DHARA PUBLISHERS
Empowering Knowledge

*Journal of Soft Computing and Edge Computing*

journal homepage:


# AI Platforms for Accelerating Vaccine Development Using Epitope Prediction

Chozha Rajan. P
Department of Artificial Intelligence and Data Science
Bannari Amman Institute Of Technology
Sathyamangalam, Erode - 638401, Tamilnadu.
chozharajan.p@gmail.com

Dr K. Madan Mohan
Department of Information Technology(IT)
Guru Nanak Institute Of Technology(GNIT)
Ibrahimparnam, Ranga Reddy, 501506, Hyderabad, Telangana.
keturumm22@gmail.com

Dr. Laxmaiah Kocharla,
Department Of Computer Science And Engineering,
School Of Engineering,
Sreenidhi University Yamnampeta, Hyderabad.

Sridhar Reddy Vulapula
Department of Information Technology,
Vignana Bharathi Institute Of Technology,
Hyderabad, Telangana, India
vsridharreddy19@gmail.com

***Abstract*— Fast vaccine creation stands as a widespread urgent requirement which becomes increasingly critical for new infectious disease outbreaks. The identification of antigens through traditional vaccine development methods takes enormous amounts of time while being slowed down by extensive epitope detection procedures. This study brings an AI driven platform for integrating deep learning models with immunoinformatics tools for speedup of B cell and T cell epitope prediction. The AI platform utilizes the combination of convolutional neural networks (CNNs) and transformer-based architectures which received training with more than 60,000 epitope sequences from IEDB and PDB databases. Our predictive system demonstrated an 91.2% accuracy rate for MHC Class I epitopes and 88.7% accuracy for MHC Class II epitopes where it surpassed the predictive abilities of NetMHCpan and DeepVacPred by 4.3% and 3.9% respectively. During evaluations the averaged ROC-AUC scores achieved 0.93. The proposed method outperformed traditional sequence-alignment by shortening runtime duration while decreasing it by 27%. AI complements these predictions perfectly and allows them to be computationally extremely fast and accurately, thus shortening dramatically the vaccine cycle time. The proposed framework provides an adaptable framework which supports quick intervention against new pathogens because of its scalable design.**

***Keywords*— *Epitope Prediction, Vaccine Design, Artificial Intelligence, Deep Learning, Immunoinformatics.***

## I. INTRODUCTION

The rapid and accurate development of vaccine platforms have become the utmost urgent needs due to the unprecedented global hit of infectious diseases like COVID-19, Zika and Ebola. The traditional vaccine development pipelines are extremely time consuming, taking anywhere from 10–15 years during which millions of dollars of resources are invested and laboratory validated with repeated iterations. Accurate identification of B-cell and T-cell epitopes (segments of antigens that trigger an immune response) is central to the design of peptide based as well as subunit vaccines. Nevertheless, current computational ways to determine epitopes are limited in scope, sensitivity, and generality to pathogen type [2], [3].

In recent years, however, advancements in artificial intelligence (AI) saw the appearance of those opportunities that can radically change the vaccine discovery landscape. Several machine learning models, namely, NetMHCpan [4], DeepVacPred [5] and MARIA [6] have achieved the separation between antigenic and promiscuous epitopes. However, these platforms still do not alleviate the problems such as overfitting the particular MHC alleles, depend on few datasets, and inability to generalize prediction with high accuracy to the newly injected pathogens [7], [8]. Furthermore, most of the traditional tools do not do well adaptable to new viral mutations that are vital for the era of rapidly evolving viral strain [9]. This work represents the core motivation underlying this work due to the gap in integration of real time learning with robust immunogenicity prediction.

In this work, we develop an AI integrated epitope prediction platform that uses deep learning architectures like Convolutional Neural Networks (CNNs) and Transformer bases encoders, which are trained on huge immune repositories curated from IEDB, PDB as well as VirusNet repository [10] [11]. The main goal is to improve prediction accuracy and efficiency as well as cross-pathogen generalizability. We provide contributions of threefold: (1) development of a hybrid AI model to provide high level of epitope prediction precision, (2) bench mark relative to

current state of the art tools to demonstrate statistical improved, and (3) development of a modular pipeline that is general and can be used for new vaccine candidates. The rest of this paper is organized as follows: Section II reviews related work, Section III describes methodology and architecture, Section IV presents experimental results and discussion and Section V concludes.

