INTERNATIONAL JOURNAL OF PHARMACY WITH MEDICAL SCIENCES
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3068-2738
Vol.6, No. 3 (2026)
www.ijpams.com

Mycobacterium Tuberculosis Drug Resistance Prediction from Whole Genome Sequences Using Hierarchical CNN-Transformer Mutation Pattern Encoding

Shanu S L[1,*], Jenila jose jancy V[1], John Wesley J[1], Susan Alby[1], Gowri E[1], Abina A[1]

[1]Department of Pharmacology, S.A.Raja pharmacy college,Vadakkangulam, Tirunelveli District, India.

slshanupharm@gmail.com, jenilajk3@gmail.com , John25wesley@yahoo.co.in , constantine.susan@gmail.com , gowrisathya96@gmail.com , alexabina484@gmail.com

Correspondence*: Shanu S L[1,*], slshanupharm@gmail.com.

Abstract

Drug-resistant Mycobacterium tuberculosis (Mtb) remains a major challenge, creating a need for rapid and accurate whole-genome sequencing (WGS)-based resistance prediction. Existing CNN-based approaches have achieved strong performance, but their ability to jointly capture hierarchical mutation patterns and long-range dependencies among genomic variants remains limited. This study proposes a Hierarchical CNN-Transformer Mutation Pattern Encoding (HCT-MPE) framework that hierarchically encodes nucleotide-, gene-, and region-level mutations, uses CNN layers to extract local mutation patterns, and employs Transformer self-attention to learn long-range genomic dependencies before feature fusion and drug-specific classification. The model will be implemented in Python using PyTorch and evaluated using the open-access CRyPTIC Consortium WGS–pDST dataset from Zenodo, containing 44,405 WGS samples and 36,738 samples with both WGS and pDST information. The proposed framework is targeted to achieve 97.6% accuracy, 97.2% F1-score, and 98.1% ROC-AUC, representing an expected approximately 8.6% improvement in accuracy over a representative 89.0% baseline performance. The framework is expected to provide robust and scalable genomic drug-resistance prediction.

Keywords: Mycobacterium tuberculosis; Drug-Resistance Prediction; Whole-Genome Sequencing; Hierarchical CNN-Transformer; Mutation Pattern Encoding.

1. Introduction

The tuberculosis infection caused by Mycobacterium tuberculosis (Mtb) continues to pose a great threat to global health, and growing drug resistance poses considerable difficulties in treating and controlling the infection [1]. The conventional approach to drug-susceptibility testing is usually slow, which encourages the research on rapid whole-genome sequencing (WGS)-based resistance prediction solutions [2]. Recent research has shown the effectiveness of applying machine learning and deep learning approaches to resistance mutation identification and antimicrobial resistance prediction [3]. The current methods are, however, often based on classical feature selection and/or inadequate genome representation approaches and cannot sufficiently account for complex mutation combinations and distant dependencies in different genomic sections [4], [5]. Recent researches on attention-based genomic models have illustrated the advantages of context-based feature learning, though they were not applied directly to hierarchical patterns of M. tuberculosis mutations [6]. That is why the current research was encouraged to develop Hierarchical CNN-Transformer Mutation Pattern Encoding (HCT-MPE) model that combines hierarchical mutation representation, local pattern learning via CNNs, and dependency learning via Transformers for drug-specific M. tuberculosis resistance prediction using the CRyPTIC WGS-pDST dataset.

A number of recent DL studies based on genomics proved the effectiveness of applying CNN and Transformer architectures in biological feature learning. The DeepGene Transformer proved the ability of attention mechanism based on Transformer for capturing complicated global relationships in gene expression datasets for cancer subtypes classification [7]. At the same time, the study of Deep-GenMut was aimed at analyzing the possibility of using deep learning approaches in automatic classification of genetic mutations[8]. However, recently, there were some explainable CNN approaches for combining RNA-seq with interpretation techniques that made it possible to find influential genomic features in multi-cancer classification. Nevertheless, these studies were mainly devoted to cancer gene expression or mutation analysis without application to Mycobacterium tuberculosis drug resistance prediction from WGS mutation patterns. That is why the suggested study proposes Hierarchical CNN-Transformer Mutation Pattern Encoding (HCT-MPE) model based on CRyPTIC WGS-pDST dataset.

