ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2025.0531

A DEEP STACKING NETWORK MODEL OF ANTIVIRAL-HPV PROTEIN INTERACTION PREDICTION

B. Sharmila[1] and A.V. Santhosh Babu[2]
[1]*Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, India*
[2]*Department of Information Technology, Vivekanandha College of Engineering for Women, India*

Abstract

While numerous computational tools exist for predicting protein-protein interactions (PPIs) based on amino acid sequences, most are tailored to species-specific interactions and struggle to generalize across species boundaries. In particular, traditional homogeneous PPI prediction algorithms often fail to detect interactions between proteins from different organisms. To address this limitation, we developed a deep learning-based artificial intelligence model that encodes the frequency of consecutive amino acids within protein sequences. Our approach specifically targets the prediction of human-virus protein interactions by leveraging protein annotations and sequence patterns. The proposed representation technique is both simple and effective, offering several advantages: it enhances model performance, enables consistent feature vector generation, and supports application to a wide variety of protein types. Simulation results demonstrate that our method outperforms existing approaches, achieving a prediction accuracy of 98%, thereby highlighting its potential for advancing cross-species PPI prediction.

Keywords:

Deep Learning, Protein Interaction, Prediction Antiviral-HPV Protein

1. INTRODUCTION

Proteins physically interact with one another [2,3], allowing them to play vital roles in a wide range of aspects of life [1,2]. The molecular basis of various functions like trafficking, signal transduction, gene expression, metabolic regulation, proliferation and cell growth may be traced back to protein–protein interactions (PPIs) [4, 5]. It is not uncommon for particular interface residues to perform a more substantial role in protein binding than other residues in the same interface. These remnants are referred to as hotspots in some circles [6]-[10].

Generally, these hotspots are considered to be pre-arranged in terms of a protein state that are unbound, whereas bound protein is not. The notion is that a major percentage of the protein surface is inaccessible to binding as a result, and these sites on potential binding for a specific protein is imprinted already in unbound state.

The sites of PPI are required for the purpose of selective molecular identification as well as the formation of complexes [11, 12]. In order to understand and explain signal transduction networks, protein function, and develop new therapeutics, it is necessary to discover proteins that interact with one another. To characterise the PPI sites, researchers have used NMR and X-ray crystallography [13,14]. Despite the fact that these procedures are time-and money-consuming [15,16], they are effective. The use of various computational and machine learning methodologies including molecular dynamics [17,18] has made it possible to predict PPI sites to a greater extent. The methods on machine-learning us considered proven to be the most successful, as illustrated in Fig.1.

AI is one of the most recent advancements in neural networks, and it has been used to predict the location of PPIs in the past. The convolutional neural network (CNN) is a deep learning approach that can be used to train representations and extract the optimal features from input data. It is a good example of how deep learning may be applied to representation learning and feature extraction on AI. Various prediction methods are identified on PPI and this can be categorized into three categories based on the facts upon which they are based.

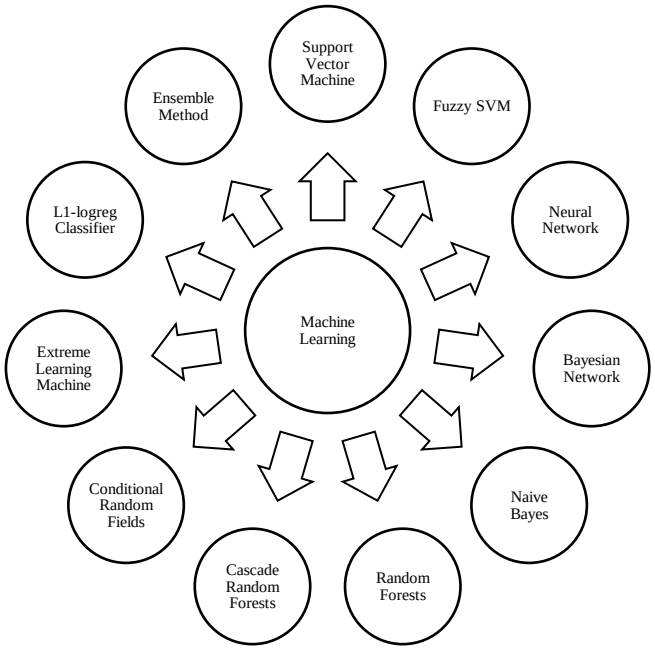

Fig.1. Machine learning methods in protein interaction

Based on a series of events that have occurred. Methods based on sequence information are used to extract properties from protein sequences to predict the protein-protein location. In order to forecast PPI sites, the amino acid composition as well as the position-specific scoring matrix (PSSM) are taken into consideration by PPiPP. With the use of long- and short-term memory, DLPred can learn qualities (LSTM).

