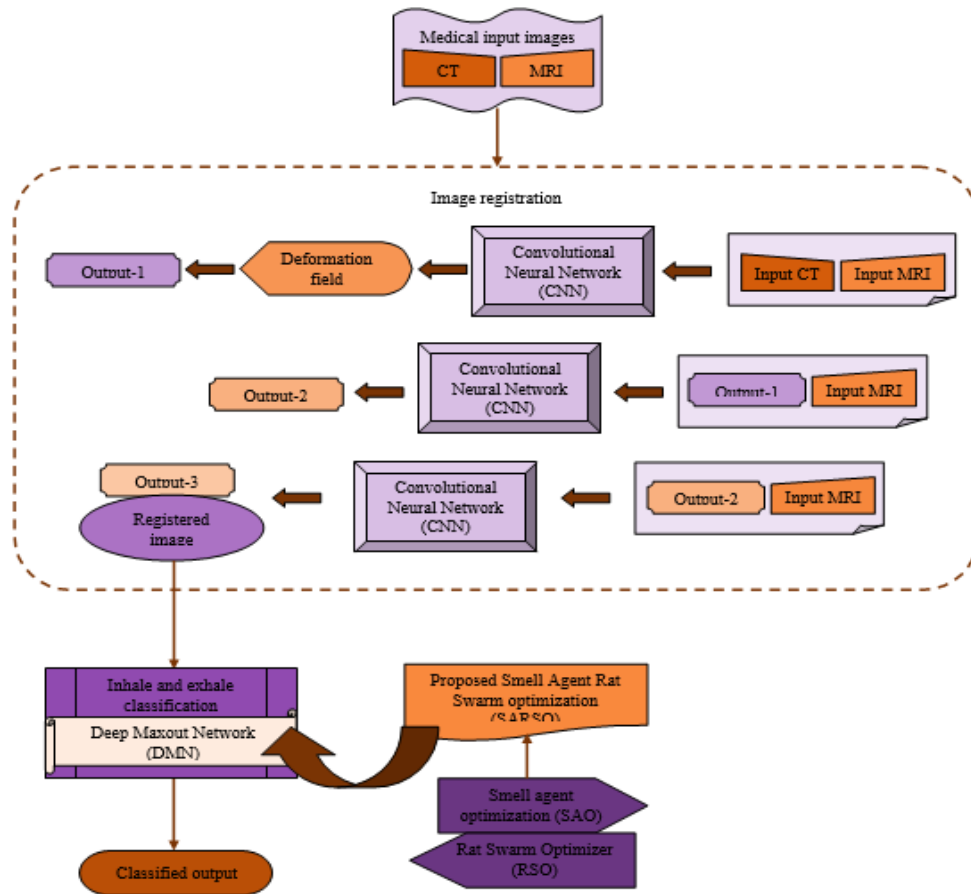

Figure 1. Diagrammatical illustration of SARSO-DMN for MIR and classification

2.1. Acquisition of input MRI and CT image

An input MRI and CT images considered for MIR and classification are obtained from definite database [18], which is specified by,

$$R = \{R_1, R_2\} \tag{1}$$

Here, R indicates database comprises of CT and MRI images whereas R_1 and R_2 can be given as,

$$R_1 = \{M_1, M_2, \ldots, M_m, \ldots, M_i\} \tag{2}$$

$$R_2 = \{C_1, C_2, \ldots, C_m, \ldots, C_i\} \tag{3}$$

where, M_m and M_i represents m^{th} input MRI image and overall MRI images in database whereas C_m and C_i indicates m^{th} CT image and total CT images in database.

2.2. Image registration using CNN

MIR is a primary processing stage in the clinical image processing. It is a process of assigning two or several input images into one image that offer much information. The registration procedure involves an evaluation of optimum transformations, which superiorly allocates objects of interest in input images. Here, MRI and CT images signified as M_m and C_m are taken as input, where input MRI image is fixed image. Initially, both input images are fed to CNN and then deformation field is applied to get output-1 denoted as D_m. The deformation field [19] from every previous coarse level are utilized as preliminary supposition by incorporating them by means of functional composition as well as warp a moving image on high resolution. Thereafter, input MRI image and first output specified by M_m and D_m are passed to CNN. Thus, output-2 represented by X_m is obtained. Afterwards, X_m and M_m are subjected to CNN to achieve registered output.