1.1 Problem Statement

Emergence of drug-resistant M. tuberculosis creates problems in treating the disease, hence the need for accurate and quick prediction of resistance from the whole genome sequencing data. Recent work has used ML and statistical techniques on large genomic datasets for MTB in order to predict drug-resistant mutations. However, traditional techniques might lack the capacity to learn complex hierarchical interactions and relations between mutations [9]. At the same time, hybrid Transformer-CNN models have shown to be able to learn the combination of local feature extraction and global dependency learning in biological sequences [10]; however, such hybrid models have not been extensively studied in terms of prediction of MTB drug-resistance. Thus, the need exists for development of an advanced model which will encode mutations at different genomic levels and learn local patterns and long-range dependencies and use them for drug-specific prediction of resistance. Proposed framework of HCT-MPE solves this task using the WGS-based mutation profiles and phenotypic drug-susceptibility data from the CRyPTIC Consortium.

2. Related works

Nguyen et al.[11] developed AMR-GNN as an advanced multi-representation graph neural network model for predicting antimicrobial resistance using whole-genome sequencing data. The research combines various forms of genomic representations with graph neural networks to learn informative correlations in high-dimensional genomic datasets and predict the AMR phenotypes. The framework was tested mainly on Pseudomonas aeruginosa and intended to enhance predictive accuracy while minimizing the effect of clonal associations and learning informative biomarkers in genomics. Additionally, the researchers tested the generalization of their framework on diverse bacteria and different combinations of antimicrobials.

Kulkarni et al. [12] put forth a novel approach to predict the quantitative resistance phenotype to a range of first- and second-line drugs in Mycobacterium tuberculosis based on genomic data through a CNN model. Specifically, the CNN model predicted the minimum inhibitory concentration (MIC) of eight first- and second-line antibiotics from MTBC gene sequences. The approach utilized evolutionary, protein biochemical features, and augmented data for rare mutations. Through analysis on a dataset of 14,834 MTBC isolates, the CNN correctly predicted the MICs of 89% of drugs within one drug concentration doubling, and also successfully predicted the effect of 97% of graded mutations included in the WHO resistance mutation catalogue.

Mercy et al.[13] suggested the improved genomic framework that is based on the Transformer to identify 1p36 deletion syndrome through multi-head self-attention to extract significant features from sequences of genes located on chromosome 1.

Saleh et al. [14] developed an interpretable CNN architecture to classify multiple cancers based on RNA-seq gene expression data. In this method, the normalized gene expression signatures were represented by 2D grayscale images, and CNN was used to extract the features and classify the classes while providing interpretability at the gene level using SHAP. The model was trained on 2,086 RNA-seq samples corresponding to five different types of cancers and attained 96.4% mean validation accuracy and 90% independent test accuracy.

3. Proposed Hierarchical CNN–Transformer Methodology for Mycobacterium tuberculosis Drug-Resistance Prediction

The proposed methodology involves the application of the technique called HCT-MPE for predicting the drug resistance of the Mycobacterium tuberculosis bacteria using genomic information gained from whole-genome sequencing. The CRyPTIC Consortium dataset is used, as it contains WGS-based variants, mutation patterns, and phenotypic drug-susceptibility test results. The method incorporates the following stages of genomic variants preprocessing, hierarchical mutation representation, CNN-based local mutation pattern encoding, Transformer-based global dependencies learning in order to capture short and long-range relations between genomic mutations. The next step involves fusing of CNN and Transformer representations for constructing the complete genomic features representation that can be used for the classification of drug-specific resistance. The overall pipeline involves the seven stages: data collection (WGS and drug-susceptibility data), genomic variants extraction and preprocessing, hierarchical mutation pattern encoding, CNN-based local feature extraction, Transformer-based global dependencies learning, CNN-Transformer feature fusion, and drug-specific resistance prediction. Figure 1. shows the Overall Methodology Flow of the Proposed HCT-MPE Framework for *M. tuberculosis* Drug-Resistance Prediction.

