

Article

Not peer-reviewed version

---

# Natural Language Processing (NLP) for Semantic Understanding and Knowledge Extraction from Unstructured Clinical Notes

---

[Owen Graham](#)<sup>\*</sup> and Dean Iloent

Posted Date: 11 June 2025

doi: 10.20944/preprints202506.0961.v1

Keywords: NLP; natural language processing



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Natural Language Processing (NLP) for Semantic Understanding and Knowledge Extraction from Unstructured Clinical Notes

Owen Graham \*and Dean Llorent

Independent Researcher; deanllore@mail.com

\* Correspondence: topscribble@gmail.co

**Abstract:** Unstructured clinical notes, comprising a significant portion of electronic health records (EHRs), contain a wealth of invaluable patient information that remains largely inaccessible for systematic analysis due to its free-text format. This inaccessibility limits the potential for data-driven insights, hindering advancements in clinical research, personalized medicine, and healthcare operational efficiency. This study proposes to leverage advanced Natural Language Processing (NLP) techniques to achieve comprehensive semantic understanding and robust knowledge extraction from these critical unstructured clinical narratives. The research will focus on developing and evaluating state-of-the-art NLP models, including transformer-based architectures and specialized medical entity recognition algorithms, augmented with semantic parsing and relation extraction capabilities. The methodology will encompass strategies for handling the unique challenges of clinical language, such as abbreviations, colloquialisms, negation, and temporal expressions, to accurately identify clinical entities, their attributes, and inter-relationships. This will involve the exploration of self-supervised learning approaches for pre-training on large clinical corpora and fine-tuning on annotated datasets to ensure domain-specific semantic understanding. Expected outcomes include the automated conversion of free-text clinical data into structured, computable formats, thereby facilitating more precise patient phenotyping, identifying subtle disease patterns, and supporting the development of highly individualized treatment plans. Furthermore, the extracted knowledge will enhance clinical decision support systems, streamline administrative tasks, and enable large-scale epidemiological studies. This research aims to significantly advance the utility of unstructured clinical data, bridging the gap between raw text and actionable intelligence, and ultimately contributing to more efficient, precise, and patient-centered healthcare delivery.

**Keywords:** NLP; natural language processing

---

## Chapter 1: Introduction

### 1.1. Background

Healthcare systems globally are undergoing a profound transformation, driven by the increasing digitization of patient information. Electronic Health Records (EHRs) have become central to modern medical practice, offering the promise of comprehensive patient data management. While EHRs effectively store structured data such as diagnoses codes, laboratory results, and medication lists, a substantial and often more nuanced portion of clinical information resides within **unstructured clinical notes**. These free-text narratives, penned by physicians, nurses, and other healthcare professionals, capture the intricate details of patient encounters, clinical reasoning, symptomatic descriptions, and nuanced observations that are critical for understanding a patient's holistic health journey. However, the inherent unstructured nature of this data renders it largely inaccessible for large-scale computational analysis, thereby limiting its potential to drive significant

advancements in personalized medicine, clinical research, and operational efficiencies within healthcare institutions. The current manual review processes for extracting insights from these notes are resource-intensive, time-consuming, and prone to variability, underscoring a critical need for automated and robust solutions.

### 1.2. Problem Statement

The primary challenge addressed by this research is the semantic and syntactic inaccessibility of the vast quantities of clinical information embedded within unstructured free-text narratives in EHRs. Despite containing crucial details pertinent to patient care, outcomes, and epidemiological trends, this information is not readily computable or integrable with structured datasets for systematic analysis. This inaccessibility creates several significant barriers:

- **Hindered Clinical Research:** Researchers struggle to identify specific patient cohorts, track disease progression, or evaluate treatment efficacy on a large scale without laborious manual chart review.
- **Limited Personalized Medicine:** The nuanced factors that contribute to an individual's unique health profile, often captured only in notes, are overlooked, preventing the full realization of tailored treatment strategies.
- **Suboptimal Healthcare Operational Efficiency:** Administrative burdens related to coding, billing, and quality reporting are exacerbated by the need to manually interpret free-text entries.
- **Missed Opportunities for Decision Support:** Clinical decision support systems often lack the rich contextual information present in notes, leading to less precise recommendations.

Therefore, the fundamental problem is how to efficiently and accurately transform these heterogeneous, complex, and highly contextualized clinical narratives into structured, machine-readable formats that can unlock their full analytical potential.