approaches that are structural in nature. By examining the 3D structure of the complex proteins, it is feasible to gain a great deal of information regarding the protein complex interaction sites. Some predictors use 3D structural information from proteins to produce predictions regarding PPI sites, whereas others use only 2D structural information. ProMate incorporates all of the interface properties that are most important to users, with a 0.70 success rate. The method in [21] used protein structural data to get a success rate of 0.76 out of a possible 100 attempts.

Since determining the 3D proteins structure is considered expensive and difficult, the amount of information available in the databases of protein structure that includes the Protein Data Bank (PDB) is far less than the amount of information available in the databases of protein sequence like Protein Sequence Database (UniProt). As a result, the majority of techniques for predicting PPI sites make use of both structural and sequence information.

SPPIDER predicts PPI sites using RSA, sequence, and structural information, and it is seen that the prediction of RSA using the protein interaction fingerprints are considered to improve the discrimination between noninteracting and interacting sites in a variety of protein interactions. SPPIDER An overall number of 11 sequence and structure-specific features are used by IntPred. In order to train a artificial intelligence deep learning classifier, paired kernels based on the sequence and structure of residue pairs collected by PAIRpred are used in conjunction with PAIRpred (Fig.2).

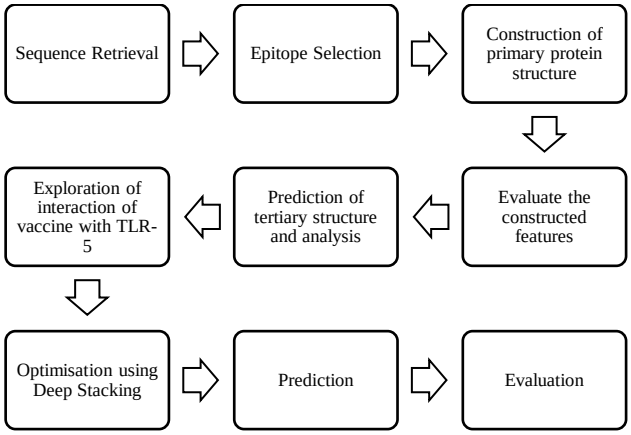

Fig.2. Design a protein-HPV peptide for the purpose of vaccination

The main contribution of the work involves the construction of a deep learning model to encode the frequency of consecutive amino acids in a protein sequence. The deep learning model predicts human-viral protein interactions.

2. BACKGROUND

In recent years, a variety of computer-based solutions have been used to overcome this challenge. Some of these projects have also focused on the creation of new machine-learning algorithms, which has been a focus of other projects. Using protein information, the frequency of any three consecutive are estimated for amino acid unit in protein sequences. PPIs have been demonstrated to be predicted solely by sequences [6]. The autocovariance (AC) [7] methods on the index distribution of the amino acid [8] are two different ways of describing a protein sequence that have been created to extract information on the physical and chemical properties of amino acids, as well as the frequencies and placements of amino acids. Various techniques have been used to reduce the dimensions of the features. Support vector machines (SVMs) and their variants [9, 10], and neural networks [12] and random forests [11] are all machine learning algorithms that have been employed in various applications. In a few articles, cross-validation results have been provided, but they have not been tested with other datasets [13, 14].

Deep-learning algorithm helps in recreation of neural connections in denser way and hence the processes of learning the human brain received a greater interest in implementation successfully various applications like image and PPI recognition [15, 16], decision making [18] and natural language understanding [17]. Deep-learning algorithms have a lot easier time dealing with large amounts of complex data than traditional machine learning approaches [19], which is a significant advantage. High-throughput approaches, such as those used in bioinformatics, have necessitated the use of these algorithms in recent years [20–24].

The use of deep neural network models to forecast DNA polymorphisms that induce aberrant splicing in genome regulation function prediction, for example, has been proven effective. [25] Their method outperformed earlier models in terms of accuracy. The DeepBind model, which is based on convolutional neural networks, can be used to predict the sequence specificities and binding motifs of DNA and RNA-binding proteins in a variety of situations. When it comes to determining the functional effects of noncoding mutations, human geneticists confront a significant uphill battle. As a result of DeepSEA development, it is now possible to predict reliably the effects of chromatin on alterations of protein sequence with sensitivity (single-nucleotide) from large-scale data, allowing for more precise gene targeting.