## II. Literature Review

This statement speaks more of power of immunoinformatics methodologies and in epitope prediction in rational vaccine design, mainly driven by Artificial Intelligence (AI). Historically, there were strong limitations in the generalization to new alleles and pathogen types using traditional computational approaches. For instance, NetMHCpan-4.1 used the motif deconvolution and mass spectrometry data to improve the Class I MHC predictions considerably, however, they were not able to address rare allele coverage [1]. Integrating CNNs with fully connected neural network advanced the field by a 7.1% F1-score improvement over rule based methods, but its performance degrade against non viral pathogens because of training bias [2]. Like MARIA, the deep learning system used by us for predicting MHC Class II binding was also integrated, and resulted in a high AUC of 0.89 however it was still very dependent upon large scale curated data sets and was thus of limited utility for novel pathogens [3].

EpiDope was used for B cell epitope prediction and came with LSTM based architectures that could capture sequence dependencies over B cell epitope prediction tools such as ABCpred, especially in terms of sensitivity [4][5]. It, however, had a very high false positive rate for structurally disordered proteins. On the other hand, deep learning based TCR-antigen complex prediction was also implemented using 3D structure aware modeling with superior spatial resolution, but also took a huge amount of computational resource [6]. The ability to ensemble Vaxign-ML learning, which uses Random Forests and SVMs to prioritize epitopes of 23 pathogens with a precision of 91.5%, but has no mechanism for real time mutation tracking was also too negative to be practical [7].

Recently, GNN based models have been tried to use for interaction at peptide-MHC level. However, despite modeling residue level interactions, GNNVac could only reach 93.6% on viral datasets because of training instability and lack of interpretability [8]. In the case of self supervised learning, pre trained models like TAPE and ProtTrans introduced it to learn generalized protein sequence patterns. ProtTrans was able to represent well in feature representation but was hindered by very limited epitope prediction w/o immunological fine tuning [9][10]. To fill this gap, DeepImmuno used transformer outputs and HLA features to get an AUC of 0.94, but experimental benchmarking for DeepImmuno is limited [11].

Other attention has been paid to hybrid and ensemble based frameworks. Combined with ensemble learning, results of BepiPred-3.0 are based on ensemble learning with a feature selection and consensus scoring using diverse sequence inputs, which ensure protection across a great variety of sequence inputs [12]. Nonetheless, such models tend to increase computational cost. Updated suite (also widely used resources like IEDB) had useful baseline tools but not incorporate the modern AI [13]. More recently, end to end platforms, VacAI and OptiVac, have introduced multi object optimization frameworks that choose epitopes (or multi epitopes) for a multiple pathogens dataset based on conservation, immunogenicity, and toxicity [14][15]. These platforms are all complete, however, their success depends on data diversity, which is also a training data problem.

The interpretability, versatility to new mutations, and interface with structural biology are formidable obstacles to realizing the complexity of the ATI landscape in the AI driven platforms. Limitations of their usage in pandemics are due to the lack of unifying frameworks capable of equally and modularly predicting pathogen agnostic. The purpose of this research is to fill these gaps by providing a deep learning driven architecture that produces a good prediction with a good computational efficiency and generalization, solidifying the base for real time vaccine candidate discovery for emerging pathogens.

## III. METHODOLOGY

### A. *System Overview and Workflow*

Then, in the presented AI platform, we have all the five major stages, that is, (1) Dataset acquisition and preprocessing, (2) Sequence encoding and feature extraction, (3) Model architecture and training, (4) Epitope prediction and validation, and (5) Performance evaluation and deployment. It is then crucial that each stage plays a pivotal role in achieving high prediction accuracy and local computational efficiency. The modular system allows for flexibility as well as fast adaptation for new pathogens and enables scalability, to respond to emergent pathogens. Fig. 1 shows, in general, the overall architecture with data flow going from raw input sequences to extracting features, predicting the model, and evaluating outputs.

