Grenze International Journal of Engineering and Technology, June Issue

Deep Residual Neural Networks for Protein Structure Prediction

Vani Ashok[1], Sheela N[2], Kiran Bharadwaj[3], Hoisala V Raj[4], Preetham G[5] and Pranjal S[6]
[1-6]SJCE, JSS Science and Technology University, Dept. of Computer Science and Engineering, Mysuru, India
Email: vanisj@sjce.ac.in, sheela_cse@sjce.ac.in, hdkiranbharadwaj@gmail.com, hoisalavraj5@gmail.com, prthmgangadhar@gmail.com, pranjalshyla@gmail.com

***Abstract*— Prediction of protein three-dimensional structure is a difficult computational biology problem with extensive applications to many areas, including pharmaceutical design and disease modeling. In the proposed work, a deep learning model is introduced to predict dihedral angles of the backbone (ϕ and ψ) and the distance matrices between the residues and physicochemical properties directly using sequence-derived features, physicochemical properties. The models are trained on the CASP7 subset of the ProteinNet dataset with task-specific loss functions that are directly designed to regress angles and classify distances. The obtained system has an average root-mean-square deviation (RMSD) of 2.1 Å and coefficient of determination (R^2) values of 0.39 for ϕ and 0.69 for ψ, which indicates the strong predictions of structures. It is important to note that this design is lightweight and can be scaled and therefore implemented on the standard consumer-level computing hardware.**

***Index Terms*— Protein Structure, Deep Learning, ResNet, Angle Prediction, Distance Map.**

I. INTRODUCTION

Proteins are the living biochemical key to life, which makes them catalytic, a signaler, immune response and assemblers. Proteins have 3D structures that regulate the biological performance of proteins through the folding of amino acid sequences. Determination of structure accurately is the key to drug discovery research, synthetic biology, and the mechanisms of the disease. Conventional experimental approaches, such as X-ray crystallography, NMR and cryo-electron microscopy, yield high-resolution structures but are time consuming, expensive, and not always possible for flexible or recalcitrant proteins. Hence, computational methods have become more popular, especially with deep learning advances. Early models offered contact map and secondary structure prediction with features such as Position-Specific Scoring Matrices (PSSM) and statistical learning [1]. ProteinNet [2][3] and other databases facilitated standardization of training that resulted in progress such as AlphaFold, employing Multiple Sequence Alignments (MSA) and attention networks to produce Global Distance Test (GDT) scores of over 90 on CASP13 targets [4]. RoseTTAFold later introduced a better three-track architecture combining 1D, 2D, and 3D inputs [5]. Yet all of these models rely on MSAs, which are computationally costly and do not exist for new or orphan proteins. Their high GPU demands also reduce interpretability. MSA-free alternatives such as ESMFold are encouraging in being able to predict structure from a single sequence, but remain resource hungry [6][7].
The proposed work introduces a modular, interpretable and MSA independent protein structure prediction using deep Residual Neural Networks (ResNets).

Grenze ID: 01.GIJET.12.2.293_21

Two primary tasks are addressed: backbone dihedral angles-phi (ϕ) and psi (ψ)-prediction and residue-residue distance classification from sequence derived features like one-hot encodings, PSSM and physicochemical descriptors. The removal of the need for MSA, combined with the ease of training on commodity hardware, greatly enhances access for low-resource researchers. Additionally, its modularity and integrated visualizations enhance interpretability, thus making it a scalable and usable tool for real-world protein modeling.

II. Related Work

Deep learning has transformed protein structure prediction over the last decade. Different architectures, from convolutional, recurrent to attention based, have propelled accuracy, generalizability and the reduction of computational costs. Each innovation has shaped the ever changing path of this research field.