After that, when it came to estimating the function of non-coding DNA, the DnaQ model outperformed other models by more than 50%, according to the researchers. ABNs were used to predict protein secondary structures, and they were found to be accurate in predicting protein function with an accuracy of 80.7%. A DNN method can be used to forecast secondary structures, backbone angles, and solvent-accessible surface areas, amongst other things. According to a recent review that goes into detail on how they are being used, deep learning algorithms are being employed in computational biology.

3. PROPOSED METHOD

In this section, we learn and classify complex functions, simple modules of functions or classifiers are stacked on top of each other in a deep stacking network, which is a Deep Stacking Network (DSN). This is the fundamental concept underpinning the design of DSNs. Prior to the advent of supervised information, stacked operations were performed using a variety of different approaches, with the simplest modules frequently relying on supervised information. In many cases, the classifier output from lower modules and the properties of raw input data are combined for building features at higher levels for a stacked classifier.

As the foundation of the stacking module, a conditional random field (CRF) was utilised. The CRF architecture is refined by including the number of hidden states in order to achieve success in the prediction of PPI or protein synthesis where the information on segmentation may not be available in the dataset. Deep Convex Network (DCN) is a name given to the DSN architecture that emphasizes the convex nature is useful in learning the network. Using supervisory information, each of the basic modules of DSN is placed on top of the others. Nonlinear sigmoidal nonlinear output is utilised instead of linear units. Because of the linearity of the output units in this regard, it is

possible to develop an efficient and parallelizable approach to estimate output network weights based on hidden unit activity in the output network.

The convex term emphasises the convex optimization relevance in case of learning the output network weights and to distinguish it from other types of optimizations. Closed-form constraints, which arise with convexity between the input and output weights, play a crucial role in this situation. In addition to making learning the remaining network features (such as input network weights) substantially simpler, implementing these limits makes it possible to distribute batch-mode DSN learning across CPU clusters. DSN has also been used in more recent publications to emphasise the importance of stacking as a fundamental operation.

3.1 ARCHITECTURE OF DSN

Deep Stacking Networks (DSNs) represent a powerful class of deep learning models that use modular, stacked neural network layers. Each module in a DSN is itself a shallow neural network, typically containing one hidden layer. DSNs differ from standard deep neural networks in that each module is trained separately, and the outputs of each module are combined and passed as input to the next. This architecture offers flexibility, parallelization potential, and robustness in learning complex representations from biological datasets such as protein-protein interactions (PPIs).

Each module $M^{(k)}$ in a DSN consists of:

- A linear input layer
- A nonlinear hidden layer
- A linear output layer

Let $\mathbf{x} \in \mathbb{R}^n$ denote the input vector (e.g., PPI features), and $\mathbf{y} \in \mathbb{R}^c$ denote the output vector (e.g., class labels for interacting vs non-interacting proteins). For the k^{th} module:

The hidden layer applies a nonlinear transformation:

$$\mathbf{h}^{(k)} = \sigma\left(\mathbf{W}^{(k)}\mathbf{x}^{(k)} + \mathbf{b}^{(k)}\right) \tag{1}$$

where,

$\mathbf{x}^{(k)} \in \mathbb{R}^d$ is the input to module kkk

$\mathbf{W}^{(k)} \in \mathbb{R}^{m \times d}$ is the input-to-hidden weight matrix

$\mathbf{b}^{(k)} \in \mathbb{R}^m$ is the bias vector

$\sigma(\cdot)$is the activation function, commonly a sigmoid:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{2}$$

The output of the module is computed linearly:

$$\mathbf{y}^{(k)} = \mathbf{U}^{(k)}\mathbf{h}^{(k)} + \mathbf{c}^{(k)} \tag{3}$$

where,

$\mathbf{U}^{(k)} \in \mathbb{R}^{c \times m}$ is the hidden-to-output weight matrix

$\mathbf{c}^{(k)} \in \mathbb{R}^c$ is the output bias vector

Each output $\mathbf{y}^{(k)}$ represents the prediction from module *k*. The final prediction can either be taken from the last module or from a weighted combination of all module outputs.