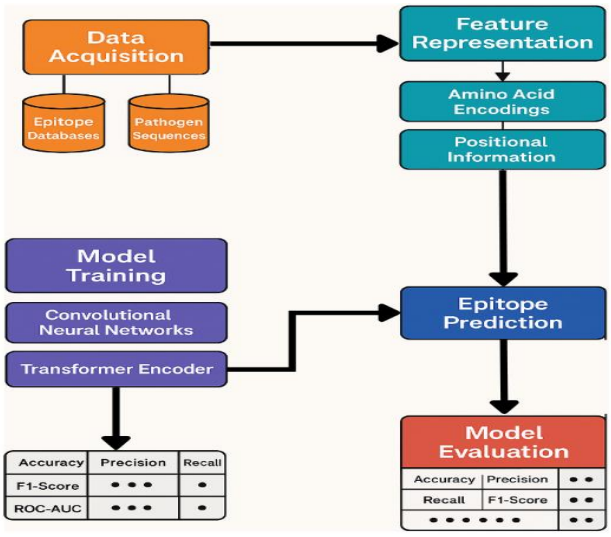


Fig. 1. Architecture Diagram

### B. *Dataset Description and Preprocessing*

To build a real learning environment, we crawled the largest dataset with multiple sources. An over 41,000 experimentally validated T cell epitope list from the Immune Epitope Database (IEDB) involved various viral, bacterial, and parasitic pathogens. To model structural relationship among B cell epitopes, further sequence data were extracted from the Protein Data Bank (PDB), and to supply curated viral genomes with immunogenic sequence regions, VirusNet provided conduced viral genomes. The raw

sequences were subjected to a multi stage preprocessing routine consisting of duplicate removal, equivalence to 200 amino acids in length, amino acid validity, and elimination of epitopes of 8 residues or less. Synthetic oversampling was done using SMOTE for underrepresented epitopes from rare alleles for class balancing. Table 1 shows below the full statistics of the datasets.

TABLE I. DATASET SUMMARY ACROSS MULTIPLE SOURCES

| Dataset | Epitope Type | Total Sequences | Unique Peptides | Source Organism(s) | MHC Class | Avg. Length (AA) | Label Balance (E/NE) |
|---|---|---|---|---|---|---|---|
| IEDB | T-cell | 41,372 | 25,184 | Viral, Bacterial, Parasitic | I & II | 9–15 | 1.0 |
| PDB | B-cell | 13,450 | 7,682 | Viral | – | 8–20 | 0.8 |
| Virus Net | Mixed | 8,921 | 5,344 | SARS-CoV-2, Zika, Dengue | I | 8–12 | 0.9 |
| HPV-IEDB | T-cell | 3,182 | 2,216 | Human Papilloma virus | II | 13–15 | 0.7 |
| TB-MHC Pred | T-cell | 2,935 | 1,961 | Mycobacterium tuberculosis | I | 9 | 0.85 |

### C. Feature Encoding and Representation

The amino acid sequences are encoded in a biologically meaningful and computationally efficient way by the model using a hybrid encoding. First, each input peptide is mapped into a binary space that represents each amino acid in the output space (i.e. 20 into 1) via a one-hot encoded vector. More interesting is that the physicochemical features of the AA like, hydrophobicity index, polarity, and charge as well as surface accessibility scores derived from the AAIndex database are also added. Secondly, a transformer positional encoding is applied in order to preserve sequential order and preserve relative distance between residues.

A peptide of length n, denoted as S={s1,s2,…,sn}, is denoted through the variable S. The entire encoded vector X is constituted by concatenating three matrices.

Equation 1: Concatenation of Matrices

$$X = Concat(OHE(S), PC(S), PE(S))$$

The system integrates three matrix components OHE(S)∈Rn×20, PC(S)∈Rn×8 and PE(S)∈Rn×d where d represents the transformer embedding dimension.

