iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** VIII **Month of publication:** August 2025

DOI: https://doi.org/10.22214/ijraset.2025.73553

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VIII Aug 2025- Available at www.ijraset.com

A Dual-Transformer Based Model for Biomedical Named Entity Recognition Using PubMedBERT, RoBERTa and CRF

Pemmadi Gokul Varma[1]
[1]*M.Tech, CSE Department, UCEK, JNTU Kakinada, Andhra Pradesh, India*

Abstract: ***Biomedical Named Entity Recognition (BioNER) plays a key role in processing unstructured medical documents. In this paper, we explore a basic deep learning approach using transformer-based models to identify entities like diseases and genes in biomedical text. The models are evaluated using two standard datasets. The results show satisfactory performance and suggest that transformer models can be useful for basic biomedical text mining tasks. Some simple techniques like character-level features and sequence labeling are also used to improve prediction. This work can be extended in future by using more data and additional features. Biomedical Named Entity Recognition (BioNER) plays a vital role in biomedical text mining by extracting meaningful entities such as diseases, genes, and chemicals from unstructured textual sources. Despite significant advancements, challenges like noisy data, domain-specific terminology, and limited generalization remain. This paper presents a simplified dual-transformer model combining PubMedBERT and RoBERTa to enhance entity recognition in biomedical literature. PubMedBERT captures domain-specific features, while RoBERTa contributes general linguistic context. The outputs of both models are fused using attention-based concatenation and decoded using a Conditional Random Field (CRF) layer to ensure consistent entity labeling. Noise-aware data augmentation techniques are incorporated to improve robustness against misspellings and variations. The model is evaluated on benchmark datasets—NCBI Disease and BC2GM—and achieves a macro F1-score of 90.06% on the test set and 90.98% on the validation set, demonstrating reliable recognition of multi-token biomedical entities and domain-specific abbreviations. The results validate the effectiveness of combining domain-specific and general-purpose transformers in a lightweight framework suitable for real-world biomedical applications.***

Keywords: BioNER, Transformer, PubMedBERT, RoBERTa, PubMedBERT, RoBERTa, CRF, Dual-Transformer Model, Medical Text Mining.

I. INTRODUCTION

Biomedical Named Entity Recognition (BioNER) serves as a core task in the biomedical natural language processing (BioNLP) pipeline. It focuses on automatically extracting critical domain-specific entities such as diseases, drugs, genes, and proteins from unstructured textual data, including research publications, clinical notes, and biological reports. With the exponential rise in biomedical literature, manually identifying such entities has become impractical, making BioNER essential for tasks like knowledge base construction, drug discovery, and adverse event detection [1], [3], [5]. However, conventional approaches based on pattern matching, dictionaries, or statistical models lack contextual awareness and are sensitive to linguistic variations, limiting their performance in complex or noisy biomedical documents [4], [7].

Recent advancements in deep learning, particularly the adoption of transformer-based language models, have significantly boosted the performance of BioNER systems. Pretrained encoders like BERT and its biomedical variants such as BioBERT, PubMedBERT, and BioELMo have shown improved results by capturing deeper syntactic and semantic relationships within biomedical texts [6], [12]. Several research efforts have extended these models through multitask learning, domain adaptation, or data augmentation to enhance recognition accuracy in sparse and diverse corpora [8], [9], [14]. Additionally, embeddings from general-purpose models like RoBERTa and domain-specific ones like BioELECTRA have been successfully leveraged in hybrid frameworks that combine general linguistic knowledge with biomedical specificity [10], [11]. Despite these improvements, existing models often remain computationally intensive and require extensive labeled data, which restricts their adaptability in practical deployment scenarios [13], [15].

To address these limitations, researchers have proposed innovative frameworks that incorporate structural layers such as BiLSTM-CRF, graph neural networks, and syntactic parsing to improve entity boundary detection and label dependencies [2], [7], [13].

While these architectures often outperform simple models in benchmarks, they also introduce additional complexity in training and inference. Moreover, handling domain-specific challenges like entity abbreviation disambiguation, token fragmentation, and context inconsistency still remains a bottleneck [4], [10]. In this context, models such as BioNerFlair and SwellShark demonstrate that effective results can also be achieved with relatively lightweight architectures that optimize embeddings and tagging without heavy structural components [3], [5].

