The Fourteenth International Conference on Information, Intelligence, Systems and Applications

IISA2023

Deep Fuzzy Cognitive Map methodology for Non-Small Cell Lung Cancer diagnosis based on Positron Emission Tomography imaging

Anna Feleki, Ioannis D. Apostolopoulos, Elpiniki Papageorgiou, Nikolaos D. Papathanasiou, Dimitris J. Apostolopoulos, Nikolaos Papandrianos

Department of Energy Systems, University of Thessaly, Larisa, Greece

ΕΛ.Ι.Δ.Ε.Κ.
Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας

The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers" (Project Number: 3656).

UNIVERSITY OF THESSALY

EMERALD

Contents

- Introduction of the problem
- Related work
- Methodology
- Integration of clinical and image data
- Metrics & Results
- Produced interconnections
- Conclusions

Introduction of the problem

Non-Small Cell Lung Cancer (NSCLC)

- **Objective**: Classify 2D SPN(Solitary Pulmonary Nodule) image representations into benign and malignant for NSCLC data.
- **Motivation**: Early characterization of SPNs enables early treatment and can increase the survival rate.
- **Methodology**: Fuzzy Cognitive Maps, Particle Swarm Optimization and Deep learning image classification.

Related work

- In [1] Apostolopoulos et al:
  - Focused on the SPNs (Solitary Pulmonary Nodules) detection and developed VGG-19 reaching 84.3% accuracy in CT scans.
- In [2] Apostolopoulos et al:
  - Attained accuracy of 94% with VGG-16 with PET/CT data.
- In [3] Apostolopoulos et al:
  - Developed a 3D CNN, which achieved an accuracy of 89.68%.
- In [4] Salihoğlu et al:
  - Developed a DNN (Deep Neural Network) and an Extreme gradient boosting (XGB) algorithm to detect SPNs in PET/CT scans. XGB obtained 79% accuracy and CNN 80%.
- In [5] Shao et al:
  - Implemented a 3D-CNN based to distinguish benign lesions and invasive adenocarcinoma (IAC) in ground-glass nodules (GGNs) based on PET/CT scans.

[1]: I. D. Apostolopoulos, N. D. Papathanasiou, and G. S. Panayiotakis, "Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning," Biocybern. Biomed. Eng., vol. 41, no. 4, pp. 1243–1257, Oct. 2021, doi: 10.1016/j.bbe.2021.08.006.

[2]:I. D. Apostolopoulos et al., "Automatic classification of solitary pulmonary nodules in PET/CT imaging employing transfer learning techniques," Med. Biol. Eng. Comput., vol. 59, no. 6, pp. 1299–1310, Jun. 2021, doi: 10.1007/s11517-021-02378-y.

[3]: I. D. Apostolopoulos, D. J. Apostolopoulos, and G. S. Panayiotakis, "Solitary Pulmonary Nodule malignancy classification utilising 3D features and semi-supervised Deep Learning," in 2022 13th International Conference on Information, Intelligence, Systems & Applications (IISA), Jul. 2022, pp. 1–6. doi: 10.1109/IISA56318.2022.9904334.

[4] Y. S. Salihoğlu et al., "Diagnostic Performance of Machine Learning Models Based on 18F-FDG PET/CT Radiomic Features in the Classification of Solitary Pulmonary Nodules," Mol. Imaging Radionucl. Ther., vol. 31, no. 2, pp. 82–88, Jun. 2022, doi: 10.4274/mirt.galenos.2021.43760.

[5]: X. Shao et al., "Application of dual-stream 3D convolutional neural network based on 18F-FDG PET/CT in distinguishing benign and invasive adenocarcinoma in ground-glass lung nodules," EJNMMI Phys., vol. 8, no. 1, p. 74, Nov. 2021, doi: 10.1186/s40658-021-00423-1.