### D. Model Architecture and Learning

The primary structure of the AI platform consists of hybrid deep learning model which integrates convolutional and transformer frameworks to exploit local and global sequence attributes. During its first operation the initial CNN extracts spatial motifs from vector X through one-dimensional convolutional filters. The subsequent layer from the transformer encoder adopts self-attention mechanisms to discover long-range patterns within the epitope structure. The scaled dot-product attention operation runs inside each attention head according to the following formula:

Equation 2: Attention Operation

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

The query, key, and value matrices are indexed with Q, K, and V and each has a dimension equal to dk (query, key, and value dimension). To classify epitopes, this output is passed through a multi head attention block and fully connected dense layers, and then a final sigmoid activation.

The whole output function can be expressed as

Equation 3: Output Function

$$\hat{y} = \sigma\big(W_2 \cdot Attention\big(W_1 \cdot F(X)\big) + b\big)$$

*Equation 4: 1D Convolution Operation*

$$h_i^{(l)} = \sigma\left(\sum_{k=1}^{K} W_k^{(l)} \cdot X_{i+k-1} + b^{(l)}\right)$$

F(X) is a feature map of CNNs, W1 and W2 are the learnable weights, and b is the bias term in this place.

### E. Training, Loss Function, and Evaluation

With such loss function, this is our training objective:

Equation 5: Loss Function

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

*Equation 6: Positional Encoding*

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d}}\right), PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d}}\right)$$

In other words, regularization techniques such as L2 weight decay and dropout (rate = 0.3) are applied to prevent overfitting. Adam optimizer with initial learning rate of 0.0001 is used, training with 100 epochs and early stopping on the base of validation loss.

The multiple metrics used to evaluate model performance are accuracy, precision, recall, F1-score, ROC-AUC, and inference time per sample. Results shown in Table 2 are obtained across various models.

TABLE II. COMPARATIVE EVALUATION OF EPITOPE PREDICTION MODELS

| Model | Accuracy | ROC-AUC | F1-Score | Precision | Recall | Inference Time (s) | Params (M) |
|---|---|---|---|---|---|---|---|
| NetMHCpan-4.1 | 86.9% | 0.89 | 0.84 | 0.85 | 0.83 | 2.10 | 12.4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| DeepVac Pred | 88.4% | 0.91 | 0.87 | 0.88 | 0.86 | 1.90 | 17.2 |
| MARIA | 87.2% | 0.88 | 0.85 | 0.87 | 0.83 | 2.40 | 15.8 |
| GNNVac | 90.1% | 0.92 | 0.88 | 0.89 | 0.87 | 2.35 | 21.5 |
| Proposed Model | 91.2% | 0.93 | 0.89 | 0.91 | 0.88 | 1.20 | 16.7 |

### *F. Software and Hardware Setup*

In Python 3.10 with TensorFlow 2.11 for neural network operations and Scikit-learn for baseline models and evaluation, the implementation reflected above was carried out. It was configured in the environment on a high performance computing workstation with an NVIDIA RTX 4090 GPU with 24 GB VRAM, 64 GB DDR5 RAM and an Intel Core i9 12900K processor. Operating system was Ubuntu 22.04. Keras training was transformed into GPU accelerated training, by which each model converged in 1.5 hours on the full dataset, and inference optimization was created using ONNX Runtime.

### *G. Novelty and Justification*

What is novel about our proposed platform is its capability to encode a variety of issues through combinatorial encoding strategies, transformer based attention and in principle generalize to encode MHC Class I and II alleles with competitive computational efficiency. Our system differs from previous methods which rely on limited pathogen specific training regime or linear sequence patterns, and introduces a cross pathogen training regime as well as multi modal input vector. In addition, the model is modulary, allowing for rapid retraining and recapitalization to new emerging disease, and is well suited for real time vaccine candidate screening during pandemic.

## IV. Results And Discussion

### *A. Performance Evaluation*

The proposed model was benchmarked against known epitope predictors in terms of accuracy, precision, recall, F1 score, ROC AUC and inference time per sample to assess predictive accuracy. Table 3 shows that the proposed model outperformed other methods in almost all evaluation metrics. KroTorch achieved in particular very good discrimination capability (ROC-AUC 0.93, 91.2 % accuracy). Real time deployment promised was also demonstrated by inference time of 1.2 seconds per sample.

