International Journal of Data Science and Artificial Intelligence (IJDSAI)
Volume 01, Issue 02, November – December (2023)

RESEARCH ARTICLE

DINGO OPTIMIZED FUZZY CNN TECHNIQUE FOR EFFICIENT PROTEIN STRUCTURE PREDICTION

P. G. Sreelekshmi [1, *] and S. C. Ramesh [2]

[1] Lecturer in Computer science, Department of BCA, University Institute of Technology, Malayinkeezhu Under University of Kerala, India
[2] Associate Professor, Department of Electronics and Communication Engineering, PSN College of Engineering and Technology, Tirunelveli, India.

[*]*Corresponding e-mail: Sreelekshmi.prf12@outlook.com*

Abstract – Protein is made up of a variety of molecules that are required by living organisms, such as enzymes, hormones, and antibodies. In step 2, the max-pooling layer and the convolutional layer evaluate the input data to create the finest feature map F1, which is half the image size in both horizontal and vertical directions. The full feature is then retrieved in step 2 using the max pooling layer and the residual block at the proper resolution. In this paper, we introduce Di-Fuzzy CNN (Fuzzy Convolutional Neural Network with Dingo optimizer), a novel technique for predicting protein activities that incorporates two types of information they are protein sequence and protein structure. We extract diverse features at different scales utilizing convolutional neural networks to provide comprehensive information for feature segmentation. To handle a variety of uncertainties in feature selection and produce segmentation results that are more dependable, fuzzy logic modules are employed. Finally, we employ Dingo optimization to boost the suggested method's effectiveness and speed in order to produce the best outcomes. Using a variety of datasets, the suggested model has been tested (HSSP, PDB, UGR14b, DSSP). Tests demonstrate that our approach can decrease FPR, increase protein structure accuracy, decrease prediction time, and increase TPR for feature selection. Our predictive model performs better than most state-of-the-art techniques.

Keywords *–Amino Acid Features, Protein Structure, Convolutional Neural Network (CNN), Fuzzy logic, Dingo Optimization algorithm.*

1. INTRODUCTION

Polymeric macromolecules known as proteins are made up of linear chains of amino acid building blocks connected by peptide bonds. Different biological mechanisms in living things produce proteins. [1] The structure of a protein and the chemical characteristics of its amino acids determine its activity. The genetic sequence of a protein can be used to determine its structure. Additional information about protein structure can be obtained by predicting the primary, quaternary, secondary, and tertiary structures [2]. Put differently, protein structure prediction refers to the process of estimating a protein's three-dimensional structure from its fundamental structure [3].

Protein structure prediction is the process of forecasting a protein's different amino acid sequences from its three-dimensional structure [4]. Its folding, secondary, tertiary, and quaternary structures may all be predicted from its fundamental structure. The issues of protein design and protein structure prediction are essentially different. One of the main objectives of theory and bioinformatics in medicine (e.g., drug design) is the prediction of protein structure.

The basic structure of the amino acid series is depicted in Figure 1. A protein's matching gene determines its main structure. α-helix could be the typical secondary structural state. Since hydrogen bonds develop within the chain, they are entropically benifical than beta sheets, despite the fact that their potential energy is not as low as that of beta sheets. The three-dimensional structure of a protein is mostly determined by the interactions between the R groups of the amino acids that make up the protein. Proteins must be sampled in various experimental conditions in order to identify the quaternary structure of proteins, which can be done using a range of experimental techniques [5], [18].

For the purpose of estimating the three-dimensional structure, protein structure prediction is crucial. There is a widespread misperception that it is hard to infer a protein's structure from its amino acid sequence since the amino acid sequence contains sufficient information to reveal a protein's three-dimensional structure. Extrapolating characteristics from amino acid sequences is a crucial step in increasing the precision of protein structure prediction [6, 7]. Utilizing sizable protein databases, protein structure prediction ascertains if a query sequence, in whole or in part, resembles a known structure [8], [19].

Finding a protein's structure from a collection of amino acids is a difficult task in molecular biology and bioinformatics. Many studies use different data mining

techniques to predict the structure of proteins. Nevertheless, the computing time and forecast accuracy of earlier methods were inadequate [20]. For over ten years, protein structures have been predicted through the application of neural networks. Inspired by neural networks' recent success, DL networks have been utilized in several articles to predict protein shapes. We suggested a Di-Fuzzy CNN to improve the accuracy and speed of protein structure prediction in order to solve the current problems [21-24].