Dataset

132 cases of Malignant

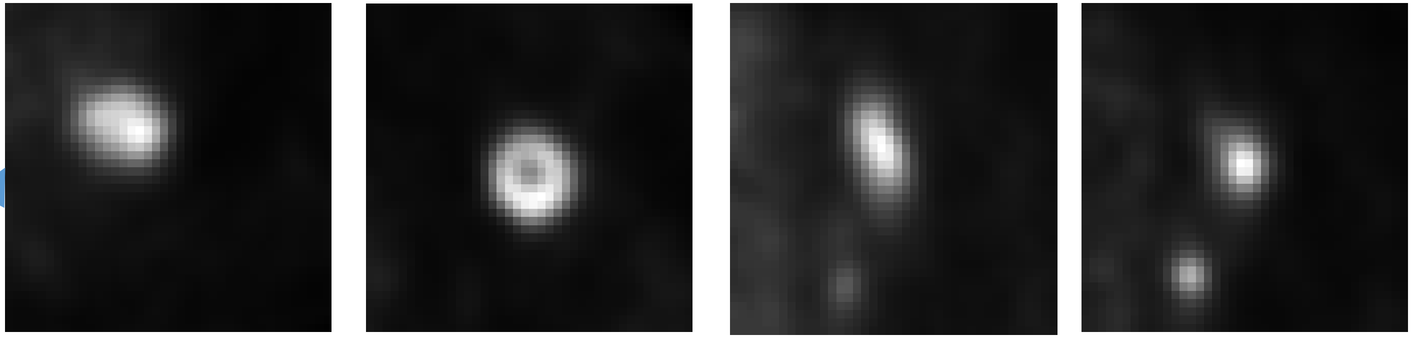
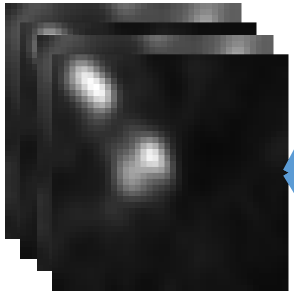
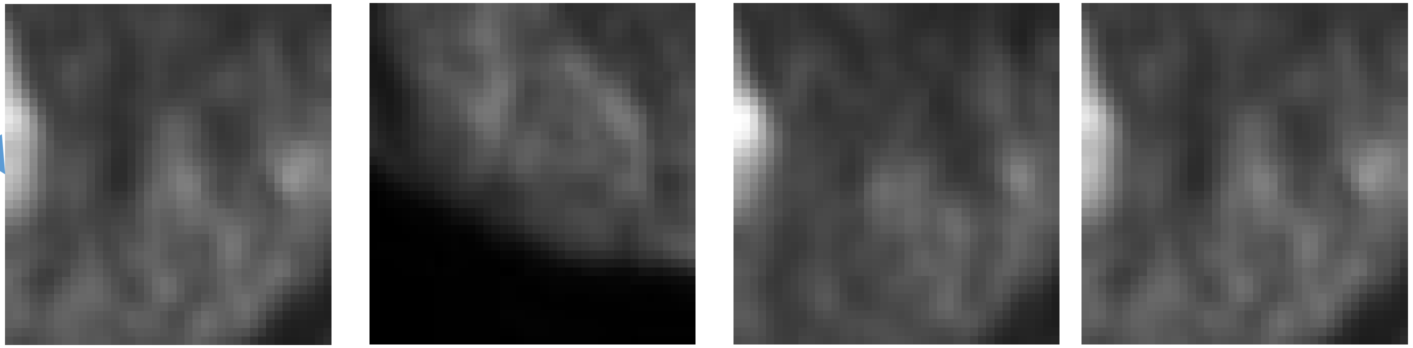