3. METHOD

CNN [20] is defined as a category of artificial neural network (ANN), which needs convolutional layer and also consists of other layers like convolutional layer, fully connected (FC) and pooling layers to form deep CNN.

− Convolutional layer

A convolutional layer intends to learn the feature depictions of inputs. This layer is comprised of various convolutional kernels that are utilized for computing numerous feature maps. Particularly, individual neuron of feature map is joined to an area of neighborhood neurons in prior layer. An entire feature maps are acquired utilizing various diverse kernels. A feature value at position (b, c) in s^{th} feature map of ωth layer can be mathematically calculated by,

$$l_{b,c,s}^{\omega} = g_s^{\omega Z} \varsigma_{b,c}^{\omega} + a_s^{\omega} \tag{4}$$

here, g_s^{ω} and a_s^{ω} indicates weight vector as well as bias of sth filter of ωth layer whereas $\varsigma_{b,c}^{\omega}$ specifies input patch that is centered at position (b, c) of ωth layer.

− Activation function

Activation function presents non-linearity to CNN that are desirable for the multi-layer networks for detecting non-linear features. It can be calculated as (5).

$$v_{b,c,s}^{\omega} = v\left(l_{b,c,s}^{\omega}\right) \tag{5}$$

Where, $v(.)$ implies non-linear activation function. A general activation functions are namely rectified linear unit (ReLU), tanh and sigmoid.

− Pooling layer

This layer intends for achieving shift-invariance by decreasing feature maps resolution. It is normally located among two convolutional layers. Individual feature map of pooling layer is associated to its related feature map of previous convolutional layer. The pooling function can be represented as (6),

$$\rho_{b,c,s}^{\omega} = pool\left(v_{\alpha,\beta,s}^{\omega}\right), \forall(\alpha, \beta) \in N_{b,c} \tag{6}$$

here, pooling function is denoted by $pool(.)$ and $N_{b,c}$ specifies local neighbor across position (b, c). An average pooling and max-pooling are common pooling functions.

− FC layer

Following various convolutional layers and pooling layers, CNN structure consists of one or numerous FC layers that intends to execute higher-level reasoning. The FC layers take every neuron in prior layer and link them to all single neuron of present layer for generating global semantic information.

− Output layer

The last CNN layer is output layer and SoftMax function is generally utilized for several tasks.

3.3. Inhale and exhale classification utilizing DMN

The inhale and exhale classification are performed for validating a deformable registration. Here, DMN is utilized to carry out inhale and exhale performance by concerning G_m as input. The tuning of DMN is accomplished by SARSO, which is formed by joining smell agent optimization (SAO) with rat swarm optimizer (RSO).

3.3.1. Architecture of DMN

Maxout model refers to feed-forward configuration namely deep CNN or else multilayer perceptron's that make use of newer variety of an activation function specified as Maxout unit. DMN [21], [22] contains numerous layers that generate hidden activations using Maxout function.

− ReLU

Initially, ReLU is engaged in the restricted Boltzmann machines (RBM) that can be illustrated as,

$$\varphi_v = \begin{cases} K_v, if\, K_v \geq 0 \\ 0, if\, K_v < 0 \end{cases} \tag{7}$$

here, K_v symbolizes input subjected to neuron and φ_v denotes output.

− Maxout

It is an actual sort of ReLU operation that obtains max function on $\lambda(\lambda = 2)$ trainable linear functions. Assume particular input $K \in A^\partial$, where K implies state vector or else raw input vector of a hidden layer and Maxout unit output can be formulated by,

$$Q_v(K) = \max_{\varpi \in [1,\lambda]} B_{v\varpi} \tag{8}$$

where, $B_{v\varpi} = K^O X_{\ldots v\varpi} + H_{v\varpi}, X \in A^{\partial \times \delta \times \lambda}$ signifies trainable parameters whereas λ signifies count of linear sub-hidden units. A Maxout unit activation balances uttermost across λ feature maps in CNN. Furthermore, Maxout unit is analogous to normally employing spatial max-pooling in CNN configuration. This attains higher value more than equal input, where spatial max-pooling is associated to λ numerous inputs.