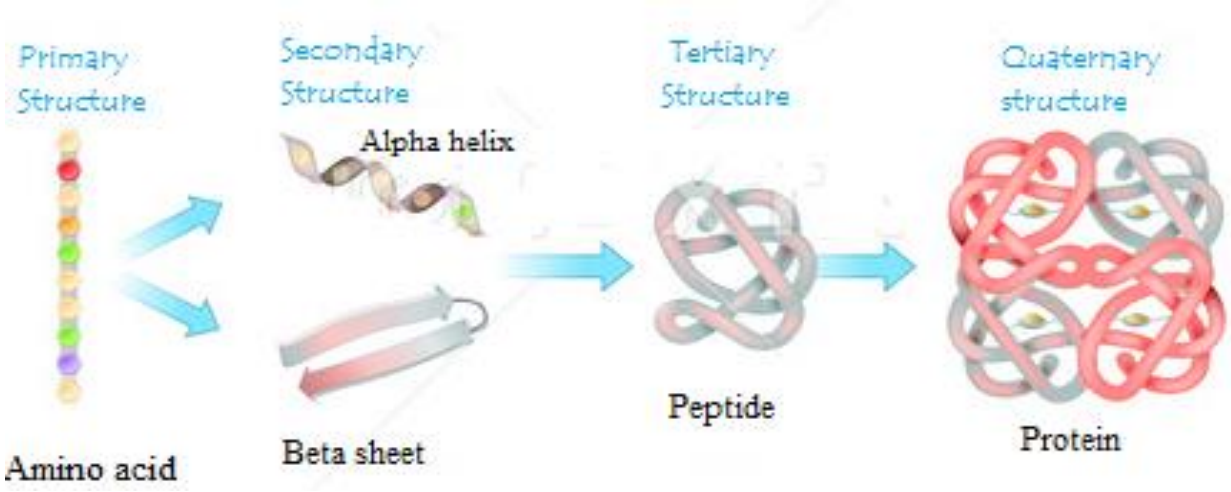

Figure 1. Resultant Graph of the Proposed System

The creation and application of protein structure predictions is explained in this article. Section II discusses related studies on protein structures that make use of different deep learning approaches. Section III presents the proposed Di-Fuzzy CNN (Dingo Fuzzy Convolutional Neural Network), along with a description and associated algorithms. Section IV contains performance results and associated analysis. Section V concludes with additional work and conclusions.

2. LITERATURE SURVEY

In 2020 Giri, et al [9] proposed, a MultiPredGO a novel multimodal technique for predicting protein functions using two separate types of information like protein secondary structure and protein sequence. We evaluated our findings to a variety of unimodal approaches in addition to two multi-modal protein function prediction methods like DeepGO and INGA. Our proposed method achieves better score. For cellular component and molecular comparisons, there were 13.05 percent and 30.87 percent enhancements over the best available comparing method.

In 2020 Zhou et al., [10] proposed, a monitored learning technique known as combining deep neural networks (CDNN) for protein structure prediction. The RMSProp optimizer uses the crossentropy error to educate the CDNN architecture. With multiple CNNs, the suggested design may remove amino acid data and mix them with raw features to train massive LSTM networks. In comparison to other approaches, the output suggest that the suggested techniques can achieve reasonable enactment with the suitable parameters.

In 2020 Bingzhen, et al., [11] suggested utilizing a confusion matrix to choose a random forest classification model. Using the "remove poor models" strategy, the forest models are selected at random. In three various data sets, the new outcomes reveal that the new technique has greater average classification accuracy and stability than the prior iteration. As a result, the confusion matrix-based random forest image classification model can increase random forest classification ability.

In 2019 Akter, and Holder, [12] proposed, a graphical feature-based framework that derives graphical features from sensor network data and employs feature selection approaches to choose more valuable features for such a classifier to have in prediction problems. Using movement data from smart home motion sensors and mobile phone GPS sensor information, as well as demographic data from smartphone GPS sensor data, the researchers used the suggested approach to forecast activity. additionally forecast. We discovered that when non-graph-based features are added to graphical feature-based frameworks, the outcomes improve.

In 2019 Gao et al., [13] proposed, A novel method for predicting equilibrium contacts in proteins is called Deep Structural Inference for Proteins (DESTINI), which blends template-based structural models with DL algorithms. DESTINI accurately predicts the tertiary structure 4 times for "hard" targets while simultaneously improving model quality for "easy" targets. DESTINI's much improved performance is partly due to the introduction of improved contact prediction template model. This paper outlines a viable technique for resolving the prediction problem in protein structure.

In 2018 Zamil, and Rahman, [14] proposed a multiscale local descriptor (MLD) to retrieve multiscale local information by feature extraction from a set of proteins. Decision trees, random forests, and bootstrap aggregation are used in classification approaches. Several algorithms produce diverse outcomes, and ensemble classification performs more accurately than current techniques. It is discovered that random forests and bootstrap aggregation are at least 10% more accurate than decision tree methods.

In 2018 Yavuz, et al., [15] proposed, An MLP methodology for predicting the secondary structure of proteins. The amino acid sequence was used to estimate protein secondary structure. There are two steps to the classification: MLP and direct MLP with CSA enhancement. The success rate for direct MLP categorization is 84.01 percent. For various numbers of rounds and hidden layers, the MLP with CSA arrangement achievement is investigated. To summarise, using CSA prior to categorization is advised for better prediction accuracy.

In 2018 Xie, et al., [16] proposed, A fuzzy support vector machine for secondary structural identification for predicting the amino acid features. Agreeing to the K-nearest neighbour algorithm, hyperplanes are assigned big membership values, whereas outliers are assigned small membership values. To test this strategy, we employed 3 databanks (e.g., CB513, data11996and RS12). Overall, our results for secondary structure prediction are better than regularly used approaches.

In 2017 Wang et al., [17] proposed, A deep recurring encoding-decoder network, secondary structure recurrent encoder. suggested using a decoder network. The CB513 and CullPDB public datasets are used to test the suggested model. Especially well-suited to modeling sequence and structural links between input protein attributes and secondary structure are encoder/decoder designs used in conjunction with GRUs. It also performs better than the opposition in terms of Q8 and Q3 accuracy. We outperformed earlier methods in predicting Q3 and Q8 with 68.20 percent and 73.1 percent accuracy in fewer epochs on the CB513 and CullPDB datasets.