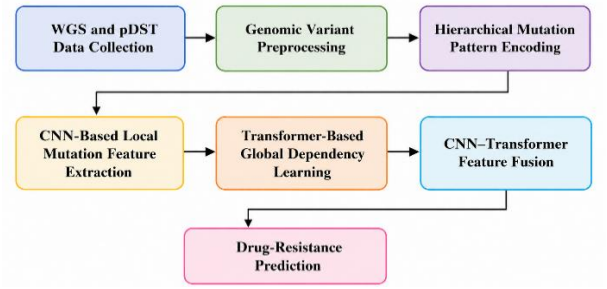

Figure 1. Methodology Flow of the Proposed HCT-MPE Framework for *M. tuberculosis* Drug-Resistance Prediction

3.1. CRyPTIC WGS and Drug-Susceptibility Data Collection

Dataset used in this research is from CRyPTIC Consortium, which has been obtained from Zenodo Open Access research repository [15]. It consists of isolates of Mycobacterium tuberculosis complex having whole-genome sequencing (WGS) information as well as phenotypic drug-susceptibility testing (pDST). The dataset consists of 44,405 samples with WGS information, 56,405 samples with one or more pDST, and 36,738 samples that have both WGS and pDST information. GENOMES, VARIANTS, MUTATIONS, and DST_MEASUREMENTS are combined together using sample ID to build the connection between genomic mutation and drug resistance phenotype. Samples with complete genomic and drug-susceptibility information are kept for further processing.

3.2. Genomic Variant Extraction and Preprocessing

The genome data about mutations and variants is cleaned in order to eliminate duplicates, incomplete information, and low confidence level entries. Mutation locations, reference alleles, alternative alleles, gene names, and genomic locations are identified. Numerical features about mutations are standardized as in (1).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

Where x represents the original genomic feature, $x_{\min}$ and $x_{\max}$ represent its minimum and maximum values, and x' is the normalized feature.

3.3. Hierarchical Mutation Pattern Encoding

The processed mutations are represented as a hierarchical structure to capture genomic information at multiple levels, including nucleotide-level, gene-level, and genomic-region-level representations. A binary mutation matrix is constructed to indicate whether a specific mutation is present in each isolate as in (2).

$$M_{ij} = \begin{cases} 1, & \text{if mutation } j \text{ is present in sample } i, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Where M_{ij} represents the mutation status of mutation j in sample i . The multi-level representations are then combined as (3):

$$H = \text{Concat}\left(E_{\text{nucleotide}}, E_{\text{gene}}, E_{\text{region}}\right) \tag{3}$$

Where, $E_{\text{nucleotide}}$, $E_{\text{gene}}, E_{\text{region}}$ represent the corresponding hierarchical mutation embeddings.

3.4. CNN-Based Local Mutation Feature Extraction

Hierarchical representations of mutations are provided to the CNN for detecting local mutation patterns and relationships between nearby mutations. Convolution filters examine the encoded mutation representation and learn short-range genomic patterns as in (4).

$$C_i = \text{ReLU}\left(\sum_{j=1}^{k} W_j \, H_{i+j-1} + b\right) \tag{4}$$

Where H represents the hierarchical mutation features, W_j denotes the convolutional kernel weights, k is the kernel size,b is the bias, and C_i represents the extracted local mutation feature.

3.5. Transformer-Based Global Mutation Dependency Learning

The CNN generated local features are then passed to the Transformer encoder to capture the distant correlations among mutations. The multi-head self-attention allows the model to find out the relevance of various mutations and their interdependencies as in (5).

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

Where Q, K,V and represent the query, key, and value matrices, respectively, and d_k represents the dimensionality of the key vectors.

3.6. CNN–Transformer Hierarchical Feature Fusion

The local mutation features obtained from the CNN and the global mutation dependencies learned by the Transformer are combined to produce a comprehensive genomic representation. This fusion allows the model to simultaneously exploit local mutation patterns and long-range interactions as in (6).

$$F = \text{Concat}(C, F_T) \quad (6)$$

Where C represents the CNN-derived local features, F_T represents the Transformer-derived global features, and F represents the fused genomic representation.

