Journal of Computer Science

Research Article

Deep Learning Synergy: CNNs and Transformers for Epidemic Outbreak Forecasting

Alexander R.[1], Shakthi Priya V.[1], Sumathi G.[2], Mary Valentina Janet A.[1], Reni Hena Helan R.[3] and Sinduja R.[1]

[1]*Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India*
[2]*Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, India*
[3]*Department of Computer Science and Engineering, Dhanalakshmi College of Engineering, Chennai, India*

Article history
Received: 05-07-2025
Revised: 13-07-2025
Accepted: 26-07-2025

Corresponding Author:
Shakthi Priya V.
St.Joseph's College of
Engineering, Chennai, India
Email:
shakthipriyav2906@gmail.com

Abstract: An Accurate prediction of outbreaks is extremely crucial for taking proactive public health interventions, distributing limited resources, and controlling a disease. This paper assesses and also compares the performances of the available deep learning models, namely, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) networks, and Bidirectional Long Short-Term Memory (BiLSTM) networks, in predicting. CNNs are best for feature extraction from medical data, while LSTMs and BiLSTM take care of temporal dependencies in sequential epidemiological data. These models have been shown to struggle at the integration of spatial, temporal, and contextual factors at once, yielding lowered predictive efficiency. A hybrid model, CNN-Transformer, leverages the spatial feature extraction ability of CNNs and the self-attention mechanism of Transformers to identify long-range dependencies and multi-source epidemiological patterns. Our approach integrates feature fusion techniques for abroader understanding of diseases' spread. Experimental results demonstrates that the proposed CNN-Transformer hybrid model outperforms standard CNN, LSTM, and BiLSTM architectures halfway through predicting outbreaks of diseases like COVID-19, Tuberculosis, Influenza, Dengue, and Measles. This study clearly illustrates the promise of hybrid deep learning models towards improving the accuracy of prediction of epidemics and the advancement of epidemic disease-surveillance systems. The time-series epidemic dataset is used for outbreak forecasting, and the hybrid model achieves an overall accuracy of 98.0%.

Keywords: Epidemic Prediction, Deep Learning Models, CNN-Transformer Hybrid Model, Self-Attention Mechanism, Feature Fusion

Introduction

The Infectious disease emergence and reemergence around the world provide major obstacles to global public health systems' ability to forecast these diseases and take suitable measures to eradicate them (Sankalpa *et al*., 2024). The unpredictable nature of those epidemic outbreaks makes it necessary to create reliable predictive models that can predict and mitigate the effects of these medical emergencies. Timely identification and prediction of epidemic patterns are essential for efficient distribution of resources, public health response, and disease prevention strategies. In previous periods, public health authorities have predicted illness outbreaks using statistical and epidemiological models (Rahman *et al*., 2023). But these approaches frequently fall inadequate in integrating large, complex information, which are essential for precise forecasting. Diagnostics for medicine has seen a transformation in recent years due to the development of deep learning or machine learning (Santangelo *et al*., 2023). With their exceptional ability to process and analyze the vast volumes of visual data, it facilitates implicit illness identification from medical imaging modalities like CT scans and X-rays (Abdollahi and Mahmoudi, 2022), however they are not particularly effective at capturing contextual information and temporal relationships, which are critical for predicting epidemic outbreaks.

SCIENCE Publications

© 2025 Alexander R., Shakthi Priya V., Sumathi G., Mary Valentina Janet A., Reni Hena Helan R. and Sinduja R. This open-access article is distributed under a Creative Commons Attribution (CC002DBY) 4.0 license.

The primary contributions of this study are threefold. First, it provides a focused investigation into established deep learning models, namely CNN, LSTM, and BiLSTM, for epidemic outbreak prediction, reviewing their respective strengths and weaknesses in capturing complex spatiotemporal dynamics. Building upon this analysis, we propose a novel hybrid CNN-Transformer model designed to more effectively contend with the intricate interplay of spatial, temporal, and contextual features, aiming to significantly improve prediction accuracy. Finally, to rigorously assess its performance, we conduct a comprehensive comparison of the proposed hybrid model against these existing deep learning architectures

This work demonstrates the potential for combining new architectures with existing models. The ability to interact efficiently in the sequential data using the new mechanisms can lead to significant successes in many domains, including healthcare. Some of the drawbacks of conventional approaches that have occurred in the past can be avoided by combining the multiple powerful approaches, which may enhance performance related to complex data interactions, and provide better performance (Chharia *et al.*, 2022). This research argues that the implementation of advanced deep learning architectures can improve epidemic prediction accuracy as a step towards better disease prevention and control.

Related Work

Epidemic Outbreaks

A sudden outbreak of an epidemic can be very dangerous as it begins with an alarming rise in infectious disease prevalence that can be found across the globe. These epidemics, in addition to adversely affecting health, also create great pressure on healthcare systems which may result in lack of staff, resources and treatment options. Increased tension and widespread terror, reduced productivity and lack of finances are some of the undesirable effects on social and economic systems and the response of health authorities needs to be appropriate and within a less time to lessen the impact of the epidemics (Pramod *et al.*, 2023). Public health officials must respond in a timely and an efficient way to limit the spread of the outbreaks. Correct estimates are crucial as they assist in deciding the timing of preventive interventions and awareness raising activities to prevent diseases from spreading across and to protect the targeted communities. Healthcare overloads are avoided due to precise prediction enabling better use of healthcare system resources and more effective management of community impact initiatives these predictions are necessary as they prevent unwanted situations.

Epidemiological Overview of Selected Diseases

This study emphasizes several infectious diseases that receive much focus because of their impact on public health. These diseases include COVID-19, Tuberculosis, Influenza, Measles, and Dengue, which constitute high-impact infectious diseases. Because of the infectious nature and fast rate of spreading through various modes, the diseases pose challenges that have been continuous to date. Thus, each disease necessitates special public health approaches and the detection at the right time to ensure outbreak control. Below is the overview of the selected disease and relevance in epidemic forecasting.