In 2017 Liu, et al., [18] proposed, two-dimensional deep convolutional neural network, the protein's secondary structure was predicted. Based on a two-dimensional input matrix, two-dimensional CNNs are better at extracting sequence interaction features and storing unique amino acid position data. Our predictive model performs better than most state-of-the-art techniques.

3. PROPOSED METHOD

A crucial step in theoretical chemistry and bioinformatics is the prediction of protein structures. Because predicting protein structure is a critical procedure in medicine (for example, medication design) and biotechnology. A large number of research projects are currently underway with the goal of determining the protein structure using various classification techniques. However, existing categorization techniques performed poorly. The Di-Fuzzy CNN Technique was created to address these restrictions.

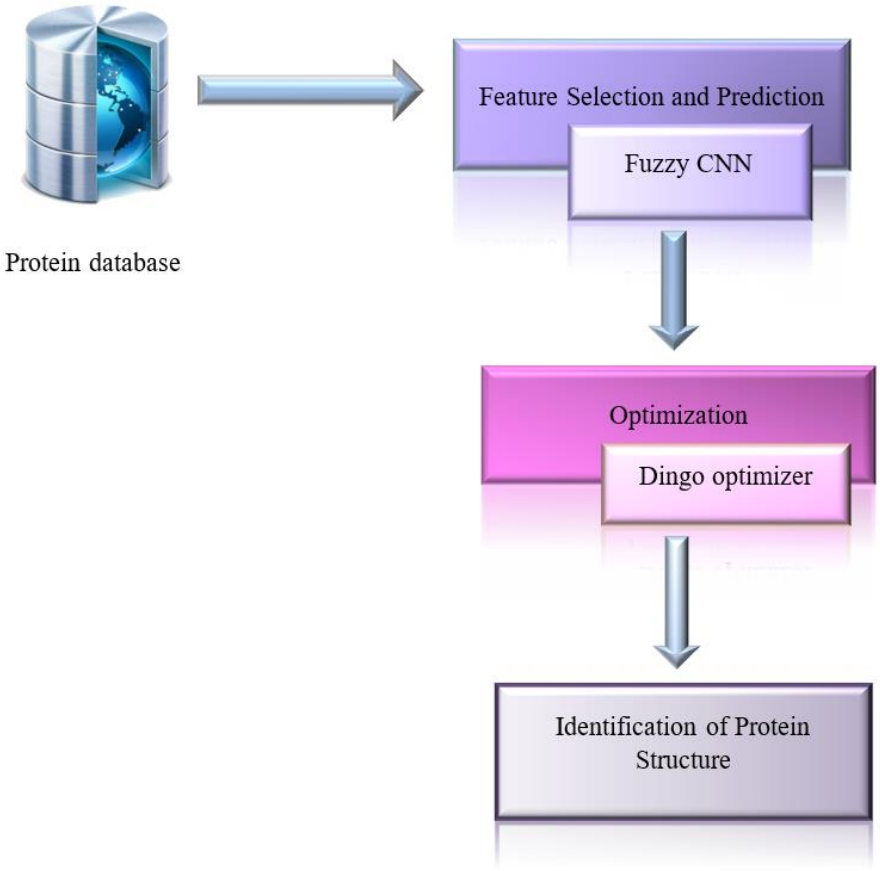

Figure 2. Proposed Methodology

Figure 2 depicts the overall process of using the Di-Fuzzy CNN technique to predict protein structure more efficiently and in less time. In the above figure, Fuzzy CNN technique is used to identify relevant amino acid characteristics from a large protein dataset and predict the protein structure. The dingo optimizer algorithm improves the accuracy and reduces the computational time as a result.

3.1. Feature Selection and Prediction

Convolutional neural networks and fuzzy learning are combined to form a new neural network structure for selection and prediction called as Fuzzy Convolutional Neural Network (Fuzzy CNN). The proposed technique uses fuzzy learning to deal with protein structure uncertainty, and the fuzzy logical units are smoothly combined into the Neural Network. Each fuzzy logical units are guided by a feature map at a specified balance, with the goal of establishing a link between the segmentation and the features outcome. A robust and accurate outcome can be obtained by taking into account, the result of the fuzzy logical units at multiple stages. The settings for the convolutional network and fuzzy logical units are integrated after end-to-end learning using training samples with physically labelled protein structures.

CNNs have gained popularity recently in the domains of bioinformatics and allied ones. Three principals have drawn researchers' attention to convolutional neural networks: shared weights, spatial subsampling, and local receptive fields. Convolutional neural networks perform several functions, including scale distortion and covariance shift, to varying degrees. Because of these characteristics, convolutional neural networks are frequently employed in research domains including segmentation and prediction. Genetic diversity is assumed to originate from proteins.

To forecast the secondary structure of proteins, features are extracted from sets of amino acids. CNN is unquestionably one of the greatest options for prediction when combined with its capacity to handle massive volumes of training samples. CNN shortens computation times while simultaneously capturing information from a large number of samples related to proteins. CNN Architecture for protein structure prediction shown in Figure 3.