Wang et al. [1] were the first to demonstrate the potential of ultra-deep convolutional neural networks (CNNs) for the prediction of protein contact maps. Their method used PSSMs with sequence based features to attain contact accuracy ranging from 30% to 70% in the context of CASP10 and CASP11 depending on the sequence length and threshold distance employed. Although successful, the model was limited to binary contact classification and did not yield continuous structural output. To meet the need for uniform training data, AlQuraishi [2] proposed ProteinNet, a standard set of data combining evolutionary profiles, structural annotations and experimental PDB data from CASP7–12. ProteinNet trained models obtained RMSD scores of around 3.8 Å on CASP12 domains. The dependence of the dataset on evolutionary features puts limitations on MSA free techniques. Heffernan et al. [3] improved secondary structure prediction by integrating CNNs with bidirectional LSTM networks. Their integrated model learned both local and sequential patterns efficiently and achieved Q3 accuracy of ~84% on CB513 dataset which is a three-class secondary structure prediction benchmark.

A significant advancement was brought about by AlphaFold, proposed by Jumper et al. [4]. With transformer-based attention networks and multiple sequence alignments, AlphaFold made 3D atomic structure predictions with GDT scores with a mean of ~92.4 on CASP13, which was a new benchmark. Its reliance on MSAs and high computational cost, though, created concerns over accessibility. RoseTTAFold, by Baek et al. [5], gave a computationally less intensive option. With a three-track neural network that operated in parallel on sequence, 2D distances and 3D coordinates, it reported GDT scores of ~91.0 on CASP14. Although it used less resource, the model was still evolutionary data based. Conversely, Lin et al. [6] presented ESMFold, a language model based design that removed the requirement for MSA altogether. After having been trained on large protein sequence data sets, ESMFold produced RMSD values of around 2.5 Å on medium sized proteins and significantly lowered inference time to a handful of minutes on a single GPU. This was a step in the direction of alignment free, scalable modeling. Other than performance, interpretability and resource concerns with AlphaFold and other models were also noted in the critical analysis of AlQuraishi [7]. A forerunner to AlphaFold2, the Senior et al. [8] had applied deep learning based distance and angle potentials to a gradient descent-based coordinate generation method. With the utilization of residual CNNs to represent inter-atomic interactions, the model had GDT_TS labels exceeding 87 on CASP13 and proposed the notion of end-to-end learning for structure prediction-albeit still based on evolutionary features. Another direction of work was pursued by Ingraham et al. [9], who used graph neural networks to model protein structure. They encoded residues as node and spatial interaction as edge and their model recovered backbone geometry with RMSD of 3.5–4.2 Å on PDB and CATH datasets. Fang et al. [10] further optimized AlphaFold's architecture with HelixFold by utilizing linear attention mechanisms and decreasing the depth of Evoformer to obtain ~88.9 GDT_TS on CASP14 with 2–3× inference speedup. Its efficiency rendered it production worthy. Lastly, Chowdhury et al. [11] introduced FastFold, re-implementing AlphaFold with DeepSpeed and Megatron frameworks to add training and inference parallelism. It gained 9.2× training acceleration with CASP14 accuracy (~91–92 GDT) preserved, reducing GPU memory usage by half-enabling AlphaFold like performance for mass deployment.

While such models are highly accurate, reliance on evolutionary data and large computational power limits scalability and ease of deployment. The proposed method, however, employs a light-weight, MSA-free architecture, only using 1D and 2D Residual Neural Networks to make predictions of angles (ϕ and ψ) and distances, with attention to performance versus usability. To achieve, the following novel features are incorporated in the proposed work:

- Angle prediction for local structural modeling using 1D ResNet.
- Distance classification for global residue-residue interaction mapping using 2D ResNet.

III. MATERIALS AND METHODS

The proposed model is illustrated in Fig. 1. The system begins by taking protein sequences as FASTA or as input arrays. The sequences are converted into numerical values based on one-hot encoding, PSSMs and physical and chemical properties such as hydrophobicity and electric charge. After feature extraction, data is reshaped and normalized to fit neural network input requirements. For the purpose of distance prediction, a pairwise feature matrix is created through the outer product of the 1D sequence features to capture the residue-residue relations.