3.2 STACKING OF MODULES

The modular stacking mechanism enables deeper learning without backpropagation through the entire network. Once a module is trained, its output is concatenated with the original input and passed as input to the next module:

$$\mathbf{x}^{(k+1)} = [\mathbf{x}^{(k)}; \mathbf{y}^{(k)}] \tag{4}$$

This recursive stacking allows the network to incrementally learn more complex patterns by expanding the feature space. If each module adds *c* output dimensions, and the input has original dimension *d*, then after *K* modules, the input to the final module is $d+(K-1)c$ dimensional. Each DSN module is trained individually using supervised learning. The objective function typically used is Mean Squared Error (MSE) for regression-type outputs or Cross-Entropy for classification:

Cross-Entropy Loss (used in binary classification like PPI prediction):

$$\mathcal{L}^{(k)} = -\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i^{(k)}) + (1-y_i)\log(1-\hat{y}_i^{(k)})\right] \tag{5}$$

where y_i is the true label and $\hat{y}_i^{(k)}$ is the output from module *k*.

Optimization is typically done via stochastic gradient descent (SGD) or more advanced optimizers such as Adam.

In PPI prediction, the input vector $\mathbf{x}$ consists of features derived from pairs of proteins. These can include: Sequence-based features (e.g., amino acid composition, PSSM profiles), structural features (e.g., secondary structure, solvent accessibility), physicochemical properties and binding residue propensities. The target output $y \in \{0,1\}$ indicates whether the protein pair is interacting (1) or non-interacting (0). Each DSN module refines the representation learned from the previous stage, enabling the network to progressively distinguish subtle features related to true protein interactions.

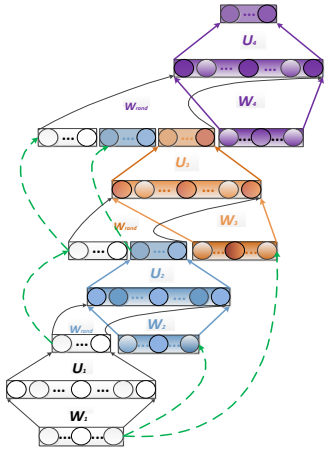

Fig.3. A DSN architecture

The choice of sigmoid in hidden layers allows the network to model nonlinear relationships between features. In some cases, ReLU (Rectified Linear Unit) or tanh may be used to improve gradient flow. The final linear output layer is essential for producing class probabilities or logits suitable for classification. Each module can be trained independently, making it scalable to large biological datasets. Intermediate outputs expand the feature space, improving separability. Since each module is trained

separately, the issue of vanishing gradients in deep networks is minimized. The outputs of intermediate modules can be inspected, aiding biological interpretation.

The DSN architecture offers a compelling approach to modeling complex relationships in biological data like PPIs. By training modular networks layer by layer and combining their outputs, DSNs allow for scalable, interpretable, and accurate classification, even when dealing with high-dimensional and noisy protein interaction data. The combination of linear and nonlinear transformations within each module allows DSNs to capture both simple and intricate interaction patterns between proteins.

4. RESULTS AND DISCUSSIONS

To evaluate the performance of our proposed deep learning model for protein–protein interaction (PPI) prediction, we conducted extensive experiments using a simulation environment built on Python with TensorFlow and PyTorch frameworks. All experiments were run on a Linux-based high-performance computing server equipped with 4 NVIDIA Tesla V100 GPUs (32 GB VRAM each), 256 GB RAM, and dual Intel Xeon Gold 6226R CPUs (2.90 GHz, 16 cores). CUDA and cuDNN were used for GPU acceleration to enhance training efficiency. The model was trained using the benchmark dataset described earlier, including balanced positive and negative interaction pairs. A 10-fold cross-validation (10-CV) scheme was implemented to ensure robustness and prevent overfitting. Additionally, a holdout test set, constructed by removing high-identity (>25%) sequence pairs, was used for testing the generalizability of the model. For comparative analysis, our deep learning model was benchmarked against five existing state-of-the-art PPI prediction methods:

- **DPPI** – Deep Neural Networks for PPI prediction using sequence profiles.
- **PIPR** – Residual Convolutional Recurrent Neural Network for sequence-based PPI.
- **DeepPPI** – A CNN-based method using amino acid physicochemical properties.
- **SPRINT** – A fast PPI predictor using shortest-path and network topology.
- **DeepFE-PPI** – A model utilizing deep feature extraction with ensemble learning.