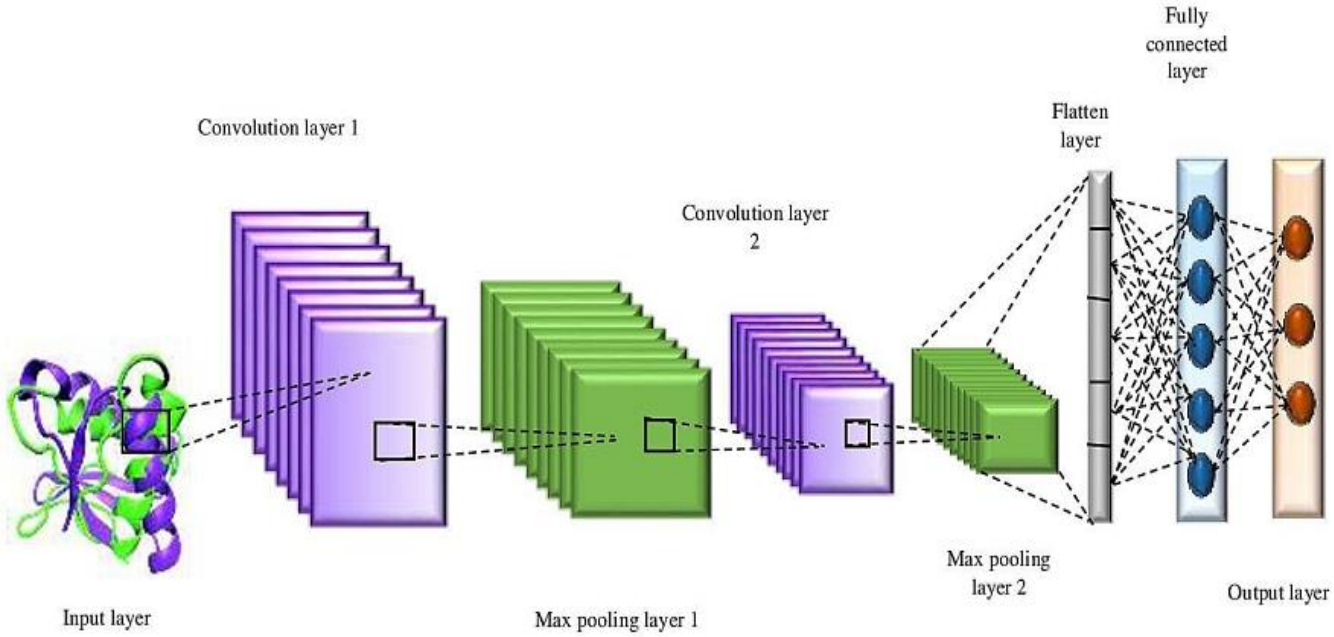

Figure 3. CNN Architecture for protein structure prediction

Step 2 involves an inspection of the input data by the convolution and max-pooling layers, which then produce the finest feature map F1, which is half the size of the image in both horizontal and vertical dimensions. Then, in step 2, a max-pooling layer and a residual block are used to extract all of the features at the appropriate resolution. Two max-pooling layer steps come after the remaining blocks. There is utilization of fully bonded layers. The mapping of representations between inputs and outputs is aided by the FC layer.

To build the finest feature for protein structure prediction, it selects partially the size in both the vertical and horizontal sides from the given input and processes it run over 2 convolutional layers first, then using a max-pooling level at a pace of two steps. The entire set of features in the matching resolution are then extracted using a remaining block is followed by a stride of two max-pooling layers. Fully connected layers are used. The FC layer assists in the mapping of representations between input and output.

Every membership function labels the feature points with a fuzzy linguistic term, and all of the Gaussian function membership are provided as

$$V_{x,y,k,c} = e^{-(\frac{k_{x,y,c}-\mu_{k,c}}{\sigma_{k,c}}).^2}, x = 1 \ldots W, y = 1 \ldots W, k = 1..M \ldots \quad (1)$$

Where,

H, W The feature's height and width.

M Membership function applied for each function.

(x, y) coordinate feature point of $F_{x,y,c}$

$\mu_{k,c}$ and $\sigma_{k,c}$ The gaussian member function's mean and standard deviation

$V_{x,y,k,c}$ In channel c, the fuzzy logic feature's k-th output (x, y).

3.2. Optimization

For better accuracy in Fuzzy CNN, the Dingo Optimization Algorithm is used. This optimization uses a unique technique to tackle specific problems by altering the network's variables and derivatives to find the best solution. DOA's main concept is as follows: a rapid sequence to initialise the speed and position of the search agent, consequently increasing the rate of search agents, generating a huge numbers of search agents, and eventually finding the best agent. To achieve the best results, the algorithm employs 3 strategies to optimise the categorised output.

The first strategy is encircling, the search agents (Dingo) often seek nominal objectives while alone, but create groups while hunting substantial goals. Flow chart for Dingo Optimization Algorithm shown in Figure 4

$$\overrightarrow{m_i}(p+1) = \beta_1 \sum_{a=0}^{N} \frac{[\overrightarrow{\varphi a(p)} - \overrightarrow{m_i}(p)]}{N} - \overrightarrow{m_*}(p) \quad (2)$$

Where, $\vec{m}(p+1)$ search agent's new position.