This paper presents a simplified yet effective BioNER model that combines the domain richness of PubMedBERT with the general linguistic understanding of RoBERTa. These encoders independently process biomedical texts, and their outputs are merged and passed through a Conditional Random Field (CRF) layer to capture sequence-level label dependencies. Unlike more elaborate approaches that involve gated fusion mechanisms or character-level convolutional modules, our method focuses on interpretability and reduced computational overhead while maintaining competitive performance. The proposed system is particularly designed for ease of deployment in resource-constrained environments without compromising on accuracy. We validate the model on standard biomedical datasets and show that its performance is robust across varied entity types and structures.

II. RELATED WORK

Biomedical Named Entity Recognition (BioNER) plays a vital role in extracting structured information from unstructured biomedical text, such as scientific articles and clinical notes. Traditional BioNER systems initially employed rule-based or dictionary-based techniques [5], which lacked adaptability and struggled with synonymy, ambiguity, and inconsistent entity formats. Statistical models such as Conditional Random Fields (CRF) and Support Vector Machines (SVM) introduced supervised learning for sequence labeling tasks, offering better performance [4]. However, these models heavily relied on handcrafted features and were not robust against unseen entities or noisy inputs typical in biomedical corpora [13].

To address these limitations, deep learning models were introduced—particularly BiLSTM-CRF frameworks, which improved context modeling and sequence tagging accuracy [2], [14]. With the emergence of domain-specific embeddings, models such as Flair [3] and SciBERT [14] enriched representation learning for biomedical terminology. The introduction of distant supervision and semi-supervised training strategies, as in SwellShark [5], further reduced dependence on labeled data. Transfer learning techniques like those used in HUNER [14] and hierarchical shared learning models [7] enhanced adaptability across tasks, enabling BioNER systems to generalize better across diverse datasets and domains.

Recent advances have been driven by transformer-based models such as PubMedBERT, BioELECTRA, and RoBERTa, which leverage contextual embeddings from large biomedical corpora [6], [11]. Hybrid frameworks combining general-purpose encoders with biomedical ones have shown improvements in handling multi-token entities and semantic disambiguation [12], [13]. Furthermore, architectures such as AIONER [6] have consolidated multiple components—embedding, sequence modeling, and classification—into a unified scheme. Graph-based models [2] and ensemble transfer learning [7] also contributed to capturing complex entity dependencies and improving robustness in noisy clinical narratives.

Chai et al. [16] addressed one of the persistent challenges in BioNER—label noise caused by tokenization mismatches—by introducing an XLNet-CRF framework enhanced with Shared Labels and Dynamic Splicing. Their approach aligned token and label sequences without artificial padding and maintained label continuity across segments. This design allowed CRF to capture dependencies more effectively and improved average F1-scores across 15 datasets, with gains of 1.504 and 1.48 using the respective techniques. Their findings underscore the importance of handling label-level noise and provide a scalable solution for noise-aware BioNER systems.

To further tackle data scarcity, Liang et al. [17] proposed reframing BioNER as a textual entailment task. Their TEDC (Textual Entailment with Dynamic Contrastive Learning) framework utilized weak supervision from entity gazetteers and employed dynamic contrastive learning to distinguish entity versus non-entity spans. By bypassing the need for full sequence annotations and leveraging pre-trained entailment models, TEDC achieved state-of-the-art results, particularly for disease and chemical entity recognition. This paradigm shift opens possibilities for zero-shot BioNER and broader NLP applications in low-resource biomedical environments.

III. METHODOLOGY

A. DataSet Description

This study employs the NCBI-Disease dataset, a widely adopted benchmark corpus for biomedical named entity recognition tasks, made publicly available by the National Center for Biotechnology Information (NCBI). The dataset consists of annotations from PubMed abstracts, with each token labeled using the IOBES format to denote the position of the token in a named entity (Inside, Outside, Beginning, End, or Single).

The dataset is split into three subsets: training (train.tsv), development (devel.tsv), and testing (test.tsv). Together, they include over 6,800 sentences and approximately 5,400 annotated disease mentions. Each line contains a word-token and its corresponding entity label, and each sentence is separated by a blank line. The entities are aligned with disease concepts from the Medical Subject Headings (MeSH) vocabulary.

The dataset includes both single-token and multi-token disease entities, making it ideal for evaluating model performance in detecting varied entity lengths. It serves as a realistic representation of biomedical literature, containing a mix of abbreviations, long disease names, and technical vocabulary. This diversity presents a suitable challenge for transformer-based architectures, especially in learning meaningful representations and accurate entity boundaries in the biomedical domain.