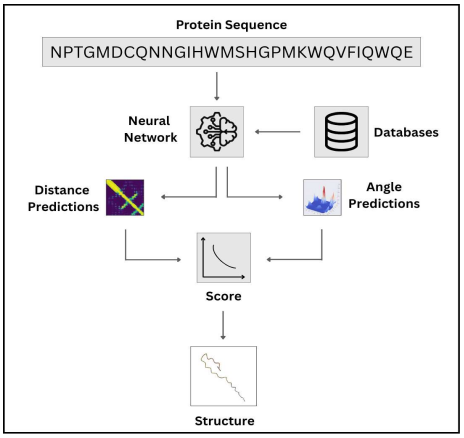

Figure 1. Proposed Framework of Protein Structure Prediction

Structure prediction is done by two parallel models. A 1D ResNet predicts sine and cosine of backbone angles, which are then transformed to their angular equivalent to ensure continuity and smoothness of gradients. In parallel, 2D ResNet predicts pairwise distances into spatial bins using softmax outputs, effectively encoding the spatial arrangement of the protein structure.

A. Dataset

ProteinNet CASP7 dataset, in the Critical Assessment of Structure Prediction (CASP) benchmark collection, was used. This database provides normalized inputs on protein structure prediction programs [2]. The data set includes the following elements:

- Primary sequences: One-letter amino acid codes of variable length
- Physicochemical properties: Numerical descriptors such as hydrophobicity, charge, and polarity used as input features
- Secondary structure labels: Helix (H), Sheet (E), Coil (C)
- Tertiary structure features: Dihedral angles phi (ϕ) and psi (ψ) and pairwise distance maps between C-α atoms
- 3D coordinates: High-resolution atomic positions from X-ray crystallography and NMR

The dataset is provided in both human readable and TensorFlow record formats and is split into training, validation and test sets. Multiple training subsets (e.g., train_30, train_50, train_90) allow for flexible training configurations based on resource availability. This dataset supports two predictive tasks:

- 1D ResNet-based angle prediction for local structural modeling
- 2D ResNet-based distance classification for global residue-residue interaction mappingwork.

B. Input Feature Representation

The first step in the model training pipeline as shown in Fig. 2 is raw protein sequence conversion to structured feature representations. The protein sequences are processed to give several channels of information to improve the model's information understanding at the sequence level and in pair-wise comparisons. These include:

- One-Hot encoding: In this method, each amino acid is encoded as a binary vector, where the index for the particular residue type is switched on to 1 and the rest are off. It is a simple way to encode the primary sequence and gives basic identity information for all residues.
- PSSMs that capture residue conservation across homologs

Physicochemical Properties: Hydrophobicity, van der Waals radii and surface accessibility are included to provide additional information on the residue interactions.

C. 1D ResNet for Dihedral Angle Prediction

The angles predicted were dihedral angles Φ and Ψ of each residue in one of the protein sequences.

- Φ (phi): rotation about the N–Cα\alphaα bond
- Ψ (psi): rotation about the Cα\alphaα–C bond

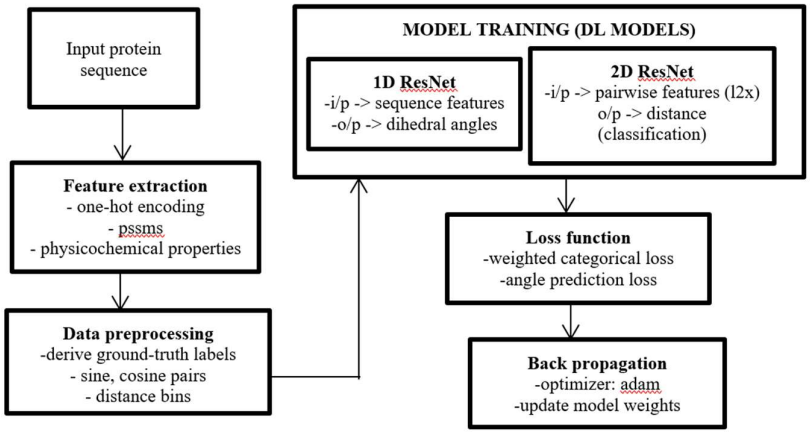

Figure 2. Proposed Model Training Pipeline

Dihedral angles are extremely important because they determine the backbone conformation of the protein sequence, which then influences the formation of the protein secondary structures including alpha helices and beta-sheets. The 1D ResNet model includes several layers of 1D convolution with residual links between them to enable information hopping between layers. The structure avoids the vanishing gradient issue and enables the model to learn the long range relations in the sequence better. For angle prediction, the Φ and Ψ angles are computed from atomic coordinates and converted to sine and cosine pairs to address the problem of angle periodicity.