Image data

243 PET scans





111 cases of Benign

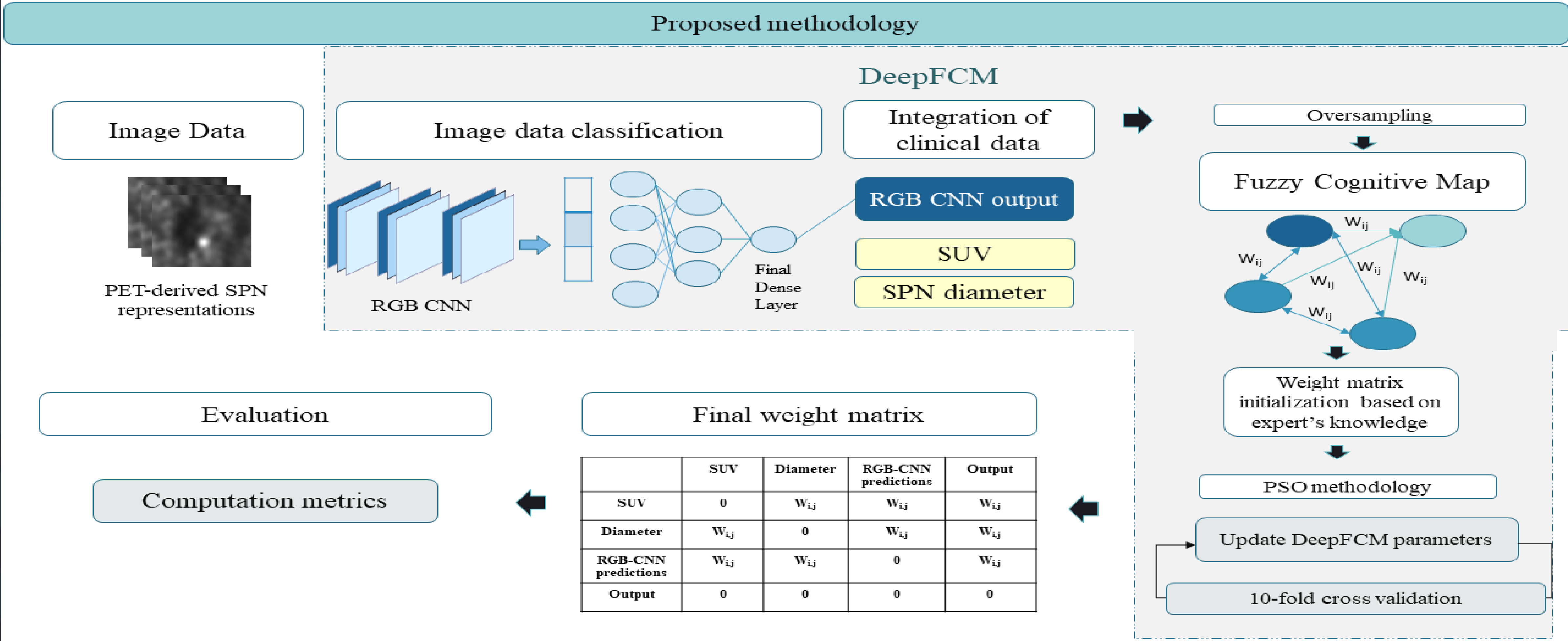
+

With the addition of:

Clinical data:
- ❑ SUVmax
- ❑ Diameter

Methodology

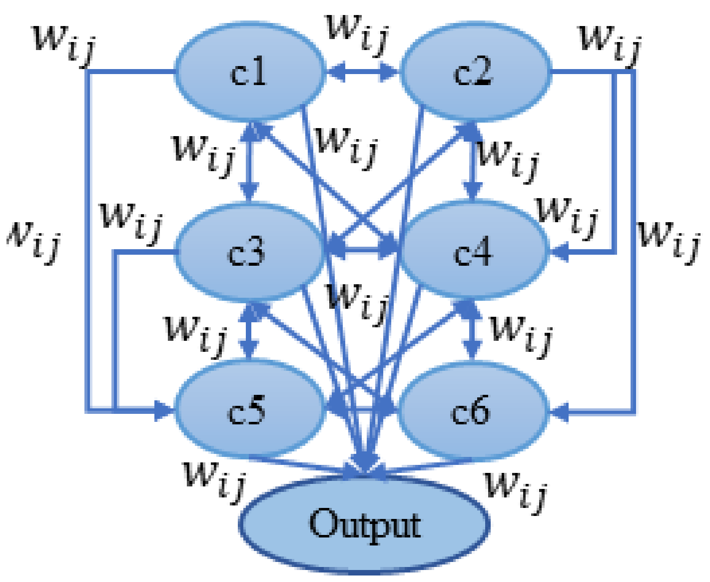
DeepFCM methodology



	SUV	Diameter	RGB-CNN predictions	Output
SUV	0	W_{ij}	W_{ij}	W_{ij}
Diameter	W_{ij}	0	W_{ij}	W_{ij}
RGB-CNN predictions	W_{ij}	W_{ij}	0	W_{ij}
Output	0	0	0	0