− DMN

This is the sort of trainable activation function with numerous layer arrangements. For an input $K \in A^\partial$, hidden unit activation is modeled as,

$$M_{v,\varpi}^1 = \max_{\varpi \in [1,\lambda_1]} K^O X_{\ldots v\varpi} + H_{v\varpi} \tag{9}$$

$$M_{v,\varpi}^2 = \max_{\varpi \in [1,\lambda_2]} M_{v,\varpi}^1{}^O X_{\ldots v\varpi} + H_{v\varpi} \tag{10}$$

$$M_{v,\varpi}^\delta = \max_{\varpi \in [1,\lambda_\delta]} M_{v,\varpi}^{\delta-1}{}^O X_{\ldots v\varpi} + H_{v\varpi} \tag{11}$$

$$M_{v,\varpi}^q = \max_{\varpi \in [1,\lambda_q]} M_{v,\varpi}^{q-1}{}^O X_{\ldots v\varpi} + H_{v\varpi} \tag{12}$$

$$Q_v = \max_{\varpi \in [1,\lambda_q]} M_{v,\varpi}^q \tag{13}$$

here, λ_δ implies total units in δ^{th} layer and q signifies entire layers consisted in DMN. The DMN activation is much stronger for evaluating an arbitrary recurrent activation function except non-convex activation operations. A conventional non-linear activation functions such as rectified linear and an absolute value rectifier are assessed by DMN when λ is not lower than 2. However, an effectual feature extraction requires a large amount difficult non-linear operation. DMN is capable for estimating an arbitrary activation operation during parameter augmentation.

3.3.2. Training of DMN utilizing SARSO

SAO [23] utilizes phenomena of the smell and an instinctive trailing characteristic of agent for identifying a smelling source. It was proven as effective approach in resolving unimodal as well as multimodal operations. RSO [24] is designed to solve challenging optimization issues. A major stimulation of this approach is attacking and chasing characteristics of rats naturally. This algorithm has higher performance ability in unidentified search spaces. Here, SAO and RSO are incorporated to form SARSO that is an effectual approach to tune DMN for inhale and exhale classification.

A fitness computation is done by identifying a variation amid targeted output and DMN output, which is calculated as,

$$\lambda = \frac{1}{i} \sum_{m=1}^{i} [T_m - S_m]^2 \tag{14}$$

where, i signifies overall CT and MRI images whereas T_m and S_m represents targeted and DMN outputs. SARSO performs beneath elucidated steps to get superior solution.
−   Step 1: Initialize the solution
        SAO process is initiated by generating initial location (population) of the smell molecules randomly. The initialization process can be formulated as,

$$\Re = \{\Re_1, \Re_2, \ldots, \Re_n, \ldots, \Re_\xi\} \tag{15}$$

where, $\Re_\xi$ represents total variables, $\Re_n$ denotes n^{th} candidate solution and $\Re$ specifies population.
−   Step 2: Calculation of objective function
        The calculation of objective function is performed by computing difference amid targeted and DMN outputs using (14).
−   Step 3: Sniffing method
        Assume smell molecules signified as L and hyperspace, wherein smell molecules are evaporating is given by H. The smell molecules population is allocated a location as,

$$\Re_n^\mu = [\Re_{L,1}^\mu, \Re_{L,2}^\mu, \ldots, \Re_{L,H}^\mu] \tag{16}$$

where, $n = \{1, 2, \ldots, L\}$ and μ indicates smell molecules position.
        A velocity vector T is considered as diffusion vector that is the displacement of smell molecules from smell source or origin. The smell molecules velocity can be computed by (17).

$$t_n^\mu = [t_{L,1}^\mu, t_{L,2}^\mu, \ldots, t_{L,H}^\mu] \tag{17}$$

        The location of individual candidate solution can be determined by location vector $J_n^\mu \in P^L$ and molecule velocity $T_n^\mu \in P^L$. As the smell molecules movements are non-uniform in six-dimensional coordinates, a velocity can be illustrated by,

$$\frac{\Delta J_{(u,v,w)}^\mu}{\Delta\mu} = T_{(x,y,z)}^\mu \tag{18}$$

here, $\Delta J_{(u,v,w)}^\mu$ implies displacement, $\Delta\mu$ indicates change in time and $T_{(x,y,z)}^\mu$ specifies velocity.