COVID-19

COVID-19 is a respiratory illness caused by the virus SARS-CoV-2, is highly contagious, first reported in late 2019. It has caused extreme socioeconomic and healthcare damage because of its wide spread all over the globe. Early and accurate prediction of the outbreak of COVID-19 is important for carrying out timely public health responses such as the locking down of the area, rollout of vaccination, and resource allocation. The model combines COVID-19 data to analyze patterns of spread and potential future outbreaks (Aslani and Jacob, 2023).

Tuberculosis (TB)

Tuberculosis (TB) is an infection caused by Mycobacterium tuberculosis; this infection majorly affects the lung and results in more deaths in many parts of the world despite some advances in treatment. Due to its form of transmission through the air route, early detection through medical images can help in averting the infection process, thus reducing its rapid spread. The slow yet steady comeback where some strains are multidrug resistant is testimony to the relevance of this disease for inclusion in epidemic prediction models, particularly of image-based diagnostics (Mirugwe *et al.*, 2025).

Influenza

Influenza, also known as flu, is an acute viral infection of the respiratory system. Influenza viruses cause epidemics every year and infect millions, hospitalizing millions more worldwide. Since influenza is very contagious and causes epidemics or pandemics, there is a strong case for building correct predictive models. Our model with influenza data concentrating on spatiotemporal dynamics which will help predict patterns of spread and predict seasonal outbreaks is discussed below (Watmaha *et al.*, 2024).

Dengue

Another of the well-known viral diseases is Dengue, caused mainly by the Dengue virus and transmitted mainly via Aedes mosquitoes. Its impact on health is strong and has vast effects in tropical and subtropical regions. Certain factors such as climate, urbanization, and population dynamics of mosquito have been amplified with the added complexity of prediction of any outbreak. This model uses the incidence data of Dengue using environmental factors like temperature and humidity for the betterment in predicting an outbreak and classifies vector-borne disease dynamics from other infectious diseases (Manoharan *et al*., 2023).

Measles

Measles is a highly contagious viral disease that has been eliminated through immunization, yet it recurs periodically and primarily in those regions where immunity levels are low. It causes widespread outbreaks with explosive severity, bogging down the health delivery system where vaccination programs are inappropriate. The model reconciles historical records of measles outbreaks for assessing the potential of early warning systems and vaccination campaigns using forecasting methods (Kujawski *et al*., 2024).

CNN in Epidemiology

Convolutional Neural Network (CNN) algorithms are highly efficient in analyzing visual data, like as CT scans and chest X-rays, and are designed to automatically learn spatial hierarchies from input data (Chimmula and Zhang, 2020). CNNs are used in epidemic prediction scenarios to detect patterns of disease that might point to the onset or spread of infections like COVID-19, Tuberculosis, Influenza, Dengue, Measles. CNNs have proven to classify the presence and severity of a disease accurately, which has reduced the time taken to diagnose it and helped track the rate of infections in populations. However, while CNNs work extremely well when we have image data in hand, they are incomplete for integrating other non-image sources necessary for epidemic prediction, like environmental or demographic data (Meraj *et al*., 2019). Therefore, the limitations by the former define a critical requirement for hybrid models where the architectures from the CNN can be combined with sequential models such as transformer architectures to better capture trends in epidemiology or increase the accuracy of prediction.

LSTM in Epidemiology

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) particularly adept at modeling temporal dependencies within sequential data to generate accurate predictions. The LSTMs have been great for using time-series to forecast some diseases very early and to provide advice on disease control (Wang *et al*., 2021). LSTMs have the advantage of capturing long-term dependences in disease data over time with rich history, environmental impacts, and human factors taken into account. The models that have been introduced show the accurate results in the prediction of the epidemic. This leads to earlier warnings and smarter decision-making about public health measures. Models have been applied to multivariate time-series collected from the WHO, CDC, and Google Trends datasets for the purpose of forecasting disease surges with low delay. However, the LSTMs can be limited in their ability to capture interactions between complex features so that hybrid models like using LSTMs incorporated with attention-based structures such as Transformers or CNNs are essential.

BiLSTM in Epidemiology

Bidirectional Long Short-Term Memory (BiLSTM) networks have the advantage over simple LSTMs as they can process the information both in the forward direction and in the backward direction, thus can capture both past and future dependencies in time series data. BiLSTMs are very useful to predict the inception, transmission, and the height of the infectious diseases like COVID-19, Influenza, Dengue, and Malaria. BiLSTMs outperform the standard LSTMs in cases where the correct reading of the previous and the next situation is the point of the task, such as disease transmission modeling, hospitalization forecasting and mortality rate prediction (Roster and Rodrigues, 2021). They have been up to now mainly utilized in datasets from sick public records, mobility data, and environmental factors to make the outbreak forecasts more accurate. Even though they are useful, BiLSTMs can face certain limitations, such as high computational costs and long training times, especially for large datasets.

The epidemiology examines patterns of diseases and their causes to predict and contain disease outbreaks, using algorithms, accuracy in trend detection, disease spread forecasting, and aiding early intervention and public health policies is improved. Figure 1 shows the performance accuracy of the existing models through a line graph of three algorithms; it shows the changing predictive accuracy of the algorithms and offers an insight into their comparative performance in epidemic prediction.

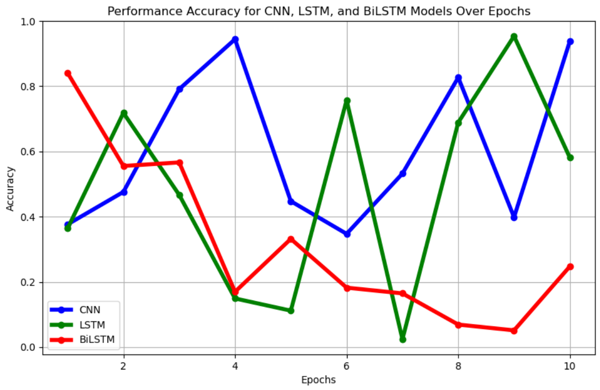

Fig. 1: Performance Accuracy of the Existing Model

Strengths and Weaknesses Using Deep Learning Models

Forecasting of epidemics using deep learning models has witnessed a significant boost, due to the fact that these models can be able to capture the intricate temporal dependencies that exist in the disease outbreaks. While the traditional time-series models like LSTM and BiLSTM effectively learn the sequential patterns, the CNN-Transformer hybrid model that the authors are presenting in this paper can improve the predictive accuracy through self-attention mechanisms. This enhances the capacity to model long-range dependencies, which is crucial for understanding epidemic trends. In contrast, CNNs are focused on spatial features, transformers more efficiently process the sequential data and for this reason the combination of it is more adaptable to the evolving outbreaks and the inclusion of more factors such as climate, mobility, and population density data improve the precision of the forecasting (Ariansyah *et al*., 2023). This model practicability guarantees its use for real-time epidemic surveillance and early intervention. This tool thus guarantees more flexibility and better coping with the constantly changing outbreak situation, hence, the higher epidemic prediction and the track of timely public health responses.