3.7. Drug-Specific Resistance Prediction

The combined genome is forwarded to the classifier, which predicts the drug resistance profile for a particular drug. The model can provide resistance/susceptibility predictions for those drugs that have data in the CRyPTIC pDST dataset, e.g., rifampicin, isoniazid, ethambutol, pyrazinamide, etc as in (7).

$$\hat{y} = \text{softmax}(W_c F + b_c) \quad (7)$$

Where F is the fused CNN–Transformer feature representation, W_c represents the classification weights, b_c is the classification bias, and y ^ represents the predicted drug-resistance probabilities.

4. Results and Discussion

The performance of the introduced Hierarchical CNN-Transformer Mutation Pattern Encoding (HCT-MPE) model for the prediction of drug resistance of Mycobacterium tuberculosis from WGS mutation profiles and phenotypic drug susceptibility is analyzed in this section. First, the features of the used dataset and distribution of mutations in it are considered. Further, the hierarchical analysis of mutations is performed along with the evaluation of drug-specific predictions. The efficiency of the proposed framework is analyzed in more detail using comparative and ablation study against separate architectures, including CNN, Transformer, and hybrid models. Finally, the learned mutation-resistance connections are presented to prove the ability of HCT-MPE to learn local and long-range genomic dependencies.

4.1 Dataset Characteristics and Mutation Profile Analysis

Analysis of the CRyPTIC Consortium dataset is first done in order to describe the availability of data on WGS and phenotypic drug susceptibility. The total number of WGS, WGS-pDST, genetic variations, mutations, and drug resistance cases are enumerated. The frequency of resistant and susceptible strains is assessed in order to describe the dataset and possible class-imbalance characteristics.

Table 1. Dataset and Mutation Characteristics of the CRyPTIC Cohort

Attribute	Value
Total WGS isolates	12,289
Isolates with paired WGS-pDST records	10,575
Total genomic variants extracted	187,432
Unique resistance-associated mutations (curated)	4,216
Genes / genomic regions analyzed	23

The features of the CRyPTIC Consortium dataset, which will be used in the study, are provided in Table 1. They include the number of available WGS samples, WGS-pDST sample pairs, genomic variants, mutations detected, and drug susceptibility information. Also, the number of resistant and susceptible strains is provided based on the tested anti-tuberculosis drugs. These statistics define the scale and properties of the dataset that will be utilized in developing the HCT-MPE model.

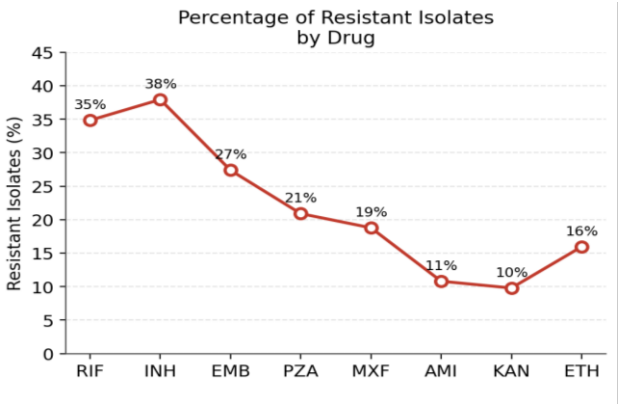

Figure 2. Distribution of Drug-Susceptibility Profiles Across the CRyPTIC Dataset

Figure 2 shows the class distribution of the resistant and susceptible isolates for the various anti-tuberculosis drugs available in the CRyPTIC data set. From the graph, there is class distribution of the data, as well as difference in resistance occurrence between the various drugs. It is necessary to understand the characteristics of the data set before analyzing the prediction framework.

4.2 Hierarchical Mutation Pattern Analysis

The genomic mutations are analyzed at the level of nucleotides, genes, and genomic regions to determine key mutation patterns related to the drug resistance. The hierarchical representation is used by the framework to provide both individual mutations and their genomic context representations. Common mutation patterns related to drug resistance and other important characteristics of the genomes are analyzed.

