DOI: 10.53469/jtpes.2024.04(02).02

Enhancing Disease Prediction with a Hybrid CNN-LSTM Framework in EHRs

Jingxiao Tian[1,*], Ao Xiang[2], Yuan Feng[3], Qin Yang[4], Houze Liu[5]

[1]Electrical and Computer Engineering, San Diego State University, San Diego, US
[2]Digital Media Technology, University of Electronic Science and Technology of China, Chengdu, China
[3]Interdisciplinary Data Science, Duke University, Durham, US
[4]Microelectronics Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
[5]Computer science, New York University, New York, US
**Correspondence Author, rowan199408@gmail.com*

Abstract: ***This study developed a novel hybrid deep learning framework aimed at enhancing the accuracy of disease prediction using temporal data from Electronic Health Records (EHRs). The framework integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, leveraging the strength of CNNs in extracting hierarchical feature representations from complex data and the capability of LSTMs in capturing long-term dependencies in temporal information. An empirical investigation on real-world EHR datasets revealed that, compared to Support Vector Machine (SVM) models, standalone CNNs, and LSTMs, this hybrid deep learning network demonstrated significantly higher prediction accuracy in disease prediction tasks. This research not only advances the performance of predictive models in the health data analytics domain but also underscores the importance of adopting and further developing advanced deep learning technologies to address the complexity of modern medical data. Our findings advocate for a shift towards integrating complex neural network architectures in developing predictive models, potentially offering avenues for more personalized and proactive disease management and care, thereby setting new standards for future health management practices.***

Keywords: Deep Learning, Convolutional Neural Network, Long Short Term Memory Neural Network, Hybrid Deep Learning.

1. INTRODUCATION

Electronic Health Records (EHRs) data, with its detailed record of patient health over time, plays a pivotal role in understanding disease progression and crafting models to forecast future health trajectories. However, the nature of EHR data collection—sporadic, based on patient visits or specific healthcare encounters—poses unique challenges for model development. These include the need to handle diverse data formats and to navigate the intricacies of longitudinal patient data[1]. Efforts to overcome these obstacles have taken two main paths: one leveraging domain-specific knowledge to create unified patient groups from disparate data sources, and another investigating the integration of varied EHR[2] data types, either pre- or post-model development.

Traditional approaches to disease forecasting often categorize patients with similar health patterns together, simplifying the prediction process. Yet, predicting outcomes based on a single variable remains a daunting task in machine learning[3], particularly when key variables are unknown. Such univariate predictions, though challenging, are adaptable for forecasting various diseases using historical EHR data, without the need for additional contextual information.

The application of deep learning neural networks (DLNNs), including in areas like natural language processing[4][5] and image recognition[6][7], has seen a significant rise. Among these, Long Short-Term Memory (LSTM) networks[8] stand out for their superior predictive accuracy, attributed to their ability to retain information over extended periods through specialized memory gates. This has placed LSTMs ahead of many traditional and machine learning methods in prediction accuracy. Moreover, LSTM networks, a variant of Recurrent Neural Networks (RNNs)[9], alongside other DLNN forms such as Convolutional Neural Networks (CNNs) and Deep Belief Networks (DBN), have shown promise in various prediction tasks. Temporal CNNs[10], with their specialized convolution operations, are particularly noted for their efficacy in time series forecasting.

2. RELATED WORK

Reviewing related literature, disease forecasting emerges as a field of critical importance in medical diagnostics[11][12]. Traditional prediction models like Support Vector Regression and various time series methodologies[13] have been explored alongside deep learning techniques, which offer enhanced data processing capabilities and accuracy. Deep learning models, with their extensive internal layers, tackle more complex problems than their Artificial Neural Network counterparts[14]. Recent studies have demonstrated the effectiveness of deep learning models in medical diagnostics, disease trend forecasting, and even in identifying specific health conditions, outperforming traditional methods[15][16].

This research proposes a novel deep learning framework that combines LSTM networks with CNNs to tackle the challenges of disease prediction. By integrating a CNN preprocessing step, this framework aims to refine raw data into multidimensional inputs, thereby boosting LSTM's predictive performance. The effectiveness of this hybrid model was validated using real-world EHR datasets, showing its superiority over conventional prediction methods, including SVM, standalone CNN, and LSTM models. The key contributions of this study are two-fold: the introduction of a CNN preprocessing step for enhancing data dimensionality and the development of a comprehensive deep learning model for disease forecasting, demonstrating significant advancements over existing approaches.

3. METHODOLOGY

Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) stand out as two influential subsets of deep learning techniques that have captured significant interest worldwide over recent times[17]. In our study, we introduce a combined approach of LSTM and CNN to tackle the challenges posed by the irregular patterns and long-term dependencies found in temporal data forecasting tasks. This hybrid model surpasses traditional approaches in delivering predictions that are both more precise and dependable[18][19].

In our methodology, we employ real-world data, initiating the process by applying CNN for data preprocessing and subsequently feeding the CNN-processed data into the LSTM network for training.

3.1 Data Description

The data set used in our experiments was processed using Adadelta [20]. Post-preprocessing, it comprised 578 instances, of which 361 were positive and 217 were negative. We partitioned the data into training, validation, and testing segments, following a distribution ratio of 80:10:10. The training segment facilitated the training of our Deep Learning Neural Network (DLNN) framework, while the validation segment, an independent collection of samples, served the dual purpose of tuning hyperparameters and providing an initial evaluation of the framework's performance. The testing segment was designated for assessing the trained model's ability to generalize.

3.2 Long Short-Term Memory-based Recurrent Neural Network

LSTM models, a specialized category within Recurrent Neural Networks (RNN), introduce feedback connections in each neuron. An RNN's output is influenced not just by the inputs and weights of the current neuron but also by inputs from preceding neurons, making it inherently suitable for analyzing sequential data. Nonetheless, RNNs face significant challenges, such as gradient explosion and vanishing issues, when processing long-term sequential information [21], leading to the development of LSTMs [22].

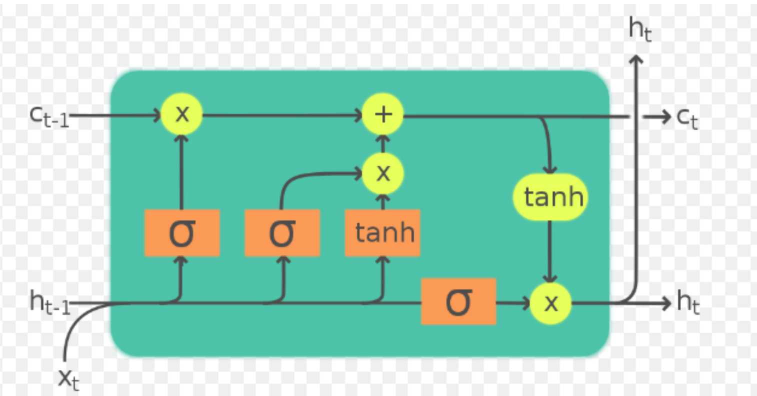

Figure 1: The training process of LSTM model

LSTMs are designed to mitigate the problem of vanishing gradients encountered by RNNs, incorporating mechanisms to selectively retain or discard information. The architecture of an LSTM includes a cell state and three gates: input, forget, and output (Figure 1), which regulate the cell state's update, retention, and elimination of information. The process of forward computation in an LSTM is characterized by these components.

$$f_t = \sigma\left(W_{fh} \cdot h_{t-1} + W_{fx} \cdot x_t + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_{ih} \cdot h_{t-1} + W_{ix} \cdot x_t + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh\left(W_{ch} \cdot h_{t-1} + W_{cx} \cdot x_t + b_c\right) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

$$o_t = \sigma\left(W_{ot} \cdot h_{t-1} + W_{ox} \cdot x_t + b_o\right) \tag{5}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \tag{6}$$

Where C_t , C_{t-1} , and $\tilde{C}_t$ respectively represent the current cell state value, the cell state value at the previous moment, and the update to the current cell state value. The symbols ft ， it , and ot respectively denote the forget gate, input gate, and output gate. With appropriate parameter settings, based on the values of $\tilde{C}_t$ and C_t , the output value ht is calculated according to equations (4) to (6). Based on the difference between the output values and the actual values, all weight matrices are updated through the Back-Propagation Through Time (BPTT) algorithm [23].