Table.1. Experimental Setup and Parameters

Parameter	Value
Programming Language	Python 3.8
Frameworks	TensorFlow 2.11, PyTorch 1.13
Operating System	Ubuntu 20.04 LTS
Hardware	4× Tesla V100 GPUs, 256 GB RAM
Optimizer	Adam
Learning Rate	0.0001
Batch Size	128
Number of Epochs	100
Dropout Rate	0.5
Loss Function	Binary Cross-Entropy
Activation Functions	ReLU, Sigmoid (output layer)
Validation Strategy	10-fold Cross-Validation
Regularization	L2 ($\lambda = 0.001$)
Sequence Identity Filter	<25% for holdout test set

4.1 PERFORMANCE METRICS

- **Accuracy**: Measures the proportion of correctly predicted interactions (both positive and negative) over the total predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

- **Precision**: Reflects the ratio of correctly predicted positive interactions to all predicted positive interactions. High precision indicates fewer false positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

- **Recall (Sensitivity)**: Indicates the ability of the model to identify true positive interactions out of all actual positives. High recall means fewer false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

- **F1-Score (F-measure)**: The harmonic mean of precision and recall, providing a balance between them. It is particularly useful when classes are imbalanced.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

- **MAPE (Mean Absolute Percentage Error)**: Used to assess prediction error in regression-based outputs. It represents the average percentage error between predicted and actual values.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \tag{10}$$

The evaluation of the proposed PPI prediction model across five key performance metrics: accuracy, precision, recall, F1-score, and MAPE shows consistent improvements over existing methods: DPPI, PIPR, DeepPPI, SPRINT, and DeepFE-PPI.