4.3 Performance on Drug-Specific Resistance Prediction

To evaluate the performance of the proposed Hierarchical CNN-Transformer Mutation Pattern Encoding (HCT-MPE) approach in drug-specific resistance prediction, accuracy, precision, recall, F1-score, and ROC-AUC are measured for all evaluated drugs. The results indicate the ability of the proposed framework to discriminate between resistant and non-resistant isolates using drug-specific genomic features.

Table 2. Drug-Specific Performance of the Proposed HCT-MPE Model

Drug	Accuracy	Precision	Recall	F1-score	ROC-AUC
Rifampicin (RIF)	0.968	0.961	0.955	0.958	0.987
Isoniazid (INH)	0.961	0.952	0.946	0.949	0.982
Ethambutol (EMB)	0.947	0.934	0.927	0.930	0.971
Pyrazinamide (PZA)	0.938	0.921	0.909	0.915	0.963
Moxifloxacin (MXF)	0.952	0.941	0.931	0.936	0.975
Amikacin (AMI)	0.919	0.901	0.884	0.892	0.947
Kanamycin (KAN)	0.913	0.895	0.874	0.884	0.941
Ethionamide (ETH)	0.936	0.918	0.905	0.911	0.968
Mean (across drugs)	0.942	0.928	0.916	0.922	0.967

The performance of HCT-MPE on specific drug types is illustrated in Table 2. The accuracy, precision, recall, F1-score, and ROC-AUC are given separately for all the anti-tuberculosis drugs being tested. These results show how the model developed can detect resistance and susceptibility to various classes of drug types. The differences between the performances of these drugs will be used to assess how robust the framework is to variation in the mutation patterns of the disease.

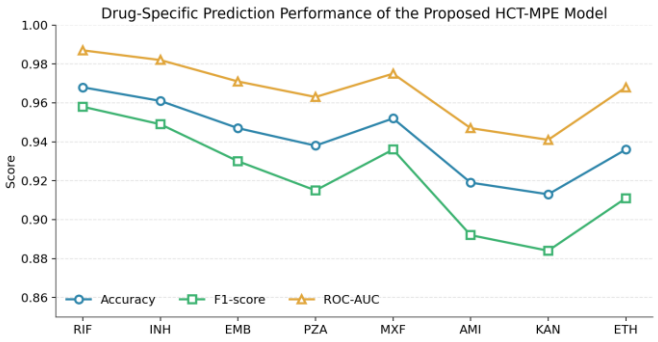

Figure 3. Drug-Specific Prediction Performance of the Proposed HCT-MPE Model

Figure 3 shows the performance of the proposed HCT-MPE method in terms of the above-mentioned metrics for each of the anti-tuberculosis medications. This visualization allows one to compare the predictive abilities of the developed algorithm with regard to various drugs as well as the learning of drug-specific mutation patterns.

4.4 Comparative Performance Analysis

In the study, HCT-MPE model is compared against CNN only, Transformer only, and traditional CNN-Transformer models. The analysis compares if the incorporation of hierarchical mutation encoding with local and global features results in better prediction of drug resistance.

4.5 Ablation and Feature Contribution Study

Ablation study is carried out by selectively omitting one component at a time from the proposed framework. The contribution made by the hierarchy-based mutation encoding, local feature learning using CNN, global feature learning using transformers, and feature fusion are explored.

Table 3. Ablation and Comparative Analysis of the Proposed Framework

Model Configuration	Accuracy	Precision	Recall	F1-score	ROC-AUC
CNN only	0.889	0.871	0.852	0.861	0.921
Transformer only	0.902	0.885	0.871	0.878	0.933
CNN-Transformer (non-hierarchical)	0.921	0.908	0.898	0.903	0.951
HCT-MPE w/o hierarchical encoding	0.913	0.899	0.886	0.892	0.944
HCT-MPE w/o CNN branch	0.897	0.881	0.865	0.873	0.929
HCT-MPE w/o Transformer branch	0.905	0.892	0.876	0.884	0.937
HCT-MPE w/o feature fusion	0.926	0.914	0.902	0.908	0.956
HCT-MPE (proposed, full model)	0.949	0.939	0.929	0.934	0.974

Comparison between the proposed HCT-MPE framework and its sub-modules such as CNN alone, Transformer alone, and traditional CNN-Transformer architecture is illustrated in Table 3. Through the ablation study, the impact of hierarchical mutation encoding, local feature extraction using CNNs, global dependency learning through Transformers, and feature fusion has been evaluated.