Fuzzy cognitive maps

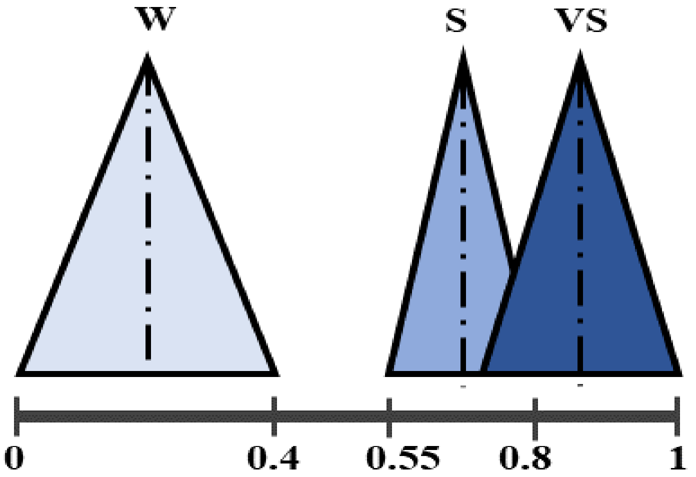
- Concepts in a Fuzzy Cognitive Map (FCM) are representations of the key factors or variables within the system that capture different aspects of the problem domain. These concepts are derived from a combination of historical data and expert knowledge.
- FCM transforms the input knowledge into casual relationships among concepts of the system.
- Regarding the interconnections of the system, they rely on the range [-1, 1], and whether an interconnection has a positive or negative, or zero value depends on the kind of connection.
    - $W_{ij} > 0$, expresses positive causality,
    - $W_{ij} < 0$, expresses negative causality.
    - $W_{ij} = 0$, expresses no causality.



	c1	c2	c3	c4	Output
c1	0	w_{ij}	0	w_{ij}	w_{ij}
c2	w_{ij}	0	w_{ij}	0	w_{ij}
c3	0	w_{ij}	0	0	w_{ij}
c4	w_{ij}	0	0	0	w_{ij}
Output	0	0	0	0	0

Experts' knowledge

- Nuclear experts were asked to describe the relationships between input concepts with the output and they provided linguistic values in a fuzzy set format.
- Fuzzy logic was employed to capture the vagueness and fuzziness of the relationships between the variables in the FCM. Traditional logic relies on binary values (true/false), while fuzzy logic extends this by allowing variables to have degrees of membership between 0 and 1. This flexibility enables the representation of uncertainty and imprecision in the clinical data, enhancing the accuracy of the classification process.
- For the interconnections, that experts did not provide the initial values were assigned randomly from the range [-1, 1].