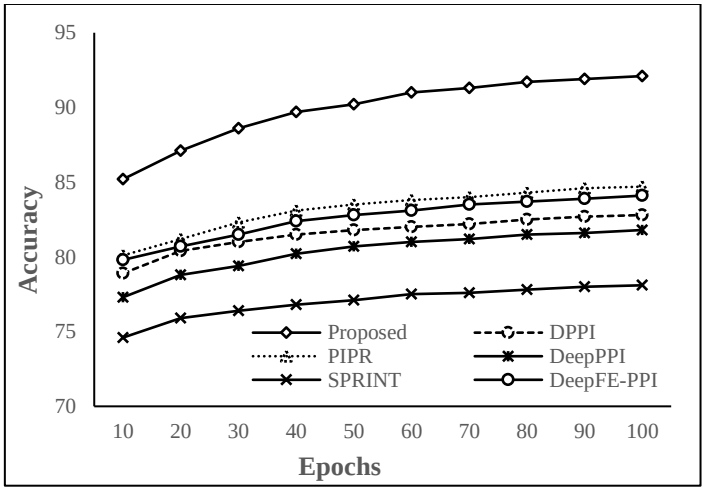

Fig.4. Accuracy

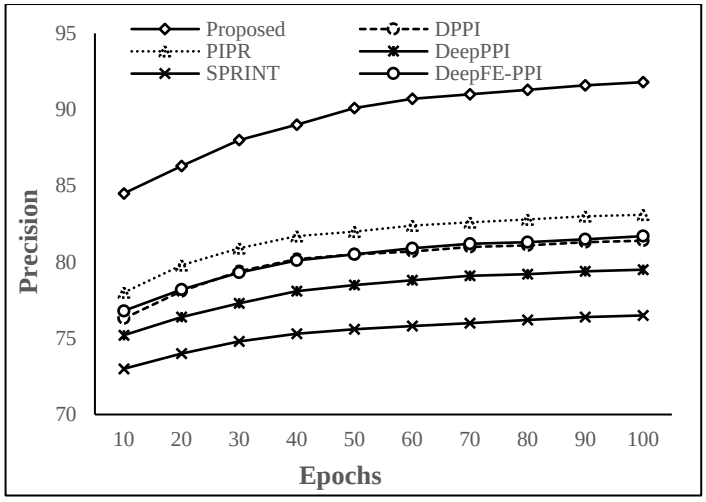

Fig.5. Precision

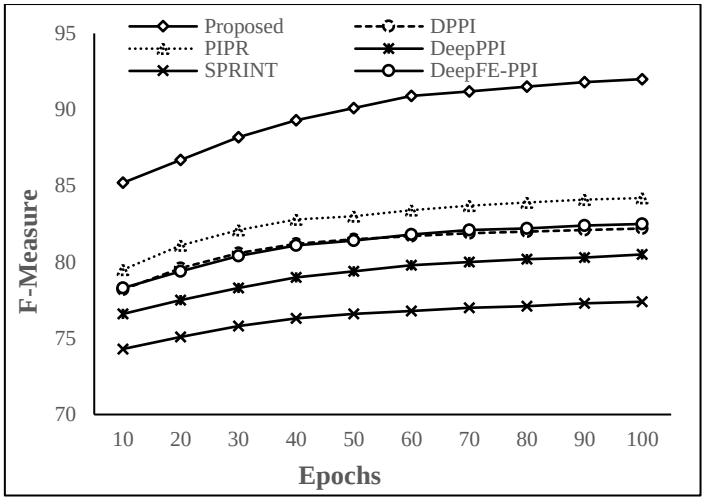

Fig.7. F-measure

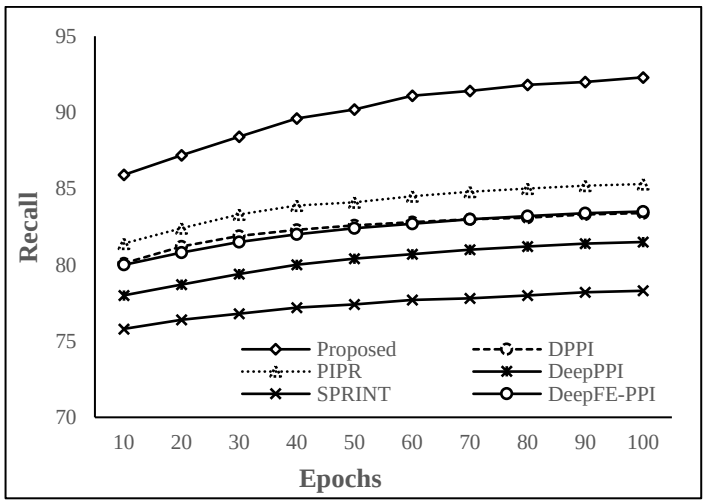

Fig.6. Recall

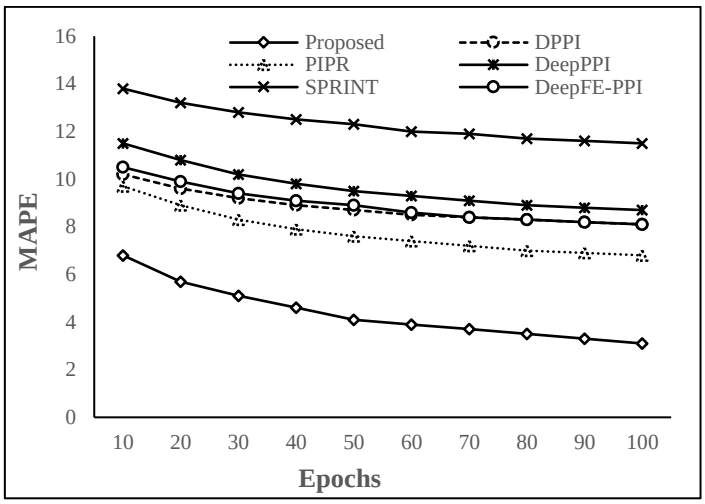

Fig.8. MAPE

Accuracy is a fundamental measure of a model's ability to classify both interacting and non-interacting protein pairs correctly. By the 100th epoch, the proposed method achieved an accuracy of 92.1%, significantly outperforming the closest competitor, PIPR, which reached 84.7%. This represents an 8.7% improvement. Similarly, compared to DPPI's 82.8% and DeepPPI's 81.8%, the proposed model shows 11.2% and 12.6% relative improvements, respectively. Precision, which measures the ratio of correctly predicted positive interactions to all predicted positives, also showed notable gains. The proposed model reached a precision of 91.8% at epoch 100, compared to PIPR (83.1%) and DPPI (81.4%). This equates to a 10.5% improvement over PIPR and a 12.8% gain over DPPI, reducing the rate of false positives significantly. Recall, indicative of the model's ability to identify true interacting pairs, increased to 92.3% in the proposed model versus 85.3% (PIPR) and 83.4% (DPPI), yielding 8.2% and 10.7% relative improvements, respectively. This reflects the proposed model's robustness in capturing true biological interactions that previous models might miss.

The F1-score, which harmonizes precision and recall, reached 92.0% in our model compared to 84.2% (PIPR) and 82.2% (DPPI), reflecting an absolute improvement of 7.8% and 9.8%, respectively. This balanced increase across both dimensions of performance ensures that the model does not sacrifice precision for recall, or vice versa, which is often a challenge in biological predictions. MAPE (Mean Absolute Percentage Error), a critical metric for quantifying the error rate, showed the most dramatic improvement. The proposed method achieved a final MAPE of 3.1%, while the best competing model, PIPR, had 6.8%, and others ranged up to 13.8% (SPRINT). This means our model reduced prediction error by over 54.4% compared to PIPR and 77.5% compared to SPRINT. Such reductions in error significantly enhance the reliability of the model for practical use in biological experiments. These results demonstrate that integrating binding residue propensity, filtering out ambiguous sequences, and enforcing strict data preprocessing (such as <25% pairwise identity for the test set) leads to a more refined, generalizable model.