$\overrightarrow{m_*}(p)$ Best search agent.

$\overrightarrow{m_i}(p)$ Current search agent.

N Random Integer Numbers.

The second strategy is persecution, in which the search agent pursues the small prey separately until it is trapped.

$$\overrightarrow{m_i}(p+1) = \overrightarrow{m_i}(p) + \beta_1 * e^{\beta_2} * \left(\overrightarrow{m_s}(p) - \overrightarrow{m_i}(p)\right) \quad (3)$$

Where, $\vec{m}(p+1)$ dingo movement

s From 1 to the maximum size, a random number will be created.

The third strategy is scavenger, the search agent comes finds carrion to eat, they engage in scavenging activity while moving around their habitat at random, and then the fitness is determined.

$$\overrightarrow{m_i}(p+1) = \frac{1}{2}[e^{\beta_2} * \overrightarrow{m_s}(p) - (-1)^{\sigma} * \overrightarrow{m_i}(p)] \quad (4)$$

In addition, with 3 strategies, the chances of dingoes surviving are taken into account.

$$SR(i) = \frac{fitness() - fitness(i)}{fitness() - fitness(min)} \quad (5)$$

Where, *fitness (max)* is the greatest fitness

fitness (min) is the inferior fitness ratios

The low survival rate is given by,

$$\overrightarrow{m_i}(p) = \overrightarrow{m_*}(p) + \frac{1}{2}[\overrightarrow{m_s}(p) - (-1)^{\sigma} * \overrightarrow{m_t}(p)] \quad (6)$$

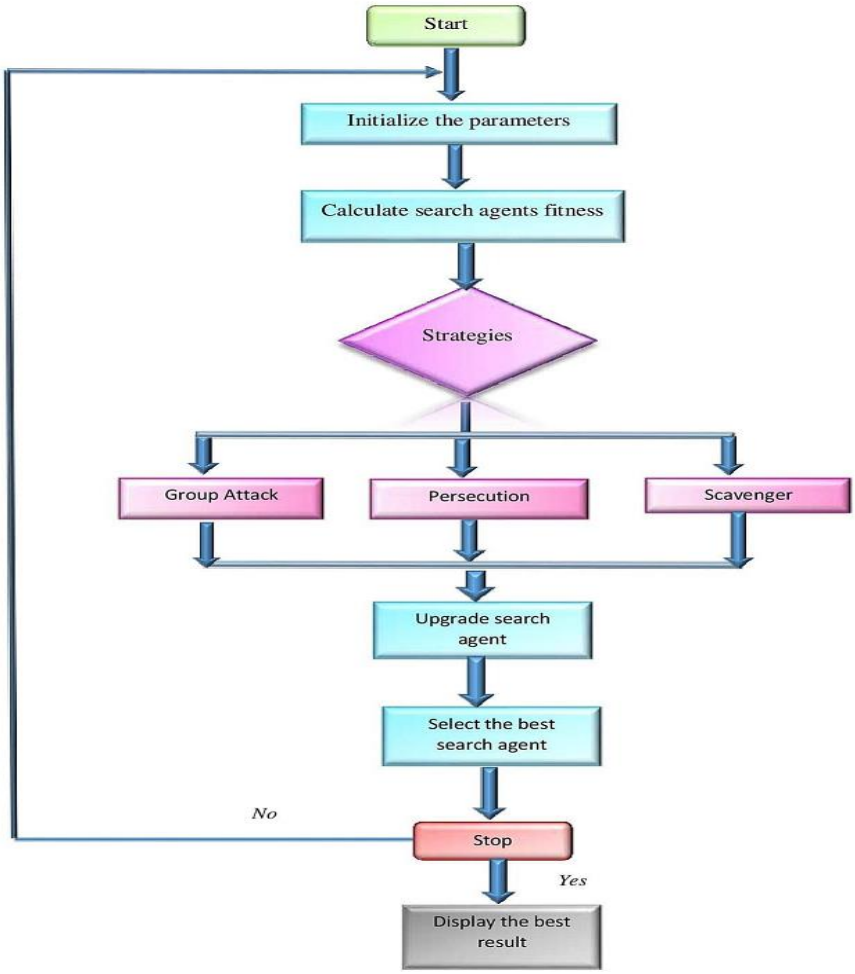

Figure 4. Flow chart for Dingo Optimization Algorithm

4. RESULTS AND DISCUSSIONS

The effectiveness of the Di-Fuzzy CNN technique is demonstrated in FPR (false positive rate), TPR (true positive rate), and is measured at the level of utilizing several datasets, including HSSP, PDB (Protein Data Bank), UGR14b, and DSSP (Secondary Structure Protein Data Bank). Protein structure prediction rate, PSPA (protein structure prediction accuracy), and PSPT (protein structure prediction time) are examined using tables and graphs in this section.