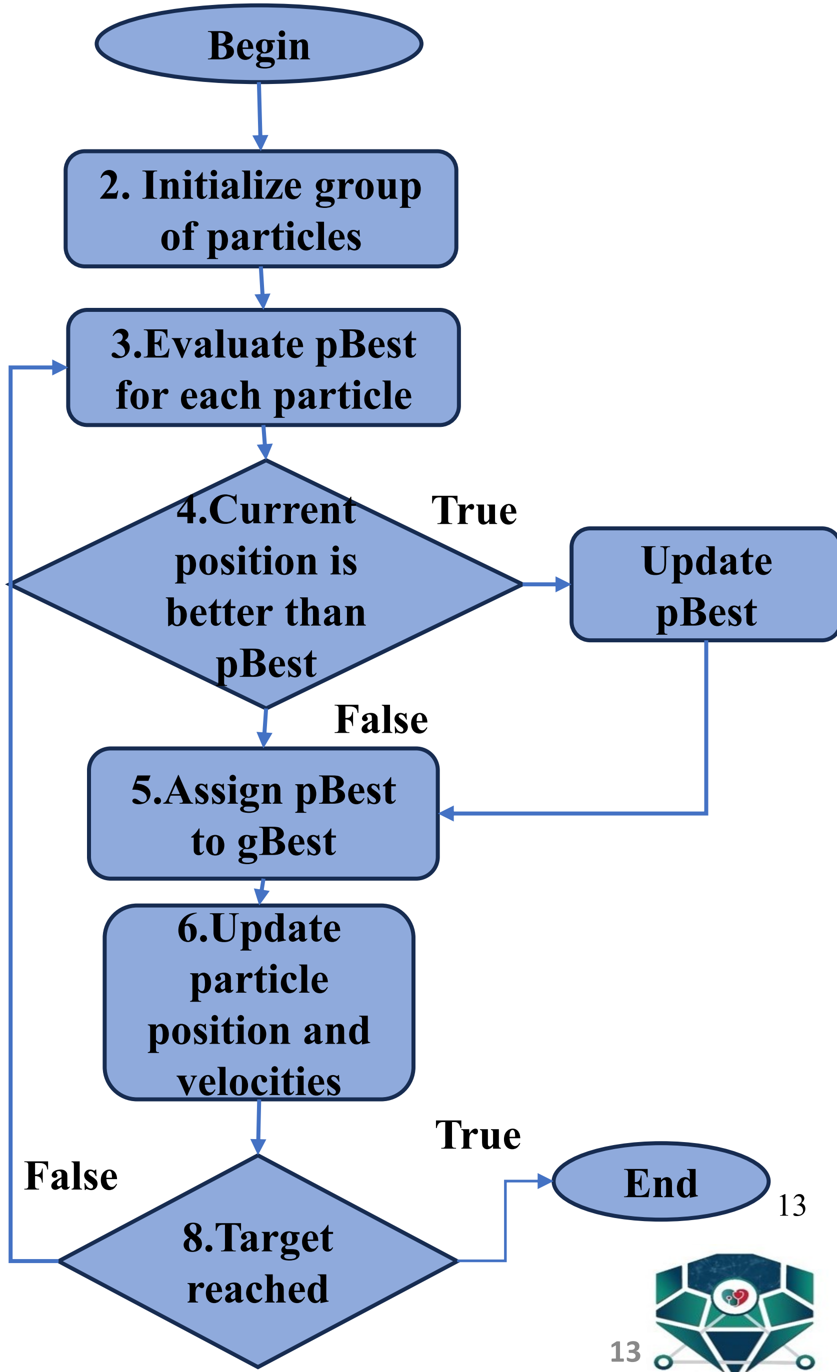
Input Concept	Suggested linguistic values assigned by experts	Linguistic values transformed to ranges
SUVmax->Output	Strong	[0.55 -0.8]
Diameter->Output	Weak	[0-0.4]
RGB-CNN predictions->Output	Very Strong	[0.7-1]

Particle swarm optimization

- Particle Swarm Optimization (PSO) is a computation method based on the social behavior of birds included in a flock. PSO extracts the optimal solution to the problem among many candidate solutions. In our case candidate solutions are the weight matrices.
- PSO calculations are based on the interconnection's values initialized by experts.
- PSO allows the FCM to adapt and optimize its interconnections based on the data. By iteratively adjusting the weights in the FCM, PSO can better represent the complex relationships between the concepts and improve its overall performance.
- PSO efficiently explores the solution space to find the optimal set of weights for the FCM. It leverages the collective behavior of particles in the swarm to search for promising regions in the solution space, guiding the FCM towards better solutions.