$$\Delta J_{(u,v,w)}^\mu = T_{(x,y,z)}^\mu \times \Delta\mu \tag{19}$$

        Thus, a generalized form of above equation can be modeled as,

$$\Delta J_{(u,v,w)}^{(\mu)+1} - \Delta J_{(u,v,w)}^\mu = T_{(x,y,z)}^\mu \times \Delta\mu \tag{20}$$

the aforementioned equation can be rearranged as,

$$\Delta J_{(u,v,w)}^{(\mu)+1} = T_{(x,y,z)}^\mu \times \Delta\mu + \Delta J_{(u,v,w)}^\mu \tag{21}$$

hence, $\Delta\mu$ in above equations is taken as 1. Now, substitute $\Delta\mu$ in above equation and thus equation becomes,

$$\Delta J_{(u,v,w)}^{(\mu)+1} = T_{(x,y,z)}^\mu + J_{(u,v,w)}^\mu \tag{22}$$

here, (u, v, w) and (x, y, z) are location and velocity coordinates of smell molecules in search space.
−   Step 4: Smell velocity updating and location
        All smell molecules consist of related velocities, wherein they move and updates their location in search space. As gas molecules travel in non-uniformity manner, the gas velocity is acquired from hydrostatic pressure of gas that can be given by,

$$p = \frac{1}{2}(lkt^2) = \frac{3}{2}(\tau lZ) \tag{23}$$

here, l denotes quantity of substance in gas. Thus, velocity of non-uniformity gas molecules in search space is formulated by,

$$t_n^{(\mu+1)} = \sqrt{\frac{3\tau Z}{k}} \tag{24}$$

from (24), smell molecules velocity is updated as,

$$t_n^{(\mu+1)} = t_n^\mu + d_0 \times \sqrt{\frac{3\tau Z}{k}} \tag{25}$$

− Step 5: Trailing method

For an agent to effectively trail a smell path, agent must be capable to keep in mind about molecule position with finest smell attentiveness and location with worse attentiveness of smell [25]. An agent utilizes this information while training to evade moving towards a domain with poor fitness. A trailing characteristic of an agent can be mathematically modeled as,

$$\mathfrak{R}_n^{(\mu+1)} = \mathfrak{R}_n^\mu + d_1 \times f \times \left(\mathfrak{R}_\vartheta^{(\mu)} - \mathfrak{R}_n^{(\mu)}\right) - d_2 \times f \times \left(\mathfrak{R}_\chi^{(\mu)} - \mathfrak{R}_n^{(\mu)}\right) \tag{26}$$

$$\mathfrak{R}_n^{(\mu+1)} = \mathfrak{R}_n^\mu [1 - d_1 \times f + d_2 \times f] + d_1 \times f \times \mathfrak{R}_\vartheta^{(\mu)} - d_2 \times f \times \mathfrak{R}_\chi^{(\mu)} \tag{27}$$

the standard equation of RSO can be modeled by,

$$\overrightarrow{U_n}(r+1) = \left|\overrightarrow{U_e}(r) - \overrightarrow{U}\right| \tag{28}$$

− Step 6: Termination

The process above mentioned for SARSO performs continually until optimum solution is obtained.

4. RESULTS AND DISCUSSION

The best achievements attained by SARSO-DMN for MIR and classification are expounded in below sub-sections.

4.1. Experimentation setup

An experimental execution of the SARSO-DMN for MIR and classification was carried out using Python. Multimodal ground truth datasets for abdominal MIR were utilized in this study. The dataset includes T1-weighted MRI, CT, and cone beam (CB) CT images that are intrinsically co-registered. The MRI, CT, and CBCT images were provided in .zip file format and were extracted for use in the experiments. These datasets provide ground truth for image registration and segmentation, making them ideal for evaluating the performance of the proposed method. The .zip files containing the MRI, CT, and CBCT images were unzipped, and the images were loaded into the Python environment. The intensity values of the images were normalized to ensure uniformity and to enhance the performance of the neural networks. All images were resized to a uniform dimension suitable for the CNN input layer.