4.1. True Positive Rate experiment results

The amount of amino acid characteristics that are accurately chosen as related to the total amount of features are calculated in TPR. It is employed in the prediction of protein structure for feature selection. TPR is expressed as a % and can be computed using the equation below.

$$TPR = \frac{Amount\ of\ features\ selected}{Total\ amount\ of\ features} * 100 \quad (7)$$

Table 1. The Result for TPR

Dataset	True positive Rate in percentage			
	BFO	PROTEUS	WPC - IRFC	Di-Fuzzy CNN
HSSP	82	83	92	94
PDB	84	87	95	97
UGR14b	80	82.7	90	93.4
DSSP	80.6	82	86	90

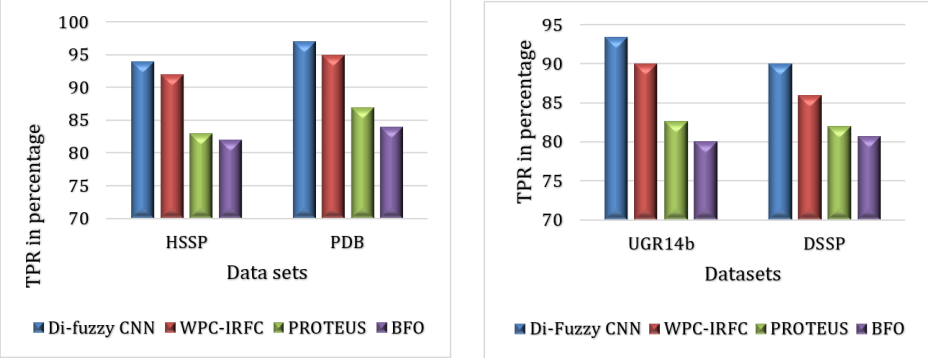

Figure 5. Graphical representation for TPR using different techniques and datasets

The outcomes of TPR in Di-Fuzzy CNN methods are compared to current methods such as WPC-IRFC, PROTEUS, and BFO for estimating protein structures from large protein datasets. When employing 100-550 amino acid characteristics from the PDB dataset in a study, the Di-Fuzzy CNN approach achieves a TPR of 97 percent, whereas previous methods reach only 95 percent, 87 percent, and 84 percent. Therefore, the proposed method outperforms existing methods. Graphical representation for TPR using different techniques and datasets shown in Figure 5. The Result for TPR shows in Table 1.

4.2. Rate of False Positives

The potential of mistakenly rejecting the null hypothesis for a test when making many comparisons is known as a false positive ratio. The amount of amino acid characteristics that are wrongly identified as significant to the total no. of features obtained as input is calculated. By using proposed technique FPR is low. Graphical representation for FPR using different techniques and datasets Figure 6. Comparison table for FPR Shows in Table 2

$$FPR = \frac{the\ no.\ of\ features\ that\ were\ incorrectly\ picked}{Total\ number\ of\ Features} * 100 \quad (8)$$

Table 2. Comparison table for FPR

Dataset	False Positive Rate in percentage			
	BFO	PROTEUS	WPC-IRFC	Di-Fuzzy CNN
HSSP	17.4	17	7	6
PDB	16	13	5	3
UGR14b	15	12	6	5
DSSP	16.7	14	9	7

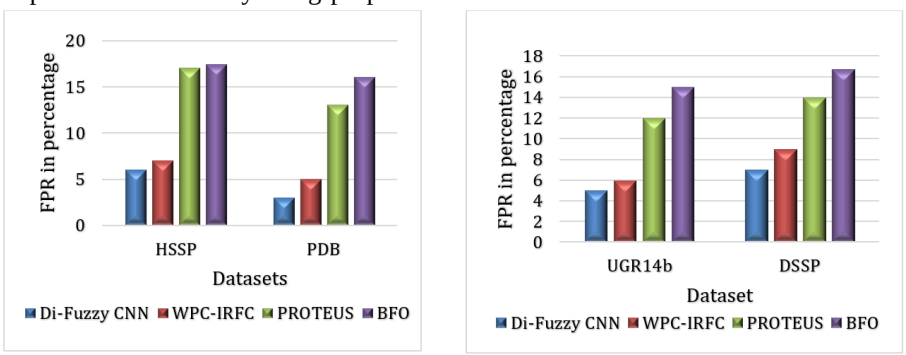

Figure 6. Graphical representation for FPR using different techniques and datasets

4.3. Structure of Proteins Experiment on prediction accuracy

It can be defined as the proportion of the total amount of protein structures that can be accurately predicted using particular features of amino acids.

$$PSPA = \frac{amount\ of\ properly\ predicted\ protein\ structure}{Total\ amount\ of\ protein\ structures} * 100 \quad (9)$$

By using Di-Fuzzy CNN methods the results of PAPR is compared with different existing techniques such as WPC-IRFC, PROTEUS, and BFO for accuracy.