FCM-PSO

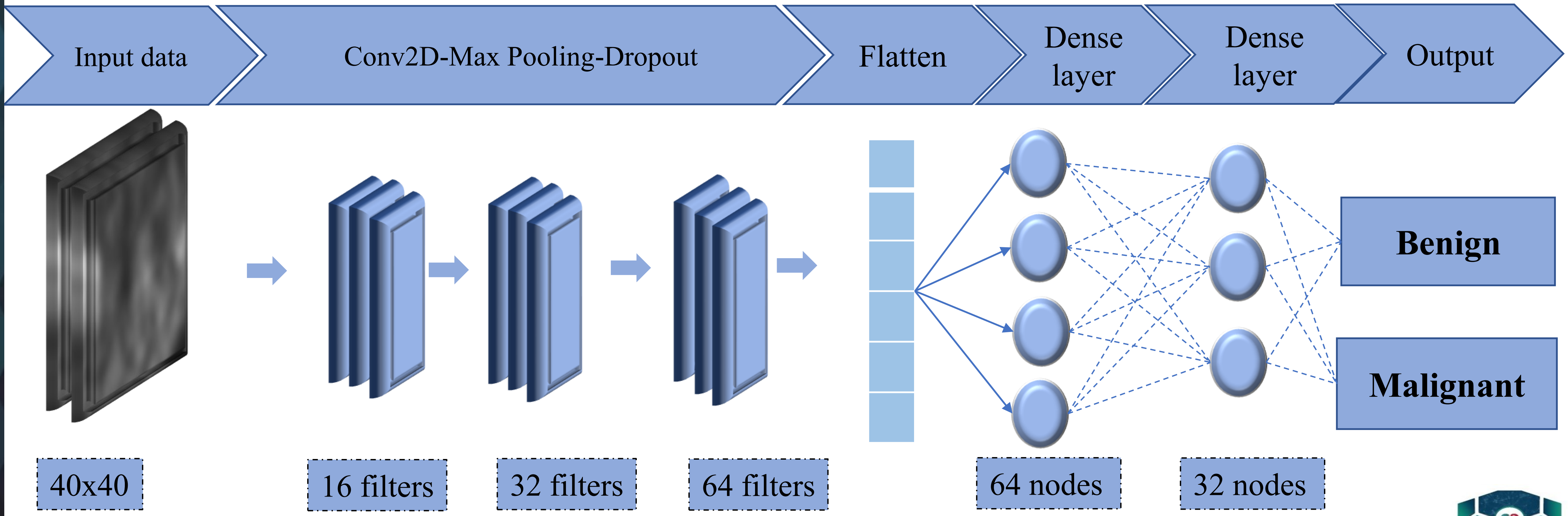
1. Definition of the FCM structure with nodes and initial values of interconnections, based on experts' knowledge.
2. Initialization of a population of particles representing potential solutions. Each particle represents a weight matrix.
3. Evaluation of the fitness function by calculating the error among the FCM predicted values and the output.
4. Update Personal Best: For each particle, the personal best fitness value is computed, and the weight matrix is stored.
5. Update Global Best: A comparison is conducted among the personal best fitness values of all particles and the global best fitness value is computed among all particles.
6. Update velocities and positions based on PSO mathematical equations to adjust the weights to global best.
7. Repeat steps 3 to 6 for a specified number of iterations or until FCM reaches convergence.
8. When the iterations are through or FCM terminates the weights correspond to the global best solutions are attained.
9. The final weight matrix is evaluated in the testing dataset.



RGB-CNN

Preprocessing of image data:

- Data shuffle
- Data split
- Data augmentation (Width_shift_range, height_shift_range,shear_range, zoom_range)

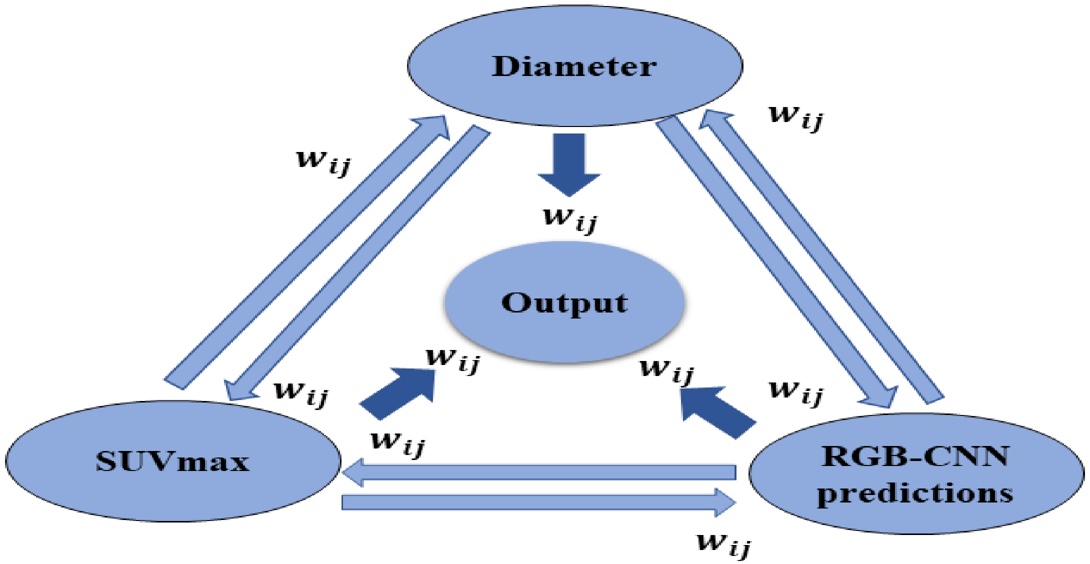


Integration of clinical and image data

Integration of clinical and image data

Integration approach includes:

- Fill missing values of clinical data.
- Each column of the clinical data represents an input concept.
- The extracted predictions from RGB-CNN are added as extra concept to the DeepFCM.
- Oversampling with sampling strategy all to generate data.
- One output concept to characterize the classification output



Metrics & results

Metrics

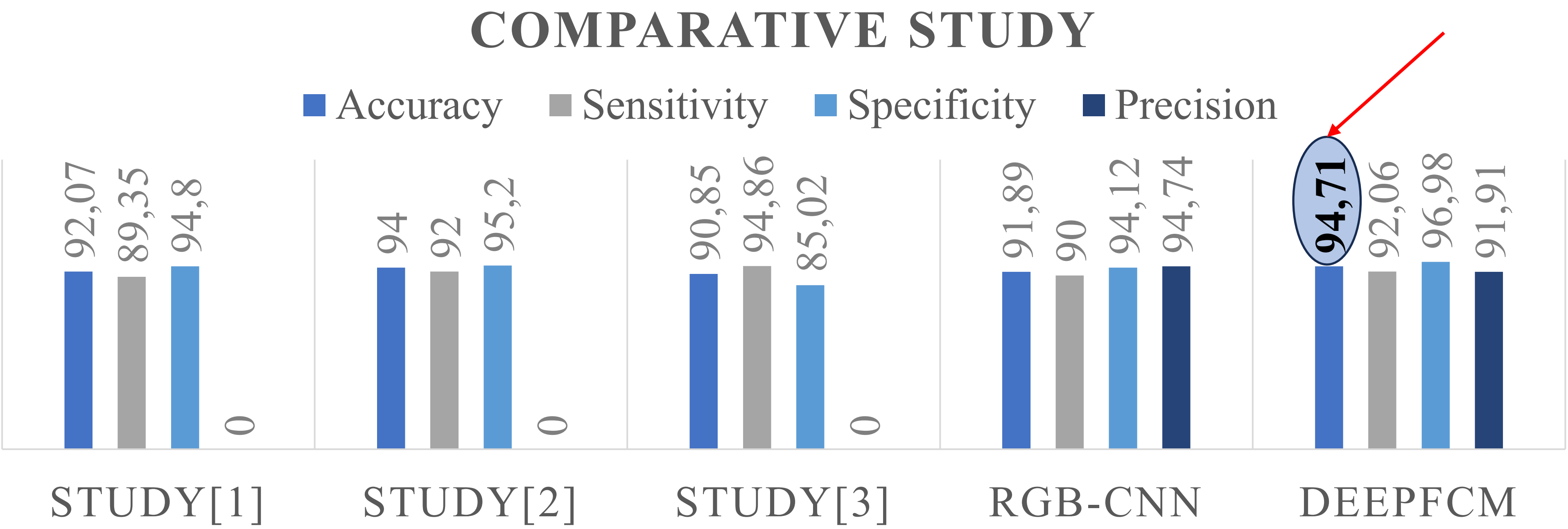
True Positives (TP): The number of positive instances correctly predicted as positive.
False Positives (FP): The number of negative instances incorrectly predicted as positive.
True Negatives (TN): The number of negative instances correctly predicted as negative.
False Negatives (FN): The number of positive instances incorrectly predicted as negative.

- **Accuracy**: Accuracy is a measure of how well a classification model correctly predicts the labels of the data points. It is calculated as the ratio of correctly predicted instances to the total number of instances.
- **Loss:** Loss is a value that quantifies the error between the predicted output and the actual target.
- **Sensitivity:** (Recall or True Positive Rate): Sensitivity measures the proportion of positive instances correctly identified by the model.
- **Specificity:** Specificity measures the proportion of negative instances correctly identified by the model.
- **Precision:** Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive by the model.