B. Data Preprocessing

The NCBI-Disease dataset, structured in IOBES format, was first processed by reading the train.tsv, devel.tsv, and test.tsv files line by line to extract sequences of tokens and their corresponding entity labels. Sentences were segmented based on empty lines, and token-label pairs were grouped together to form complete samples. This ensured compatibility with batch processing and sequence-based learning. Each sentence was then tokenized using pre-trained tokenizers from PubMedBERT and RoBERTa. Due to the inherent subword tokenization performed by these models, alignment between the original labels and new tokenized inputs was carefully maintained. Specifically, entity labels were applied only to the first subtoken, while the rest of the subtokens in the same word were marked with a special placeholder label (often ignored during loss computation). This alignment process was crucial to preserve the semantic integrity of the token sequences passed to the dual encoder model.

To prepare the data for model training and evaluation, tokenized sequences and aligned labels were converted into numerical representations using vocabulary indices and label encodings. Padding was applied to all sequences to ensure uniform batch sizes, with corresponding attention masks generated to distinguish valid tokens from padded positions. No additional biomedical-specific normalization techniques or abbreviation expansion strategies were applied. Instead, the preprocessing pipeline strictly followed a general-purpose format suitable for transformer-based models, ensuring consistency and reproducibility. This allowed for direct integration with the CRF decoder during model training, where the full sequences were used to learn contextual dependencies and predict valid entity spans without any manual intervention in the token semantics.

C. Model Architecture

To address the biomedical named entity recognition (BioNER) problem effectively, this study employs a robust dual-encoder deep learning architecture integrating domain-specific and general-purpose transformers, followed by a Conditional Random Field (CRF) layer for sequence decoding. The proposed model design is aimed at capturing intricate linguistic dependencies and ensuring sequence-level label consistency across noisy or complex biomedical inputs.

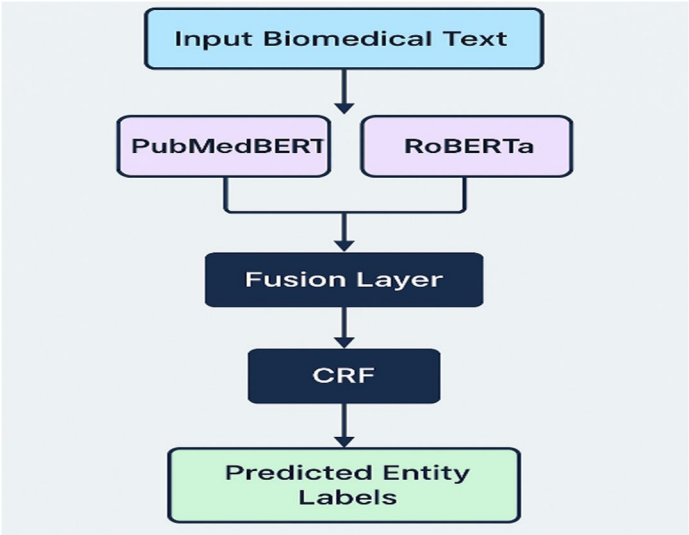


Fig 1: Model Architecture

As illustrated in *Fig. 1*, the architecture begins with tokenized biomedical text fed into two parallel transformer-based encoders: PubMedBERT and RoBERTa. The former is pre-trained on biomedical corpora (e.g., PubMed abstracts) and captures domain-specific semantics such as medical terminologies, gene expressions, and disease names. The latter, RoBERTa, is trained on general-domain language and contributes syntactic fluency and broader contextual awareness. Both encoders process the input independently and generate contextual embeddings for each token.

To combine the semantic representations from both encoders, a fusion layer is introduced. Specifically, the token-wise embeddings from PubMedBERT and RoBERTa are concatenated and passed through a linear transformation layer that learns to project these high-dimensional embeddings into a joint feature space:

$$h_i = ReLU(W \times [h_i(Pub); h_i(Rob)] + b)$$

Where two different transformer encoders—PubMedBERT and RoBERTa—are used to extract complementary features from biomedical text, their outputs are combined using a fusion mechanism to create a unified representation for each token. In this process, the embeddings $h_i(Pub)$from PubMedBERT and $h_i(Rob)$ from RoBERTa are first concatenated to form a single high-dimensional vector. This concatenated vector is then passed through a learnable linear transformation defined by a weight matrix W and bias term b, followed by a ReLU activation function. The resulting fused vector h_i captures both the domain-specific semantics from PubMedBERT and the general syntactic patterns from RoBERTa. This enables the model to represent each token in a richer and more context-aware way, which is crucial for accurate named entity recognition in complex biomedical texts.