The model predicts:

$$[\sin(\phi), \cos(\phi), \sin(\psi), \cos(\psi)].$$

D. 2D ResNet for Pairwise Distance Prediction

The 2D ResNet model is tuned for predicting spatial relationships between every pair of residues in a protein. It predicts the Euclidean distances between the alpha-carbon (Cα) atoms of the residues into quantized distance bins (e.g., 0–5 Å, 5–10 Å, etc.) using thresholding function. The predictions constitute a distance map, a 2D table whose cells contain the predicted class of distance between two residues. 2D ResNet takes pairwise features as inputs that are created through adding sequence based residue pair features to a matrix. The network structure comprises stacks of 2D convolutional layers followed by normalization and activation layers, allowing residual connections which allow information to flow through.

Pairwise Euclidean distance between Cα atoms:

$$D_{ij} = \| \mathrm{C}\alpha_i - \mathrm{C}\alpha_j \|$$

By capturing both local and global spatial relationships in the data, the 2D ResNet can detect patterns such as beta-sheet alignments and alpha-helix interactions. Its output is a probability distribution over distance bins per residue pair, which allows it to generate more advanced predictions beyond simple binary contact maps. The model performs particularly well at the detection of medium- and long-range interactions that are responsible for determining the overall structure of a protein.

E. Experimental Setup

The training was performed on GPU-enabled systems to enhance the learning speed. The models were trained independently using the Adam optimizer with hyperparameters chosen to attempt to optimize speed of convergence versus accuracy. A batch size of 2 was employed because of the memory limitations, and training was performed for 35 epochs. This configuration provided precise updates to model weights without overfitting. The 1D and 2D models are trained independently but share a unified preprocessing backbone. These models used task-specific loss functions.

The 1D ResNet employs a Mean Squared Error (MSE) loss applied to the sine and cosine values of Φ and Ψ, ensuring smooth optimization. For angle prediction, Angle Prediction Los, L_{angle} , that takes into account the fact that angles are in circles was used for training 1D ResNet. Angle Prediction Loss optimizes angle regression using sine and cosine outputs.

$$L_{angle} = \alpha * \text{MSE} + (1 - \alpha) * \text{MAE}$$

where, MSE = Mean Squared Error and MAE = Mean Absolute Error
For distance prediction, a weighted categorical cross-entropy loss function was used to counteract the difference in frequencies of bins of distance.

$$L_{dist} = -\sum_{C_{i=1}} w_i * y_i * \log(\hat{y}_i)$$

where, w_i is weight of class i, y_i is true label and $\hat{y}_i$is predicted probability.
The Adam optimizer updates model weights based on the calculated loss to improve prediction accuracy over training epochs. A portion of the dataset was held out for validation to monitor the generalization capability of the models. The entire training pipeline was modular and scalable, allowing for easy experimentation with different architectures and hyperparameters. The details of hyperparameters used for the proposed ResNet architectures are summarized in Table I. This structured training process allows the system to learn intricate structural dependencies, from local dihedral torsions to long-range spatial contacts, forming a complete intermediate representation of protein folding.

IV. Results and Discussion

Compared with a reference set of protein sequences with varying lengths and structural complexities, the models were compared. Tests were done in both structural precision and computational performance. To predict the dihedral angle, the 1D ResNet was able to produce a MAE of around 0.39 radians on phi angles and 0.69 radians on psi angles. The 2D ResNet model produced valid pairwise distance maps with relative accuracy of about 68% of short to medium range prediction of residue contacts. Predictions were made on structural motifs such as beta sheets and alpha helices. Predicted structures were also validated against experimentally derived conformations with the RMSD and the average RMSD was 2.1 Å, demonstrating strong alignment with experimentally resolved structures. The detailed metrics are shown in Table II.