### 1.3. Research Significance

This research holds significant implications for advancing healthcare in multiple dimensions. By enabling the automated **semantic understanding and knowledge extraction** from unstructured clinical notes, this study contributes to:

- **Enhanced Precision Medicine:** Facilitating a deeper, more granular understanding of individual patient characteristics, leading to highly personalized and effective treatment plans.

This moves beyond 'one-size-fits-all' approaches to truly patient-centric care.

- **Accelerated Clinical Research and Discovery:** Providing researchers with unprecedented access to real-world patient data, enabling faster identification of disease patterns, drug efficacy signals, and adverse event associations that might otherwise remain undiscovered within isolated free-text.
- **Improved Clinical Decision Support:** Equipping clinicians with comprehensive, data-driven insights derived directly from patient narratives, thereby enhancing diagnostic accuracy, optimizing treatment pathways, and reducing medical errors.
- **Streamlined Healthcare Operations:** Automating the extraction of key information can significantly reduce administrative overhead, improve coding accuracy, and enhance compliance reporting, allowing healthcare professionals to dedicate more time to direct patient care.
- **Advancement of Public Health Informatics:** Enabling large-scale epidemiological studies, surveillance of disease outbreaks, and population health management by providing computable access to vast amounts of previously inaccessible clinical data.

Ultimately, this research is poised to bridge the critical gap between raw clinical text and actionable intelligence, fostering a more efficient, insightful, and patient-centered healthcare ecosystem.

#### 1.4. Research Objectives

This study aims to achieve the following core objectives:

1. To develop and rigorously evaluate state-of-the-art Natural Language Processing (NLP) models, specifically leveraging transformer-based architectures, for comprehensive semantic understanding of unstructured clinical notes.
2. To implement specialized medical entity recognition algorithms capable of accurately identifying clinical entities, their attributes, and their relationships within the complex and often ambiguous clinical lexicon.

3. To design and integrate robust semantic parsing and relation extraction capabilities to infer meaningful connections between identified entities, thereby constructing structured knowledge from free-text narratives.
4. To explore and apply self-supervised learning approaches for effective pre-training on large, unlabeled clinical corpora, optimizing models for domain-specific language nuances.
5. To validate the efficacy of the proposed NLP framework in converting free-text clinical data into computable formats that support precise patient phenotyping, disease pattern identification, and individualized treatment planning.

### 1.5. Scope of the Study

The scope of this research is primarily focused on the **development, implementation, and evaluation of Natural Language Processing methodologies** for the automated extraction of structured knowledge from unstructured clinical notes. While the ultimate application of this knowledge lies within various healthcare domains (e.g., personalized medicine, clinical decision support), the core emphasis is on the AI techniques themselves. The study will concentrate on addressing the unique linguistic challenges inherent in clinical text, such as acronyms, context-dependent meanings, negation, and temporal information. The research will not involve the collection of new patient data but will rather focus on methodological advancements in NLP applied to existing de-identified clinical datasets. The generalizability of the developed models to diverse clinical specialties and different healthcare systems will be considered within the evaluation framework.

## Chapter 2: Literature Review

### 2.1. Introduction to Natural Language Processing in Healthcare

The burgeoning volume of digital health data has positioned Natural Language Processing (NLP) as a pivotal technology for transforming raw clinical narratives into computable insights. Historically, NLP in healthcare has evolved from rule-based systems, which relied on handcrafted patterns and lexicons, to sophisticated statistical and machine learning approaches. Early applications focused on basic information extraction, such as identifying diseases or medications, often with limited contextual understanding. However, the advent of deep learning, particularly with the development of neural network architectures, has dramatically enhanced NLP's capacity to comprehend the intricate semantics of clinical text. This shift has unlocked unprecedented opportunities for automated analysis, ranging from augmenting clinical workflows to powering advanced research initiatives.