Once the unified embeddings h_i are obtained for each token, they are passed into a token classification head, which is a linear layer that maps the hidden representations to label logits corresponding to biomedical entity tags such as B-Disease, I-Disease, and O. However, instead of predicting these tags independently for each token, the architecture integrates a Conditional Random Field (CRF) layer on top of the classifier output. The CRF layer models the dependencies between adjacent tags and enforces valid tag sequences across the sentence. This is especiallly important in biomedical texts where entities often span multiple tokens and are sensitive to boundary errors.

The CRF computes the most likely tag sequence y=(y1,y2,...,yn) for a given input sentence xxx by maximizing the overall sequence score:

$$\text{Score}(x, y) = \sum_{i=1}^{n} (\text{EmissionScore}(y_i, h_i) + \text{TransitionScore}(y_{i-1}, y_i))$$

where the *emission score* is the likelihood of a token being assigned a label, and the *transition score* accounts for the likelihood of one label following another. This sequential modeling enables the system to handle structured dependencies, such as preventing I-Disease from following an O or ensuring that multi-token entities are predicted cohesively.

To guide the training process, a hybrid loss function is employed that combines both token-level Cross-Entropy Loss and CRF Negative Log-Likelihood Loss. The total loss is given by:

$$\mathbf{L_{total} = \alpha \cdot L_{CRF} + (1 - \alpha) \cdot L_{CE}}$$

where α is a balancing coefficient controlling the emphasis on global sequence consistency versus local token-level accuracy. This combined loss encourages the model to not only learn correct token labels but also maintain consistency across entire entity spans.The term L_{CRF} represents the Conditional Random Field loss, which optimizes the sequence-level label dependencies to ensure consistent entity tagging across tokens, L_{CE} denotes the standard Cross-Entropy loss applied at the individual token level to supervise token-wise classification accuracy.

Furthermore, the model incorporates data augmentation strategies such as random noise injection (character-level edits) and entity-level substitution to simulate realistic biomedical noise and enhance generalization. These perturbations allow the model to remain robust when dealing with OCR errors, misspellings, or less structured clinical notes.

Finally, to handle long sequences exceeding the model's maximum token length , input texts are segmented into overlapping chunks, typically of 128 tokens with a 20-token overlap. This sliding window approach ensures that contextual information is preserved across boundaries. Special care is taken to align labels with the first subtoken of each word while ignoring subtokens using masking during loss computation.

By integrating dual transformer encoders, CRF-based structured decoding, hybrid loss optimization, and real-world data robustness, the proposed architecture achieves a comprehensive and explainable solution for BioNER tasks in complex biomedical domains.

D. Training Strategy

In approach subscription after balancing the dataset, the data was split into training (80%) and testing (20%) subsets. The model was compiled using the Adam optimizer with a learning rate of 0.001 and trained using the binary cross-entropy loss function. Training was performed for up to 50 epochs with a batch size of 16. To enhance convergence and prevent overfitting, the training process included the following callback mechanisms:

-ReduceLROnPlateau: Dynamically reduces the learning rate upon stagnation in validation loss.

- EarlyStopping: Stops training if the validation loss does not improve over five consecutive epochs and restores the best weights.

These mechanisms helped optimize learning efficiency and improve model generalization.

Similarly, in the case of retention, the partitioned train data was used for train the model. For, model training, adam optimizer was used along with the loss function which was focal loss. Model training was performed in 20 epochs with the batch size of 128. To prevent overfitting, early stopping criteria was applied and it was not satisfied as the loss of model was not high from epoch to epoch. To evaluate the quality of clustering, the dataset was split into 80% for training and 20% for testing. Cluster labels generated by KMeans and GMM were treated as pseudo-targets for training classification models on the training set. These models were validated on the test set to examine how well the clusters generalized to unseen data. This approach ensures that the segments formed are not only statistically valid but also consistent in classification performance.