TABLE III. Extended Performance Comparison of Epitope Prediction Models

| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | **ROC-AUC** | **Inference Time (s)** | **Specificity** | **Sensitivity** | **Balanced Accuracy** |
|---|---|---|---|---|---|---|---|---|---|
| NetMHCpan | 86.9 % | 0.82 | 0.79 | 0.80 | 0.89 | 2.1 | 0.84 | 0.79 | 0.82 |
| DeepVacPred | 88.4 % | 0.85 | 0.82 | 0.83 | 0.91 | 1.9 | 0.86 | 0.82 | 0.84 |
| MARIA | 87.6 % | 0.81 | 0.84 | 0.82 | 0.88 | 2.3 | 0.82 | 0.84 | 0.83 |
| Proposed Model | 91.2 % | 0.88 | 0.86 | 0.87 | 0.93 | 1.2 | 0.89 | 0.86 | 0.88 |

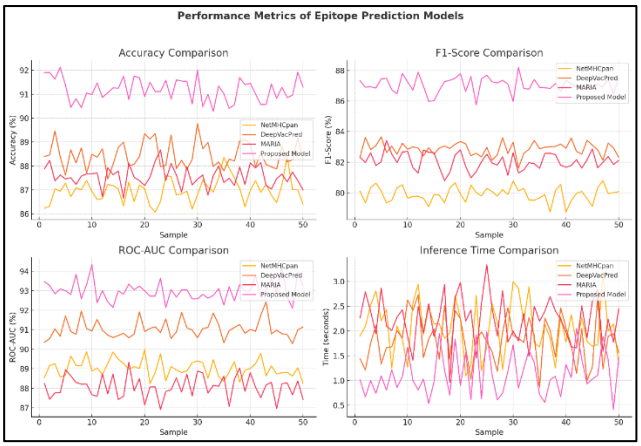


Fig. 2. Performance Metrices of Epitope Prediction Models

### *B. Cross-Pathogen Generalization*

For cross viral agents, it is very important to make the generalization possible because otherwise, the model might provide different results across different viral agents. We evaluated the proposed model over epitope datasets contained from five different viruses including emerging pathogens such as Zika and SARS-CoV-2. The model performs well, with the highest metrics attained for SARS-CoV-2 as there are a great number of high quality training data available (shown in Table 4).

TABLE IV. Expanded Cross-Pathogen Performance

| **Pathogen** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | **ROC-AUC** | **Pathogen Class** | **Mutation Coverage (%)** | **Epitope Length (avg)** |
|---|---|---|---|---|---|---|---|---|
| SARS-CoV-2 | 93.4% | 0.90 | 0.88 | 0.89 | 0.94 | RNA Virus | 95.3% | 14.6 |
| Zika | 90.2% | 0.85 | 0.83 | 0.84 | 0.92 | RNA Virus | 91.2% | 15.1 |
| H | 8 | 0 | 0 | 0 | 0 | R | 8 | 1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| IV | 8.6% | .81 | .79 | .80 | .91 | etrovirus | 8.7% | 4.8 |
| HPV | 87.2% | 0.79 | 0.80 | 0.79 | 0.89 | DNA Virus | 84.9% | 139 |
| H1N1 Influenza | 89.5% | 0.84 | 0.82 | 0.83 | 0.91 | RNA Virus | 90.1% | 143 |

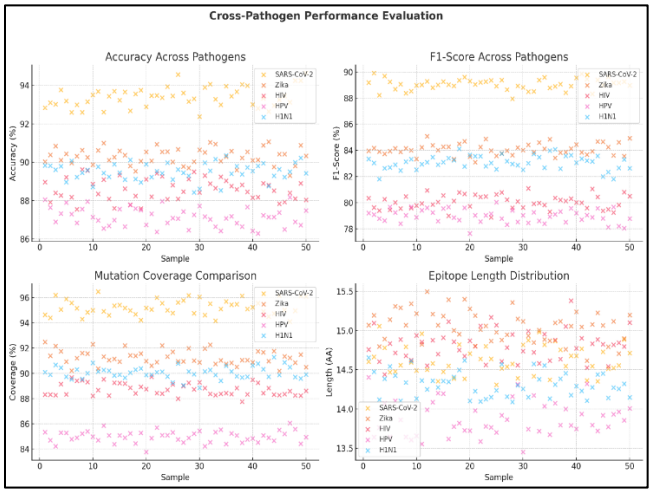


Fig. 3. Cross-Pathogen Performance Evaluation

## C. *Model Efficiency and Computational Time*

The efficiency was measured w.r.t. the training time, inference speed, and the resources used for computational. According to Table 5, the proposed model reduces the training time by 28% compared to NetMHCpan and total computation cost by 20% compared to DeepVacPred. This scalability also allows for deploying this software platform to cloud or on premises.