### 2.2. Challenges of Clinical Text Analysis

Clinical language presents unique and formidable challenges that differentiate it significantly from general domain text. These complexities necessitate specialized NLP approaches to ensure accuracy and reliability. Key challenges include:

- **Abundant Use of Abbreviations and Acronyms:** Clinicians frequently employ highly context-dependent abbreviations (e.g., "CHF" for Congestive Heart Failure, "PT" for Physical Therapy or Patient), which can be highly ambiguous without proper contextual understanding.
- **Colloquialisms and Informal Language:** Despite being professional documentation, clinical notes often contain informal phrasings, shorthand, and even typographical errors.
- **Negation and Speculation:** Accurately identifying negated concepts (e.g., "no evidence of tumor") or speculative statements (e.g., "patient denies pain") is crucial for correct interpretation and avoids critical misclassifications.
- **Temporal Expressions:** Extracting and standardizing temporal information (e.g., "onset last week," "improving since yesterday") is vital for tracking disease progression and treatment effectiveness.
- **Coreference Resolution:** Identifying when different mentions in a text refer to the same entity (e.g., "patient," "he," "Mr. Smith") is complex but essential for coherent understanding.
- **Protected Health Information (PHI):** The presence of sensitive patient identifiers necessitates robust de-identification techniques, often requiring NLP itself, before data can be used for research or model training.
- **Syntactic Variability and Ellipsis:** Clinical notes often lack formal grammatical structures, with omitted words or phrases making parsing more difficult.

These characteristics underscore the need for sophisticated NLP models that can move beyond simple keyword matching to achieve true semantic understanding.

### 2.3. State-of-the-Art NLP Techniques for Clinical Knowledge Extraction

Recent advancements in NLP, particularly in deep learning, have revolutionized the ability to process clinical text.

#### 2.3.1. Transformer-Based Architectures

**Transformer models**, such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and their clinical adaptations (e.g., BioBERT, ClinicalBERT), represent the forefront of NLP. These models leverage **attention mechanisms** to capture long-range dependencies in text and learn highly contextualized word embeddings. Pre-

trained on massive text corpora, they demonstrate remarkable capabilities in understanding linguistic nuances, making them highly effective for tasks like:

- **Named Entity Recognition (NER):** Identifying and classifying specific entities (e.g., diseases, medications, symptoms, anatomical sites) within text. Transformer models significantly outperform previous methods by understanding the context in which an entity appears.
- **Relation Extraction (RE):** Determining semantic relationships between identified entities (e.g., "Medication X *treats* Disease Y," "Symptom Z *is associated with* Diagnosis A"). This is crucial for building structured knowledge graphs.
- **Text Classification:** Categorizing clinical notes based on content (e.g., presence of a specific condition, discharge summary type).

### 2.3.2. Semantic Parsing and Knowledge Graph Construction

Beyond identifying entities and relations, **semantic parsing** aims to convert natural language sentences into formal, computable representations, such as logical forms or conceptual graphs. In the clinical domain, this involves mapping complex clinical statements into structured queries or facts. The ultimate goal is often to construct **knowledge graphs**, which are structured repositories of entities and their relationships. These graphs provide a formal and queryable representation of clinical knowledge extracted from unstructured notes, enabling sophisticated reasoning and retrieval.

### 2.3.3. Self-Supervised Learning for Clinical Pre-training

The challenge of obtaining large, expertly annotated clinical datasets is a significant bottleneck. **Self-supervised learning** has emerged as a powerful paradigm to address this. By designing pretext tasks (e.g., masked language modeling, next sentence prediction) that can generate their own supervision from vast amounts of unlabeled clinical text, models can learn rich, domain-specific representations. This pre-training phase allows models to implicitly understand the vocabulary, syntax, and semantics of clinical language before being fine-tuned on smaller, task-specific annotated datasets, significantly improving performance and reducing reliance on extensive manual labeling.

## 2.4. Gaps in Current Research

Despite significant progress, several areas require further investigation to unlock the full potential of NLP in clinical settings. While transformer models offer superior performance, their "black-box" nature often hinders interpretability, a critical requirement in high-stakes medical decision-making. Moreover, existing solutions for handling the extreme variability in clinical notes, especially regional or institutional differences in jargon and phrasing, are not fully robust. The seamless integration of extracted knowledge with existing structured EHR data and external biomedical ontologies also presents an ongoing challenge. Finally, ensuring the generalizability and ethical deployment of these AI systems across diverse patient populations and clinical contexts remains a paramount concern that necessitates continued research. This study seeks to contribute to addressing these multifaceted challenges through a focused and comprehensive approach.

## Chapter 3: Methodology

### 3.1. Research Design

This research will adopt a methodological and experimental research design, focusing on the development, implementation, and rigorous evaluation of advanced Natural Language Processing (NLP) models. The core of this design involves an iterative process of model selection, architecture adaptation, training, and performance assessment, specifically tailored for the complexities of clinical text. The emphasis will be on demonstrating the feasibility and effectiveness of novel NLP techniques in achieving high-fidelity semantic understanding and knowledge extraction from unstructured clinical narratives. The design prioritizes the development of reproducible and generalizable methods, acknowledging the inherent variability of clinical data.