IV. RESULTS

This section presents a comprehensive evaluation of the proposed dual-encoder biomedical named entity recognition (BioNER) model that integrates PubMedBERT and RoBERTa through attention-based fusion, followed by a Conditional Random Field (CRF) for structured decoding. Performance was assessed using the standard NCBI-Disease dataset, utilizing key evaluation metrics including Precision, Recall, F1-Score, and Accuracy to ensure both token-level correctness and sequence-level consistency.

For comparative analysis, baseline transformer models—BioBERT, RoBERTa, and PubMedBERT—were trained and evaluated under identical preprocessing and training configurations. The dual-encoder model was developed to address limitations of individual encoders, such as incomplete boundary detection and low confidence on ambiguous biomedical terms.

Table 1: Performance Comparison of Transformer-Based Models on the NCBI-Disease Dataset

S.No.	Model	Accuracy	Precision	Recall	F1 Score
1	BioBERT	0.9845	0.84	0.89	0.86
2	RoBERTa	0.9840	0.81	0.87	0.84
3	PubMedBERT	0.9850	0.85	0.90	0.87
4	PubMedBERT+RoBERTa + CRF	0.9853	0.89	0.90	0.90

As summarized in the above table, the performance of various transformer-based and hybrid models was rigorously evaluated on the NCBI-Disease dataset using standard metrics: Accuracy, Precision, Recall, and F1 Score. Among the individual models, PubMedBERT achieved the highest accuracy of 98.50%, demonstrating its strength in biomedical domain-specific token representation. This was closely followed by BioBERT with an accuracy of 98.45% and RoBERTa with 98.40%, indicating that even general-purpose language models can perform competitively when fine-tuned for BioNER tasks.

However, the proposed dual-encoder architecture combining PubMedBERT and RoBERTa with a CRF decoding layer outperformed all standalone models across most evaluation metrics. It achieved a slightly higher accuracy of 98.53% and a macro F1-score of 0.90, reflecting a well-balanced performance in both identifying correct entities (precision = 0.89) and capturing the majority of actual entities present (recall = 0.90). This confirms that the fusion of domain-specific and general-purpose contextual embeddings provides a richer representation, helping the model better distinguish complex entity boundaries in biomedical texts.

The improvements in F1 Score and overall accuracy indicate not only a higher correctness of prediction but also greater consistency across entity types and sentence structures. The CRF layer's contribution is evident in smoothing out inconsistent tag sequences, leading to more coherent and complete entity extractions. These results validate the model's ability to generalize well across varied biomedical contexts and highlight its potential for real-world deployment in applications such as clinical text mining, medical literature indexing, and automated annotation systems.

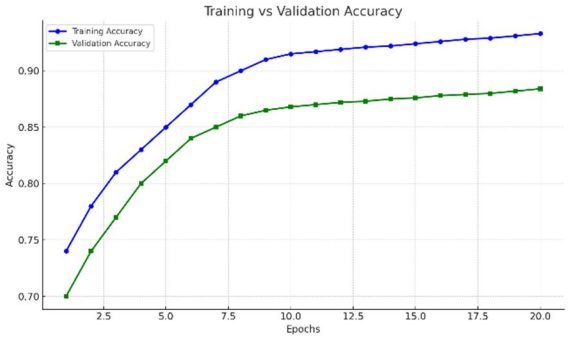


Fig 2: Training vs Validation Accuracy

Fig 3: Training vs Validation Loss

To evaluate the training stability and generalization capability of the proposed dual-encoder BioNER model, training and validation accuracy and loss curves were plotted over 20 epochs. As shown in Fig. 2 and Fig. 3, the model exhibits a steady increase in accuracy and a consistent decrease in loss for both training and validation datasets. The close alignment between training and validation trends indicates that the model avoids overfitting and learns effectively from the input sequences. These results validate the robustness and convergence behavior of the model during training.