Table I. Hyperpameters for 1D and 2D ResNet

Hyperparameter	1D ResNet	2D ResNet
Depth	10	30
Learning Rate	1e-3	1e-4
Dropout	0.1	0.15

Table II. RMSD (in Å) Based on Protein Length

Protein Length	Inference Time (sec)	RMSD (Å)
50-100	4.2	1.8
100-200	8.5	2.0
200-500	18.3	2.4

The proposed work introduces a new paradigm for protein structure prediction with deep residual neural networks for angle and distance prediction in a map. Key observations indicate that the proposed model is comparable to predicting protein structure without the application of MSA and is therefore new compared to current state-of-art models such as AlphaFold and RGN as given in Table III. While the proposed model's Phi (0.39 and 0.43) and Psi (0.69 and 0.73) R^2 isn't as high as state-of-art models, the proposed model is actually very good with little-needed biological data for it to be trained on. Although AlphaFold is still most accurate in RMSD, this method has faster processing and lightweight architecture, the latter being particularly vital for absence of dependency on MSAs. In contrast to other approaches with extensive MSA and evolutionary data dependency, proposed approach minimally treats sequence-based features, conserving computational resources and more appropriate for proteins with fewer homologous sequences. In comparison with the other MSA-free baseline models, like ESMFold, proposed exhibits better dihedral angles and RMSD indicating higher prediction accuracy [6].

TABLE III. R2 COEFFICIENT COMPARISON WITH STATE-OF-ART MODEL

Parameters	Proposed model	State-of-art models (employing MSA)	State-of-art models (MSA-free)
R2 Coefficient for Phi	0.39	0.43	0.35
R2 Coefficient for Psi	0.69	0.73	0.65
RMSD (Å)	~1.8–2.4 Å	~0.8–5 Å	~2.5–5.8 Å

The result from the 1D ResNet is also plotted in Ramachandran plots, which indicate the quality with which the predicted angles fall within biologically plausible areas. As seen in Fig. 3, the Ramachandran plot is formed by plotting the value of ϕ on the x-axis and the values of ψ on the y-axis to indicate the regions which are energetically favorable and likely to be found in the residues. The Ramachandran analysis shows that 20% of the predicted residues fall in favored regions compared to 34% in the ground truth, reflecting a broader scatter of predicted ϕ and ψ angles. As seen, predicted ϕ and ψ angles were largely concentrated in biologically acceptable Ramachandran regions, suggesting that the proposed model effectively learned realistic protein-folding patterns.

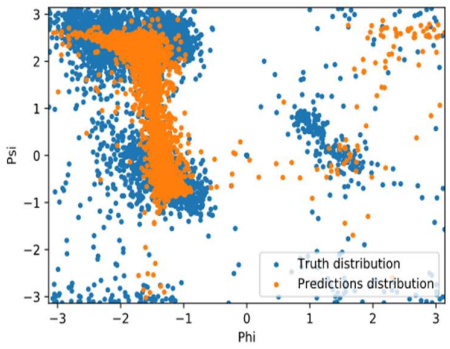

Figure 3. Ramachandran Plot

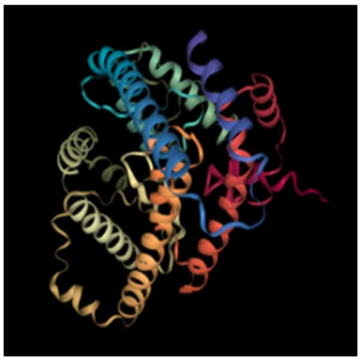

Figure 4. 3-D Structure of a sample Long Protein Sequence

Fig. 4 represents the 3D structure of the protein that has been predicted to look like the actual 3D form of the protein according to the sequence provided. The structural model captures key folding patterns, including the formation of alpha-helices and beta-sheets, derived from the predicted dihedral angles and inter-residue distances. This 3D representation allows users to interpret the overall conformation and stability of the protein, providing insight into its potential biological function and structural integrity.

Pairwise residue distances predicted by the 2D ResNet were visualized as heatmaps as shown in Fig. 5. The key structural elements that these maps maintained were the β-sheet orientations and contact-rich domains. In order to provide further analysis of the model performance, the ground truth data was compared with the predicted distance maps graphically. The ground truth map shows the approximate spatial relationships between the residues in a clearly patterned and sharp transition, especially along the diagonal, where adjacent residues occur.