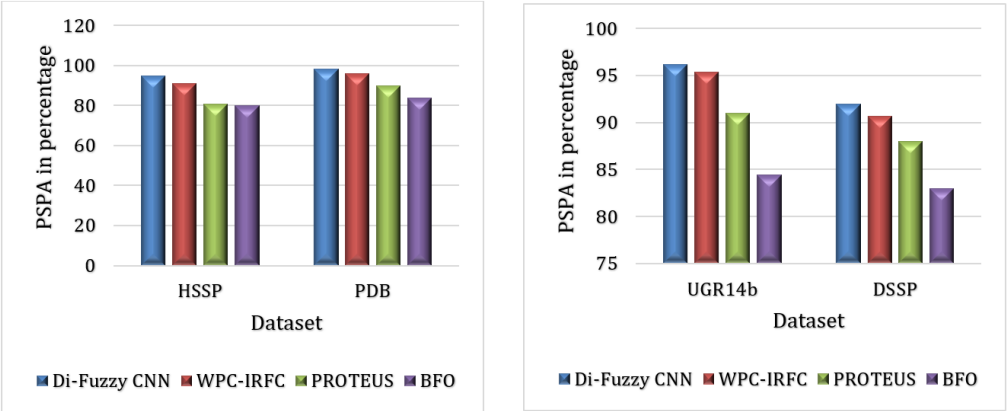

Figure 7. PSPAs graphical representation using different techniques and datasets

Table 3. Accuracy for predicting protein structure

Dataset	Accuracy rate in%			
	BFO	PROTEUS	WPC-IRFC	Di-Fuzzy CNN
HSSP	80	81	91	93
PDB	84	90	96	98
UGR14b	84.5	91	95.4	96.2
DSSP	83	88	90.7	92

PSPAs graphical representation using different techniques and datasets shows in Figure 7. Accuracy for predicting protein structure shows in Table 3.

By using PDB dataset the proposed technique achieves 98.2 percent, while previous methods reach only 96%, 90%, 84%. Therefore, the proposed method gives more accuracy compared to others.

4.4. Protein structure prediction time

PSPT is a metric that evaluates how long it takes to categorize the structure of protein from a large protein dataset file. It is measured in milliseconds (ms). Table for predicting time shown in Table 4.

$$PSPT = N * time(predicting\ the\ protein\ structure) \quad (10)$$

Table 4. Table for predicting time

Dataset	Time taken for predicting protein structure in ms			
	BFO	PROTEUS	WPC-IRFC	Di-Fuzzy CNN
HSSP	28	27	16	14
PDB	25	26	13	11
UGR14b	24	22	19	16
DSSP	23	22.8	20	16.4

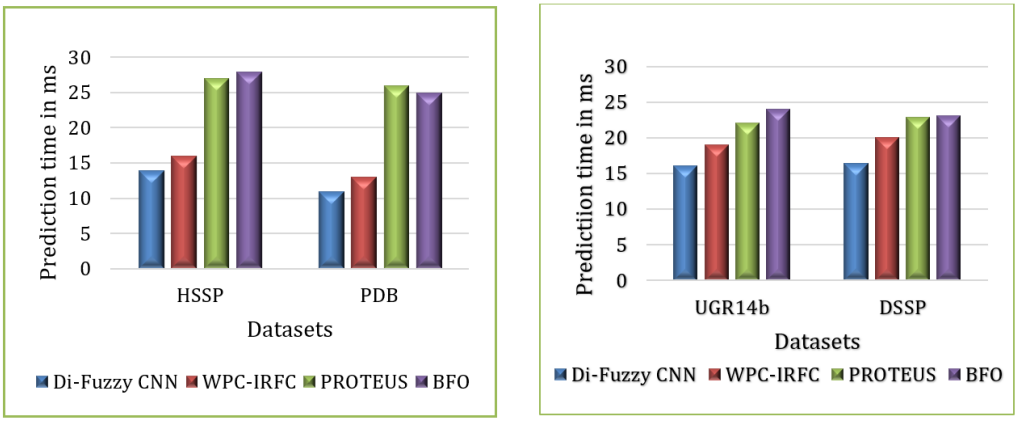

Figure 8. PSPTs graphical representation using different techniques and datasets

PSPTs graphical representation using different techniques and datasets Figure 8. The outcomes of PSPT in Di-Fuzzy CNN methods are compared to current methods such as WPC-IRFC, PROTEUS, and BFO for estimating

protein structures from large protein datasets. When employing 100-550 amino acid characteristics from the PDB dataset in a study, the Di-Fuzzy CNN decrease the prediction time as 11 milliseconds, whereas previous methods reach only in 13 ms, 26 ms, 25 ms respectively. Therefore, the proposed method outperforms existing methods.

5. CONCLUSION

In this study, amino acid sequences were used to deduce the structure of proteins. In this article, we introduce a Di-Fizzy CNN that can accurately predict proteins from amino acids. It is proposed to learn high-level semantics through end-to-end supervised learning of convolutional structures in neural networks combined with fuzzy logic. Fuzzy allows convolutional neural networks to focus more on protein structure. The proposed method is tested on a large-scale protein dataset and compared using metrics such as TPR, FPR, PSPA, and PSPT. Compared with state-of-the-art features, Di-Fuzzy CNN result analysis achieves higher performance in terms of PSPA and PSPT, resulting in effective disease diagnosis. Future research should expand the dataset size and investigate the accuracy rates of deep learning models (e.g., RNNs, capsule networks) used to extract features from protein datasets. Explore spatial complexity and predict protein structure using large datasets.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