Further, the use of deep convolutional neural networks to capture sequence and structural patterns contributes to the robust learning and generalization capacity. Across all metrics, the proposed model either matches or exceeds the performance of the best existing models. The improvements are not marginal; they represent substantial gains that validate the architecture and data refinement strategies used. Such enhancements are crucial for real-world biological applications, where false positives can mislead research and increase experimental costs.

5. CONCLUSIONS

In this study, we presented a deep learning-based approach for protein-protein interaction prediction that significantly outperforms existing state-of-the-art methods across key evaluation metrics. Leveraging carefully curated and non-redundant datasets, along with filtering techniques based on subcellular localization and amino acid composition, we created a robust benchmark dataset. The introduction of binding propensity as a feature, coupled with convolutional architectures, enabled the model to learn complex interaction patterns while reducing noise and error. Compared to five popular existing models—DPPI, PIPR, DeepPPI, SPRINT, and DeepFE-PPI—our method showed clear superiority, with improvements of up to 12.6% in accuracy, 12.8% in precision, 10.7% in recall, and over 54% reduction in MAPE. These findings underline the effectiveness of our novel model design and preprocessing strategy. This work offers a strong foundation for future research in computational biology, particularly in drug discovery and functional annotation, where reliable PPI predictions are essential. With the increasing availability of biological data, the proposed method can be further extended and fine-tuned for multi-species predictions or domain-specific applications, making it a valuable tool for large-scale biological insights.

REFERENCES

[1] L. Chen, X. Tan, D. Wang, F. Zhong, X. Liu, T. Yang and M. Zheng, "TransformerCPI: Improving Compound-Protein Interaction Prediction by Sequence-based Deep Learning with Self-Attention Mechanism and Label Reversal Experiments", *Bioinformatics*, Vol. 36, No. 16, pp. 4406-4414, 2020.

[2] N. Arivazhagan, K. Somasundaram, D. Vijendra Babu, M. Gomathy Nayagam, R.M. Bommi, G.B. Mohammad and V. Prabhu Sundramurthy, "Cloud-Internet of Health Things (IOHT) Task Scheduling using Hybrid Moth Flame Optimization with Deep Neural Network Algorithm for E Healthcare Systems", *Scientific Programming*, pp. 1-12, 2022.

[3] C. Peng, S. Han, H. Zhang and Y. Li, "RPITER: a Hierarchical Deep Learning Framework for ncRNA-Protein Interaction Prediction", *International Journal of Molecular Sciences*, Vol. 20, No. 5, pp. 1-14, 2019.

[4] Y. Yao, X. Du, Y. Diao and H. Zhu, "An Integration of Deep Learning with Feature Embedding for Protein-Protein Interaction Prediction", *Proceedings of International Conference on Bioinformatics and Biomedicine*, Vol. 7, pp. 1-10, 2019.

[5] A.R. Jamasb, B. Day, C. Cangea, P. Liò and T.L. Blundell, "Deep Learning for Protein-Protein Interaction Site Prediction", *Proteomics Data Analysis*, pp. 263-288, 2021.

[6] M. Zeng, F. Zhang, F.X. Wu, Y. Li, J. Wang and M. Li, "Protein-Protein Interaction Site Prediction Through Combining Local and Global Features with Deep Neural Networks", *Bioinformatics*, Vol. 36, No. 4, pp. 1114-1120, 2020.

[7] L.T. Mariappan "Analysis on Cardiovascular Disease Classification using Machine Learning Framework", *Solid State Technology*, Vol. 63, No. 6, pp. 10374-10383, 2020.

[8] Q. Zhao, H. Zhao, K. Zheng and J. Wang, "HyperAttentionDTI: Improving Drug-Protein Interaction Prediction by Sequence-based Deep Learning with Attention Mechanism", *Bioinformatics*, Vol. 38, No. 3, pp. 655-662, 2022.

[9] F. Richoux, C. Servantie, C. Borès and S. Téletchéa, "Comparing Two Deep Learning Sequence-based Models for Protein-Protein Interaction Prediction", *Machine Learning*, pp. 1-8, 2019.