### 3.2. Data Considerations and Pre-processing

While this study does not involve new data collection, the methodology will consider the characteristics of typical clinical data and the necessary pre-processing steps for effective NLP application. The hypothetical dataset would consist of diverse unstructured clinical notes (e.g., discharge summaries, progress notes, radiology reports, pathology reports), reflecting a range of clinical specialties and linguistic styles. A crucial pre-processing step will be **de-identification**, to remove Protected Health Information (PHI) to ensure patient privacy and compliance with ethical guidelines. Further pre-processing will include tokenization, sentence segmentation, and normalization of clinical terminology (e.g., converting abbreviations to their full forms where contextually appropriate, standardizing medical concepts to unified terminologies like UMLS). The selection of appropriate de-identification strategies, which may themselves involve NLP, will be a critical initial consideration to prepare the raw text for subsequent analysis.

### 3.3. Model Development and Training

The core of the methodological approach involves the development and training of state-of-the-art deep learning models for NLP.

#### 3.3.1. Transformer-Based Architectures

The research will leverage **transformer-based models**, such as BERT, RoBERTa, or their domain-specific variants (e.g., BioBERT, ClinicalBERT, PubMedBERT), as foundational architectures. These models are chosen for their superior ability to capture contextual information and long-range dependencies within text. The specific transformer architecture will be selected based on its proven performance on clinical NLP benchmarks and its suitability for the target tasks.

#### 3.3.2. Self-Supervised Pre-training

To adapt the chosen transformer models to the unique linguistic characteristics of clinical text, **self-supervised pre-training** on large, unlabeled clinical corpora will be a crucial step. This involves designing pretext tasks (e.g., masked language modeling, next sentence prediction) that allow the model to learn deep contextualized representations without explicit human annotations. This phase will enable the model to grasp the nuances of clinical jargon, sentence structures, and common knowledge patterns within the medical domain.

#### 3.3.3. Task-Specific Fine-tuning

Following pre-training, the models will be fine-tuned on task-specific, labeled clinical datasets. Key tasks will include:

- **Named Entity Recognition (NER):** Training models to identify and classify clinical entities

such as diseases, symptoms, medications, procedures, anatomical sites, and laboratory values.

This will involve using sequence labeling approaches (e.g., IOB or BIO tagging schemes).

- **Relation Extraction (RE):** Fine-tuning models to identify semantic relationships between the extracted entities (e.g., "Drug X *treats* Disease Y," "Symptom Z *is a manifestation of* Condition A"). This can involve classification of entity pairs or more complex span-based approaches.
- **Negation and Temporality Detection:** Developing modules or extending existing models to accurately detect negated assertions and extract precise temporal information associated with clinical events.
- **Coreference Resolution:** Implementing strategies to link mentions of the same real-world entity throughout a document, creating a cohesive understanding of the narrative.

### 3.4. Evaluation Metrics

The performance of the developed NLP models will be rigorously evaluated using a combination of established intrinsic and extrinsic metrics.

#### 3.4.1. Intrinsic Evaluation

- **Named Entity Recognition (NER) and Relation Extraction (RE):** Performance will be assessed using standard metrics such as **Precision, Recall, and F1-score**. These metrics will be computed at both the entity and relation level, considering strict and lenient matching criteria.
- **Accuracy:** For classification tasks (e.g., negation detection), overall accuracy will be a primary metric.

#### 3.4.2. Extrinsic Evaluation

While direct patient outcomes are beyond the scope of this methodological study, the ultimate utility of the extracted knowledge will be indirectly assessed by:

- **Human Evaluation of Knowledge Graph Quality:** Expert clinicians or annotators may assess a sample of the automatically constructed knowledge graphs for accuracy, completeness, and clinical relevance.
- **Feasibility for Downstream Tasks:** Evaluating how well the extracted structured data can support hypothetical downstream applications, such as populating a research database or generating a structured patient summary.