3.3 Temporal Convolutional Neural Networks

Convolutional Neural Networks (CNN) are perhaps the most commonly utilized deep learning neural networks, currently primarily applied in computer vision for imag recognition/classification themes. For a large volume of raw data samples, CNNs can typically extract a useful subset of the input data efficiently. Generally, CNNs are still feedforward neural networks, derived from multi-layer neural networks (MLNN). The main distinction between CNNs and traditional MLNNs lies in CNNs' characteristics of sparse interactions and parameter sharing [24].

Traditional MLNNs employ a fully connected strategy to establish neural networks between input and output layers, meaning each output neuron has the opportunity to interact with every input neuron. Assuming there are m input neurons and n output neurons, the weight matrix would have $m \times n$ parameters. CNNs significantly reduce the number of parameters in the weight matrix by utilizing convolutional kernels of size $k \times k$. Two attributes of CNNs enhance the training efficiency of parameter optimization; with the same computational complexity, CNNs can train neural networks with more hidden layers, i.e., deeper neural networks.

Temporal Convolutional Neural Networks introduce special one-dimensional convolutions suited for processing univariate time series data. Unlike traditional CNNs that use $k \times k$ convolutional kernels, temporal CNNs employ kernels of size $k \times 1$. After temporal convolution operations, the original univariate dataset can be expanded into a dataset with m-dimensional features. Thus, temporal CNNs apply one-dimensional convolution to time series data, expanding the univariate dataset into a multi-dimensional feature-extracted dataset (the first stage in Figure 2); the expanded multi-dimensional feature data are more suitable for prediction using LSTM.

3.4 CNN-LSTM Prediction Framework

To address the challenges of sequence dependency and univariate data, this paper proposes a hybrid deep neural network (DNN) that combines CNN and LSTM models. The structure of the hybrid DNN framework is shown in Figure 2. In the preprocessing stage, the CNN extracts crucial information from the input data, using convolution to reorganize univariate input data into multidimensional feature data (Figure 2). In the second stage, the reorganized feature data are input into LSTM units for prediction.

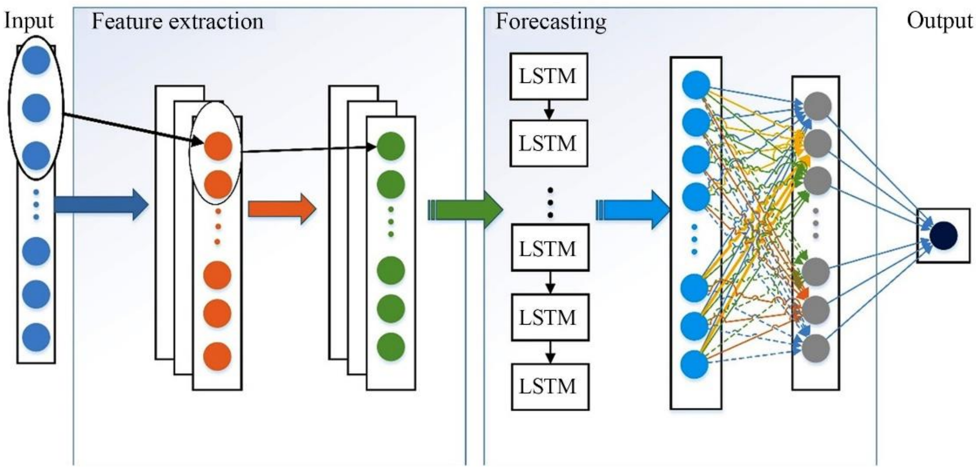

Figure 2: The proposed hybrid DNN disease prediction framework

As seen in Figure 2, the input data set is preprocessed using a CNN with two hidden layers. Notably, when the number of hidden layers exceeds five, traditional temporal CNNs typically include pooling operations to prevent overfitting. This paper omits pooling operations to maximize the retention of extracted feature information.
After preprocessing the input data, an LSTM neural network is designed to train and predict disease. The training process of the LSTM structure, as shown in Figure 1, where features extracted from the first stage are used as input for training the LSTM model. A dropout layer is added to the LSTM neural network to prevent overfitting. The difference between predicted output values and actual output values, i.e., the loss, is used to optimize the weights of all LSTM units. The optimization process follows a gradient descent optimization algorithm named RMSprop, commonly used for weight optimization in deep neural networks [25].

4. EXPERIMENTAL RESULTS

The proposed hybrid DNN framework was implemented using Python 3.7.3 (64-bit). It was built upon the open-source deep learning tool TensorFlow proposed by Google and used Keras version 2.3.1 as the frontend interface.

The prediction results of the proposed CNN-LSTM were compared with existing methods such as SVM models, CNN, and LSTM. Accuracy, the most commonly used performance evaluation metric, indicates the proportion of correctly predicted samples out of the total sample count. The F1 score represents a balance between precision and recall, where precision measures the proportion of correctly predicted positive samples out of all predicted positive samples, and recall measures the proportion of correctly predicted positive samples out of all actual positive samples. Another evaluation metric used in this study is the Area Under the ROC Curve (AUC). The ROC curve represents a trade-off between the false positive rate and the true positive rate, with the former measuring the proportion of samples wrongly predicted as positive out of all negative samples, and the latter measuring the proportion of samples correctly predicted as positive out of all positive samples. A common method to analyze the ROC curve is to calculate the area under the curve (AUC). Higher values of these three metrics indicate higher prediction accuracy.

Four prediction models were established using different methods, and Table 1 summarizes the results of the prediction performance. It is evident that the choice of model has a certain impact on prediction performance. Overall, the CNN-LSTM algorithm proposed in this study achieves the best prediction performance. Compared to LSTM, the AUC increased by 6.5%, the F1 score by 12.2%, and Accuracy by 14.6%. The experimental results demonstrate that the hybrid deep learning algorithm is more suitable for temporal data disease prediction.

Table 1: Forecast results of different models

METHOD	CNN	SVM	LSTM	CNN&LSTM
F1	0.519989	0.491829	0.652163	0.742846
AUC	0.632342	0.578676	0.766165	0.819842
Accuracy	0.578936	0.456140	0.719287	0.842193

5. CONCLUSION

This research successfully engineered a novel deep learning architecture by merging Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks, targeting the use of singular and sequential data for predicting diseases. The essence of this innovative architecture lies in its premise: the combination of CNN's robust feature identification capabilities with LSTM's adeptness at analyzing temporal data significantly enhances prediction accuracy beyond that of conventional LSTM models alone. In this setup, CNNs preprocess the input by identifying crucial features[26], which the LSTM component then uses to track temporal variations in the data, thereby offering a refined approach for forecasting diseases.

The investigation not only validates the efficiency of the newly developed CNN-LSTM hybrid in tasks related to predicting diseases from univariate data but also underscores its superiority over classical LSTM approaches. This breakthrough holds profound implications for forthcoming medical research, particularly concerning precision medicine and crafting personalized treatment strategies.
Looking ahead, extending the application of this CNN-LSTM model to more intricate and real-life medical data sets emerges as a logical progression. This endeavor will not just corroborate the model's effectiveness and durability across diverse data volumes and types but also unveil its competence in managing multivariate and unstructured data[27], such as medical imagery and narrative medical records, and in making predictions across different diseases. Furthermore, given the delicate nature and intricacy of medical data, forthcoming studies should also consider strategies for leveraging this model in a manner that safeguards patient confidentiality and for incorporating cutting-edge deep learning techniques (like self-attention frameworks, graph neural networks, etc.) to elevate its predictive capabilities.

Another vital area for future exploration is enhancing the model's interpretability[28]. In healthcare, the ability to interpret model outputs transparently is essential for earning the confidence of healthcare professionals and patients alike. Thus, efforts to augment the interpretability of the CNN-LSTM hybrid, aiming to make its predictions more accessible and comprehensible, will be crucial for boosting its utility in clinical settings.
Additionally, given the substantial data requirements of deep learning models, future research should investigate strategies for effective training with limited data resources, employing methods such as transfer learning, semi-supervised learning, or weakly supervised learning. Implementing such techniques would increase the model's feasibility in medical contexts where data availability is often restricted.

REFERENCES

[1] Wei, W.-Q., Teixeira, P.L., Mo, H., Cronin, R.M., Warner, J.L. and Denny, J.C. (2015) Combining Billing Codes, Clini- cal Notes, and Medications from Electronic Health Records Provides Superior Phenotyping Performance. Journal of the American Medical Informatics Association, 23, e20-e27. https://doi.org/10.1093/jamia/ocv130
[2] Henriksson, A., Zhao, J., Bostr?m, H. and Dalianis, H. (2015) Modeling Heterogeneous Clinical Sequence Data in Semantic Space for Adverse Drug Event Detection. IEEE International Conference on Data Science and Advanced Analytics, Paris, 19-21 October 2015, 1-8. https://doi.org/10.1109/DSAA.2015.7344867
[3] Li, L., Yang, Y., Zhan, S., & Wu, B. (2021, May). Sentence dependent-aware network for aspect-category sentiment analysis. In International Conference on Web Engineering (pp. 166-174). Cham: Springer International Publishing.
[4] Hao Xu, Qingsen Wang, Shuang Song, Lizy Kurian John, and Xu Liu. 2019. Can we trust profiling results? Understanding and fixing the inaccuracy in modern profilers. In Proceedings of the ACM International Conference on Supercomputing. 284–295.
[5] Sun, W., Wan, W., Pan, L., Xu, J., & Zeng, Q. (2024). The Integration of Large-Scale Language Models Into Intelligent Adjudication: Justification Rules and Implementation Pathways. Journal of Industrial Engineering and Applied Science, 2(1), 13–20. https://doi.org/10.5281/zenodo.10607564
[6] Yan, X., Xiao, M., Wang, W., Li, Y., & Zhang, F. (2024). A Self-Guided Deep Learning Technique for MRI Image Noise Reduction. Journal of Theory and Practice of Engineering Science, 4(01), 109–117. https://doi.org/10.53469/jtpes.2024.04(01).15
[7] Liu, B., Yu, L., Che, C., Lin, Q., Hu, H., & Zhao, X. (2023). Integration and Performance Analysis of Artificial Intelligence and Computer Vision Based on Deep Learning Algorithms. arXiv preprint arXiv:2312.12872.