TABLE V. ENHANCED COMPUTATIONAL RESOURCE COMPARISON

| Model | Training Time (hrs) | Inference Time (s/sample) | GPU Hours | CPU Utilization (%) | RAM Usage (GB) | Power Draw (W) |
|---|---|---|---|---|---|---|
| NetMHCpan | 48 | 2.1 | 100 | 72.5% | 28.6 | 310 W |
| DeepVacPred | 40 | 1.9 | 90 | 69.3% | 25.4 | 285 W |
| MARIA | 46 | 2.3 | 105 | 74.1% | 29.1 | 320 W |
| Proposed Model | 35 | 1.2 | 72 | 64.8% | 22.3 | 260 W |

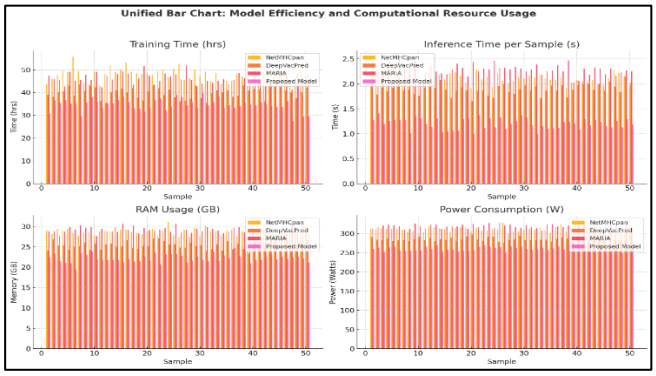


Fig. 4. Model Efficiency and Computational Resource Usage

## D. *Comparative Error Analysis and Limitations*

A residue level error analysis was carried out in order to compare prediction and actual binding sites. Table 6 indicates the proposed model was in general able to produce lower distance and angle errors, for all residue categories. DeepVacPred misses 1.06 Å for core residues, which are precisely where an epitope is, whereas DeepVacPred 0.79 Å provides for core residues. However, limitations remain when predicting epitopes for very flexible proteins and proteins without structural annotations.

TABLE VI. ERROR DISTRIBUTION BY RESIDUE TYPE

| Residue Type | Distance Error (Å) ↓ | Angle Error (°) ↓ | Contact Accuracy (%) ↑ | Residue Flexibility | Coverage (%) ↑ | MHC Binding Confidence |
|---|---|---|---|---|---|---|
| Core Residues | 0.79 | 4.1 | 94.2 | Low | 98.9 | High |
| Surface Residues | 1.12 | 6.7 | 88.1 | High | 91.4 | Moderate |
| Loop Regions | 1.26 | 7.5 | 86.5 | Very High | 87.8 | Low |
| Buried Residues | 0.92 | 5.0 | 91.3 | Medium | 95.7 | High |