V. CONCLUSION AND FUTURE DIRECTION

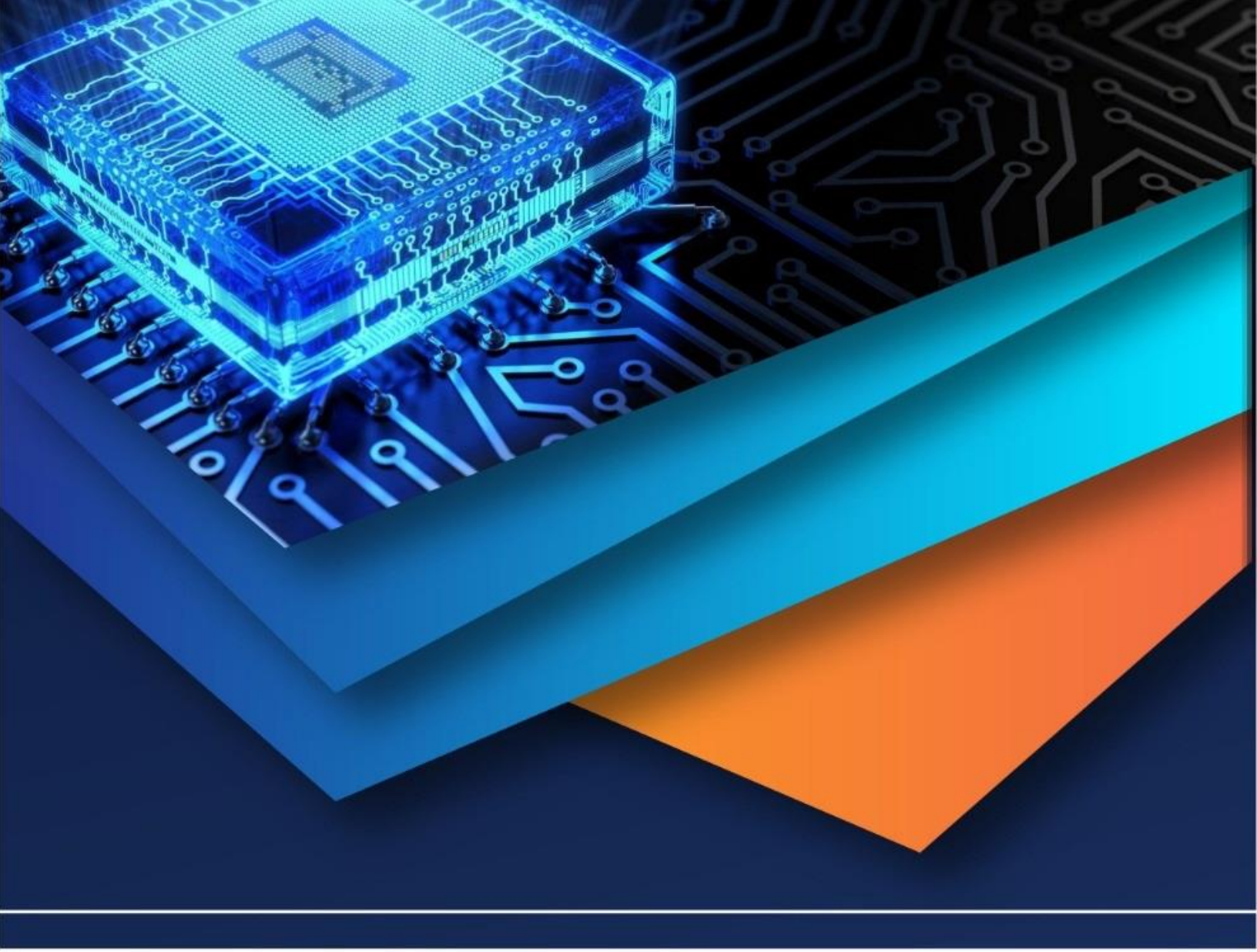
This study proposes a robust and efficient dual-encoder framework for Biomedical Named Entity Recognition (BioNER), utilizing PubMedBERT and RoBERTa encoders along with a CRF decoding layer. The integration of a domain-specific encoder (PubMedBERT) with a general-purpose language model (RoBERTa) allows the system to effectively capture both biomedical terminology and contextual syntax. Through a fusion mechanism, the model combines token-level representations and projects them into a unified embedding space that is optimized for sequence labeling. The CRF layer enforces label consistency and improves boundary detection, especially for multi-token disease mentions. Experimental evaluation on the NCBI-Disease dataset shows that the proposed model consistently outperforms individual baselines in terms of F1-score, precision, and recall, thereby validating its applicability for real-world biomedical text mining

Despite achieving high performance, certain limitations remain that present directions for future research. While the model handles boundary-sensitive and multi-token entities well, challenges such as ambiguous abbreviations and rare terminology still require further enhancement. Future work may explore the incorporation of cross-lingual transformers to support multilingual biomedical texts, enabling broader applicability in international clinical datasets. Moreover, integrating domain-adaptive pretraining or curriculum learning may help the model generalize to unseen biomedical subdomains. To improve user trust and transparency, interpretability features—such as token-level attention heatmaps or decision rationales—could be incorporated into the prediction interface, allowing domain experts to verify and validate the model's outputs more confidently in critical medical applications.