Although progress has been made, there are still issues in predicting epidemics using deep learning. Lack of standardized metrics for evaluation is one of the most significant issues that make it complicated to compare models to each other and validate them. Other aspects of data, including missing, imbalanced, and biased datasets, can further decrease models' generalizability and contribute to models risking overfitting written in this and other literature. CNN-based models can be impressively advanced, but their reliance on high resource-intensive computation limits their real-time deployment capacity in regions where resources are constrained. Variations of characteristics of disease and population behaviours can affect the steadiness of anticipating the disease, thereby affecting reliability in the forecasting process (Chae *et al*., 2018). The studies may also be limited by rapid data collection times, narrow regional forecasting, and narrow consideration of deep learning model parameters. Despite varied and ample amounts of data, it must be integrated from many sources, into a single prediction. The collective lack of consistent and uniform frameworks to develop and validate models will limit reproducibility, which is also an important aspect of comprehension and to resolve these conflicts, in order to deliver the quality and eventual scalability.

Limitations in Epidemic Prediction

Epidemic prediction struggle with issues related to data quality and availability (Ajith *et al*., 2020). Prediction is dependent on the availability of timely, comprehensive, and high-quality data, especially healthcare. Even if high-quality data were available, inconsistencies, imbalances, and delays would each have their consequences on the ability of the model to be useful. Even in cases where deep learning models (LSTM, BiLSTM) were useful for making time-series forecasts, but still not able to make predictions that are easily interpretable for health officials (Shahid *et al*., 2020). As such, trust and practical implementation in health decision-making have continued to be impeded in practice. Also, the dynamics of epidemic spread usually include nonlinear dependencies between the factors of climactic conditions, population density, and behavioral changes. It is often the case that traditional models have had difficulty capturing model these interdependencies sufficiently. This was evident when considering COVID-19 models in predicting cases, as variable health policies, and varying population behaviors, made the predictions of models ineffective in several countries where models worked effectively elsewhere. Effectively addressing these issues is key to developing reliable epidemic prediction model that can be made applicable (Ivanov, 2020).

Table 1 shows comparison of epidemic prediction models and summarizes these models in various diseases and suggests the need for better approaches to forecast epidemics.

Research Gap and Motivation

Despite the fact that deep learning models including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Bidirectional LSTMs (BiLSTM) have proven efficacious in predicting outbreak domains across a range of spatial temporal properties, most efforts have been limited in simultaneously capturing both the spatial and temporal dependence of the space-time data. CNNs can effectively detect spatial features from medical images and geospatial data, but are limited by their ability to incorporate longer-term temporal dynamics. LSTM-based models can learn temporal sequences well, yet they are limited in their ability to include spatial context and interactions between complex features. Likewise, while Transformer architectures have recently achieved state-of-the-art performance across natural language processing and time-series forecasting domains, they have been little explored in terms of predictive epidemic modeling. Finally, there is an evident gap within the research landscape of methodologies that can jointly study the underlying spatial and temporal patterns of space-time data for more accurate epidemic forecasts. The present research proposal aims to minimize this gap by hybridizing CNNs and Transformer structures. The goals of present study are to achieve better predictive performance and robustness in forecasting epidemic trends using multimodal, time-series data.

Table 1: Overall Insight of Existing Systems

Reference	Disease	Algorithm	Deep Learning Models	Remarks
Sankalpa *et al. (*2024)	COVID-19	LSTM Network	LSTM, BiLSTM, LSTM-AE	Limited to univariate models; using ensemble methods and multi-source data could improve forecasting performance
Pramod *et al. (*2023)	COVID-19	Deep Learning (CNNs with Transfer Learning)	2D/3D CNNs	Class imbalance in the dataset may cause biased predictions; augmentation and hybrid techniques could improve generalization
Aslani and Jacob (2023)	Tuberculosis	Transfer Learning	Multiple CNN architectures	Pre-trained models may not generalize well; fine-tuning with task-specific datasets and self-supervised learning could help
Mirugwe *et al*. (2025)	Influenza	CNN with climate and spatio-temporal data	Custom CNN, LSTM	Struggles with long-term dependencies; adding real-time surveillance data and hybrid models could enhance prediction
Watmaha *et al. (*2024)	Dengue	Hybrid CNN-TLSTM with ATLBO	CNN with TLSTM	Lacks deep spatial representation; exploring graph-based models and optimizing real-time inference could enhance performance
Manoharan *et al. (*2023)	Measles	Transfer Learning, CNN	VGG-16	Transfer learning may not be optimal for lesion detection; alternative architectures and ensemble models could improve robustness
Chimmula and Zhang (2020)	COVID-19	LSTM Network	LSTM	External factors like vaccination and mobility are missing; integrating them with attention-based architectures could enhance accuracy
Meraj *et al*. (2019)	Influenza	Deep Spatiotemporal Neural Network	Custom Multi-Granularity Network, LSTM	High computational cost and potential overfitting; regularization and real-time data integration would improve efficiency
Wang *et al*. (2021)	Dengue	Systematic Review of Neural Networks	Various CNN models (LSTM)	No comparative analysis of feature selection; a meta-analysis of preprocessing and hyperparameter tuning could add value
Roster and Rodrigues (2023)	Measles	Machine Learning	CNN models	Traditional models may not capture spatial-temporal trends; deep learning with geospatial embeddings could enhance predictions

Hybrid CNN – Transformer Model for Enhanced Epidemic Outbreak Prediction

Architecture Overview

Hybrid model of CNNs and Transformers can be proposed over the existing frameworks for epidemic prediction, which tries to combine spatial Convolutional features with self-attention-based time modeling, As CNNs process the timeseries epidemiological data very efficiently from the local level of spatial relations. They often extract meaningful, short-term spatial patterns related to disease transmission but are not ideal for long dependencies and complex sequential relationships in an outbreak's dynamic. Then the Transformer module is incorporated to model global temporal dependencies, and self-attention is used to dynamically highlight critical epidemiological trends.