- Accuracy: $\frac{TP+TN}{TP+FP+TN+FN}$
- Sensitivity: $\frac{TP}{TP+FN}$
- Specificity: $\frac{TN}{TN+FP}$
- Precision: $\frac{TP}{TP+FP}$

Results

10-fold cross validation was applied to ensure DeepFCM's generalization to results.

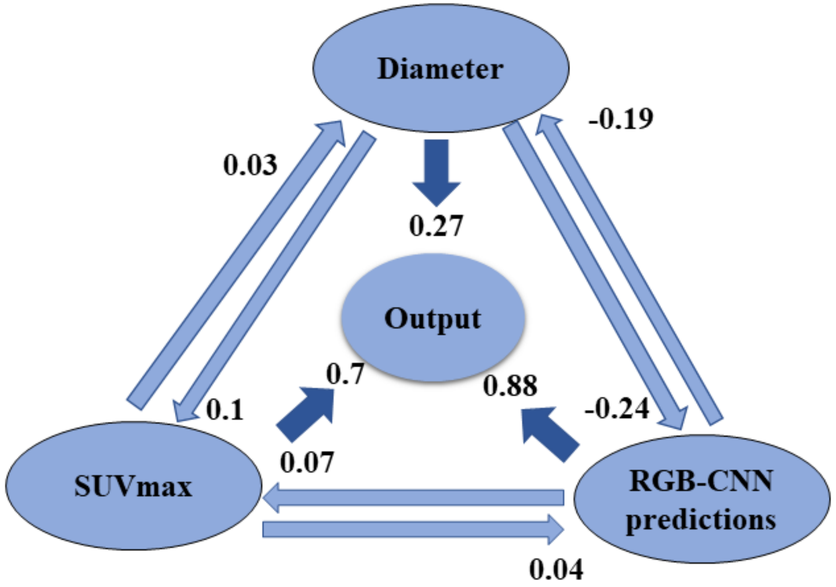


[1]: I. D. Apostolopoulos, N. D. Papathanasiou, and G. S. Panayiotakis, "Classification of lung nodule malignancy in computed tomography imaging utilising generative adversarial networks and semi-supervised transfer learning," Biocybern. Biomed. Eng., vol. 41, no. 4, pp. 1243–1257, Oct. 2021, doi: 10.1016/j.bbe.2021.08.006.
[2]:I. D. Apostolopoulos et al., "Automatic classification of solitary pulmonary nodules in PET/CT imaging employing transfer learning techniques," Med. Biol. Eng. Comput., vol. 59, no. 6, pp. 1299–1310, Jun. 2021, doi: 10.1007/s11517-021-02378-y.
[3]: I. D. Apostolopoulos, D. J. Apostolopoulos, and G. S. Panayiotakis, "Solitary Pulmonary Nodule malignancy classification utilising 3D features and semi-supervised Deep Learning," in 2022 13th International Conference on Information, Intelligence, Systems & Applications (IISA), Jul. 2022, pp. 1–6. doi: 10.1109/IISA56318.2022.9904334.

Produced DeepFCM model

- CNN prediction has the strongest relationship with the output, meaning that the output is highly dependent on the prediction value produced by the RGB-CNN.
- The SUVmax variable has a strong causality with the output, as well, meaning that SUVmax has a high impact on the classification outcome.
- Diameter has a weak relationship with the output, which indicates that the diameter by itself does not provide adequate information about the patient’s status regarding NSCLC.

DeepFCM's produced interconnections between input and output concepts are in agreement with the nuclear physicians.



Input Concept	Linguistic values transformed to ranges
SUVmax->Output	[0.55 -0.8]
Diameter->Output	[0-0.4]
RGB-CNN predictions->Output	[0.7-1]

Conclusions

Conclusion of DeepFCM implementation

- The research study demonstrates impressive outcomes using the DeepFCM model, which can serve as a comprehensive tool for supporting decision-making in nuclear medicine and especially SPN malignancy characterization for early detection of NSCLC using both image and clinical data.
- DeepFCM provides:
    - ➢ Integration of both clinical and imaging data.
    - ➢ Explainability with providing interconnections among concepts.
    - ➢ High performance metrics.

Future work

➢ Future work will involve integrating state equations for FCM learning.

➢ The integration of the CT image may provide better results, as it can capture more geometrical information regarding the suspicious SPNs.

Thank you for your time!
Any Questions ?