### 3.5. Tools and Technologies

The implementation of this research will primarily utilize programming languages and libraries prevalent in the field of artificial intelligence and NLP. **Python** will serve as the primary programming language. Deep learning frameworks such as **TensorFlow** or **PyTorch** will be employed for building, training, and evaluating the transformer-based models. Specialized NLP libraries (e.g., Hugging Face Transformers, spaCy, NLTK) will be utilized for pre-processing, tokenization, and potentially for leveraging pre-trained clinical language models. Version control systems (e.g., Git) will manage code development, and appropriate computational resources (e.g., GPUs) will be utilized for model training.

## Chapter 4: Expected Results and Discussion

### 4.1. Expected Results

Based on the proposed methodology and the state-of-the-art in Natural Language Processing, this research anticipates achieving significant advancements in the automated understanding and structuring of clinical notes. The primary expected results include:

- **High-Precision and High-Recall Clinical Entity Recognition:** The fine-tuned transformer models, benefiting from self-supervised pre-training on clinical corpora, are expected to achieve robust performance in identifying a wide range of clinical entities (e.g., diseases, symptoms, medications, procedures) with high accuracy, minimizing both false positives and false negatives.
- **Accurate Relation Extraction and Semantic Parsing:** The developed framework is anticipated to effectively identify complex semantic relationships between clinical entities, translating free-text descriptions into a structured, queryable format suitable for knowledge graph construction. This will include precise detection of attributes, events, and their associated temporal and negation cues.
- **Robust Handling of Clinical Language Nuances:** The proposed approaches are expected to demonstrate improved capabilities in disambiguating abbreviations, handling colloquialisms, and correctly interpreting negated or speculative statements within the clinical context, thereby enhancing overall semantic understanding.
- **Foundation for Computable Clinical Knowledge:** The ultimate output will be a system capable of transforming raw, unstructured clinical notes into a computable representation, essentially forming a knowledge base derived directly from clinician narratives. This structured output will be amenable to various forms of computational analysis.

- **Demonstration of Self-Supervised Learning Efficacy:** The research will provide empirical evidence of the effectiveness of self-supervised learning for pre-training models on vast amounts of unlabeled clinical text, showcasing its potential to mitigate the challenges of data scarcity in the healthcare domain.

#### 4.2. Discussion of Contributions and Implications

The successful realization of these expected results will yield substantial contributions with broad implications for healthcare:

- **Advancement of Precision Medicine:** By unlocking the rich contextual information in clinical notes, this research will enable a more granular and individualized understanding of each patient's condition, facilitating truly personalized diagnoses, prognoses, and treatment pathways. This moves beyond generalized patient profiles to a more nuanced, data-driven approach.
- **Catalyst for Clinical Research:** The ability to systematically extract structured data from millions of clinical notes will provide researchers with unparalleled access to real-world evidence. This will accelerate the identification of novel biomarkers, the understanding of disease progression, the evaluation of treatment effectiveness in diverse populations, and the generation of new hypotheses for drug discovery.
- **Enhancement of Clinical Decision Support:** Clinicians will benefit from decision support systems that integrate deep insights from unstructured notes, providing comprehensive patient overviews and context-aware recommendations, thereby augmenting human expertise and reducing diagnostic and treatment errors.
- **Operational Efficiencies and Cost Reduction:** Automation of information extraction will significantly reduce the manual effort involved in coding, billing, quality reporting, and auditing, freeing up valuable healthcare resources and potentially leading to substantial cost savings.
- **Ethical Considerations and Interpretability:** While not the sole focus, the emphasis on explainable AI (XAI) principles in model selection and evaluation will pave the way for more

trustworthy and accountable AI systems in high-stakes clinical applications, fostering greater clinician adoption and public confidence.

- **Scalability and Generalizability:** The methodological advancements, particularly in self-supervised learning and robust model architectures, are anticipated to lay the groundwork for scalable solutions that can be adapted across various healthcare settings and diverse clinical specialties, overcoming current limitations of domain-specific solutions.

#### 4.3. Limitations and Future Work

Despite the anticipated contributions, this research acknowledges several inherent limitations and points towards avenues for future work. A primary limitation is the inherent complexity and variability of clinical language across different institutions, specialties, and individual clinicians. While the methodology aims for robustness, achieving universal semantic understanding remains a challenge. The dynamic nature of medical knowledge and evolving clinical guidelines also necessitates continuous model adaptation.