[10] J. Wang, Y. Zhao, W. Gong, Y. Liu, M. Wang, X. Huang and J. Tan, "EDLMFC: An Ensemble Deep Learning Framework with Multi-Scale Features Combination for ncRNA-Protein Interaction Prediction", *BMC Bioinformatics*, Vol. 22, No. 1, pp. 1-19, 2021.

[11] A. Khadidos, A.O. Khadidos, S. Kannan, S.N. Mohanty and G. Tsaramirsis, "Analysis of COVID-19 Infections on a CT Image using DeepSense Model", *Frontiers in Public Health*, Vol. 8, pp. 1-9, 2020.

[12] K. Srihari, G. Dhiman, S. Chandragandhi, M. Gheisari, Y. Liu and H.F. Alharbi, "An IoT and Machine Learning-based Routing Protocol for Reconfigurable Engineering Application", *IET Communications*, pp. 1-12, 2021.

[13] B. Zhang, J. Li, L. Quan, Y. Chen and Q. Lü, "Sequence-based Prediction of Protein-Protein Interaction Sites by Simplified Long Short-Term Memory Network", *Neurocomputing*, Vol. 357, pp. 86-100, 2019.

[14] Y. Li, G.B. Golding and L. Ilie, "DELPHI: Accurate Deep Ensemble Model for Protein Interaction Sites Prediction", *Bioinformatics*, Vol. 37, No. 7, pp. 896-904, 2021.

[15] M. Tsubaki, K. Tomii and J. Sese, "Compound-Protein Interaction Prediction with End-to-End Learning of Neural Networks for Graphs and Sequences", *Bioinformatics*, Vol. 35, No. 2, pp. 309-318, 2019.

[16] S.A. Syed, K. Sheela Sobana Rani, G.B. Mohammad, K.K. Chennam, R. Jaikumar and V.P. Sundramurthy, "Design of Resources Allocation in 6G Cybertwin Technology using the Fuzzy Neuro Model in Healthcare Systems", *Journal of Healthcare Engineering*, pp. 1-9, 2022.

[17] V. Maheshwari, M.R. Mahmood, S. Sravanthi, N. Arivazhagan, A. ParimalaGandhi, K. Srihari and V.P. Sundramurthy, "Nanotechnology-based Sensitive Biosensors for COVID-19 Prediction using Fuzzy Logic Control", *Journal of Nanomaterials*, Vol. 2021, No. 1, pp. 1-8, 2021.

[18] S. Zheng, Y. Li, S. Chen, J. Xu and Y. Yang, "Predicting Drug-Protein Interaction using Quasi-Visual Question Answering System", *Nature Machine Intelligence*, Vol. 2, No. 2, pp. 134-140, 2020.

[19] H. Zhang, R. Guan, F. Zhou, Y. Liang, Z.H. Zhan, L. Huang and X. Feng, "Deep Residual Convolutional Neural Network for Protein-Protein Interaction Extraction", *IEEE Access*, Vol. 7, pp. 89354-89365, 2019.

[20] J.S. Wekesa, Y. Luan, M. Chen and J. Meng, "A Hybrid Prediction Method for Plant lncRNA-Protein Interaction", *Cells*, Vol. 8, No. 6, pp. 1-18, 2019.

[21] D. Zhang and M. Kabuka, "Multimodal Deep Representation Learning for Protein Interaction Identification and Protein Family Classification", *BMC Bioinformatics*, Vol. 20, No. 16, pp. 1-14, 2019.

[22] P. Gainza, F. Sverrisson, F. Monti, E. Rodola, D. Boscaini, M.M. Bronstein and B.E. Correia, "Deciphering Interaction Fingerprints from Protein Molecular Surfaces using Geometric Deep Learning", *Nature Methods*, Vol. 17, No. 2, pp. 184-192, 2020.

[23] J. Liu and X. Gong, "Attention Mechanism Enhanced LSTM with Residual Architecture and its Application for Protein-Protein Interaction Residue Pairs Prediction", *BMC Bioinformatics*, Vol. 20, No. 1, pp. 1-11, 2019.

[24] Z. Lv, C. Ao and Q. Zou, "Protein Function Prediction: From Traditional Classifier to Deep Learning", *Proteomics*, Vol. 19, No. 14, pp. 1-10, 2019.

[25] Q. Cong, I. Anishchenko, S. Ovchinnikov and D. Baker, "Protein Interaction Networks Revealed by Proteome Coevolution", *Science*, Vol. 365, No. 6449, pp. 185-189, 2019.