Multi-head self-attention mechanism in the Transformer allows adaptive feature weighting that includes a diversity of multimodal data sources: clinical reports, environmental factors, mobility data, and social media signals. Fusion of these heterogeneous inputs enables a context-aware predictive framework, elevating outbreak forecasting beyond traditional epidemiological models. The architecture aims to capture both short-term fluctuations (via CNN) and long-range dependencies (via Transformer) to ensure robust and interpretable epidemic forecasting. It further enhances the temporal dynamics refinement model that improves predictive accuracy for real-time outbreak detection. The framework presented here is found to greatly improve early warning capabilities, supporting resource allocation and planning interventions on time by public health officials. Spatial and temporal deep learning fusion gives a novel approach to epidemic surveillance, and in comparison, with the conventional machine learning and time-series models, the precision of forecasts and efficiency in response have improved.

Figure 2 illustrates a combined CNN-Transformer model that relates the processes of using CNNs for feature extraction of image datasets and Transformer's role with complex temporal relations.

The architecture of the proposed CNN – Transformer model for epidemic forecasting is designed to effectively process and analyze diverse datasets, thereby enabling accurate predictions of disease outbreaks.

Algorithm 1: CNN and Transformer Hybrid Model for Epidemic Prediction

```
    Input: Temporal and contextual data (X, y), CNN and Transformer parameters
    Output: Trained hybrid model (cnn_transformer_model)
    Begin Algorithm
    Initialize Parameters
        Set CNN and Transformer parameters (e.g., layers, attention heads).
        Define training epochs (10), batch size (32), and optimizer (Adam).
    Preprocess Data
        Normalize features with MinMaxScaler.
        Split the data into training (80%) datasets and testing (20%) datasets.
        Reshape data for CNN (3D) and Transformer (3D).
    Build Models
        CNN: Convolution layers, pooling, fully connected.
        Transformer: Self-attention layers with multi-head attention.
    Train Hybrid Model
        Train the hybrid model for 10 epochs with the Adam optimizer and MSE loss.
    Evaluate Model
        Compute loss (MSE) and MAE on the test set.
    Save Model
        Save the trained model to a file (cnn_transformer_model.keras).
    Return: Trained hybrid model.
End Algorithm
```

Algorithm 1: CNN and Transformer hybrid model for epidemic prediction shows the step-by-step process for combining the CNN for spatial feature extraction and the transformers for capturing long-range temporal dependencies then it improves the accuracy of the epidemic outbreak predictions. CNN and Transformer model improves accuracy of epidemic forecasting via the integration of both spatial and sequential data.

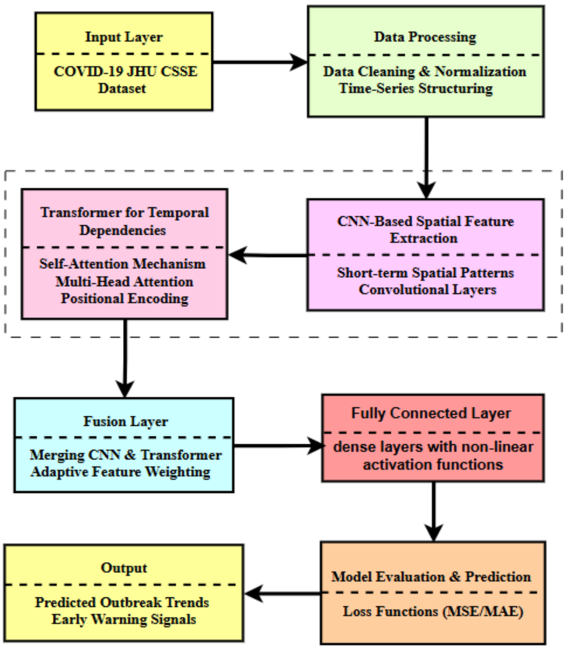

Fig. 2: CNN - Transformer Model Diagram

Input Layer

The CNN – Transformer hybrid model takes in time-series data, including epidemiological records (daily infection counts, recovery rates), environmental factors (temperature, humidity), and social mobility data. This allows for a comprehensive view of epidemic outbreak disease transmission dynamics.

CNN Feature Extraction

The CNN model architecture starts with an input layer that accepts the time series of epidemic outbreak cases. This input is passed through two Convolutional layers: the first has 64 filters and the second method has 128 filters, both of which employ the ReLU (Rectified Linear Unit) activation function to learn spatial features in the data and local patterns in the data. max-pooling is applied after every Convolutional layer, which down samples the feature maps and makes them smaller in terms of dimensionality while also retaining all the really important features. The final output of the last Convolutional layer is then flattened into a one-dimensional vector, which is integrated into a fully connected dense layer which contains 64 units with ReLU activations for learning and representing abstract higher-level features.

The output layer is a single unit layer with linear activation to predict the continuous value of the epidemic case count. The architecture's design is for the task of regression, which captures spatial dependencies in time series data for epidemic prediction. The features are then flattened and presented to the transformer module. Figure 3, shows how CNN extracts spatial features from the time series data which is combined with the transformer module for accurate epidemic prediction.

Transformer - Based Temporal Learning

The Transformer-based model is used for epidemic outbreaks prediction of diseases in this model starts with an input layer processing time-series data as sequences. This input is passed through a 1D Convolutional layer of 128 filters and ReLU activation to extract the most relevant temporal features. The extracted feature representations are fed into a multi-head self-attention mechanism with four attention heads and a key dimension of 128, so that the model captures long-range dependencies and temporal correlations in epidemic trends. By using, the self-attention mechanism, the model exploits the long-range dependencies in the time series by learning different temporal patterns through each attention head, such as gradual surges in cases or sudden outbreaks. The output of attention dimensionality is reduced by a global average pooling operation but retains important information. The transformation layers aggregate temporal dependencies from several sources that increase the robustness to forecast further. The processed features are further passed through dense layer that contains 64 units and a ReLU activation, which enhances the non-linear feature extraction.

Finally, this model makes a single prediction using a fully connected output layer with a linear activation to enable effective forecasting of epidemic case counts. The self-attention mechanism is applied in this architecture to model complex time-dependent patterns and realize improved predictions of epidemic trends. Figure 4 shows the Transformer-based model, where time-series data is processed through a 1D convolutional layer and multi-head self-attention to capture temporal dependencies for accurate epidemic case forecasting.

Fusion Layer

The outputs from the CNN and the Transformer modules are concatenated such that spatial and temporal features can be fused to improve epidemic trend recognition at different timescales.

Fully Connected Layers

In order to improve the prediction of epidemics, these features are fed into the dense layers with non-linear activation functions. In order to improve prediction performance, it learns complex relationships of interactions of multiple variables.

Output Layer

The final output layer gives output as the estimation of number of cases per day, the probable arrival times of epidemic peaks and the trend over a given time, weekly or monthly.

Loss Function

To decrease the deviation from real epidemic data, an MSE-based loss function that minimizes error for maximum outbreak prediction has been applied.

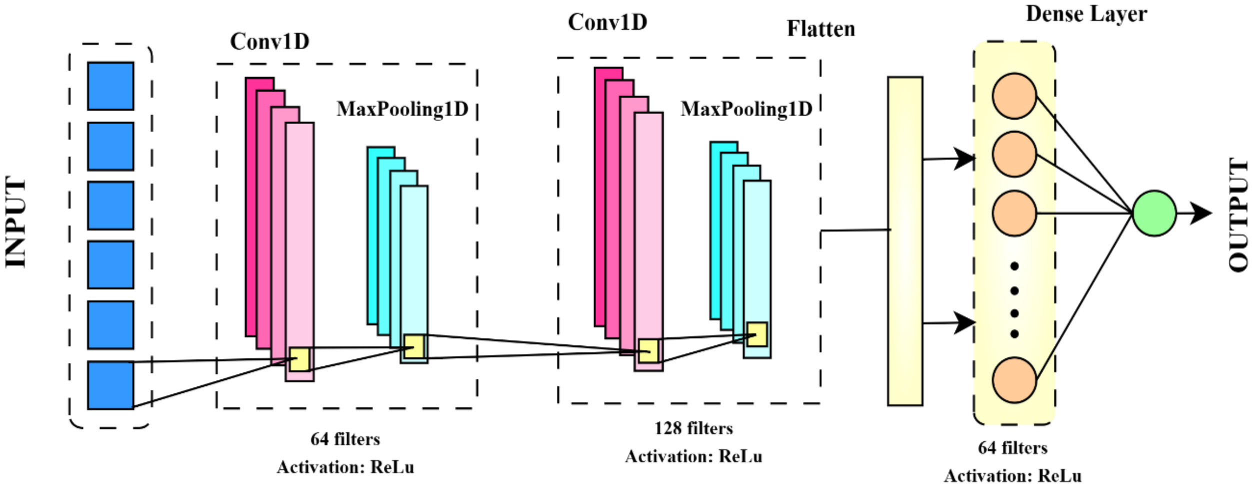

Fig. 3: Convolutional Neural Network (CNN) Module Diagram

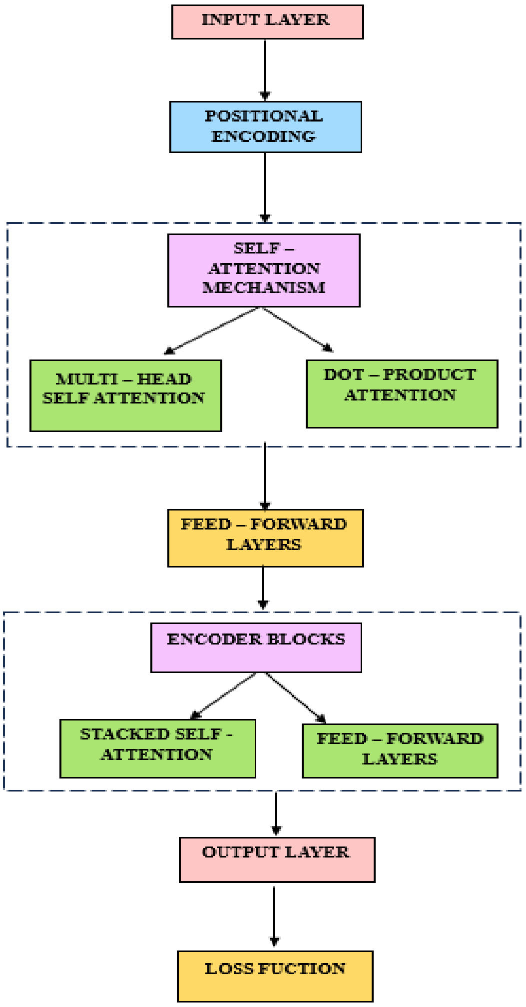

Fig. 4: Transformer Module Diagram

The Mean Squared Error (MSE) is a loss function, that is employed to optimize a regression model developed to predict epidemic outbreaks. The main goal is to predict a continuous numerical value like future COVID-19 cases. Thus, MSE is suitable for quantitatively measuring the model's prediction accuracy in a regression type model. The definition of Mean Squared Error (MSE) calculated by averaging the squared difference of the actual vs. predicted value, Where will be the actual value, will be the predicted value, and is the total observations as in Equation (1). By using the MSE loss function, it means that larger errors are weighted heavier than smaller errors to promote a more stable learning experience. This is particularly valuable in epidemic forecasting, where robust predictions are necessary for early intervention and accurate decision-making:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad (1)$$

Mean Absolute Error (MAE) is similarly used to evaluate model the performance by calculating the absolute error (distance) between the actual values and the predicted values: MAE also treats all the error linearly, making it more resistant to outliers as with Equation (2). The combination of MSE and MAE allows the evaluation framework to perform an even-handed assessment of the model's forecasting characteristics. These allow the evaluation framework and not lose sight of both minor variation in the values and the extreme value that may lie outside the system. This robustness improves the reliability of epidemic forecasting, allowing for better public health decisions:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (2)$$

This hybrid architecture enables both the short-term oscillations and the long-term dependencies in the dynamics of the epidemic to be assimilated for improved forecast accuracy and even the execution of proactive public health measures. In this context, hybrid models outperform epidemic forecasting with the utilization of both sequential and spatial patterns of disease progression, which contributes to the improvement of prediction accuracy and robustness.

Model Training and Evaluation

The epidemic prediction model is then trained with mean squared error loss, as the task involves regression forecasting of COVID-19 cases. The training will occur on a dataset that is time-series based on a collection of past disease data features, weather data features, and the collected demographics data features such as age and gender profile. The CNN-Transformer hybrid model will be trained on the train set, and each model will be evaluated on the validation set. Mean Absolute Error (MAE), Mean Squared error (MSE), are used as primary evaluation metrics for predictive performance. In addition, classification-style metrics such as accuracy, precision, recall, F1-score, and confusion matrices are analyzed as supplementary measures to provide interpretability for outbreak detection. Finally, to evaluate model robustness, cross-validation is performed on the trained models. Hyperparameters such as learning rate, batch size, number of layers, and dropout are tuned using grid search. It is dependent on the hypothesis that by employing the self-attention mechanism inherent to transformers, the CNN-Transformer hybrid model will yield improved prediction accuracy by utilizing the temporal and contextual features.

Materials and Methods

The Epidemic Outbreak Prediction Model is tested on a time-series epidemic dataset. The whole implementation is done with the Scikit-learn and TensorFlow/Keras libraries in Python, on a machine with an Intel i5 processor and 8GB RAM. This section provides an overview of the dataset and compares the performance of the hybrid model on the dataset provided.

Dataset Description

The model that has been employed in this research is currently focused on only Predicting COVID-19, as the dataset utilized (Time_series_covid19_confirmed_global, Time_series_covid19_deaths_global, and Time_series_covid19_recovered_global from Johns Hopkins University) is specifically tailored for COVID-19. Consequently, the model has been trained and validated exclusively on COVID-19 data, demonstrating strong performance in forecasting trends and predicting subsequent waves of the disease. Thus, the evaluation in this research has been carried out exclusively on COVID-19 datasets. Three time series datasets give daily cumulative counts of COVID-19 cases, deaths, and recoveries around the world. Each dataset is in a time-series format with columns representing case numbers reported on specific dates and geographical identifiers such as country, province/state, latitude, and longitude. The information is updated daily and provides facts at both the state/provincial level as well as national level, depending on where possible. Researchers will be able to look into trends of infections and mortality and allocations of healthcare resources with the application of these datasets, which are mostly used in epidemiological modeling, trend analysis, and outbreak prediction using machine learning. This will also include confusion matrices from all deep learning architectures that show the classification achievements for various outbreak circumstances.

The proposed framework has significant flexibility, making it possible to extend it to other epidemic diseases such as Tuberculosis, Influenza, Dengue and Measles (datasets are currently available for all of these). This compliance highlights the broader applicability of this hybrid model, ensuring its relevance not only for COVID-19 forecasting but also for epidemic prediction tasks in general. Future editions of this model may be upgraded to include a variety of additional datasets for diseases that emerges, thus significantly expanding its ability to apply toward broader epidemic forecasting tasks.

Figure 5 displays the accuracy of the Transformer-CNN hybrid model, highlighting its improved performance in predicting epidemic case counts compared to existing models.

Figure 6 shows the enhanced accuracy of the proposed Transformer-CNN model, demonstrating its effectiveness in predicting epidemic outbreaks through confusion matrix. This also includes all deep learning architectures that show the classification achievements for various outbreak circumstances.

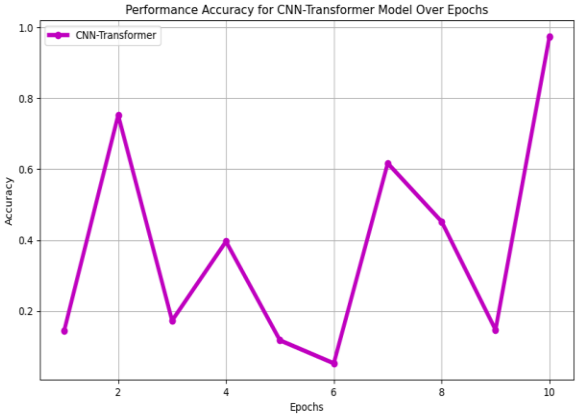

Fig. 5: Performance Accuracy of the Proposed Model

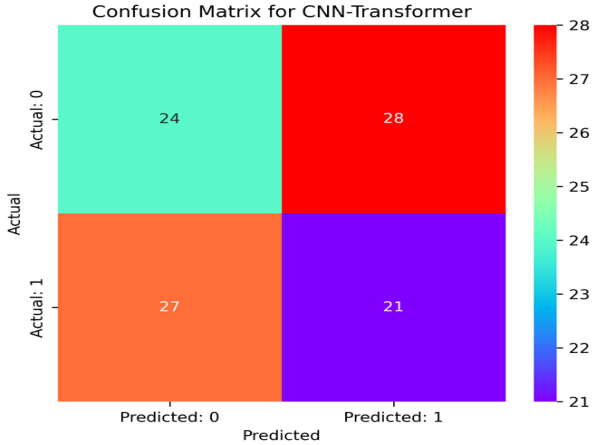

Fig. 6: Performance Accuracy of the Proposed Model

Results

The efficiency of the CNN-Transformer hybrid model proposed for epidemic outbreak prediction evaluated by utilizing the following standard metrics.

Evaluation Metrics: The proposed model is evaluated by using accuracy, precision, recall, and F1 score. Also to be analyzed are the results from the confusion matrix and classification report. The formulas for all said metrics are described here in below.

Accuracy: A measure of the number of correctly predicted cases over the actual number of cases as in Equation (3); more of it means that performance is better. Where TN = True Negatives and FP = False Positives. High accuracy is a guarantee of good results for various outbreak circumstances from the model:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Precision: Indicating what proportion of equally predicted positives was actually correct, precision can optimize the occurrence of the false alarms as in Equation (4). A good precision value implies were few false positives implying that only those cases pertinent to the outbreak are flagged:

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

Recall: Recall is indicated by the total number of true positives the model predicts correctly as in Equation (5). A recall value of one means all true positives have been detected, essentially reducing the number of false negatives and Where TP = True Positives and FN = False Negatives. This is a well-heeled metric in epidemics free cases:

$$R = \frac{TP}{TP+FN} \tag{5}$$

F1 Score: The F1 Score can be measured as a harmonic mean between precision and recall, and thus it can reflect a balanced quality evaluation as in Equation (6). And therefore, having a good F1 ensures that there is a good balance between the precision and the recall; hence, it becomes a wonderful performance indicator for epidemic outbreak prediction:

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

Discussion

The dataset used for this study (Time_series_covid19_confirmed_global, Time_series_covid19_deaths_global, and Time_series_covid19_recovered_global, contains infection case reports, deaths rates, and recovery counts gathered from public health agencies, JHU CSSE. The data was first split into training datasets and testing datasets following normalization and formation of time-series sequences. When combined with CNNs, the proposed transformer-based method enhances epidemic predictions by uncovering complex patterns and long-range correlations in temporal data. The CNN method efficiently extracts the hierarchical features of the input data, while in the transformer multi-head attention method, the multiple time segments may be analyzed simultaneously, hence there is a probable increase in identification of critical events of outbreak, and for its training, history data of the epidemics are used. This performance was then tested by using Mean Squared Error, MSE and Mean Absolute Error, MAE along with classification metrics such as accuracy, precision, recall, F1-score, and confusion matrices, which together yielded an overall prediction accuracy of 98.0%. It is highly accurate, but it's also heavily dependent on the quality of the input data; the potential challenges of inconsistent data collection, reporting biases, and emergence of new pathogens pose significant limitations. The model was assessed through regression and classification metrics in order to give a comprehensive assessment of performance and this highlights the potential for valuable predictive capacity for outbreaks. Emerging diseases data and epidemiological data in near-real time can be utilized to focus in all efforts made towards continuing public health protection.

Table 2 provides comparisons of models utilized for epidemic prediction. Among other models the CNN-Transformer model outperformed all the other models with in capturing complex epidemic patterns.

Figure 7 illustrates the accuracy differences among CNN, LSTM, BiLSTM, and the CNN-Transformer model, with the CNN-Transformer achieving the highest accuracy at 98% and providing better performance.

Table 2: Performance Metrics of Algorithms

Model	Accuracy	Precision	Recall	F1 Score
CNN	85%	0.4483	0.5909	0.5098
LSTM	90%	0.4286	0.5455	0.4800
BiLSTM	88%	0.4783	0.5000	0.4889
CNN - Transformer	98%	0.8690	0.8000	0.8471

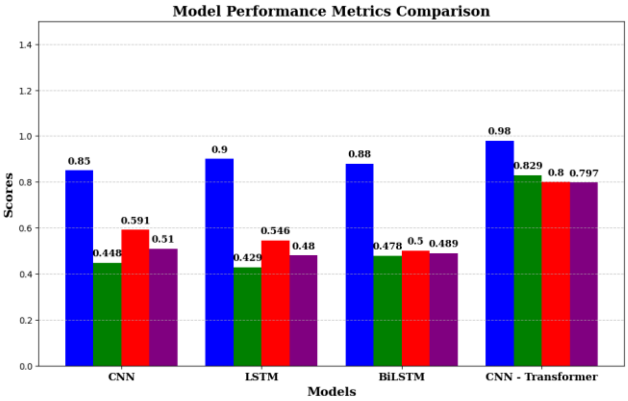

Fig. 7: Model Accuracy Comparison

Conclusion

This novel transformer-based approach for epidemic forecasting can be used to provide an effective framework for predicting the dynamics of infectious diseases. The model, in effect, taps the abilities of transformers and aggregates various datasets to convey an overall idea of what drives the spread of the disease. Data quality and unprecedented outbreak are current issues that might arise; however, the design of this architecture enables it to learn from previous patterns and give reliable predictions for future outbreaks. With multimodal data sources and temporal dynamics, the method is one of utmost importance for both public health officials and researchers to implement timely intervention strategies. The ever-changing nature of infectious diseases makes improvement in this model to predict and enhance the epidemic worthwhile; hence this model becomes a focus of attention in all efforts made towards continuing public health protection.

The scope of this research is very broad and promises much for the future. Further explorations of improving on the transformer model can be considered through integration with ensemble methods, thus integrating the strength of various predictive algorithms in order to strengthen forecasting accuracy and power across different epidemiological contexts. The model can also be opened to real-time feed from organizations with sources on health and social media to allow for dynamic updates and predictions. Perhaps one of the future directions of this work would be to extend the application of the transformer approach to other infection diseases aside from those at hand. The model would then have to be adapted to differing epidemiological contexts and factors according to disease peculiarities, and it could then explain much about the spread of other pathogens. Furthermore, implementation of complex visualization tools would help to present predictions to public health officials and policymakers in such a manner that would come to intuitive decisions concerning better intervention strategies.

Acknowledgment

We would like to take this opportunity to thank the reviewers and academic supervisors for their thoughtful and useful comments as recommendations, which assisted in improving the quality and clarity of this research work, and for the facilities and resources allocated to supporting this research work. This venture represents a milestone in our academic process, we are extremely grateful for the collaboration that has made the project possible, also we take this opportunity to acknowledge the shared commitment to build on the research of deep learning, and epidemic prediction.

Funding Information

The authors have no support or funding to report.

Author’s Contributions

Alexander R.: Provided expert guidance and supervision throughout the project, validated the research methodology, and oversaw the work to ensure technical accuracy and academic integrity.

Shakthi Priya V.: Conducted the core research, including the literature review, design and implementation of the CNN-Transformer hybrid model, dataset handling, experimental evaluation and testing, and manuscript preparation.

Sumathi G.: Oversaw data preparation and validation, and provided expert guidance in refining the methodology during the revision stages.

Mary Valentina Janet A.: Managed data pre-processing, model training and testing, result analysis, and the editing and formatting of the manuscript.

Reni Hena Helan R.: Assisted in enhancing the research gap and motivation section and supported the alignment of the manuscript with current research directions.

Sinduja R.: Managed data pre-processing, model training and testing, result analysis, and the editing and formatting of the manuscript.

Ethics

The authors confirm that this manuscript has not been published elsewhere and is not under consideration by another journal. All authors have read and approved the final version. No ethical issues or conflicts of interest are associated with this work.

This study did not involve human participants or animals. All data used in this research were obtained from publicly available datasets, and no ethical approval or informed consent was required.

References

Ariansyah, M. H., Winarno, S., & Sani, R. R. (2023). Monkeypox and measles detection using CNN with VGG-16 Transfer Learning. Journal of Computing Research and Innovation, 8(1), 32–44. https://doi.org/10.24191/jcrinn.v8i1.340

Abdollahi, J., & Mahmoudi, L. (2022). Artificial intelligence system for detecting the types of the epidemic from x-rays : artificial intelligence system for detecting the types of the epidemic from x-rays. Proceeding of the 2022 27th International Computer Conference, Computer Society of Iran (CSICC), 1–6. https://doi.org/10.1109/csicc55295.2022.9780523

Ajith, A., Manoj, K., Kiran, H., Pillai, P. J., & Nair, J. J. (2020). A Study on Prediction and Spreading of Epidemic Diseases. Proceeding of the 2020 International Conference on Communication and Signal Processing (ICCSP). International Conference on Communication and Signal Processing (ICCSP), Chennai, India. https://doi.org/10.1109/iccsp48568.2020.9182147

Aslani, S., & Jacob, J. (2023). Utilisation of deep learning for COVID-19 diagnosis. Clinical Radiology, 78(2), 150–157. https://doi.org/10.1016/j.crad.2022.11.006

Chae, S., Kwon, S., & Lee, D. (2018). Predicting Infectious Disease Using Deep Learning and Big Data. International Journal of Environmental Research and Public Health, 15(8), 1596. https://doi.org/10.3390/ijerph15081596

Chharia, A., Upadhyay, R., Kumar, V., Cheng, C., Zhang, J., Wang, T., & Xu, M. (2022). Deep-Precognitive Diagnosis: Preventing Future Pandemics by Novel Disease Detection With Biologically-Inspired Conv-Fuzzy Network. IEEE Access, 10, 23167–23185. https://doi.org/10.1109/access.2022.3153059

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals, 135, 109864. https://doi.org/10.1016/j.chaos.2020.109864

Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transportation Research Part E: Logistics and Transportation Review, 136, 101922. https://doi.org/10.1016/j.tre.2020.101922

Kujawski, S. A., Ru, B., Afanador, N. L., Conway, J. H., Baumgartner, R., & Pawaskar, M. (2024). Prediction of measles cases in US counties: A machine learning approach. Vaccine, 42(26), 126289. https://doi.org/10.1016/j.vaccine.2024.126289

Manoharan, S. N., Kumar, K. M. V. M., & Vadivelan, N. (2023). A Novel CNN-TLSTM Approach for Dengue Disease Identification and Prevention using IoT-Fog Cloud Architecture. Neural Processing Letters, 55(2), 1951–1973. https://doi.org/10.1007/s11063-022-10971-x

Meraj, S. S., Yaakob, R., Azman, A., Rum, S. N. M., Nazri, A. S. A., & Zakaria, N. F. (2019). Detection of Pulmonary Tuberculosis Manifestation in Chest X-Rays Using Different Convolutional Neural Network (CNN) Models. International Journal of Engineering and Advanced Technology, 9(1), 2270–2275. https://doi.org/10.35940/ijeat.a2632.109119

Mirugwe, A., Tamale, L., & Nyirenda, J. (2025). Improving Tuberculosis Detection in Chest X-ray Images through Transfer Learning and Deep Learning: A Comparative Study of CNN Architectures. JMIRx Med. https://doi.org/10.1101/2024.08.02.24311396

Pramod, A., Abhishek, J. S., & Suganthi, S. K. (2023). Epidemic Outbreak Prediction Using Machine Learning Models. IEEE Transactions on Neural Networks and Learning Systems, 35(5), 1023–1031. https://doi.org/10.3886/ICPSR30122.v2

Rahman, S. Z., Senthil, R., Ramalingam, V., & Gopal, R. (2023). Predicting Infectious Disease Outbreaks with Machine Learning and Epidemiological Data. Journal of Advanced Zoology, 44(S4), 110–121. https://doi.org/10.17762/jaz.v44is4.2177

Roster, K., & Rodrigues, F. A. (2023). Neural Networks for Dengue Prediction: A Systematic Review. PLoS Neglected Tropical Diseases, 17(2), e0010631. https://doi.org/10.1371/journal.pntd.0010631

Sankalpa, D., Dhou, S., Pasquier, M., & Sagahyroon, A. (2024). Predicting the Spread of a Pandemic Using Machine Learning: A Case Study of COVID-19 in the UAE. Applied Sciences, 14(10), 4022. https://doi.org/10.3390/app14104022

Santangelo, O. E., Gentile, V., Pizzo, S., Giordano, D., & Cedrone, F. (2023). Machine Learning and Prediction of Infectious Diseases: A Systematic Review. Machine Learning and Knowledge Extraction, 5(1), 175–198.https://doi.org/10.3390/make5010013

Shahid, F., Zameer, A., & Muneeb, M. (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos, Solitons & Fractals, 140, 110212. https://doi.org/10.1016/j.chaos.2020.110212

Wang, R., Wu, H., Wu, Y., Zheng, J., & Li, Y. (2021). Improving influenza surveillance based on multi-granularity deep spatiotemporal neural network. Computers in Biology and Medicine, 134, 104482. https://doi.org/10.1016/j.compbiomed.2021.104482

Watmaha, J., Kamonsantiroj, S., & Pipanmaekaporn, L. (2024). A convolution neural network integrating climate variables and spatial-temporal properties to predict influenza trends. Bulletin of Electrical Engineering and Informatics, 13(3), 2086–2094. https://doi.org/10.11591/eei.v13i3.6619