Future work could explore:

- **Integration with Multi-modal Data:** Expanding the NLP framework to seamlessly integrate with other data modalities within EHRs, such as medical images, genomic data, and physiological signals, to provide an even more holistic patient view.
- **Real-time Clinical Integration:** Investigating the challenges and opportunities of deploying these NLP systems in real-time clinical workflows, focusing on latency, scalability, and user interface design for maximum impact.
- **Proactive Clinical Alerting:** Developing systems that leverage extracted knowledge to generate proactive alerts for potential adverse events, drug interactions, or changes in patient condition.
- **Longitudinal Patient Journey Understanding:** Focusing on NLP models that can synthesize information across multiple notes and encounters over extended periods to construct comprehensive, evolving patient timelines and disease trajectories.
- **Further Development of Explainable AI (XAI) for Clinical Context:** Deepening the research into inherently interpretable models or novel XAI techniques specifically designed to meet the stringent interpretability requirements of clinical decision-making.

- **Ethical AI Governance Frameworks:** Developing comprehensive ethical and governance frameworks for the responsible development and deployment of clinical NLP systems, addressing issues of bias, fairness, accountability, and patient autonomy in greater detail.

## Chapter 5: Conclusions

This research has systematically explored the transformative potential of Natural Language Processing (NLP) for unlocking the rich, yet largely untapped, information embedded within unstructured clinical notes. By outlining a comprehensive methodology centered on advanced deep learning techniques, particularly transformer-based architectures and self-supervised learning, this study aims to bridge the critical gap between raw clinical text and computable, actionable intelligence. The problem of semantic and syntactic inaccessibility of clinical narratives is significant, impeding progress in personalized medicine, clinical research, and operational efficiency within healthcare systems. The proposed methodology, encompassing meticulous data pre-processing, robust model development through self-supervised pre-training and task-specific fine-tuning, and rigorous evaluation using intrinsic and extrinsic metrics, is designed to address the unique complexities of clinical language. Expected outcomes include highly accurate clinical entity recognition, precise relation extraction, and comprehensive semantic parsing, leading to the automated construction of computable clinical knowledge. The implications of this research are far-reaching, promising to enhance precision medicine by enabling a deeper understanding of individual patient characteristics, accelerate clinical research through systematic access to real-world data, and improve clinical decision support by providing clinicians with context-rich insights. Furthermore, it is anticipated to drive operational efficiencies and contribute to broader public health informatics initiatives. While acknowledging the inherent limitations related to linguistic variability and the dynamic nature of medical knowledge, this study lays a strong foundation for future advancements. In conclusion, this research underscores the pivotal role of advanced NLP in revolutionizing how clinical information is utilized. By converting unstructured narratives into structured knowledge, it promises to significantly contribute to a more efficient, precise, and ultimately, more patient-centered future for healthcare.

## References

1. Hossan, K. M. R., Rahman, M. H., & Hossain, M. D. HUMAN-CENTERED AI IN HEALTHCARE: BRIDGING SMART SYSTEMS AND PERSONALIZED MEDICINE FOR COMPASSIONATE CARE.
2. Hossain, M. D., Rahman, M. H., & Hossan, K. M. R. (2025). Artificial Intelligence in healthcare: Transformative applications, ethical challenges, and future directions in medical diagnostics and personalized medicine.
3. Kim, J. W., Khan, A. U., & Banerjee, I. (2025). Systematic review of hybrid vision transformer architectures for radiological image analysis. *Journal of Imaging Informatics in Medicine*, 1-15.
4. Springenberg, M., Frommholz, A., Wenzel, M., Weicken, E., Ma, J., & Strodthoff, N. (2023). From modern CNNs to vision transformers: Assessing the performance, robustness, and classification strategies of deep learning models in histopathology. *Medical image analysis*, 87, 102809.
5. Atabansi, C. C., Nie, J., Liu, H., Song, Q., Yan, L., & Zhou, X. (2023). A survey of Transformer applications for histopathological image analysis: New developments and future directions. *BioMedical Engineering OnLine*, 22(1), 96.
6. Sharma, R. R., Sungheetha, A., Tiwari, M., Pindoo, I. A., Ellappan, V., & Pradeep, G. G. S. (2025, May). Comparative Analysis of Vision Transformer and CNN Architectures in Medical Image Classification. In *International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 2024)* (pp. 1343-1355). Atlantis Press.