4.2. Comparative analysis

Deep learning [1], PACN [3], deformable deep learning [5], and unsupervised deep learning [7] are the comparative methods taken into consideration to evaluate SARSO-DMN to show its effectiveness. SARSO-DMN performs comparative assessment by varying training data and k-value concerning evaluation measures. Figure 2 expounds on the evaluation of the SARSO-DMN by varying the percentage of training data and considering various performance measures. The values acquired by the devised SARSO-DMN for 90% of the training data are detailed below. Figure 2(a) delineates the assessment of SARSO-DMN considering accuracy. SARSO-DMN attained an accuracy of 0.907, whereas deep learning, PACN, deformable deep learning, and unsupervised deep learning achieved accuracies of 0.813, 0.835, 0.861, and 0.887, respectively. This reveals a performance enhancement of 10.372%, 7.946%, 5.002%, and 2.211% by SARSO-DMN over these methods. The higher accuracy indicates that SARSO-DMN is more effective in correctly classifying the images than the other approaches. Evaluation of SARSO-DMN about false positive rate (FPR) is shown in Figure 2(b). The FPR achieved by SARSO-DMN is 0.113, while the FPRs acquired

by deep learning, PACN, deformable deep learning, and unsupervised deep learning are 0.223, 0.209, 0.179, and 0.131, respectively. This lower FPR for SARSO-DMN signifies a reduction in negative instances incorrectly classified as positive, highlighting its reliability in minimizing false alarms. Figure 2(c) demonstrates the assessment of SARSO-DMN about true positive rate (TPR). SARSO-DMN obtained a TPR of 0.912, whereas deep learning, PACN, deformable deep learning, and unsupervised deep learning achieved TPRs of 0.817, 0.837, 0.867, and 0.884, respectively. This represents a performance improvement of 10.484%, 8.292%, 5.004%, and 3.129% by SARSO-DMN. The higher TPR indicates that SARSO-DMN is more effective in correctly identifying positive instances, enhancing its utility in clinical applications where accurate positive detection is crucial. The results depicted in Figure 2 illustrate the superior performance of SARSO-DMN in terms of accuracy, FPR, and TPR compared to other state-of-the-art methods. The significant improvements in these metrics demonstrate the effectiveness of SARSO-DMN in MIR and classification tasks.