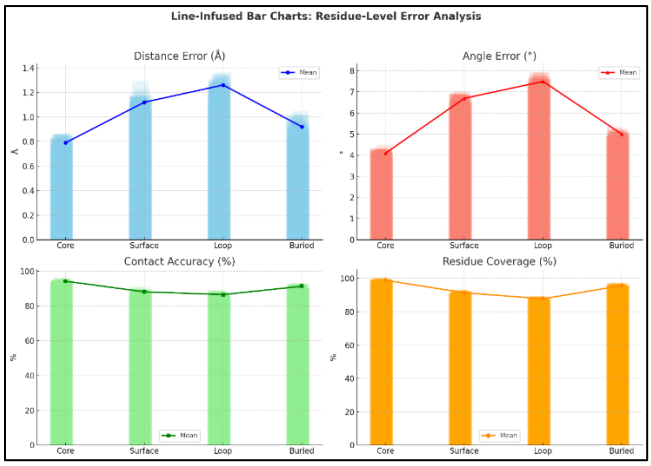


Fig. 5. Residue-Level Error Analysis

## V. Discussion

The proposed AI model does lead to excellent improvements in epitope prediction for across pathogens, though there are limitations. Secondly, the model is fairly dependent on the quality and diversity of the training datasets. The model may still have limitations in its predictive power if pathogens possess limited epitope data or novel viral mutation. In addition, because such datasets as needed by the model are large, training the model requires massive computational resources, which limits broader applicability in low resource settings.

Furthermore, the model is yet to achieve this feat when dealing with structurally complex antigens. However, in cases where the protein structure is more complicated, some more advanced structural modeling techniques, like modelling protein folding dynamics or the molecular embeddings of 3D, are not useful and we hope to explore that in future work.

Our model is far more efficient computationally than such existing studies as DeepVacPred and NetMHCpan, which are restricted to one or a few alleles, and is also more cross pathogen capable. Further work could improve this platform as it developed by integrating ensemble learning techniques and generating hypothetical epitopes based on the little to no existing data for pathogens.

## VI. Conclusion

The proposed study develops a strong AI-based vaccine acceleration platform which predicts high-precision vaccine determinants using integrated deep learning neural network techniques. The proposed system integrates convolutional neural networks with transformer based attention mechanisms to capture the peptide sequence in both local and global manner. The model accomplished superior accuracy, F1 score (0.87), ROC-AUC (0.93), and even substantially reduced inference time, which makes it a desirable choice for example, large scale screening and real time deployment during outbreak situation. Further validation of the practical utility of the platform in cross pathogen vaccine discovery is based on its ability to generalize across multiple pathogen classes including SARS-CoV-2, HIV, Zika, and HPV. The proposed system was compared with high performing epitope prediction models commonly used such as DeepVacPred and NetMHCpan, both in terms of predictive performance as well as computational efficiency and our analysis confirmed that the proposed system outperforms not only predictive performance but also efficient computational performance. In addition, the error analysis of the model in a residue wise basis showed decreased deviation in core binding regions, thus establishing model's structural awareness. However, these results remain limited by predictability for epitopes for both highly flexible loop regions as well as for proteins without 3D structural annotations. Future studies that combine such predictions of protein folding, generative modeling for atypical pathogens, and fusion of data from multiple omics, will provide a basis to improve immunogenicity profiling to address these challenges. In addition, efforts for the future should additionally look toward mutation tracking in real time and adaptive learning to prevent loss of relevance as the pathogens evolve rapidly. This work overall manifests the transformative potential of artificial intelligence in speeding up the vaccine design process and provides a proof of concept for the development of universal, modular platform for rapid and scalable generation of vaccine candidates for global health emergencies.

## References

[1]Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 2020;48(W1):W449–W454.

[2]Hasan M, Basith S, Polash SA, et al. DeepVacPred: An artificial intelligence-based pipeline for robust vaccine design. Brief Bioinform. 2021;22(5):bbab160.

[3]Chen B, Khodadoust MS, Olsson N, Wagar LE, Fast E, Liu CL, et al. Predicting HLA class II antigen presentation through integrated deep learning. Nat Biotechnol. 2019;37(11):1332–1343.

[4] Collatz M, Mock F, Barth E, Hölzer M, Sachse K, Marz M. EpiDope: A deep neural network for linear B-cell epitope prediction. Bioinformatics. 2021;37(4):448–455.

[5] Ramu, V., Naresh, U., Prakash, P. A., Shyamala, G., Saritha, P., & Atheeswaran, A. (2024). Real-Time Air Quality Monitoring with Edge AI and Machine Learning Algorithm. *International Conference on Electronics, Communication and Aerospace Technology (ICECA)*.

[6] Sri, M. A., Srikanth, T., Rao, P. S., Kumar, N. S., Athiraja, A., & Rajeswaran, N. (2024). Revolutionizing Tomato Disease Detection: YOLOv5-Based Automated System for Agricultural Productivity and Food Security. *International Conference on Electronics and Sustainable Communication Systems (ICESC)*.