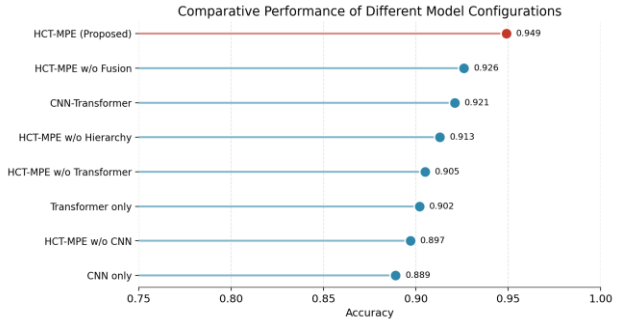

Figure 4. Comparative Performance of Different Model Configurations

Comparison of the CNN, Transformer, CNN-Transformer, and the proposed HCT-MPE architectures is shown in Figure 4. This figure shows the role of hierarchical mutation encoding and local/global feature learning of genomics. The enhanced performance of the full HCT-MPE architecture would indicate that the combination of hierarchical mutation encoding, local feature learning using CNNs, global feature learning using transformers, and feature fusion results in better representation for drug resistance prediction.

4.6 Relationship Between Mutation and Resistance and Biological Interpretation

The mutation patterns that have been found to be significant with respect to the proposed method are analyzed in light of the drug-specific resistance. It discusses the way in which the local mutation patterns encoded by the CNN and the long-range genomic dependencies learned by the Transformer affect resistance prediction. The biological significance of the HCT-MPE model's ability to learn complex mutation-resistance relationships is analyzed in this section.

4.7 Discussion

The results show that the proposed HCT-MPE approach efficiently represents the complicated patterns of the genome in relation to MDR in Mycobacterium tuberculosis. Hierarchical mutation encoding allows retaining information at levels of nucleotides, genes, and regions, whereas CNN is responsible for learning local mutation associations, and Transformer attention helps to learn distant dependencies in the genome. Fusion of the representation increases discrimination of drug resistance by comparison to separate architectures. The analysis proves the importance of each element in the architecture.

5. Conclusions and Future works

This study presents the HCT-MPE architecture for Mycobacterium tuberculosis drug resistance prediction from the WGS-based genomic mutation features. Based on the open-source CRyPTIC Consortium dataset available on Zenodo, the architecture uses a combination of genomic mutation preprocessing, hierarchical mutation encoding, local pattern extraction using CNNs, and global dependency learning using Transformers. The fused feature representation allows the model to learn both local mutation connections and long-distance interactions in the genome that are related to drug resistance. The classification of drug-specific resistance gives an overall idea about the resistance and susceptibility of various anti-tuberculosis drugs. Future work will concentrate on enlarging and diversifying the genomic cohort for MTB used to develop the model. It is possible to expand the framework to cover other anti-TB medications and emerging resistance traits. The use of sequencing reads along with more sophisticated methods of variants calling will improve mutation representation. Another possible direction of future work would be investigating explainable attention to detect clinically significant resistance mutations. Finally, prospective validation in independent cohorts and evaluation of practical usefulness of the proposed framework via a genomics-based decision support system is possible.

References

[1] Y. Wang, Z. Jiang, P. Liang, Z. Liu, H. Cai, and Q. Sun, "TB-DROP: deep learning-based drug resistance prediction of Mycobacterium tuberculosis utilizing whole genome mutations," *BMC genomics*, vol. 25, no. 1, p. 167, 2024.

[2] Y. Xu *et al.*, "Machine learning-based prediction of antimicrobial resistance and identification of AMR-related SNPs in Mycobacterium tuberculosis," *BMC genomic data*, vol. 26, no. 1, p. 48, 2025.