The classification phase employs a DMN tuned by the SARSO algorithm, which is a hybrid of SAO and RSO. This novel optimization approach enhances the DMN's performance in classification tasks. The key findings of this work underscore the substantial improvement in accuracy and reliability achieved by the SARSO-DMN approach. It attains a maximum accuracy of 90.7%, a FPR of 11.3%, and a maximum TPR of 91.2%. These results clearly outperform existing methods, demonstrating the potential of the SARSO-DMN approach in the field of medical image processing and classification. The SARSO-DMN approach presents a substantial advancement in the field of MIR and classification. Integrating advanced optimization techniques and deep learning models offers a promising solution for enhancing the precision and effectiveness of medical image processing, potentially leading to better diagnostic accuracy and patient outcomes.
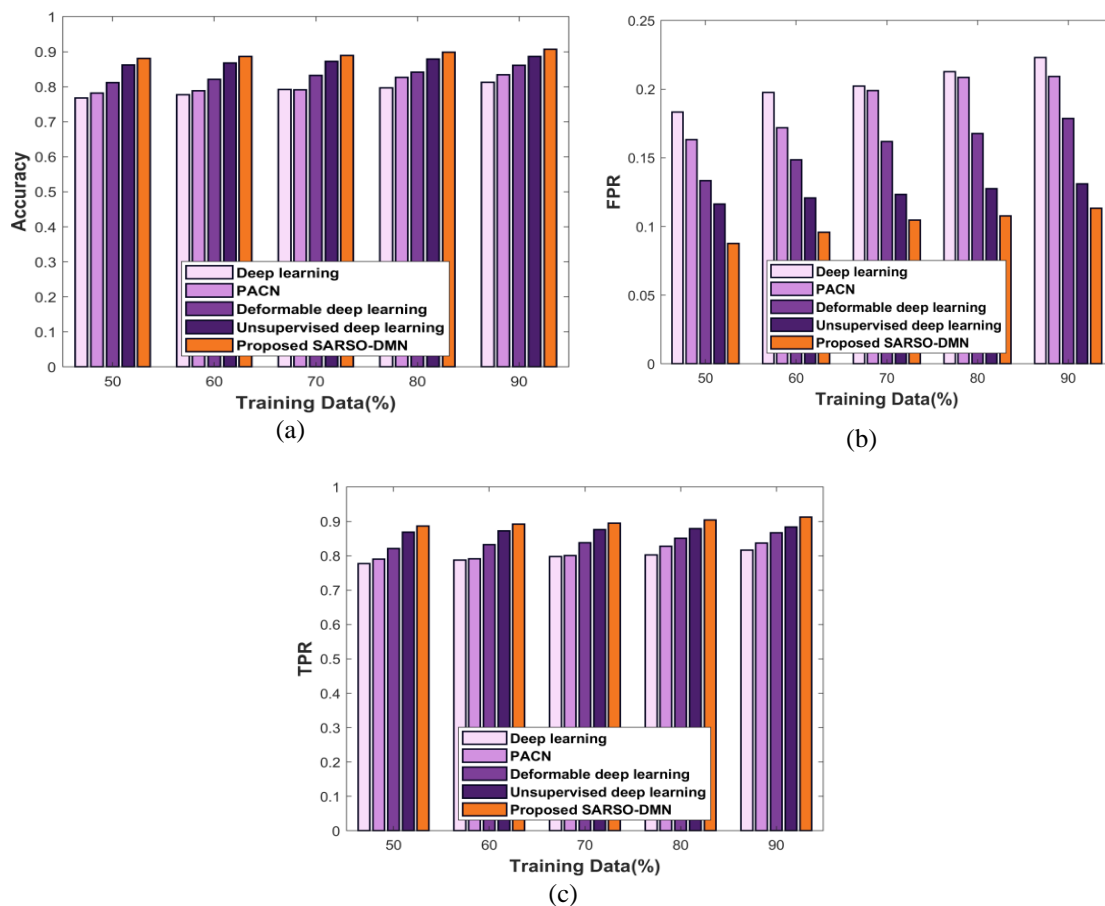

Figure 2. Comparative evaluation concerning training data; (a) accuracy, (b) FPR, and (c) TPR

5. CONCLUSION

In recent days, MIR has become a valuable assistant for clinical specialists. Its significance lies in its ability to aid in diagnosis, monitor the progression of diseases, and select appropriate therapies based on the

patient's condition. Image-based registration approaches that aggregate data from more significant parts of an image to enhance accuracy are computationally intensive and often suffer performance degradation. In this context, the SSARSO-DMN is introduced for MIR and classification. The proposed method involves two input images, such as CT and MRI images, from the dataset. The image registration process begins by passing these input images into a CNN and then applying a deformation field to obtain an intermediate output (output-1). This output and the input MRI images are then fed back into a CNN to produce output-2. Finally, output-2 and the input MRI image are subjected to another CNN to obtain the registered image. This registered image is used for inhale and exhale classification utilizing a DMN. The combination of SAO with RSO, termed SARSO, is employed to train the DMN. The SARSO-DMN achieved maximum accuracy, minimum FPR, and maximum TPR of approximately 90.7%, 11.3%, and 91.2%, respectively, when considering 90% of the training data. These findings demonstrate the effectiveness of the proposed method in improving the accuracy and reliability of MIR and classification compared to existing methods. The SARSO-DMN method can be further optimized and adapted for other medical imaging tasks, such as tumor detection, organ segmentation, and treatment planning.

REFERENCES

[1]   F. Zhu, X. Zhu, Z. Huang, M. Ding, Q. Li, and X. Zhang, "Deep learning based data-adaptive descriptor for non-rigid multi-modal medical image registration," *Signal Processing*, vol. 183, 2021, doi: 10.1016/j.sigpro.2021.108023.

[2]   M. L. Terpstra, M. Maspero, A. Sbrizzi, and C. A. T. V. D. Berg, "⊥-loss: a symmetric loss function for magnetic resonance imaging reconstruction and image registration with deep learning," *Medical Image Analysis*, vol. 80, p. 102509, 2022, doi: 10.1016/j.media.2022.102509.