FUNDING STATEMENT

Not applicable

ACKNOWLEDGEMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

REFERENCES

[1] M.N. Mounika and M.N. Chava, "Protein Structure Prediction of Stroke Disease Using Artificial Neural Networks", *Solid State Technology*, vol. 63, no. 2s, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Y.S. Ding, T.L. Zhang and K.C. Chou, "Prediction of protein structure classes with pseudo amino acid composition and fuzzy support vector machine network", *Protein and peptide letters*, vol. 14, no. 8, pp. 811-815, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[3] J. Xu, M. Mcpartlon and J. Li, "Improved protein structure prediction by deep learning irrespective of co-evolution information", *Nature Machine Intelligence*, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] H.K. Phan and T.H. Dang, "Protein structure prediction using Deep Learning", *VNU University of Engineering and Technology*, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] A.K. Mandle, P. Jain and S.K. Shrivastava, "Protein structure prediction using support vector machine", *International Journal on Soft Computing*, vol. 3, no. 1, pp. 67, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[6] S.C. Pakhrin, B. Shrestha, B. Adhikari and D.B. Kc, "Deep Learning-Based Advances in Protein Structure Prediction", *International Journal of Molecular Sciences*, vol. 22, no. 11, p.5553, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko and A. Bridgland, "Highly accurate protein structure prediction with AlphaFold", *Nature*, vol. 596, no. 7873, pp. 583-589, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] S. Montgomerie, S. Sundararaj, W.J. Gallin and D.S. Wishart, "Improving the accuracy of protein secondary structure prediction using structural alignment", *BMC bioinformatics*, vol. 7, no. 1, pp. 1-13, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[9] S.J. Giri, P. Dutta, P. Halani and S. Saha, "MultiPredGO: Deep multi-modal protein function prediction by amalgamating protein structure, sequence, and interaction information", *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1832-1838, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] S. Zhou, H. Zou, C. Liu, M. Zang and T. Liu, "Combining deep neural networks for protein secondary structure prediction", *IEEE Access*, vol. 8, pp. 84362-84370, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Z. Bingzhen, Q. Xiaoming, Y. Hemeng and Z. Zhubo, "A Random Forest Classification Model for Transmission Line Image Processing", *In 2020 15th International Conference on Computer Science & Education (ICCSE) IEEE*, pp. 613-617. 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] S. Akter and L. Holder, "Improving IoT Predictions through the Identification of Graphical Features", *Sensors*, vol. 19, no. 15, p.3250, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] M. Gao, H. Zhou and J. Skolnick, "DESTINI: A deep-learning approach to contact-driven protein structure prediction", *Scientific reports*, vol. 9, no. 1, pp. 1-13, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] K.S. Zamil and J. Rahman, "Prediction of Protein-Protein Interaction from Amino Acid Sequence Using Ensemble Classifier", *In 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2) IEEE*, pp. 1-4. 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] B.Ç. Yavuz, N. Yurtay and O. Ozkan, "Prediction of protein secondary structure with clonal selection algorithm and multilayer perceptron", *IEEE Access*, vol. 6, pp.45256-45261. , 2018 [CrossRef] [Google Scholar] [Publisher Link]

[16] S. Xie, Z. Li and H. Hu, "Protein secondary structure prediction based on the fuzzy support vector machine with the hyperplane optimization", *Gene*, vol. 642, pp. 74-83, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[17] Y. Wang, H. Mao and Z. Yi, "Protein secondary structure prediction by using deep learning method", *Knowledge-Based Systems*, vol. 118, pp.115-123, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] Y. Liu, J. Cheng, Y. Ma, and Y. Chen, "Protein secondary structure prediction based on two dimensional deep convolutional neural networks", *In 2017 3rd IEEE International Conference on Computer and Communications (ICCC) IEEE*, pp. 1995-1999, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[19] A.K. Pani, M. Manohar and R. Kumar, "An efficient algorithmic technique for feature selection in IoT based intrusion detection system", *Indian Journal of Science and Technology*, vol. 14, no. 1, pp. 76-85, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] H. Peraza-Vázquez, A.F. Peña-Delgado, G. Echavarría-Castillo, A.B. Morales-Cepeda, J. Velasco-Álvarez, and F. Ruiz-Perez, "A bio-inspired method for engineering design

optimization inspired by dingoes hunting strategies", *Mathematical Problems in Engineering*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] C. Guan, S. Wang and A.W.C. Liew, 2019, "Lip image segmentation based on a fuzzy convolutional neural network", *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 7, pp. 1242-1251. [CrossRef] [Google Scholar] [Publisher Link]

[22] D.A. Pelta and N. Krasnogor, "Multimeme algorithms using fuzzy logic-based memes for protein structure prediction", *In Recent advances in memetic algorithms* Springer, Berlin, Heidelberg, pp. 49-64, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[23] B. Kuhlman and P. Bradley, "Advances in protein structure prediction and design", *Nature Reviews Molecular Cell Biology*, vol. 20, no. 11, pp. 681-697, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] D.E. Kim, D. Chivian and D. Baker, "Protein structure prediction and analysis using the Robetta server", *Nucleic acids research*, vol. 32, no. suppl_2, pp. W526-W531, 2004. [CrossRef] [Google Scholar] [Publisher Link]

AUTHORS

P. G. Sreelekshmi currently working in University Institute of Technology Malayinkeezhu under University of Kerala as Lecturer. Prior to her recent appointment she was worked at Sivaji College of Engineering as Assistant Professor. She completed her ME in 2014 from Anna University Chennai.

S.C. Ramesh completed a bachelor of engineering in electronics and communication engineering, master of engineering in avionics and doctor of philosophy in Anna University Chennai, India. He is working as associate professor in PSN College of engineering and technology, Tirunelveli, India

Arrived: 24.11.2023
Accepted: 23.12.2023