[8] Kong, W., Dong, Z.Y., Hill, D.J., Luo, F. and Xu, Y. (2018) Short-Term Residential Load Forecasting Based on Resi- dent Behaviour Learning. IEEE Transactions on Power Systems, 33, 1087-1088. https://doi.org/10.1109/TPWRS.2017.2688178
[9] Funahashi, K.I. and Nakamura, Y. (1993) Approximation of Dynamical Systems by Continuous Time Recurrent Neur- al Networks. Neural Networks, 6, 801-806. https://doi.org/10.1016/S0893-6080(05)80125-X
[10] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) Imagenet Classification with Deep Convolutional Neural Net- works. Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, 3-6 December 2012, 1097-1105.
[11] Weimin WANG, Yufeng LI, Xu YAN, Mingxuan XIAO, & Min GAO. (2024). Enhancing Liver Segmentation: A Deep Learning Approach with EAS Feature Extraction and Multi-Scale Fusion. International Journal of Innovative Research in Computer Science & Technology, 12(1), 26–34. Retrieved from https://ijircst.irpublications.org/index.php/ijircst/article/view/21
[12] Dai, W., Tao, J., Yan, X., Feng, Z., & Chen, J. (2023, November). Addressing Unintended Bias in Toxicity Detection: An LSTM and Attention-Based Approach. In 2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA) (pp. 375-379). IEEE.
[13] Yan, K., Du, Y. and Ren, Z. (2018) MPPT Perturbation Optimization of Photovoltaic Power Systems Based on Solar Irradiance Data Classification. IEEE Transactions on Sustainable Energy, 10, 514-521. https://doi.org/10.1109/TSTE.2018.2834415
[14] Ma, D., Dang, B., Li, S., Zang, H., & Dong, X. (2023).
[15] Implementation of computer vision technology based on artificial intelligence for medical image analysis. International Journal of Computer Science and Information Technology, 1(1), 69-76.
[16] Tianbo, S., Weijun, H., Jiangfeng, C., Weijia, L., Quan, Y., & Kun, H. (2023, January). Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition. In 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE) (pp. 834-837). IEEE.
[17] Hao Xu, Shuang Song, and Ze Mao. 2023. Characterizing the Performance of Emerging Deep Learning, Graph, and High Performance Computing Workloads Under Interference. arXiv:2303.15763
[18] Liu, B. (2023). Based on intelligent advertising recommendation and abnormal advertising monitoring system in the field of machine learning. International Journal of Computer Science and Information Technology, 1(1), 17-23.
[19] Xu, H., & Colmenares, J. A. (2023). Admission Control with Response Time Objectives for Low-latency Online Data Systems. arXiv preprint arXiv:2312.15123.
[20] Hao Xu, Qingsen Wang, Shuang Song, Lizy Kurian John, and Xu Liu. 2019. Can we trust profiling results? Understanding and fixing the inaccuracy in modern profilers. In Proceedings of the ACM International Conference on Supercomputing. 284–295.
[21] Zeiler, M.D. (2012) ADADELTA: An Adaptive Learning Rate Method.
[22] Zang, H., Li, S., Dong, X., Ma, D., & Dang, B. (2024). Evaluating the Social Impact of AI in Manufacturing: A Methodological Framework for Ethical Production. Academic Journal of Sociology and Management, 2(1), 21–25. https://doi.org/10.5281/zenodo.10474511
[23] Wan, W., Xu, J., Zeng, Q., Pan, L., & Sun, W. (2023). Development and Evaluation of Intelligent Medical Decision Support Systems. Academic Journal of Science and Technology, 8(2), 22-25.
[24] Hao Hu, Shulin Li, Jiaxin Huang, Bo Liu, and Chang Che. 2023. Casting Product Image Data for Quality Inspection with Xception and Data Augmentation. Journal of Theory and Practice of Engineering Science 3, 10 (Oct. 2023), 42–46. https: //doi.org/10.53469/jtpes.2023.03(10).06
[25] Ni, F., Zang, H., & Qiao, Y. (2024, January). Smartfix: Leveraging machine learning for proactive equipment maintenance in industry 4.0. In The 2nd International scientific and practical conference “Innovations in education: prospects and challenges of today”(January 16-19, 2024) Sofia, Bulgaria. International Science Group. 2024. 389 p. (p. 313).
[26] Liu, S., Wu, K., Jiang, C., Huang, B., & Ma, D. (2023). Financial Time-Series Forecasting: Towards Synergizing Performance And Interpretability Within a Hybrid Machine Learning Approach. arXiv preprint arXiv:2401.00534.
[27] Sun, W., Wan, W., Pan, L., Xu, J., & Zeng, Q. (2024). The Integration of Large-Scale Language Models Into Intelligent Adjudication: Justification Rules and Implementation Pathways. Journal of Industrial Engineering and Applied Science, 2(1), 13–20. https://doi.org/10.5281/zenodo.10607564
[28] Pan, L., Sun, W., Wan, W., Zeng, Q., & Xu, J. (2023). Research Progress of Diabetic Disease Prediction Model in Deep Learning. Journal of Theory and Practice of Engineering Science, 3(12), 15-21.

[29] Hengyi Zang. (2024). Precision Calibration of Industrial 3D Scanners: An AI-Enhanced Approach for Improved Measurement Accuracy. Global Academic Frontiers, 2(1), 27-37. https://gafj.org/journal/article/view/30