[3]   Z. Zheng, W. Cao, Z. He, and Y. Luo, "Progressive anatomically constrained deep neural network for 3D deformable medical image registration," *Neurocomputing*, vol. 465, pp. 417–427, Nov. 2021, doi: 10.1016/j.neucom.2021.08.097.

[4]   M. Huang, G. Ren, S. Zhang, Q. Zheng, and H. Niu, "An unsupervised 3D image registration network for brain MRI deformable registration," *Computational and Mathematical Methods in Medicine*, vol. 2022, pp. 1–10, Oct. 2022, doi: 10.1155/2022/9246378.

[5]   A. Lara-Hernández *et al.*, "Deep learning-based image registration in dynamic myocardial perfusion CT imaging," *IEEE Transactions on Medical Imaging*, vol. 42, no. 3, pp. 684–696, Mar. 2023, doi: 10.1109/TMI.2022.3214380.

[6]   Q. Zheng, Q. Wang, X. Ba, S. Liu, J. Nan, and S. Zhang, "A medical image registration method based on progressive images," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1–9, Jul. 2021, doi: 10.1155/2021/4504306.

[7]   G. A. Benvenuto, M. Colnago, M. A. Dias, R. G. Negri, E. A. Silva, and W. Casaca, "A fully unsupervised deep learning framework for non-rigid fundus image registration," *Bioengineering*, vol. 9, no. 8, p. 369, Aug. 2022, doi: 10.3390/bioengineering9080369.

[8]   H. Siebert, L. Hansen, and M. P. Heinrich, "Learning a metric for multimodal medical image registration without supervision based on cycle constraints," *Sensors*, vol. 22, no. 3, p. 1107, Feb. 2022, doi: 10.3390/s22031107.

[9]   Y. Fu, Y. Lei, T. Wang, W. J. Curran, T. Liu, and X. Yang, "Deep learning in medical image registration: a review," *Physics in Medicine & Biology*, vol. 65, no. 20, p. 20TR01, Oct. 2020, doi: 10.1088/1361-6560/ab843e.

[10]   M. K. Tripathi, M. Neelakantappa, P. N. Mahalle, S. V. Channapattana, G. Deshmukh, and G. Prashant, "Detection and classification of mango fruit-based on feature extraction applying optimized hybrid LA-FF algorithms," in *Data-Centric Artificial Intelligence for Multidisciplinary Applications*, Boca Raton: Chapman and Hall/CRC, 2024, pp. 177–185.

[11]   M. K. Tripathi, P. K. Reddy, M. Neelakantappa, C. V. Andhare, and S. Shivendra, "Identification of mango variety using near infrared spectroscopy," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 3, p. 1776, Sep. 2023, doi: 10.11591/ijeecs.v31.i3.pp1776-1783.

[12]   C. Papadakis, Y. Yin, M. Danikas, and C. Charalambous, "Surface discharges and flashover voltages in Nanocomposite XLPE samples," *Engineering, Technology & Applied Science Research*, vol. 8, no. 6, pp. 3502–3504, Dec. 2018, doi: 10.48084/etasr.2186.

[13]   S. Wang, X. Shi, M. . Zhou, M. . Xu, Y. . Zhang, and H. Zhang, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 60, p. 101641, 2020.

[14]   H. R. Boveiri, R. Khayami, R. Javidan, and A. Mehdizadeh, "Medical image registration using deep neural networks: a comprehensive review," *Computers and Electrical Engineering*, vol. 87, 2020, doi: 10.1016/j.compeleceng.2020.106767.

[15]   X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: fast predictive image registration a deep learning approach," *NeuroImage*, vol. 158, pp. 378–396, 2017, doi: 10.1016/j.neuroimage.2017.07.008.

[16]   F. Zhu, M. Ding, and X. Zhang, "Self-similarity inspired local descriptor for non-rigid multi-modal image registration," *Information Sciences*, vol. 372, pp. 16–31, Dec. 2016, doi: 10.1016/j.ins.2016.08.031.

[17]   M. K. Tripathi and Shivendra, "Neutrosophic approach based intelligent system for automatic mango detection," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 41761–41783, Oct. 2023, doi: 10.1007/s11042-023-17037-7.