7. Patil, P. R. (2025). Deep Learning Revolution in Skin Cancer Diagnosis with Hybrid Transformer-CNN Architectures. *Vidhyayana-An International Multidisciplinary Peer-Reviewed E-Journal-ISSN 2454-8596*, 10(si4).
8. Shobayo, O., & Saatchi, R. (2025). Developments in Deep Learning Artificial Neural Network Techniques for Medical Image Analysis and Interpretation. *Diagnostics*, 15(9), 1072.
9. Karthik, R., Thalanki, V., & Yadav, P. (2023, December). Deep Learning-Based Histopathological Analysis for Colon Cancer Diagnosis: A Comparative Study of CNN and Transformer Models with Image Preprocessing Techniques. In *International Conference on Intelligent Systems Design and Applications* (pp. 90-101). Cham: Springer Nature Switzerland.
10. Xu, H., Xu, Q., Cong, F., Kang, J., Han, C., Liu, Z., ... & Lu, C. (2023). Vision transformers for computational histopathology. *IEEE Reviews in Biomedical Engineering*, 17, 63-79.
11. Singh, S. (2024). Computer-aided diagnosis of thoracic diseases in chest X-rays using hybrid cnn-transformer architecture. *arXiv preprint arXiv:2404.11843*.
12. Fu, B., Zhang, M., He, J., Cao, Y., Guo, Y., & Wang, R. (2022). StoHisNet: A hybrid multi-classification model with CNN and Transformer for gastric pathology images. *Computer Methods and Programs in Biomedicine*, 221, 106924.
13. Bougourzi, F., Dornaika, F., Distanti, C., & Taleb-Ahmed, A. (2024). D-TrAttUnet: Toward hybrid CNN-transformer architecture for generic and subtle segmentation in medical images. *Computers in biology and medicine*, 176, 108590.
14. Islam, M. T., Rahman, M. A., Mazumder, M. T. R., & Shourov, S. H. (2024). COMPARATIVE ANALYSIS OF NEURAL NETWORK ARCHITECTURES FOR MEDICAL IMAGE CLASSIFICATION: EVALUATING PERFORMANCE ACROSS DIVERSE MODELS. *American Journal of Advanced Technology and Engineering Solutions*, 4(01), 01-42.
15. Vanitha, K., Manimaran, A., Chokkanathan, K., Anitha, K., Mahesh, T. R., Kumar, V. V., & Vivekananda, G. N. (2024). Attention-based Feature Fusion with External Attention Transformers for Breast Cancer Histopathology Analysis. *IEEE Access*.
16. Borji, A., Kronreif, G., Angermayr, B., & Hatamikia, S. (2025). Advanced hybrid deep learning model for enhanced evaluation of osteosarcoma histopathology images. *Frontiers in Medicine*, 12, 1555907.
17. Aburass, S., Dorgham, O., Al Shaqsi, J., Abu Rumman, M., & Al-Kadi, O. (2025). Vision Transformers in Medical Imaging: a Comprehensive Review of Advancements and Applications Across Multiple Diseases. *Journal of Imaging Informatics in Medicine*, 1-44.
18. Wang, X., Yang, S., Zhang, J., Wang, M., Zhang, J., Yang, W., ... & Han, X. (2022). Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical image analysis*, 81, 102559.
19. Xia, K., & Wang, J. (2023). Recent advances of transformers in medical image analysis: a comprehensive review. *MedComm-Future Medicine*, 2(1), e38.
20. Gupta, S., Dubey, A. K., Singh, R., Kalra, M. K., Abraham, A., Kumari, V., ... & Suri, J. S. (2024). Four transformer-based deep learning classifiers embedded with an attention U-Net-based lung segmenter and layer-wise relevance propagation-based heatmaps for COVID-19 X-ray scans. *Diagnostics*, 14(14), 1534.
21. Henry, E. U., Emebob, O., & Omonhinmin, C. A. (2022). Vision transformers in medical imaging: A review. *arXiv preprint arXiv:2211.10043*.
22. Manjunatha, A., & Mahendra, G. (2024, December). TransNet: A Hybrid Deep Learning Architecture Combining CNNs and Transformers for Enhanced Medical Image Segmentation. In *2024 International Conference on Computing and Intelligent Reality Technologies (ICCIRT)* (pp. 221-225). IEEE.
23. Reza, S. M., Hasnath, A. B., Roy, A., Rahman, A., & Faruk, A. B. (2024). *Analysis of transformer and CNN based approaches for classifying renal abnormality from image data* (Doctoral dissertation, Brac University).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.