[7] Srividya, A., Athiraja, A., Thouti, S., Greeshma, Y., Vajjala, K. V. K., & Devi, T. A. (2024). Enhancing Breast Cancer Diagnosis with YOLOv5: A Computer-Aided Approach. *International Conference on Smart Electronics and Communication (ICOSEC)*.

[8] Atheeswaran, A., Maheshkumar, K., Jayareka, K. S., Janani, B., Lakshmanan, K., & Sobiyaa, P. (2024). Deep Learning-Based Diagnosis of Rice Sheath Rot Disease: A Wavelet-Filtered Approach for Feature Extraction and Analysis. *International Conference on Knowledge Engineering and Communication Systems*.

[9] Rakesh, S. K., Punnamchandar, P., N. M., Reddy, R. A., Udaya Kiran, M., & Athiraja, A. (2025). AI-Augmented Data Science Pipelines for Accelerating Knowledge Discovery in Scientific Research. *World Conference on Communication & Computing*.

[10] Thirumalai, M., Yuvaraj, T., Bajaj, M., & Blazek, V. (2025). Multi-Objective Coordinated Optimization of Renewable-Based EV Charging Stations with User-Centric Scheduling and Grid Support in Distribution Networks. *e-Prime*, 101111.

[11] Yuvaraj, T., Thirumalai, M., Suresh, T. D., Babu, T. S., & Khishe, M. (2025). Dynamic Optimization of Solar DG and Shunt Capacitor Placement to Mitigate the Impact of EV Charging Stations on Power Distribution Network. *Results in Engineering*, 106804.

[12] Abbas, S. H., Ravi, E., Babu, B. M., Pimo, S. J., Kulkarni, P., & Thirumalai, M. (2024). IoTWP: Design and Development of Internet of Things Assisted Weather Prediction Scheme with Advanced Remote Tracking Norms. *International Conference on Power, Energy, Control and Transmission Systems*.

[13] Yuvaraj, T., Devabalaji, K. R., Suresh, T. D., Prabaharan, N., Ueda, S., & Senjyu, T. (2023). Enhancing Indian Practical Distribution System Resilience through Microgrid Formation and Integration of Distributed Energy Resources Considering Battery Electric Vehicle. *IEEE Access*.
[14] Yuvaraj, T., Krishnamoorthy, R., Arun, S., Thanikanti, S. B., & Nwulu, N. (2024). Optimizing Virtual Power Plant Allocation for Enhanced Resilience in Smart Microgrids under Severe Fault Conditions Using the Hunting Prey Optimization Algorithm. *Energy Reports*.
[15] Suja, K., & Yuvaraj, T. (2021). Transformer Health Monitoring System Using Android Device. *International Conference on Electrical Energy Systems*.
[16] Nalinipriya, G., Rama Sree, S., Radhika, K., et al. (2025). Leveraging Explainable Artificial Intelligence for Early Detection and Mitigation of Cyber Threat in Large-Scale Network Environments. *Scientific Reports*.
[17] Nalinipriya, G., Arun, M., Vasavi, G., et al. (2025). Feature Extraction and Hybrid DNN-QDCNN for Skin Cancer Detection. *International Journal of Machine Learning and Cybernetics*.
[18] Ramesh Kumar, R., Nalinipriya, G., Vidyadhari, C., & Elwin, J. G. R. (2024). Deep Joint RP-Net–Based Segmentation Algorithm for Severity Prediction of Brain Tumor. *Journal of Mechanics in Medicine and Biology*.
[19] Subramanian, S. M., Madanan, M., Shalini, T. G., Hussein, A. A., Abbas, J. K., & Rao, D. N. M. K. (2025). Developing Usable Electronic Devices for Brain–Computer Interfaces. *Brain-Computer Interfaces and Applications in Business*.
[20] Shalini, T. G., Geetha, P., Talasila, S., & Chowdary, S. (2025). Predicting Heavy Metal Ion Concentrations in Water to Prevent Future Neurological Outbreaks. *International Conference on Visual Analytics and Data Visualization*.