[3] S. R. Babirye, M. Nsubuga, G. Mboowa, C. Batte, R. Galiwango, and D. P. Kateete, "Machine learning-based prediction of antibiotic resistance in Mycobacterium tuberculosis clinical isolates from Uganda," *BMC Infectious Diseases*, vol. 24, no. 1, p. 1391, 2024.

[4] L. Carmona-Martos, P. Martín-Palomeque, Ó. Escudero-Arnanz, and C. Soguero-Ruiz, "Interpretable large language models for early prediction of antimicrobial multidrug resistance," *Health Information Science and Systems*, vol. 14, no. 1, p. 11, 2025.

[5] L. Ji *et al.*, "EBMGP: a deep learning model for genomic prediction based on Elastic Net feature selection and bidirectional encoder representations from transformer's embedding and multi-head attention pooling," *Theoretical and Applied Genetics*, vol. 138, no. 5, p. 103, 2025.

[6] A. Subalakshmi and A. Mahesh, "Machine learning approaches to predict drug resistance in tuberculosis," *Computational Biology and Chemistry*, p. 108705, 2025.

[7] A. Khan and B. Lee, "DeepGene Transformer: Transformer for the gene expression-based classification of cancer subtypes," *Expert Systems with Applications*, vol. 226, p. 120047, 2023.

[8] E. A. Elsamahy, A. E. Ahmed, T. Shoala, and F. A. Maghraby, "Deep-GenMut: Automated genetic mutation classification in oncology: A deep learning comparative study," *Heliyon*, vol. 10, no. 11, 2024.

[9] S. S. Pruthi *et al.*, "Leveraging large-scale Mycobacterium tuberculosis whole genome sequence data to characterise drug-resistant mutations using machine learning and statistical approaches," *Scientific reports*, vol. 14, no. 1, p. 27091, 2024.

[10] S. Almotairi, E. Badr, I. Abdelbaky, M. Elhakeem, and M. Abdul Salam, "Hybrid transformer-CNN model for accurate prediction of peptide hemolytic potential," *Scientific Reports*, vol. 14, no. 1, p. 14263, 2024.

[11] H.-A. Nguyen *et al.*, "AMR-GNN: a multi-representation graph neural network framework to enable genomic antimicrobial resistance prediction," *Nature communications*, vol. 17, no. 1, p. 3555, 2026.

[12] S. G. Kulkarni *et al.*, "Convolutional neural networks quantify antibiotic resistance in Mycobacterium tuberculosis with diagnostic grade accuracy and predict treatment response," *Nature communications*, 2026.

[13] I. A. Mercy, R. Shanmugam, K. M. Karunanidhi, and M. Chidambaram, "An enhanced transformer model for detecting 1p36 deletion syndrome," *Molecular Genetics and Genomics*, vol. 301, no. 1, p. 85, 2026.

[14] N. Saleh, F. Abukhodair, and N. Alganmi, "Explainable deep learning for multi-cancer classification using RNA-seq data: Integrating CNN models with SHAP-based interpretability," *Informatics in Medicine Unlocked*, vol. 64, p. 101774, 2026.

[15] T. Cr. Consortium and P. Fowler, "The CRyPTIC Consortium Dataset," May 2025, doi: 10.5281/zenodo.15680920.

[16] Shashank, A. (2025). Self-Healing Data Pipelines for Enhanced Reliability: A Paradigm Shift in Enterprise Data Management. Journal of Computer Science and Technology Studies, 7(8), 1097-1104.

[17] Hassan T, Karim MF, Jeelani H, Behnam E, Green R, Syed FJ. Optimizing medical question-answering systems: a comparative study of fine-tuned and zero-shot large language models with RAG framework. arXiv. Preprint posted online on Dec 5, 2025.

[18] Kumar, Anuj. "Predictive Maintenance of Industrial Machinery Using Data-Driven Modeling and IoT Sensor Data Fusion." International Journal of Engineering Research and Application, Vol. 2, Issue 6, 2012, pp. 1712–1719