[18]   B. D. de Vos, F. F. Berendsen, M. A. Viergever, H. Sokooti, M. Staring, and I. Išgum, "A deep learning framework for unsupervised affine and deformable image registration," *Medical Image Analysis*, vol. 52, pp. 128–143, Feb. 2019, doi: 10.1016/j.media.2018.11.010.

[19]   S. Dasariraju, M. Huo, and S. McCalla, "Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm," *Bioengineering*, vol. 7, no. 4, p. 120, Oct. 2020, doi: 10.3390/bioengineering7040120.

[20]   V. Lavanya P, M. K. Tripathi, E. P. Hemand, S. K, and J. V. N. Ramesh, "Hyperspectral crop image classification via ensemble of classification model with optimal training," *Web Intelligence*, pp. 1–31, 2024, doi: 10.3233/web-230209.

[21]   A. T. Salawudeen *et al.*, "on the development of a novel smell agent optimization (SAO) for optimization problems," *i-manager's Journal on Pattern Recognition*, vol. 5, no. 4, p. 13, 2019, doi: 10.26634/jpr.5.4.15677.

[22]   G. Dhiman, M. Garg, A. Nagar, V. Kumar, and M. Dehghani, "A novel algorithm for global optimization: rat swarm optimizer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8457–8482, 2021, doi: 10.1007/s12652-020-02580-0.

[23] F. Zöllner, "Multimodal ground truth datasets for abdominal medical image registration [data]," *heiDATA*, 2022. https://heidata.uni-heidelberg.de/dataset.xhtml?persistentId=doi:10.11588/data/ICSFUS.

[24] M. K. Tripathi and Shivendra, "Improved deep belief network for estimating mango quality indices and grading: a computer vision-based neutrosophic approach," *Network: Computation in Neural Systems*, vol. 35, no. 3, pp. 249–277, Jul. 2024, doi: 10.1080/0954898X.2023.2299851.

[25] S. Albawi, O. Bayat, S. Al-Azawi, and O. N. Ucan, "Social touch gesture recognition using convolutional neural network," *Computational Intelligence and Neuroscience*, vol. 2018, 2018, doi: 10.1155/2018/6973103.

BIOGRAPHIES OF AUTHORS

Dr. Sanjeevkumar Angadi received Ph.D. from Visvesvaraya Technological University, Karnataka, India in 2023 and M. Tech in Computer Science and Engineering from Visvesvaraya Technological University, Karnataka, in 2012. At present he is working as associate professor head of Department in the Department of Computer Science and Engineering at PCET's NMVPM Nutan College of Engineering and Research, Talegaon Dabhade, Pune, India. His research interest includes computer vision, image processing, machine learning, data mining, and natural language processing. He can be contacted at email: angadi.sanjeevkumar@gmail.com.

Dr. Mukesh Kumar Tripathi received a Ph.D. in computer science and engineering from VTU, Belagavi. He has 15 years of teaching and administrative experience. He has supervised and co-supervised more than five masters and 20 B.E. students. He is working as an assistant professor with the Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India. He is authored or co-authored more than twenty publications and more than 400 citations. His research interests include soft computing, machine learning, intelligent systems, image processing, and hyperspectral. He can be contacted at email: mukeshtripathi016@gmail.com.

Dr. Chudaman Devidasrao Sukte received Ph.D. in Computer Science and Engineering from Dr Babasaheb Ambedkar Marathwada University Aurangabad, M.Tech. in Computer Engineering from Dr Babasaheb Ambedkar Technological University Lonere-Raigad and B.E. from SGGS Institute of Engineering and Technology, Nanded. He has 19 years of teaching experience. He has published 4 patents and more than 20 research papers in Journals and conferences. His area of research includes cloud computing, computer network and security, and machine learning. He can be contacted at email: rajeshsukte@gmail.com.

Dr. Shivendra received a Ph.D. in Electronics and Communication Engineering from JJT University, Rajasthan, India, in 2023. He received his M.Tech. degree in Electronics and Communication from Thrivalluavar Univesity Vellore Tamilnadu, India 2015. He received his B.E. degree in Electronics and Communication from GNDEC, Bidar, India 2011. He is working as an assistant professor in the Department of BCA at D.K. College, Dumraon, Buxar. He has supervised and co-supervised more than 20 students. He can be contacted at email: srivastavashivendra29@gmail.com.