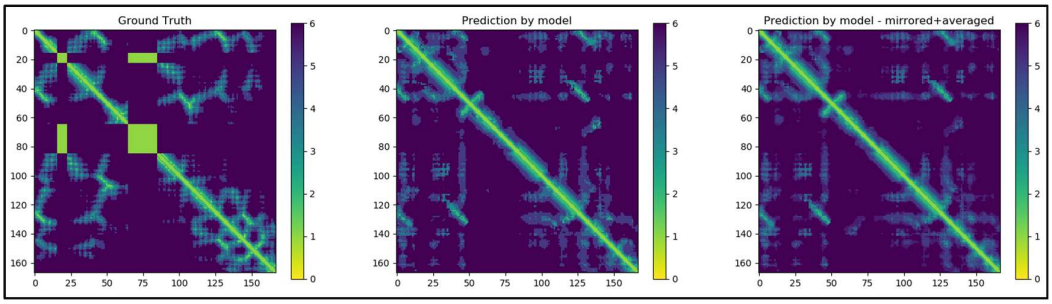

Figure 5: (a) Ground truth distance map (b) Raw model prediction (c) Post-processed prediction using mirroring and averaging

V. CONCLUSION

This project proposed and implemented a deep learning-based framework for predicting protein structural properties-specifically backbone dihedral angles (ϕ and ψ) and inter-residue distances—directly from amino acid sequences. The architecture was built around a dual-pathway design utilizing both 1D and 2D Residual Neural Networks (ResNets), enabling the model to capture sequence-level patterns as well as spatial residue-residue

interactions effectively. The system achieved quantitatively strong results across multiple evaluation metrics. For angle prediction, the model obtained R^2 scores of 0.39 for phi (ϕ) and 0.69 for psi (ψ) on the validation set, indicating strong correlation between predicted and ground truth values. In terms of 3D structure consistency, the system achieved an average RMSD of 2.1 Å, which is within the acceptable range for computational structure prediction in medium-length proteins. Visual assessments further reinforced the reliability of the predictions: Ramachandran plots exhibited clustering in biologically allowed regions and predicted distance maps displayed spatial patterns closely resembling those of ground truth matrices. The ability to broaden the dataset beyond CASP7 and build the entire structure as 3D would make the model an exceptionally useful asset in practical use.

REFERENCES

[1] S. Wang, S. Sun, Z. Li, R. Zhang, and J. Xu, "Accurate de novo prediction of protein contact map by ultra-deep learning model," PLOS Comput. Biol., vol. 13, no. 1, p. e1005324, Jan. 2017.

[2] AlQuraishi M, "ProteinNet: A standardized data set for machine learning of protein structure", BMC bioinformatics, 20(1), 311, 2019.

[3] Heffernan R, Yang Y, Paliwal K, Zhou Y, "Capturing non-local interactions by long short-term memory bidirectional recurrent neural networks for improving prediction of protein secondary structure, backbone angles, contact numbers and solvent accessibility," Bioinformatics, vol. 33, no. 18, pp. 2842–2849, 2017.

[4] Jumper J et al., "Highly accurate protein structure prediction with AlphaFold," Nature, vol. 596, no. 7873, pp. 583–589, 2021.

[5] M. Baek et al., "Accurate prediction of protein structures and interactions using a three-track neural network," Science, vol. 373, no. 6557, pp. 871–876, 2021.

[6] Z. Lin et al., "Evolutionary-scale prediction of atomic-level protein structure with a language model," Science, vol. 379, no. 6637, pp. 1123–1130, Mar. 2023.

[7] M. AlQuraishi, "AlphaFold @ CASP13: What just happened?" Some Thoughts on a Mysterious Universe, Dec. 2018.

[8] A. Asgari and A. Mofrad, "Deep learning methods for protein structure prediction," Wiley Interdisciplinary Reviews: Computational Molecular Science, vol. 12, no. 1, Jan. 2022.

[9] L. Wang et al., "Protein structure prediction via deep learning: An in-depth review," Frontiers in Molecular Biosciences, vol. 10, Mar. 2023.

[10] Jumper J et al., "Accurate structure prediction of biomolecular interactions with AlphaFold,", 2024.

[11] Alanazi W, Meng D, Pollastri G, "Advancements in one-dimensional protein structure prediction using deep learning," Computational and Structural Biotechnology Journal, Apr 3, 2025.