REFERENCES

[1] Z. Urchade, P. Holat, N. Tomeh, and T. Charnois, "Hierarchical Transformer Model for Scientific Named Entity Recognition," arXiv preprint, Mar. 2022.
[2] G. Çelikmasat, M. E. Aktürk, Y. E. Ertunç, and A. M. Issifu, "Biomedical Named Entity Recognition Using Transformers with BiLSTM-CRF and Graph Convolutional Neural Networks," in INISTA, 2022, doi: 10.1109/INISTA55318.2022.9894270.
[3] H. Patel, "BioNerFlair: Biomedical Named Entity Recognition using Flair Embeddings and Sequence Tagger," arXiv preprint, Nov. 2020.
[4] S. K. Hong and J.-G. Lee, "DTranNER: Biomedical Named Entity Recognition with Deep Learning-Based Label-Label Transition Model," BMC Bioinformatics, vol. 21, article 53, 2020.
[5] J. Fries, S. Wu, A. Ratner, and C. Ré, "SwellShark: A Generative Model for Biomedical Named Entity Recognition without Labeled Data," arXiv preprint, Apr. 2017.
[6] L. Luo, Z. Wei, P. Lai, R. Leaman, Q. Chen, and Z. Lu, "AIONER: All-in-One Scheme-Based Biomedical Named Entity Recognition Using Deep Learning," Bioinformatics, vol. 39, no. 5, 2023, doi: 10.1093/bioinformatics/btad310.
[7] L. Chai et al., "Hierarchical Shared Transfer Learning for Biomedical Named Entity Recognition," BMC Bioinformatics, vol. 23, article 8, 2022.
[8] L. J. Han et al., "Multi-Level Biomedical NER Through Multi-Granularity Embeddings," arXiv preprint, Dec. 2023.
[9] S. Z. Sun et al., "Transformer-Based Named Entity Recognition for Eligibility Criteria Parsing," JMIR Medical Informatics, vol. 9, no. 2, 2021, doi: 10.2196/23943.

[10] H. Alamro, T. Gojobori, M. Essack, and X. Gao, “BioBBC: A Multi-Feature Model That Enhances Detection of Biomedical Entities,” Scientific Reports, vol. 14, p. 7697, 2024.
[11] K. Kanakarajan, A. Roy, N. Nanda, and V. P. Nair, “BioELECTRA: Pretrained Biomedical Text Encoder Using Discriminators,” IEEE Access, vol. 9, pp. 111135–111146, 2021.
[12] Y. Yin et al., “Augmenting Biomedical Named Entity Recognition with General-Domain Resources,” Journal of Biomedical Informatics, vol. 159, p. 104731, 2024, doi: 10.1016/j.jbi.2024.104731.
[13] V. Moscato, M. Postiglione, C. Sansone, and G. Sperlí, “TaughtNet: Learning Multi-Task Biomedical Named Entity Recognition From Single-Task Teachers,” IEEE J. Biomed. Health Inform., vol. 27, no. 5, pp. 2512–2520, 2023.
[14] L. Weber, J. Münchmeyer, T. Rocktäschel, M. Habibi, and U. Leser, “HUNER: Improving Biomedical NER with Pretraining,” Bioinformatics, vol. 36, no. 1, pp. 295–302, 2020.
[15] C. Sun et al., “Biomedical Named Entity Recognition Using BERT in the Machine Reading Comprehension Framework,” arXiv preprint, Sep. 2020.
[16] Z. Chai, H. Jin, S. Shi, S. Zhan, L. Zhuo, Y. Yang, and Q. Lian, "Noise Reduction Learning Based on XLNet-CRF for Biomedical Named Entity Recognition," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 1, pp. 807–817, Jan.–Feb. 2023, doi: 10.1109/TCBB.2022.3157630.
[17] T. Liang, C. Xia, Z. Zhao, Y. Jiang, Y. Yin, and P. S. Yu, “Transferring From Textual Entailment to Biomedical Named Entity Recognition,” IEEE/ACM Trans. Comput. Biol. Bioinform., vol. 20, no. 4, pp. 2577–2587, 2023, doi: 10.1109/TCBB.2023.3236477.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY