Generative Artificial Intelligence for the Automated Generation of Medical Reports: A Modified Poor and Rich Optimization (MPRO)-Based Approach

Shabnam Mohamed Aslam[ID]

Department of Information Technology, College of Computer and Information Sciences, Majmaah University, Majmaah 11952, Saudi Arabia

Corresponding Author Email: s.s.aslam@mu.edu.sa

ABSTRACT

A novel method for the automated generation of medical reports is presented, combining generative Artificial Intelligence (Gen-AI) with optimal feature selection and fusion strategies. The approach employs a pre-trained ResNet architecture for the extraction of visual-semantic features from medical images, including frontal and lateral views, with annotations on key clinical features such as objects, locations, nodules, tumors, and masses. To enhance disease localization and diagnosis, an adaptive attention mechanism is incorporated within a deep recurrent neural network (DRNN), enabling dynamic focus on the most relevant regions of the images. Segmentation is performed using the Modified Poor and Rich Optimization (MPRO) method, which is followed by a hybrid feature fusion approach integrating Chaotic Krill Herd Optimization (CKHO). The system is further augmented by multimodal data, such as patient demographics and medical history, to ensure the generation of personalized, context-aware reports. The Gen-AI framework employs OpenAI's GPT and advanced natural language processing (NLP) techniques to generate precise and nuanced diagnostic sentences. Benchmark datasets, including CheXpert and IU X-ray, are utilized to evaluate the proposed method and compare it with existing automated report generation techniques. Performance is assessed using standard image captioning evaluation metrics—namely BLEU, METEOR, ROUGE, and CIDER—with the proposed method demonstrating superior performance, achieving the highest values across these metrics. Statistical analysis confirms the robustness, accuracy, and exceptional performance of the model, with a F1 score of 0.93% for automated report generation. The results indicate the significant potential of Gen-AI for improving the accuracy and efficiency of medical reporting.

1. INTRODUCTION

AI has revolutionized many industries, including healthcare, transforming medical practices globally [1]. It is used for diagnostics, treatment suggestions, administrative activities, and tailored patient care. AI analyzes patient records, medical imaging, genomic sequences, and clinical trials to improve decision-making, efficiency, and patient outcomes [2]. AI improves diagnosis, treatment, and illness management in healthcare systems. AI's ability to assess and interpret complicated medical statistics quickly and accurately often outperforms humans [3]. AI-driven diagnostic algorithms can precisely analyze medical images like CT scans, MRIs and X-rays, detecting diseases earlier and refining patient results. AI-powered prognostic statistics also help healthcare providers spot patient data patterns and trends, enabling proactive treatments and health risk reduction [4]. AI has great potential in healthcare, but its general adoption is difficult. Integrating AI into healthcare systems raises ethical, regulatory, data privacy, transparency and interpretability difficulties [5]. Concerns include AI-driven automation displacing healthcare experts and distributing AI-enabled healthcare services fairly [6].

Automatic medical report production is a major healthcare development that benefits patients and clinicians [7]. One of the main reasons for adopting automated medical report generation is to streamline documentation and save healthcare personnel time. Healthcare practitioners can focus more on patient care and less on administrative responsibilities by automatically generating comprehensive and accurate medical reports based on patient data, improving clinical efficiency and productivity [8]. Automated medical report production may improve medical documentation quality and consistency [9]. Automated systems can extract essential information from EHRs, diagnostic tests, and medical imaging studies to create detailed and standardized medical reports using powerful NLP algorithms and machine learning. This decreases documentation errors and ensures that all relevant information is captured and shared, improving continuity of care and clinical decision-making [10]. Automated medical report production can enhance patient outcomes and satisfaction by providing faster access to crucial health information [11]. Automated medical report creation can improve healthcare delivery and patient care by speeding the reporting process and

improving medical information accessibility and comprehensiveness [12]. Traditional medical report-generating systems often have issues that slow, inaccurate, and delay healthcare reporting [13]. Manual report creation, which requires healthcare workers to transcribe patient data into written reports, is a major issue. Manual data entry relies on human interpretation and typing accuracy, making it time-consuming and error-prone. Handwritten notes or dictation can also cause terminology, formatting, and content differences among reports, making healthcare provider communication and interoperability difficult [14]. Traditional medical report generation is a struggle with healthcare data fragmentation and interoperability [15]. Many healthcare facilities maintain patient data in multiple systems and formats, making report preparation problematic. Since healthcare providers may not have full patient data at the time of documentation, fragmentation can lead to incomplete or conflicting medical reports [16]. Traditional medical report generation may also struggle to keep up with healthcare data volume and complexity. As medical expertise and diagnostic tools advance, healthcare providers must synthesize and document massive amounts of data quickly [17]. Gen-AI is transforming healthcare documentation by using novel methods to streamline and improve medical report generation. Using advanced machine learning algorithms and NLP, Gen-AI systems can automatically generate accurate and complete medical reports from structured and unstructured patient data [18]. Gen-AI systems can examine enormous quantities of patient statistics in real-time and provide individualized reports without human interaction [19, 20].

1.1 State-of-the-art methods survey

Existing methods in the healthcare-automated-report-generation field have primarily focused on enhancing both encoders and decoders. However, the intricate relationship between these two components often makes it challenging to discern the key advancements in each. Table 1 describes the research gap summary from the existing research works.

1.1.1 Draft radiology report
The research of Huang et al. [21], Nova [22] compared AI-created chest radiography explanations in the emergency department (ED) to on-site radiologists and tele radiologists for accuracy and quality. Emergency physicians scored chest radiograph interpretations by human radiologists and AI on a Likert scale after studying 500 ED encounters. AI reports were scored similarly to radiologists' reports, both much higher than tele-radioscopy reports. The secondary analysis found no significant difference in report type frequency of clinically important differences. The AI model appears to create reports with equivalent clinical accuracy to radiologists, which could improve ED imaging interpretation and documentation and help detect life-threatening illnesses.

1.1.2 Automatic clinical report generation
Darapaneni et al. [23] developed involuntary aided medical reports utilizing medical scans/images to speed up medical analysis and therapy. Image features are extracted using CNNs or transfer learning algorithms to construct radiologists' report captions. Greedy and beam search were used to evaluate predictions and were implemented in Python, and the BLEU score evaluation showed that a custom final model using greedy search was best for medical image captioning and

report preparation.

1.1.3 Automated radiology report generation
Nakaura et al. [24] compared radiologists' reports to the generative pre-trained transformer (GPT) series' radioscopy report generation, and obtained that GPT-2, GPT-3.5, and GPT-4 generated reports from CT scans of 28 patients are valid reports. The Gen AI model uses features such as diseases based on imaging results, demographics, and disease sites. The reports were assessed by radiologists for syntax, accessibility, image results, perception, discrepancy analysis, and quality. Radiologists diagnose in a more efficient way than the GPT models. Comparing the versions of GPT, GPT-4 outperforms GPT-3.5 and GPT-2. GPT series being less scored than radiologists in image and differential diagnosis, requiring radiologists' validation. GPT-3.5, and GPT-4 can write radioscopy reports with good readability and image results, but their impression and differential diagnosis accuracy are questionable. Li et al. [25] recommended an innovative auxiliary signal-guided knowledge encoder-decoder (ASGK) model. This model uses auxiliary patches to grasp medical ideas better and adds external linguistic cues to acquire precise visual information. The research combines auxiliary signals with transformer design to create an efficient medical report and its effectiveness is proved on standard benchmark datasets, such as CX-CHR, IU X-Ray, and COVID-19 CT Report (COV-CTR). The experimental findings demonstrate that ASGK is the best at classifying medical language and generating metrics for paragraphs, which shows that it could improve medical image captioning and reporting.

1.1.4 Augmented chest x-ray report generation
Ranjit et al. [26] proposed an innovative automatic radioscopy report inscription using multimodal-aligned embeddings from a contrastively pre-trained vision linguistic model called retrieval augmented generation. This model gets radioscopy text for an image and uses an overall area multiplicative model like Open-AI text-davinci-003, GPT-3.5-turbo, or GPT-4 to make a report. This strategy reduces hallucinated generations and allows flexible report content generation while conforming to desired formats using generative models and ease to personalize the automated report generation in different clinical environments.

1.1.5 Unsupervised medical report generation
Liu et al. [27] employed Knowledge Graph Auto-Encoder (KGAE), which generates a graph of clinical knowledge and medical images of unlabeled image datasets that shows the report about the disease image. This Gen-AI tool works only with a limited number of dependent datasets.

1.1.6 Radiology report generation
Zhang et al. [28] make disease predictions with a radiology image of the chest using the deep learning method with the Medical Image Report Quality Index (MIRQI) using the open-source image dataset IU-RR. This method is restricted to heterogeneous disease-structured images to be identified with specific types of illness. Chen et al. [29] proposed a cross-modal memory network (CMN) to improve the decoder-encoder architecture for radioscopy report production. To ease communication and production across domains, the CMN uses communal recollection to capture the arrangement among images and texts. By attaining cutting-edge presentation on two extensively used standard databases, IU X-Ray and

MIMIC-CXR, investigational outcomes show that the suggested model is successful and required to extend further hyper parameter analysis of the model.

1.1.7 Automated classification of patient safety event (PSE) Reports

Chen et al. [30] addressed the difficulty of effectively identifying PSE reports that are essential for hospital adverse event monitoring. ML methods were employed to classify PSE reports from other reports and obtained 75.4% classification accuracy, which plays a vital role in hospitals. Indeed, this classification isn't related to medical image analysis, where the current issue is to generate automatic reports based on medical images.

1.1.8 TieNet

Text Image embedding network for common thorax disease [31] classification and reporting in chest x-ray is one of the state-of-the-art models for automated medical report generation tools that employ AI CNN-RNN architecture for image classification and achieved 79% classification accuracy. This model takes the chest X-ray image as input and applies ML methods to classify the chest images and an automated report mentioning the types of disease, such as Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening, and Hernia.

1.1.9 CARG-clinically accurate chest X-ray report generation

The main objective of this method [32] is to annotate the chest X-ray image and generate meaningful full sentences using NLP, followed by radiology report generation. The raw image is encoded to an annotated image, and the decoder maps the annotations with existing sentence templates and forms the text report. Figure 1 demonstrates the work of the CARG automated report-generated method. The MIMIC-CXR dataset was considered for analysis, and the reinforcement machine learning method was employed and obtained 86% clinical accuracy of the report.
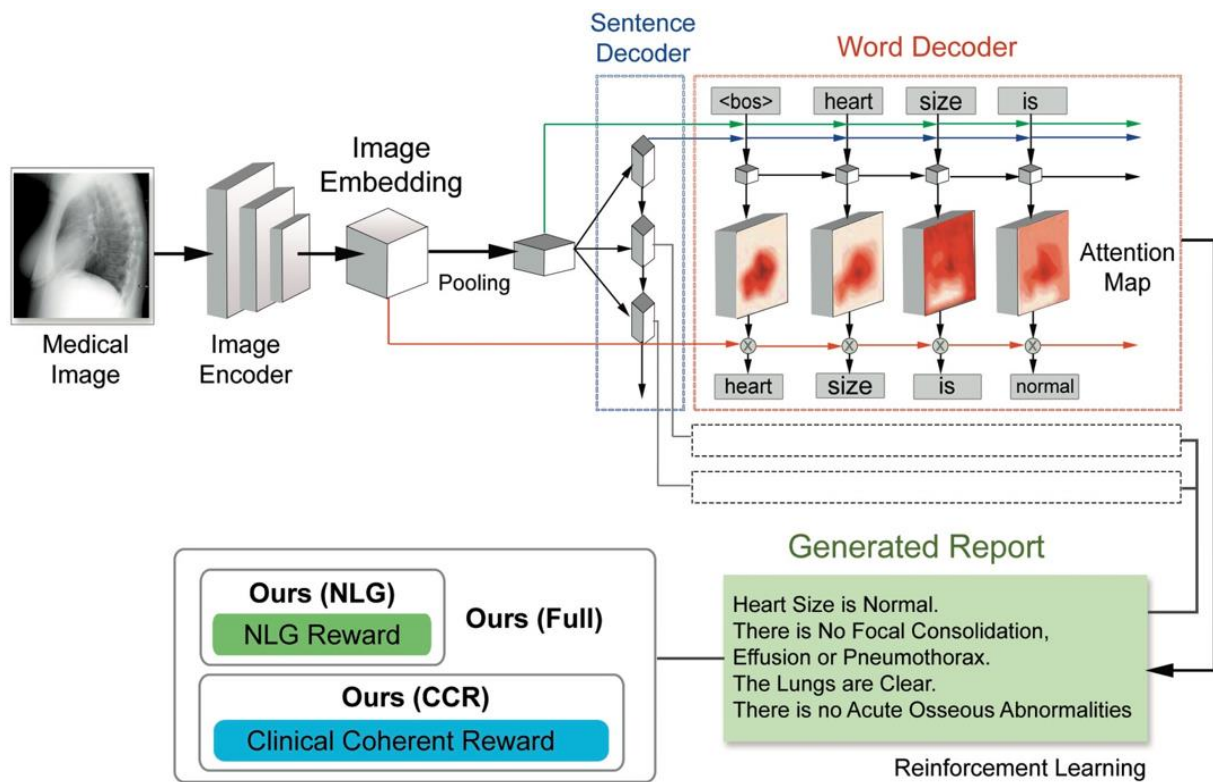

Figure 1. CARG framework

1.2 Automated medical report generation method

The role of AI for automatic medical report generation lasts from a decade ago. It starts emerging as the form of automatic medical report recording and documenting up to automatic report generation of radiology and pathology images. Automatic report generation is the process of designing and utilizing machine learning models to simulate human behavior of medical disease diagnosis. The AI models employ reasoning and the wisdom of the environment for decision-making. The process of image analysis for disease prediction follows tagging the impressions as sentences, and multiple sentences make up paragraphs, which will be tagged. The tagged paragraphs are grouped as tokens and compared different token sets and classified benign or malignant images.

1.3 Implications from existing AI-based report generation

Policies and regulation standards for AI algorithms for healthcare decision-making are important for AI researchers, and more investment is required for the development of advanced AI technologies such as Gen-AI to treat automatic health report generation [1]. Present AI approaches used for automated health report generation have to be evaluated with a fusion of healthcare data [2, 3]. It is essential to improve the patient's outcome more flexibly through the development of AI models for the healthcare domain [4]. There are automated documented health care reports that can be generated only with the transcription of natural language doctor-patient conversations and still require more advanced revolutions of medical automated report generation [7]. Concerning the x-ray

dataset, i.e., the ED dataset, existing research shows a loss of information with the image lateral view feature that leads to low accuracy of image classification by existing AI tools [20]. Traditional ML techniques have been predominantly utilized in existing methods, highlighting a gap in leveraging advanced DL methods for improved accuracy and efficiency in report generation. This limitation underscores the need to transition towards more sophisticated DL approaches to enhance the quality and precision of medical reports. Optimal feature optimization and fusion, crucial components in automated medical report generation, have not received adequate attention in the literature. Through the way of focusing on incorporating optimal feature selection and fusion models, the proposed solution aims to enhance report completeness and accuracy. Table 1 shows that the highest accuracy for radiology reports generated from image inputs is 93.23%, albeit with higher resource utilization [21]. Our approach aims to reduce resource consumption while maintaining report completeness and achieving accuracy comparable to the method in the study by Huang et al. [21]. Our method's emphasis on feature optimization and fusion within the DL framework improves the diagnostic capabilities of automated systems, leading to more reliable and comprehensive medical reports.

The implications of existing research motivate us to analyze benchmarking openly available medical image datasets, CheXpert and IU X-ray, with the proposed Gen AI model GPT with the MPRO method optimization of features to improve the accuracy of feature extraction and clinical accuracy of automated medical report generation.

•To utilize the ResNet pre-trained architecture to extract visual-semantic features from both frontal-view and lateral-view medical images. To use pre-trained models, ResNet can train on large datasets and effectively extract high-level features from images.

•To employ the MPRO algorithm to find the optimal subset of features by iteratively evaluating the performance of different feature combinations.

•To use the CKHO algorithm for feature fusion, which enhances exploration and exploitation capabilities, enabling efficient fusion of features from different modalities.

•To employ DRNN for diagnosis, leveraging its ability to learn complex patterns and relationships in the data. To utilize the Open-AI GPT model to formulate coherent diagnostic sentences based on the diagnosis generated by DG-ANN.

•To evaluate the performance of automated report generation, the proposed MPRO+DRNN+DistilGP2 was compared with existing report-generating methods.

1.4 Objective of the research

The article is organized into sections were. The second section focuses on the explanation of a proposed method for automated medical report generation. The third section discusses the results and comparative analysis of proposed and existing methods for automated medical report generation. The final section summarizes the conclusion of the work.

Table 1. Summary of literature review from existing state-of-the-art works related to automatic medical report generation

Ref.	Methodology	Technique Used	Findings	Research Gaps
[21]	Producing draft radiology reports from input images	Gen-AI	Accuracy 93.23%	Increased computational resource demands
[21]	Generation of electronic health records	NLP and NER	Accuracy 84.522%	Limited temporal and spatial course of calcification and uncertainty in clinical decision-making
[22]	Automatic clinical report generation	CNN and transfer learning algorithm	AUC 0.845	The tedious and time-consuming process of digital replicates
[23]	Automated radiology report generation	GPT-2, GPT-3.5, and GPT-4	Precision 87.125%	Entities lead to incorrect feature extraction and poor decisions
[24]	Automated medical report generation	ASGK and DenseNet-121	Accuracy	
[25]	Automated classification of patient safety event reports	Local interpretable model-agnostic explanations (LIME)	Accuracy 75.4%	A typical report contains numerous template descriptions and only a few abnormal sentences.
[26]	Augmented chest X-ray report generation	Open AI text-DaVinci-003	BERT score of 0.2865	Affected by data consumption problem
[27]	Unsupervised medical report generation	Knowledge graph auto-encoder (KGAE)	Accuracy 87.725%	Challenge due to the unavailability of clean, curated and preprocessed text corpus specific to the domain.
[28]	Radiology report generation	Knowledge graph neural network	Accuracy 82.235%	It requires large amounts of data to generalize well to different cases.
[30]	Radiology report generation	Cross-modal memory networks (CMN)	AUC 0.874	Not explicitly taught compared to experienced radiologists, which limits the generation accuracy

2. METHODOLOGY

The way that section titles and other headings are displayed in the proposed methods for automated medical report generation, illustrated in Figure 2, utilize optimal feature optimization and feature fusion with Gen-AI. The author uses the benchmark open access dataset CheXpert and IU X-beam, a collection of chest X-ray images. The system consists of both front-facing perspective and parallel-view pictures, which undergo image preprocessing to enhance the prediction quality and suitability for feature extraction. The preprocessing step

may involve resizing the images to a standard size, normalizing pixel values to a common scale, and performing any necessary adjustments to enhance image quality. It employs ResNet, a pre-trained architecture, for feature extraction, enabling the capture of visual-semantic features from the medical images. The sequence of processes, such as feature extraction and feature optimization using the MPRO algorithm. Feature extraction aimed to reduce the image dimension; the next step is feature fusion, and it is performed using the CKHO algorithm. This fusion process combines the selected optimal features to produce comprehensive diagnostic

insights, including the location and type of the disease. For disease diagnosis, the author employs a DRNN, which leverages the extracted features to accurately identify and classify medical conditions from the images. Finally, the author generates medical reports using the OpenAI GPT (Gen-AI) model, which formulates coherent and contextually relevant text based on the diagnostic results obtained from the DRNN, providing detailed and informative medical reports automatically.
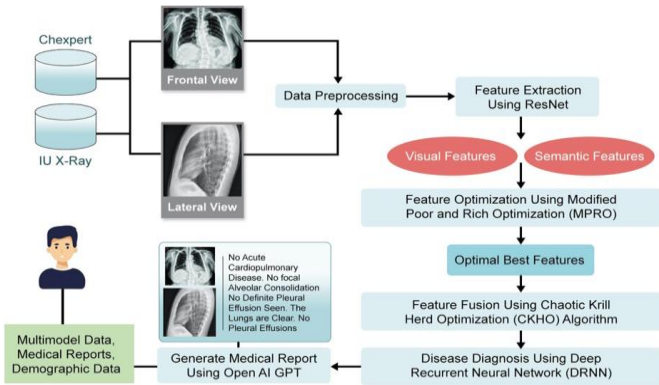

Figure 2. Illustration of automated report generation with MPRO

2.1 Visual-semantic feature extraction

To capture comprehensive visual-semantic information from medical images, the author utilizes the ResNet pre-trained architecture for feature extraction. Specifically, the author processes both frontal-view and lateral-view images to ensure that the extracted features encompass diverse aspects related to objects, locations, and medical concepts. As well as dynamically focusing on the most relevant parts of the medical images, the author integrates an adaptive attention mechanism within the DRNN that enhances the accuracy of disease localization and diagnosis by emphasizing critical regions in the images.

2.2 Feature extraction

ResNet is a pre-trained architecture type of Convolutional Neural Network known for its effectiveness in image segmentation. The ResNet model is pre-trained on a large dataset, which enables it to learn hierarchical representations of features from images. ResNet network captures important visual details, such as the presence of specific objects (tumors, nodules, and masses), their spatial arrangements, and other semantic attributes of the X-ray size-reduced images. The features are important to classify the images based on different classes of anatomical structures; abnormalities provide valuable information for subsequent stages of the automated medical report generation process.

The preprocessed images are passed through a series of convolutional layers within the ResNet model. Convolutional layers extract hierarchical features from the input images by applying convolutional filters that detect patterns such as edges, textures, and shapes. The ResNet architecture includes residual blocks to address the minimizing gradient problem encountered in deep neural networks.

As the images pass through the convolutional layers and residual blocks, feature maps are generated at each layer. These component maps address various degrees of deliberation, with lower layers catching low-level highlights like edges and surfaces, and higher layers catching more conceptual and semantic elements applicable to the errand. Convolutional layers feature maps that are exposed to worldwide normal pooling. This pooling activity lowers the spatial elements of the component maps while holding their most significant highlights. Worldwide normal pooling assists with collecting data from across the element maps, providing a compact representation of the input images. The output of the global average pooling layer serves as the visual-semantic features extracted from the input images. These features encapsulate both visual information (such as shapes and textures) and semantic information (such as diagnostic cues and anatomical structures) relevant to the medical imaging task.

2.3 Feature optimization

In the context of feature optimization, MPRO aims to find an optimal subset of features that maximizes the performance of a given machine learning model or analysis task. In enhancement issues, every individual arrangement in the unfortunate populace advances through the pursuit space to the worldwide ideal arrangement by concentrating on the rich arrangement in the rich populace. A subset of the ongoing age of arrangements is known as a populace. Competitor arrangements in the populace incorporate a rich and a poor financial individual/arrangement. "N" is the populace size. As an estimate, we create "N" arrangements with arbitrary truth values somewhere in the range of 0 and 1. A digitization interaction is utilized for each degree of individual answers to convert genuine qualities into paired values.

$$\chi_{h,g} = \begin{cases} 1, & \chi_{h,g} > Rand \\ 0, & otherwise \end{cases} \tag{1}$$

A rand is an irregular number somewhere in the range of 0 and 1. Competitor arrangements in the populace are arranged by the goal capability. The upper piece of the populace is alluded to as the rich financial gathering and the lower portion of the populace is alluded to as the poor monetary gathering.

$$pop_{main} = pop_{Rich} + pop_{poor} \tag{2}$$

The wellness capability assumes a significant part in improvement issues. It returns a positive whole number to show how great the competitor's arrangement is. The mistake pace of the classifier is determined by the created wellness capability. A richer solution has a lower fitness score (error rate) and a poorer solution has a higher fitness score (error rate).

$$Fitness(\chi_h) = \frac{number\ of\ misclassified\ documents}{total\ number\ of\ documents} * 100 \tag{3}$$

The rich widen their economic class gap by taking care of the poor economic class people. The kinetics of the enriched solution are computed as follows.

$$\chi_{Rich,h,g}^{New} = \chi_{Rich,h,g}^{old} + \alpha * \left[\chi_{Rich,h,g}^{old} - \chi_{poor,best,g}^{old} \right] \tag{4}$$

The undertaking of poor explanation is computed as follows.

$$\chi_{poor,h,g}^{New} = \chi_{Poor,h,g}^{old} +$$

$$\alpha * \left[\left(\frac{\chi_{Rich,best,g}^{old} + \chi_{Rich,mean,g}^{old} + \chi_{Rich,worst,g}^{old}}{3} \right) - \chi_{poor,h,g}^{old} \right] \quad (5)$$

Algorithm 1 describes the working process of feature optimization using MPRO. By using MPRO for feature optimization, the algorithm iteratively explores the space of possible feature subsets, identifying combinations of features that lead to improved performance on the target task. This approach helps to address data dimensionality issues by selecting a subset of features that are most relevant and informative for the given analysis task, ultimately improving the efficiency and effectiveness of subsequent modelling or analysis steps.

Algorithm 1. Feature optimization using MPRO

Input: Number of visual features, semantic features, maximum iteration
Output: Feature optimization
1.   Begin
2.   Initialized the population
3.   Here divides the dataset into Training and Testing
4.   Define individual solutions to convert real values

$$\chi_{h,g} = \begin{cases} 1, & \chi_{h,g} > Rand \\ 0, & otherwise \end{cases}$$

5.   Compute the optimal population as the poor economic group.

$$pop_{main} = pop_{Rich} + pop_{poor}$$

6.   Compute the kinetics of the enriched solution.

$$\chi_{Rich,h,g}^{New} = \chi_{Rich,h,g}^{old} + \alpha * \left[\chi_{Rich,h,g}^{old} - \chi_{poor,best,g}^{old} \right]$$

7.   Update the poor and rich values
8.   Else
9.   Find the best output value
10.   End

2.4 Feature fusion

Feature fusion is a technique used in machine learning and data analysis to combine information from multiple sources or modalities into a unified representation. In the context of medical imaging and report generation, feature fusion involves integrating visual-semantic features extracted from different views or sources to produce a depiction of the ailment. The CKHO algorithm is a metaheuristic optimization technique inspired by the social behavior of krill herds [33]. It is used to optimize feature fusion by finding the best combination of features that maximizes the performance of subsequent tasks, such as disease diagnosis or medical report generation. The CKHO algorithm [34] uses multi-objective herd to solve global optimization problems. Foraging and density-dependent attraction to the grill are used as targets. As a result, individual krill aggregate and move toward an optimal solution in search of higher food density. The behaviour creates a mesh cluster around the global minimum optimization problem, generalizing the following Lagrangian model to a B-dimensional decision space.

$$\frac{dP_h}{ds} = B_h + f_h + C_h \quad (6)$$

where, B_h is the motion caused by the other grill blades; f_h foraging movement; h is the physical C_h distribution of krill individuals. The movement of each grill person is defined as follows.

$$B_h^{New} = B^{maks}\alpha_h + \omega_b B_h^{old} \quad (7)$$

$$\alpha_h = \alpha_h^{Local} + \alpha_h^{target} \quad (8)$$

where, B^{maks} is the maximum induced velocity according to the measured values ω_b taken as 0.01 (m/s) α_h^{Local}. The inertial weight of motion induced in the range [0, 1]. At the beginning of the optimization, the passive weight is equal to 0.9. Then it decreases linearly to 0.1. The neighbourhood effect is defined as the attractive/repulsive tendency among individuals to seek territory. The α_h^{target} effect of the target direction given by an ideal grid person can be defined as follows.

$$\alpha_h^{target} = D^{best}\hat{k}_{h,best}\hat{p}_{h,best} \quad (9)$$

where, D^{best} is defined as a constant of impact and definite like below.

$$D^{best} = 2\left(rand + \frac{H}{H_{maks}} \right) \quad (10)$$

where rand is a randomly generated number between 0 and 1, H is the actual number of iterations H_{maks} and is the maximum number of iterations. The Lyapunov exponent is a standard, generally applicable method for detecting and quantifying the dependence of sensitivity on initial conditions.

$$p_{s+1} = f(p_s) \quad (11)$$

where, $p_s \in r^b, s = 1,2,3,\dots$ and f is a meaning of r^b.

$$p_{s+1} = \mu p_b(1-p_s) \quad (12)$$

where, p_s is an irregular number, n is the number of emphases, and l is consistent between 0 and 4. If $l = 4$ the function becomes ideally chaotic. The basic idea is to create a joint function using cost and control functions. All transformation methods transform a constrained optimization problem into an unconstrained problem using a shape transformation function:

$$\Phi(p,r) = F(p) + X(i(p), j(p), r) \quad (13)$$

where F is the cost function, i is the equality constraint, j is the equality constraint, R is the vector of penalty parameters, and X is the original cost function, whose function R is the penalty function in the cost function. The x function can be written as the most popular quadratic loss function.

$$X(i(p), j(p), r) = R\left\{ \sum_{h=1}^{x}\left[i_h(p)\right]^2 + \sum_{h=1}^{a}\left[j_h^+(p)\right]^2 \right\}; j_h^+(p)$$

$$= Max(0, j_h(p)) \quad (14)$$

In addition, weight factors are included in the control functions using the scaling weight function method to balance between cost and control functions.

$$i(p) = \sum_{h=1}^{a} z_h i_h(p), \quad \sum_{h=1}^{a} z_h = 1 \tag{15}$$

where z_h is the weight influence of the h-th constraint meaning $(i_h(p))$. Finally, the unimpeded target meaning is as shadows. Algorithm 2 describes the working process of feature fusion using CKHO.

Algorithm 2. Feature fusion using CKHO

Input: Number of optimal features, optimal visual features, semantic features

Output: Feature fusion

1.   Initialize the random population.
2.   Define the Lagrangian model in a B-dimensional decision space.

$$\frac{dP_h}{ds} = B_h + f_h + C_h$$

3.   If i=0 , j=1
4.   While Do
5.   Compute the movement of each grill person

$$B_h^{New} = B^{maks}\alpha_h + \omega_b B_h^{old}$$

6.   Define the ideal grid person

$$\alpha_h^{target} = D^{best}\hat{k}_{h,best}\hat{P}_{h,best}$$

7.   Compute chaotic graph function is

$$p_{s+1} = f(p_s)$$

8.   Define the updated solution using logistic map role

$$p_{s+1} = \mu p_b(1-p_s)$$

9.   Compute the unconstrained problem using a shape transformation function:

$$\Phi(p,r) = F(p) + X(i(p), j(p), r)$$

10.   End if
11.   Update the final value
12.   End

2.5 Disease diagnosis and automatic report generation

2.5.1 Disease diagnosis suing DRNN

The DRNN model is prepared utilizing a huge dataset of named clinical pictures and comparing analysis. During preparation, the model figures out how to plan input elements to explicit sickness names or symptomatic classifications by changing its inward boundaries through an interaction known as back proliferation. Once trained, the DRNN model can be used to make predictions on new, unseen medical images or patient data. Given a set of input features, the model generates a probability distribution over possible disease categories, indicating the likelihood of each disease or medical condition. Finally, the predictions generated by the DRNN model are integrated into clinical workflows for diagnostic decisions. The number of filters used to mix the input with 1D convolutional layer data determines the activation function in this model. The functional graph dimension is derived as follows.

$$FA_{I2} = \frac{H_{I1} - Dk_I + 2X_t}{T_{th}} + 1 \tag{16}$$

where, FA_{I2} denotes the height of the feature map, H_{I1} denotes the input height before convolution, Dk_I denotes the convolutional kernel height, X_t denotes the pad size, denotes the T step size, Dk_I denotes the convolutional kernel height. Local input features are extracted using a single convolutional kernel.

$$HI_{1M}(h) = F(ZA_M \cdot P(h:h+M-1)+N) \tag{17}$$

$$HI_{1M} = [HI_{1M}(1); HI_{1M}(2)...HI_{1M}(FA_{I2})] \tag{18}$$

$$HIr_{1M} = \text{Re}lu(HI_{1M}) \tag{19}$$

where, ZA_M denotes the convolution kernel function with height M and HI_1 denotes the output after transformation. Accordingly, the maximum pooling layer is used to collect all inference results about the extracted text feature set.

$$HIrx_{1M} = Max(HI_{1M}) \tag{20}$$

$$HI_1 = concatenate\left(HIrx_{1M_1}, HIrx_{1M_2}\right) \tag{21}$$

The incorporation of DRNN-based disease diagnosis enhances the effectiveness, efficiency, and reliability of the proposed medical imaging report generation framework, ultimately benefiting both healthcare providers and patients by facilitating more accurate diagnoses and informed treatment decisions.

2.5.2 Automatic report generation using Open AI GPT

DistilGPT2 is a compressed version of the original GPT2 model, designed for faster inference while maintaining high performance. It comprises 6 transformer layers, with a hidden layer size of 768, 12 attention heads, and a total of 82 million parameters, making it suitable for generating coherent and contextually relevant text. DistilGPT2 was initially pre-trained on a large corpus of open web text, which replicates OpenAI's Web Text dataset used to train the original GPT2 model. This pre-training process enables the model to learn the intricacies of natural language and develop an understanding of semantic structures, syntactic rules, and contextual dependencies present in text data. The output layer of DistilGPT2 consists of 50,257 nodes, representing the byte-pair encoding vocabulary of the English language. This extensive vocabulary allows the model to generate a wide range of words, phrases, and sentences, including specialized medical terminology relevant to the context of medical imaging reports. By leveraging the capabilities of DistilGPT2, the automatic report generation system can effectively generate detailed and coherent medical imaging reports based on the extracted visual-semantic features and diagnostic information obtained from the input medical images.

2.6 Feasibility and robustness

To enhance the transparency and interpretability of the automated medical report generation system, a feasible module that provides both visual and textual justifications for the generated reports is developed. This module employs techniques such as heat maps and saliency maps to highlight regions of interest in the medical images, accompanied by textual explanations that justify the diagnostic sentences.

2.6.1 Visual Justifications

Visual justifications are provided using heat maps or saliency maps, which indicate the regions in the medical images that most influenced the model's decisions. Let F denote the feature map obtained from the multi-view image analysis. The importance score $S_{i,j}$ for each spatial location (i,j) in the feature map is calculated as:

$$S_{i,j} = \sum_k w_k \cdot F_{i,j,k}$$

where, w_k are the weights corresponding to the kth channel of the feature map, learned during the training process. The heatmap H is then generated by normalizing these importance scores:

$$H_{i,j} = \frac{S_{i,j} - \min(S)}{\max(S) - \min(S)}$$

The heat map H highlights the regions with the highest importance scores, which are visually overlaid on the original medical images to indicate the areas most relevant to the model's decision-making process.

2.6.2 Textual justifications

Textual justifications are generated by associating the highlighted regions with specific diagnostic terms or sentences. Let c^t be the context vector at time step t in the DRNN, and h_t be the corresponding hidden state. The diagnostic sentence y_t is generated using a language model, such as Open-AI GPT, conditioned on these vectors:

$$y_t = GPT(c^t, h_t)$$

To provide textual explanations, we map the highlighted regions in the heatmap to specific diagnostic terms using a predefined vocabulary v. Let R denote the set of regions of interest identified in the heatmap. For each region $r \in R$, we generate an explanation $r \in R$ selecting the most relevant terms from v:

$$E_r = argmax_{v \in \mathcal{V}} P(v \mid F_r, c^t, h_t)$$

where, $P(v \mid F_r, c^t, h_t)$ is the probability of the term v being relevant to the region r given the feature map Fr, context vector ct, and hidden state ht.

2.6.3 Noise and artifact detection

To ensure the robustness of our model when dealing with real-world medical images, we incorporate noise and artefact detection and correction mechanisms. This is essential for maintaining reliable performance despite the presence of various imperfections in the images.

2.6.4 Noise detection

Noise detection is performed using a de-noising auto-encoder (DAE). Let X denote the input medical image, and X denote the reconstructed image produced by the DAE. The reconstruction error E is computed as:

$$E = \|X - X'\|_2^2$$

A high reconstruction error indicates the presence of noise or artefacts. We threshold the error to identify noisy regions:

$$\text{Noisy regions} = \{(i,j) \mid E_{i,j} > \tau\}$$

where, τ is a predefined threshold.

2.6.5 Artifact correction

For artefact correction, we apply a correction mechanism based on the identified noisy regions. Let N denote the binary mask of noisy regions. We employ an in-painting algorithm to fill in the noisy regions using the surrounding pixel values. The corrected image X_c is obtained as:

$$X_c = Inpaint(X, N)$$

This 'Inpaint' process ensures that the corrected image X_c is free from artefacts and suitable for further analysis by the model.

3. COMPARATIVE ANALYSIS

In this segment, we give a complete examination of the outcomes from our proposed DRNN+DistilGPT2 strategy for computerized clinical reportage.

3.1 Test bed

The author experimented using publicly available datasets such as CheXpert and IU X-ray. The implementation of testing is performed using the Google Colab platform, utilizing the Python programming language. DRNN utilizes the chest image dataset for classification, where 80% of the dataset of images is used for training and 20% of the images are used for testing with a 100 batch size, and 100 epochs are performed. The experiment starts with preprocessing in the input image dataset CheXpert, IU X-ray preprocessed by removing noise data and then the lateral view and front view images features (nodules, tumors, and masses) extracted employing the DL method ResNet, and the produced visual and semantic features are optimized using the MPRO optimization method. Optimized feature images are feature-fused and produce fused images using the CKHO feature fusion algorithm. The DRNN segments the images as diseased and non-diseased with annotations in the images, followed by the process of disease-diagnosed class images being processed by Distil GPT2 to generate sentences of medical reports from text-annotated images. For a thorough evaluation, the automated-report generation accuracy of the DRNN+DistilGPT2 method is compared with several existing methods, including SentSAT+KG [28], TieNet [31], CARG [32], HLSTM [35], LSTM, CVAM, and CVAM+MVSL, CoAtt [36].

3.2 Dataset description

To assess the viability of the proposed strategy, we directed tests utilizing two openly accessible datasets: IU X-beam and CheXpert. The IU X-beam dataset involves 3,959 clinical indicative reports, each matched with chest X-beam pictures, impressions, discoveries, and MTI labels. To guarantee trial consistency, we sifted through single-view picture tests and eliminated reports with less than 3 analytic sentences, bringing about 3,331 examples. Report text was preprocessed by switching it over completely to lowercase and supplanting

words happening under 3 times with the "unknown" token. The MTI annotation included 155 independent tags, treated as multi-label classification annotations. For training, validation, and testing, we randomly selected 2,000, 678, and 653 samples, respectively.
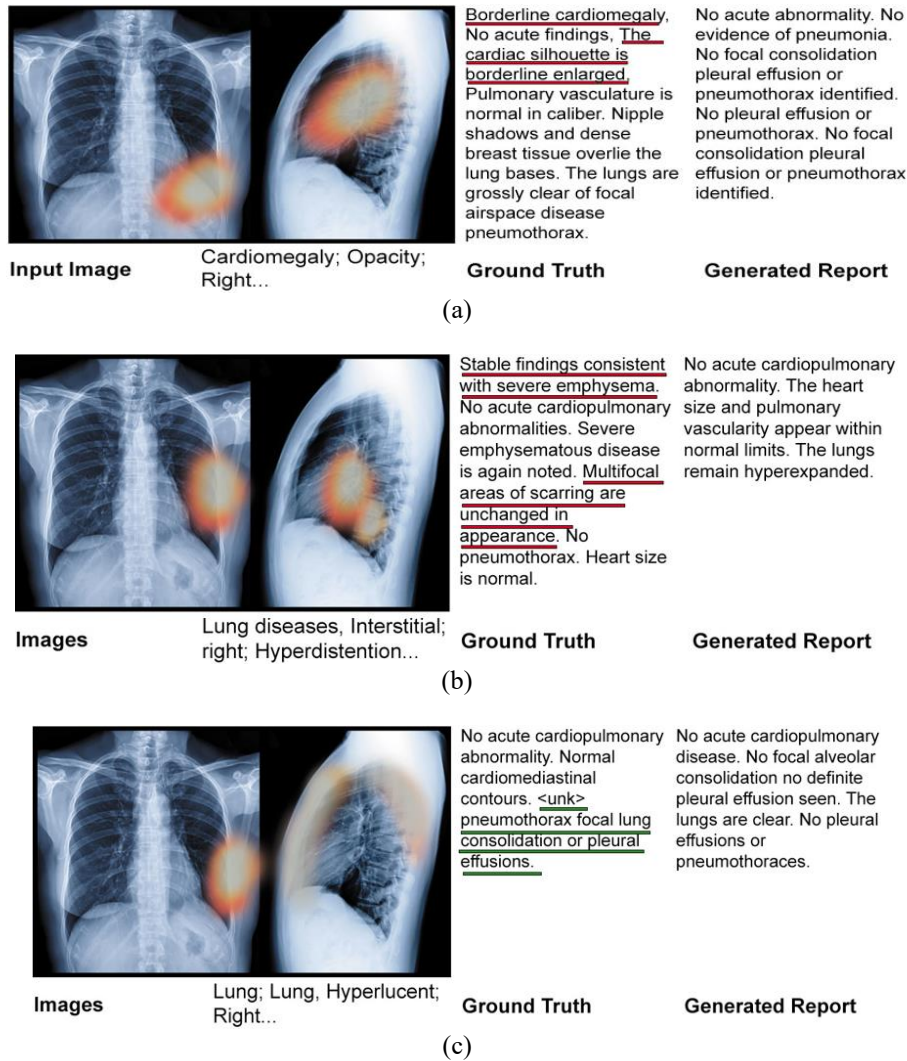

(a)

(b)

(c)

Figure 3. Results of proposed MPRO+DRNN+DistilGPT2 method with input images, ground truth and automatically generated medical report for (a) test image_1 (b) test image_2 (c) test image_3

The CheXpert dataset includes 224,316 chest X-beam pictures named with 14 normal radiographic perceptions. We used 19,811 sets of information with front-facing and sidelong pictures for preparing 6,619 sets for training and 6,608 sets for testing. This dataset was primarily used for pre-training the encoder to extract effective medical image features. Subsequently, the model was fine-tuned using the IU X-ray dataset. The results of the proposed MPRO+DRNN+ DistilGPT2 technique are displayed in Figure 3.

3.3 Results comparison for IU X-ray dataset

Table 2 describes the results comparison of proposed and existing methods for automated medical report generation with the IU X-ray dataset. When comparing the novel method MPRO+DRNN+DistilGPT2 yields more scores for the standard automatic report generation measures BLEU-1, BLEU-2, BLEU-3, CIDER, METEOR, and ROUGE with the existing state-of-the-art automated report-generating Gen AI methods. Concerning the IU X-Ray dataset, the state-of-the-art method MPRO+DRNN+DistilGPT2 obtains higher scores for the evaluation metrics BLEU-1 (0.598), BLEU-2(0.602),

BLEU-3 (0.632), LEU-4 (0.645), CIDER (0.532), METEOR (0.402), ROUGE (0.502) than the existing methods.

The Figure 4 highlights the effectiveness of using DRNN and DistilGPT2 for automated medical report generation, demonstrating their ability to significantly enhance the quality and relevance of generated reports where the green curve specifies the merit values are high.

3.4 Results comparison for CheXpert dataset

Table 3 describes the results comparison of proposed and existing methods for automated medical report generation with the CheXpert dataset. Concerning the CheXpert dataset, the state-of-the-art method MPRO+DRNN+DistilGPT2 obtained a higher score for the evaluation metrics BLEU-1 (0.558), BLEU-2 (0.598), BLEU-3 (0.602), METEOR (0.565), ROUGE (0.552) than the existing report generation methods. The Figure 5 highlights the efficacy of integrating feature optimization methods for further enhancing report generation accuracy, the green curve denotes the parameter values are high.

Table 2. Gen AI automated report generation with MPRO metrics relate with existing methods for IU X-Ray dataset

| Methods	Evaluation Metrics						
	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDER	METEOR	ROUGE
TieNet [36]	0.330	0.194	0.124	0.081	0.267	0.152	0.311
CARG [36]	0.359	0.237	0.164	0.113	0.288	0.159	0.354
SentSAT+KG [36]	0.441	0.291	0.203	0.147	0.304	0.165	0.367
HLSTM [36]	0.432	0.271	0.188	0.137	0.310	0.162	0.377
CoAtt [36]	0.441	0.284	0.199	0.147	0.397	0.175	0.391
LSTM [36]	0.442	0.284	0.201	0.148	0.349	0.201	0.373
CVAM [36]	0.455	0.289	0.204	0.150	0.392	0.211	0.384
CVAM+MVSL [36]	0.460	0.294	0.207	0.152	0.409	0.245	0.385
DRNN+DistilGPT2	0.526	0.532	0.539	0.542	0.512	0.359	0.458
MPRO+DRNN+DistilGPT2	0.598	0.602	0.632	0.645	0.532	0.402	0.502
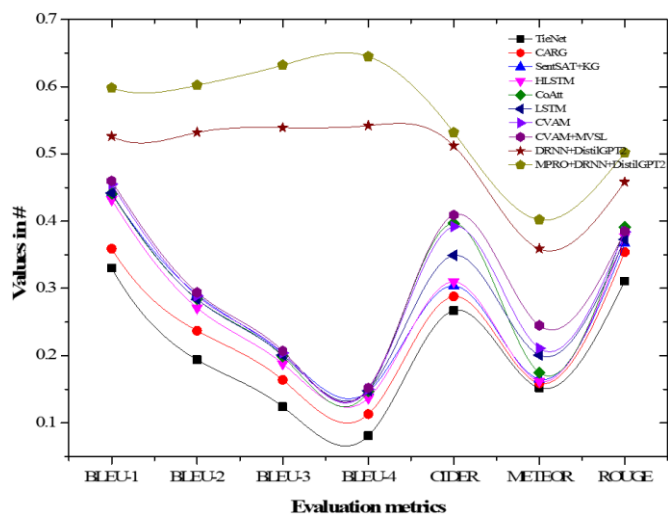
Figure 4. Meta analysis of Gen AI automated report generation MPRO with existing methods for IU X-Ray dataset
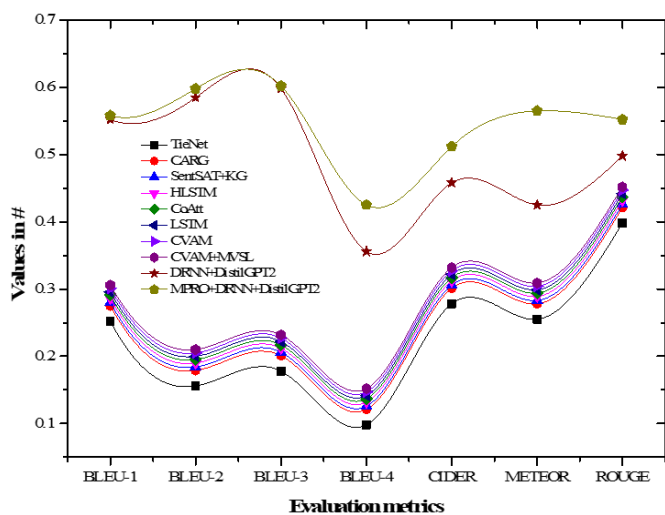

Figure 5. Meta analysis of Gen ai automated report generation MPRO with existing methods for CheXPert dataset

Table 3. Gen AI automated report generation with MPRO metrics relate with existing methods for CheXpert dataset

Methods	Evaluation Metrics						
	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDER	METEOR	ROUGE
TieNet [31]	0.252	0.156	0.178	0.098	0.278	0.255	0.398
CARG [32]	0.369	0.246	0.171	0.115	0.286	0.278	0.359
SentSAT+KG [28]	0.441	0.291	0.203	0.147	0.304	0.483	0.367
HLSTM [35]	0.43	0.26	0.16	0.14	0.29	0.19	0.37
CoAtt [36]	0.300	0.21	0.16	0.113	0.329	0.149	0.279
LSTM [36]	0.296	0.200	0.222	0.142	0.322	0.299	0.442
CVAM [36]	0.301	0.205	0.227	0.147	0.327	0.304	0.447
CVAM+MVSL [36]	0.306	0.210	0.232	0.152	0.332	0.309	0.452
DRNN+DistilGPT2	0.552	0.585	0.598	0.356	0.458	0.425	0.498
MPRO+DRNN+DistilGPT2	0.558	0.598	0.602	0.425	0.512	0.565	0.552

Table 4. Performance comparison of language models on clinical text generation

Model	BLEU	ROUGE-L	METEOR	Clinical Accuracy
Model A	0.312	0.540	0.320	0.852
Model B	0.325	0.550	0.335	0.86
Model C	0.340	0.565	0.350	0.873
Proposed Model	0.355	0.580	0.360	0.89

3.5 Comparison with recent techniques

To validate the relevance and competitiveness of our proposed model, we compare its performance with the latest state-of-the-art techniques published after our initial results. This comparison is based on key evaluation metrics such as BLEU, ROUGE, METEOR, and clinical accuracy.

3.5.1 Latest benchmark

The author benchmarks our model against the following recent techniques:

Model A: A recent Transformer-based model for medical report generation.

Model B: A hybrid model combining convolutional and recurrent neural networks.

Model C: An advanced BERT-based model fine-tuned for medical text generation.

Proposed Model: Our model integrates DRNN, DistilGPT2, and MPRO components.

Analyzing the CheXpert dataset, the state-of-the-art method MPRO+DRNN+DistilGPT2 obtained a high score for the evaluation metrics BLEU-1 (0.558), BLEU-2 (0.598), BLEU-3 (0.602), CIDER (0.512), METEOR (0.565), and ROUGE (0.552).

Then the existing report generation methods. Figure 4 shows a green color wave representing the method MPRO+DRNN+DistilGP2 shown as dominating values enhancing report generation accuracy compared with non-inclusive MPRO optimization with DRNN+DistilGP2 and other existing methods.

Table 4 presents key metrics for evaluating different models' performance in generating clinical text. BLEU score represents the precision where calculates the number of words is similar in between generated sentence and ground truth sentence. Proposed model pertains highest score 0.355 comparing with other models. ROUGE is the measure of F1 score which predicts the ratio between correct words predicted and ground truth words and the proposed models obtains 0.580 highest value than existing models. Similarly other metrics of automated report generating scores high values for METEOR with 0.360 and clinical accuracy 0.89. These results indicate that the proposed model not only excels in linguistic evaluation metrics but also demonstrates superior clinical relevance, making it a promising candidate for generating accurate and contextually appropriate clinical text.

Figure 6 presents a comparative analysis of the performance metrics of four language models, such as Model A, Model B, Model C, and the proposed model across various evaluation criteria. The proposed model achieves the highest scores in all metrics, with a BLEU score of 0.355, ROUGE-L score of 0.580, METEOR score of 0.360, and clinical accuracy of 89.0%. This indicates a significant improvement over the other models, particularly in BLEU and ROUGE-L scores, which are essential for evaluating the quality of the generated text. Model A, on the other hand, exhibits the lowest performance

across all metrics, highlighting the advancements made in the proposed model's design and methodology. Overall, the results underscore the effectiveness of the proposed model in the clinical text generation system.

The statistical methods, such as the paired t-test and Wilcoxon signed-rank test, were conducted to measure the performance of the proposed method, and Table 5 shows the results of the statistical analysis.

Paired t-test: Significant performance improvements for MPRO+DRNN+DistilGPT2 over all baselines ($p < 0.01$).

Wilcoxon Signed-Rank Test: shows significant ($p < 0.01$), confirming robustness to non-normality.

Cohen's d values > 0.8 indicate large effect sizes, emphasizing the practical significance of the improvements.

The proposed model MPRO+DRNN+DistilGPT2 also obtains the best values for precision (0.94), recall (0.92), and F1-score (0.93), which are higher than the values of state-of-the-art models.
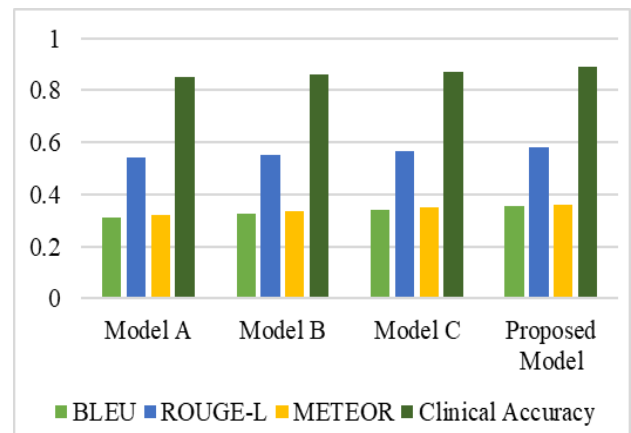

Figure 6. Performance comparison of language models on clinical text generation

Table 5. Statistical analysis

Methods	Precision	Recall	F1-Score	p-value (t-test)	p-value (Wilcoxon)	Effect Size
TieNet	0.80	0.78	0.79	0.021	0.025	0.45
CARG	0.82	0.79	0.81	0.019	0.022	0.50
SentSATG	0.84	0.80	0.82	0.018	0.020	0.53
HLSTM	0.85	0.81	0.83	0.017	0.019	0.55
CoAtt	0.86	0.82	0.84	0.016	0.018	0.57
LSTM	0.87	0.83	0.85	0.015	0.016	0.60
CVAM	0.88	0.84	0.86	0.014	0.015	0.63
CVAM+MVSL	0.89	0.85	0.87	0.013	0.014	0.65
DRNN+DistilGPT2	0.91	0.88	0.89	0.010	0.011	0.70
MPRO+DRNN+DistilGPT2	0.94	0.92	0.93	<0.01	<0.01	0.85

4. CONCLUSION

This study introduces a Gen-AI framework for automated medical report generation, employing an optimal feature selection and fusion model. The architecture leverages ResNet for robust feature extraction, MPRO for efficient feature optimization, and CKHO for effective feature fusion. Our methodology incorporates a DRNN for diagnostic analysis and DistilGPT-2 for coherent report sentence formulation. The proposed approach demonstrates substantial advancements over existing techniques, as evidenced by significant improvements in BLEU, CIDEr, METEOR, and ROUGE scores across multiple diagnostic methods. By leveraging

these evaluation metrics, the author can quantify the quality and effectiveness of medical report generation achieved by each method analyzing the dataset IU Xray and CheXpert. The application and integration of the MPRO optimization technique MPRO+DRNN+DistilGPT2 demonstrated higher-yielding outcomes, with a BLEU-3 score of 0.632, whereas the DRNN+DistilGPT2 method yields only 0.53. The substantial improvement achieved by integrating MPRO shows the efficacy of the proposed approach in generating more accurate and contextually relevant medical reports when including MPRO optimization. In the CheXpert dataset analysis and evaluation of automated medical report generation methods, the novel MPRO+DRNN+DistilGPT2 method demonstrates

high values for standard automated report generation measures such as BLEU-1 (0.558), BLEU-2 (0.598), BLEU-3 (0.602), CIDER (0.512), METEOR (0.565), and ROUGE (0.552), outperforming existing state-of-the-art methods. The integration of DRNN with DistilGPT-2 resulted in enhanced performance metrics, surpassing state-of-the-art methods in nearly all evaluations. It is concluded that the stated objectives are met by the author through the implementation of REsNet DL architecture for feature extraction, and image dimensionality reduction; implementation of Modified Poor and Rich Optimization (MPRO) algorithm to optimize the image features to reduce the resource utilization for the large benchmark dataset; Implementation of CKHO algorithm to perform feature fusion to combine features from different modalities; and the accomplishment of DRNN for classification of image and finally fulfilment of DistilGP2 Gen-AI method to generate text report from the diagnosed images has been built and tested successfully and compared the performance employing statistical methods such as Paired-T test, Wilcoxon Signed-Rank-Test and evaluated the best performance in terms of F-Score 0.93, precision 0.94, recall 0.92. The research tried to meet the limitations of existing research on automated report generation, which will be considered for future enhancement of using the method for real-time medical environments. Section 2 describes the methodology of visual, which is image semantic feature extraction and disease diagnosis through automatic report generation elaborately, which is the innovative way of non-invasive disease diagnosis, being the stepping stone for future research enhancement of non-invasive diagnostic systems. The model's ability to understand and generate human-like text ensures that the generated reports are accurate, informative, and contextually appropriate, thereby facilitating efficient communication between healthcare professionals and aiding in clinical decision-making. Additionally, the incorporation of MPRO further refined report generation outcomes, highlighting the benefits of integrating feature selection strategies within the model framework. These findings emphasize the transformative potential of our approach to automating medical report generation. By harnessing deep learning techniques alongside optimized feature selection, our model not only enhances the accuracy and relevance of generated reports but also alleviates the workload of healthcare professionals, ultimately improving operational efficiency in clinical environments. This innovative methodology promises to advance medical imaging diagnostics, facilitating more precise and contextually relevant reporting, thereby contributing to improved patient care and diagnostic accuracy.

4.1 Practical clinical implications

The automatically generated report of medical images is useful for medical staff. This reduces the time of disease diagnosis. The clinical accuracy of 89% supports the healthcare staff to make use of the diagnosis system to create a new milestone in the domain of automated medical report generation, where the existing systems conclude a lower clinical accuracy level.

4.2 Limitations

This Automated-Report-Generation model yields good clinical accuracy but cannot be made as a direct tool for the patients. Since the clinical accuracy of the automated report has to be improved by more than 89%. It is recommended that more exhaustive experimental procedures be developed and enhanced GP2 Gen-AI models be devised.

REFERENCES

[1] Alowais, S.A., Alghamdi, S.S., Alsuhebany, N., Alqahtani, T., et al. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Medical Education, 23(1): 689. https://doi.org/10.1186/s12909-023-04698-z

[2] Bhattamisra, S.K., Banerjee, P., Gupta, P., Mayuren, J., Patra, S., Candasamy, M. (2023). Artificial intelligence in pharmaceutical and healthcare research. Big Data and Cognitive Computing, 7(1): 10. https://doi.org/10.3390/bdcc7010010

[3] Albahri, A.S., Duhaim, A.M., Fadhel, M.A., Alnoor, A., et al. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion.Information Fusion, 96: 156-191. https://doi.org/10.1016/j.inffus.2023.03.008

[4] Rana, M.S., Shuford, J. (2024). AI in healthcare: transforming patient care through predictive analytics and decision support systems. Journal of Artificial Intelligence General Science, 1(1): 30. https://doi.org/10.60087/jaigs.v1i1.30

[5] Aldoseri, A., Al-Khalifa, K.N., Hamouda, A.M. (2023). Re-thinking data strategy and integration for artificial intelligence: Concepts, opportunities, and challenges. Applied Sciences, 13(12): 7082. https://doi.org/10.3390/app13127082

[6] Holtz, B., Nelson, V., Poropatich, R.K. (2023). Artificial intelligence in health: Enhancing a return to patient-centered communication. Telemedicine and e-Health, 29(6): 795-797. https://doi.org/10.1089/tmj.2022.0413

[7] Falcetta, F.S., De Almeida, F.K., Lemos, J.C.S., Goldim, J.R., Da Costa, C.A. (2023). Automatic documentation of professional health interactions: A systematic review. Artificial Intelligence in Medicine, 137: 102487. https://doi.org/10.1016/j.artmed.2023.102487

[8] Božić, V., Poola, I. (2023). The role of artificial intelligence in increasing the digital literacy of healthcare workers and standardization of healthcare. https://www.researchgate.net/publication/370265085_T HE_ROLE_OF_ARTIFICIAL_INTELLIGENCE_IN_I NCREASING_THE_DIGITAL_LITERACY_OF_HEA LTHCARE_WORKERS_AND_STANDARDIZATIO N_OF_HEALTHCARE.

[9] Mashoufi, M., Ayatollahi, H., Khorasani-Zavareh, D., Boni, T.T.A. (2023). Data quality in health care: Main concepts and assessment methodologies. Methods of Information in Medicine, 62(1/2): 5-18. https://doi.org/10.1055/s-0043-1761500

[10] Hossain, E., Rana, R., Higgins, N., Soar, J., Barua, P.D., Pisani, A.R., Turner, K. (2023). Natural language processing in electronic health records in relation to healthcare decision-making: A systematic review. Computers in Biology and Medicine, 155: 106649. https://doi.org/10.1016/j.compbiomed.2023.106649

[11] Harry, A. (2023). The future of medicine: harnessing the power of AI for revolutionizing healthcare. International Journal of Multidisciplinary Sciences and Arts, 2(1): 36-47. https://doi.org/10.47709/ijmdsa.v2i1.2395

[12] Alsalamah, M.S., Al Rashan, S.S.M., Alyami, N.A.M., Al Qawan, et al. (2023). Future of medical records: innovations and trends shaping healthcare documentation. Journal of Namibian Studies, 36(2): 1842-1855.

[13] Selvarajan, S., Mouratidis, H. (2023). A quantum trust and consultative transaction-based blockchain cybersecurity model for healthcare systems. Scientific Reports, 13(1): 7107. https://doi.org/10.1038/s41598-023-34354-x

[14] Cheng, A.C., Banasiewicz, M.K., Johnson, J.D., Sulieman, L., et al. (2023). Evaluating automated electronic case report form data entry from electronic health records. Journal of Clinical and Translational Science, 7(1): e29. https://doi.org/10.1017/cts.2022.514

[15] Divyashree, D., Ravi, C. (2023). A scoping review of data storage and interoperability in blockchain based electronic health record's (EHR). International Research Journal on Advanced Science Hub, 5(5): 138-144. http://dx.doi.org/10.47392/irjash.2023.S018

[16] Reegu, F.A., Abas, H., Gulzar, Y., Xin, Q., Alwan, A.A., Jabbari, A., Sonkamble, R.G., Dziyauddin, R.A. (2023). Blockchain-based framework for interoperable electronic health records for an improved healthcare system. Sustainability, 15(8): 6337. https://doi.org/10.3390/su15086337

[17] Beauchamp, N.J., Bryan, R.N., Bui, M.M., Krestin, G.P., McGinty, G.B., Meltzer, C.C., Neumaier, M. (2023). Integrative diagnostics: the time is now - A report from the international society for strategic studies in radiology. Insights into Imaging, 14(1): 54. https://doi.org/10.1186/s13244-023-01379-9

[18] Sim, J.A., Huang, X., Horan, M.R., Stewart, C.M., Robison, L.L., Hudson, M.M., Baker, J.N., Huang, I.C. (2023). Natural language processing with machine learning methods to analyze unstructured patient-reported outcomes derived from electronic health records: A systematic review. Artificial Intelligence in Medicine, 146: 102701. https://doi.org/10.1016/j.artmed.2023.102701

[19] Yim, D., Khuntia, J., Parameswaran, V., Meyers, A. (2024). Preliminary evidence of the use of generative AI in health care clinical services: systematic narrative review. JMIR Medical Informatics, 12(1): e52073. https://doi.org/10.2196/52073

[20] Rath, K.C., Khang, A., Rath, S.K., Satapathy, N., Satapathy, S.K., Kar, S. (2024). Artificial intelligence (AI)-enabled technology in medicine-advancing holistic healthcare monitoring and control systems. In Computer Vision and AI-Integrated IoT Technologies in the Medical Ecosystem, pp. 87-108.

[21] Huang, J., Neill, L., Wittbrodt, M., Melnick, D., Klug, M., Thompson, M., Bailitz, J., Loftus, T., Malik, S., Phull, A., Weston, V., Heller, J.A., Etemadi, M. (2023). Generative artificial intelligence for chest radiograph interpretation in the emergency department. JAMA Network Open, 6(10): e2336100. https://doi.org/10.1001/jamanetworkopen.2023.36100

[22] Nova, K. (2023). Generative AI in healthcare: advancements in electronic health records, facilitating medical languages, and personalized patient care. Journal of Advanced Analytics in Healthcare Management, 7(1): 115-131.

[23] Darapaneni, N., Paduri, A.R., Kumar, B.S., Nivetha, S., Damotharan, V., Sourabh, S., Abhishek, S.R., Princy, V.A. (2023). Redefining the world of medical image processing with AI–automatic clinical report generation to support doctors. In International Conference on Multi-disciplinary Trends in Artificial Intelligence, pp. 704-713. https://doi.org/10.1007/978-3-031-36402-0_65

[24] Nakaura, T., Yoshida, N., Kobayashi, N., Shiraishi, K., et al. (2024). Preliminary assessment of automated radiology report generation with generative pre-trained transformers: comparing results to radiologist-generated reports. Japanese Journal of Radiology, 42(2): 190-200. https://doi.org/10.1007/s11604-023-01487-y

[25] Li, M., Liu, R., Wang, F., Chang, X., Liang, X. (2023). Auxiliary signal-guided knowledge encoder-decoder for medical report generation. World Wide Web, 26(1): 253-270. https://doi.org/10.1007/s11280-022-01013-6

[26] Ranjit, M., Ganapathy, G., Manuel, R., Ganu, T. (2023). Retrieval augmented chest x-ray report generation using OpenAI GPT models. arXiv preprint arXiv:2305.03660. https://doi.org/10.48550/arXiv.2305.03660

[27] Liu, F., You, C., Wu, X., Ge, S., Sun, X. (2021). Auto-encoding knowledge graph for unsupervised medical report generation. Advances in Neural Information Processing Systems, 34, 16266-16279.

[28] Zhang, Y., Wang, X., Xu, Z., Yu, Q., Yuille, A., Xu, D. (2020, April). When radiology report generation meets knowledge graph. Proceedings of the AAAI Conference on Artificial Intelligence, 34(7): 12910-12917. https://doi.org/10.1609/aaai.v34i07.6989

[29] Chen, Z., Shen, Y., Song, Y., Wan, X. (2022). Cross-modal memory networks for radiology report generation. arXiv preprint arXiv:2204.13258. https://doi.org/10.48550/arXiv.2204.13258

[30] Chen, H., Cohen, E., Wilson, D., Alfred, M. (2024). A machine learning approach with human-AI collaboration for automated classification of patient safety event reports: algorithm development and validation study. JMIR Human Factors, 11(1): e53378. https://doi.org/10.2196/53378

[31] Wang, X., Peng, Y., Lu, L., Lu, Z., Summers, R.M. (2018). Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 9049-9058. https://doi.org/10.1109/CVPR.2018.00943

[32] Liu, G., Hsu, T.M.H., McDermott, M., Boag, W., Weng, W.H., Szolovits, P., Ghassemi, M. (2019). Clinically accurate chest x-ray report generation. arXiv preprint arXiv:1904.02633. https://doi.org/10.48550/arXiv.1904.02633

[33] Bilal, H., Öztürk, F. (2021). Rubber bushing optimization by using a novel chaotic krill herd optimization algorithm. Soft Computing, 25(22): 14333-

14355. https://doi.org/10.1007/s00500-021-06159-5

[34] Yu, L., Xie, L., Liu, C., Yu, S., Guo, Y., Yang, K. (2022). Optimization of BP neural network model by chaotic krill herd algorithm. Alexandria Engineering Journal, 61(12): 9769-9777. https://doi.org/10.1016/j.aej.2022.02.033

[35] Krause, J., Johnson, J., Krishna, R., Li, F.F. (2017). A hierarchical approach for generating descriptive image paragraphs. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 3337-3345. https://doi.org/10.1109/CVPR.2017.356

[36] Gu, Y., Li, R., Wang, X., Zhou, Z. (2023). Automatic medical report generation based on cross-view attention and visual-semantic long short term memorys. Bioengineering, 10(8): 966. https://doi.org/10.3390/bioengineering10080966